Modeling of Traffic Engineering Queues with QoS Differentiation

Oleksandr Lemeshko, Oleksandra Yeremenko and Roman Savchenko
V.V. Popovskyy Department of Infocommunication Engineering, Kharkiv National University of Radio Electronics,
Nauky Avenue 14, 61166 Kharkiv, Ukraine
oleksandr.lemeshko@nure.ua, oleksandra.yeremenko@nure.ua, roman.savchenko1@nure.ua

Keywords: Quality of Service, Traffic Engineering, Queue, Classification, Optimization, Bandwidth Allocation, Utilization.

Abstract: The paper analyzes the Class-Based Traffic Engineering Queue (CB-TEQ) model, which provides organization and management of class queues on the router interface, supporting differentiated and balanced packet service. It is demonstrated that the advantage of CB-TEQ lies in the implementation of an optimal interface bandwidth allocation between class queues based on a linear programming problem. At the same time, a shortcoming of the model has been identified, related to the need for heuristic selection of the normalization coefficient. To address this issue, an Enhanced Class-Based Traffic Engineering Queue (ECB-TEQ) model is proposed, in which the normalization coefficient is considered as a control variable. This enabled the calculation of the optimal value of the normalization coefficient by solving a nonlinear optimization problem with a modified criterion and a system of constraints. The results of experimental studies confirmed the efficiency of ECB-TEQ, in particular its ability to provide the maximum level of differentiation when servicing packets in different class queues.

1 INTRODUCTION

Quality of service (QoS) is a primary concern in modern communication networks [1], [2]. Complexity of QoS provision is due to several factors, including geographical distribution, environmental heterogeneity, network state dynamics, growth in the number of users and services, limited resources, and the need to support various QoS models [3]-[5]. Ensuring QoS requires the coordinated operation of all OSI layers protocols, particularly at the network layer, where scheduling and resource allocation mechanisms play a crucial role [4]-[7]. They determine the bandwidth allocation between packet queues on the network equipment (routers and switches) and form the basis of the IntServ and DiffServ models [8], [9]. These mechanisms must support multiple queues, provide service guarantees, ensure consistent packet processing, and be easy to implement. In practice, routers use several mechanisms simultaneously, and their effectiveness is determined by the mathematical models and methods used.

Modern queue management mechanisms are in demand in various network environments, as confirmed by studies [10]-[18]. The tasks of optimal resource allocation and QoS assurance are relevant in general-purpose networks [10], [19]-[21], industrial IoT solutions [11], specific environments [12], [13], [15], [17], and hierarchical solutions [16], [19], [20]. This highlights the importance of adaptive queue management methods for a wide range of scenarios. The purpose of this work is to optimize the queue management process to ensure differentiated packet servicing at the router interface. To this end, the existing solution will be considered, and its enhancement will be proposed, followed by proving its effectiveness on a set of computational examples.

2 BASIC CLASS-BASED TRAFFIC ENGINEERING QUEUE MODEL

Let the Class-Based Traffic Engineering Queue (CB-TEQ) model, introduced in [21], serve as the basis for further enhancement. Within the CB-TEQ model, the resource allocation problem is considered. Pre-aggregated packet flows are directed into class queues based on the proximity of their class values [21]. The following notations are used:

- B – router interface bandwidth (bits per second);
- N – number of class queues on the interface;
- b_i – interface bandwidth allocated for the ith class queue;
- r_i – average intensity of the aggregated packet flow arriving at the ith queue (bits per second), $i = \overline{1,N}$.

The variables b_i are subject to constraint:

$$b_i \geq 0, \tag{1}$$

$$\sum_{i=1}^{N} b_i = B. \tag{2}$$

Additionally, nonlinear overload prevention conditions must be met related to the optimal resource allocation and balanced bandwidth utilization among queues under the Traffic Engineering Queues [21]:

$$h_i^{\alpha} r_i \leq \alpha b_i \ \ (i = \overline{1,N}), \tag{3}$$

where α is the control variable for the upper dynamically regulated bound on queue utilization, following the condition:

$$0 < \alpha \leq 1. \tag{4}$$

The h_i^{α} is the class coefficient for balanced interface bandwidth allocation for the ith queue:

$$h_i^{\alpha} = 1 + i/(N \cdot D) \ \ (i = \overline{1,N}\,). \tag{5}$$

In addition, D is the normalization coefficient that determines both the influence of the queue class on h_i^{α} and the degree of bandwidth balancing across queues. As shown in [19], [20], the minimum allowable value D_{min} ensures the highest level of service differentiation among packets in different class queues. A higher queue class corresponds to a larger h_i^{α}, which in turn leads to lower utilization ρ_i for the same boundary value α. Within the framework of model (1)-(5), the utilization coefficient is given by:

$$\rho_i = \frac{r_i}{b_i} \ \ (\, i = \overline{1,N}). \tag{6}$$

An increase in the normalization coefficient D reduces the influence of the queue class on allocated bandwidth. Using expressions (3)-(6), the model provides differentiated allocation of router interface bandwidth among class queues. Maximum differentiation is achieved when D is set to its minimum permissible value [19], [20]. Condition (3) is nonlinear because of the product of the control variables b_i and α, and bandwidth balancing is ensured by α. As shown in [21], (3) can be transformed into a linear form:

$$\beta h_i^{\alpha} r_i \leq b_i \ \ (i = \overline{1,N}), \tag{7}$$

where β is the control variable, inversely proportional to the upper bound of queue utilization α, i.e.,

$$\beta = \frac{1}{\alpha}. \tag{8}$$

The variable β is subject to the constraints:

$$\beta > 1. \tag{9}$$

The queue management optimality criterion is given by:

$$\beta \to max. \tag{10}$$

Criterion (10) minimizes the upper bound of queue utilization, with weights determined by the class values in (3) and (5). Thus, the CB-TEQ solution (1)-(10) is a linear programming problem with criterion (10) and constraints (1), (2), (7), and (9).

3 ENHANCED CLASS-BASED TRAFFIC ENGINEERING QUEUE MODEL

A limitation of the CB-TEQ model is that it requires prior configuration, namely the heuristic selection of the normalization coefficient D. This coefficient must be set so that, first, the minimization problem (10), subject to constraints (1), (2), (7), and (9), admits a solution, and second, that a sufficient degree of differentiation is achieved in the service of packets across different queues. To address this, the present work proposes an enhancement of the CB-TEQ model in which the optimal (minimum) value of D (5) is computed jointly with the control variables β and b_i. This approach guarantees differentiation in packet servicing across queues. Then, considering (5), condition (7) is reformulated as follows:

$$\beta r_i (1 + i/(N \cdot D)) \leq b_i \ \ \left(i = \overline{1,N}\right). \tag{11}$$

Analysis of expression (11) and the results obtained in works [19]-[21] show that, with an increase in the coefficient D, there will be no differentiation in the servicing of packets across different queues, since their utilization coefficients (6) will be equal. To increase service differentiation, the normalization coefficient must be reduced to a specific, but previously unknown, value. Thus, with a very small D, it will be impossible to satisfy the requirements of condition (11) even with a small load on the queue. Therefore, there is a need for an optimization formulation of the problem to determine the minimum acceptable value of the normalization coefficient D.

Table 1: Input data variants.

Var. #	Class queue load, r_i								R	B	D_{min}	ρ_{min}	ρ_{max}
	1	2	3	4	5	6	7	8					
1	14	9	9	5	11	8	21	13	90	100	5.413	0.8441	0.9774
2	9	14	9	13	21	8	11	5	90	100	4.825	0.8283	0.9747
3	5	11	8	21	13	9	14	9	90	100	5.3	0.8413	0.9770
4	5	11	8	25	15	12	14	0	90	100	4.95	0.8498	0.9754
5	5	11	8	25	15	12	14	×	90	100	5.657	0.8498	0.9754
6	5	20	8	30	15	12	0	0	90	100	4.2	0.8485	0.9711
7	5	20	8	30	15	12	0	×	90	100	4.8	0.8485	0.9711
8	5	20	8	30	15	12	×	×	90	100	5.6	0.8485	0.9711
9	5	20	10	30	25	0	0	0	90	100	4	0.8649	0.9697
10	5	20	10	30	25	0	0	×	90	100	4.57	0.8649	0.9697
11	5	20	10	30	25	0	×	×	90	100	5.33	0.8649	0.9697
12	5	20	10	30	25	×	×	×	90	100	6.4	0.8649	0.9697
13	12	7	8	5	10	8	19	11	80	100	2.43	0.7086	0.9511
14	5	17	8	26	24	×	×	×	80	100	1.794	0.7416	0.9349
15	10	6	8	5	9	8	15	9	70	100	1.4	0.5833	0.9180
16	5	14	8	21	22	×	×	×	70	100	1.673	0.6259	0.8932
17	14	11	9	6	13	8	21	13	95	100	11.275	0.9185	0.9890
18	7	21	10	32	25	×	×	×	95	100	13.28	0.93	0.9852

In (11), the set of control variables now includes not only β and b_i, but also D. Consequently, expression (11) becomes nonlinear with respect to the specified control variables. The next constraint is imposed on the new control variable:

$$D > 0. \tag{12}$$

With the expansion of the control variable set, the optimality criterion must also be revised. In this work, the following criterion is proposed:

$$(\beta - kD) \rightarrow max, \tag{13}$$

where k is a sufficiently large positive constant introduced to prioritize the selection of the optimal (minimum) value of D. Thus, the Enhanced Class-Based Traffic Engineering Queue (ECB-TEQ) model is formulated as a nonlinear optimization problem with criterion (13) and constraints (1), (2), (9), (11), and (12).

4 NUMERICAL RESEARCH

Several examples of the queue management problem were examined using different input data variants (Table 1) to identify the factors influencing the optimal values of the normalization coefficient D.

The interface bandwidth B was fixed at 100 Mbps. For clarity, the number of class queues on the interface varied between five and eight. The load on individual class queues (r_i) was selected randomly, but the total load on the interface $R = \sum_{i=1}^{N} r_i$ ranged from 70 to 95 Mbps. Table 1 shows the results of the study for 18 input data variants, for each of which D_{min} was determined using the ECB-TEQ model, and the minimum and maximum values of utilization coefficients (6) for class queues were calculated. For variants in which fewer than the maximum allowable number (eight) of queues were organized, a mark × was placed in the corresponding columns of Table 1. If a queue was created but not loaded, i.e., remained empty, then zero is present in the corresponding column.

Analysis of Table 1 shows that the value of D_{min} is most sensitive to the interface utilization. For example, when the total interface load R was 95 Mbps, D_{min} exceeded 10 (input data variants 17 and 18), whereas with a load of 70 Mbps, D_{min} did not exceed 2 (input data variants 15 and 16). Figures 1 and 2 illustrate the dynamics of utilization coefficients across class queues for input data variants 18 and 15, respectively. As shown in Figure 1, as the coefficient D increases, the utilization coefficients of all queues converge, thereby reducing the level of differentiation in packet servicing across queues and effectively negating the purpose of organizing multiple class queues on the interface.

In addition, the number of formally organized versus actually utilized class queues also influenced the value of D_{min}, although to a much lesser extent. To illustrate this, variants 9-12 in Table 1 are examined in detail.

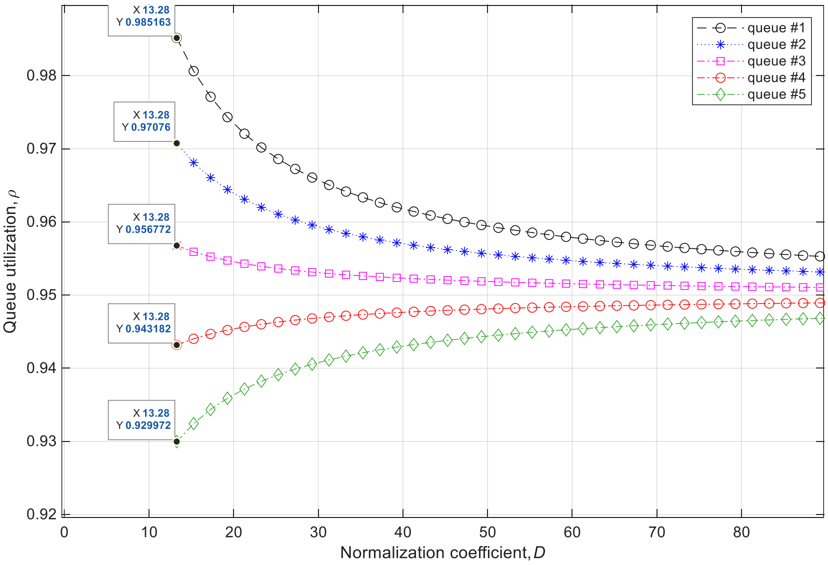

Figure 1: Variation of utilization coefficients across class queues for input data variant 18.

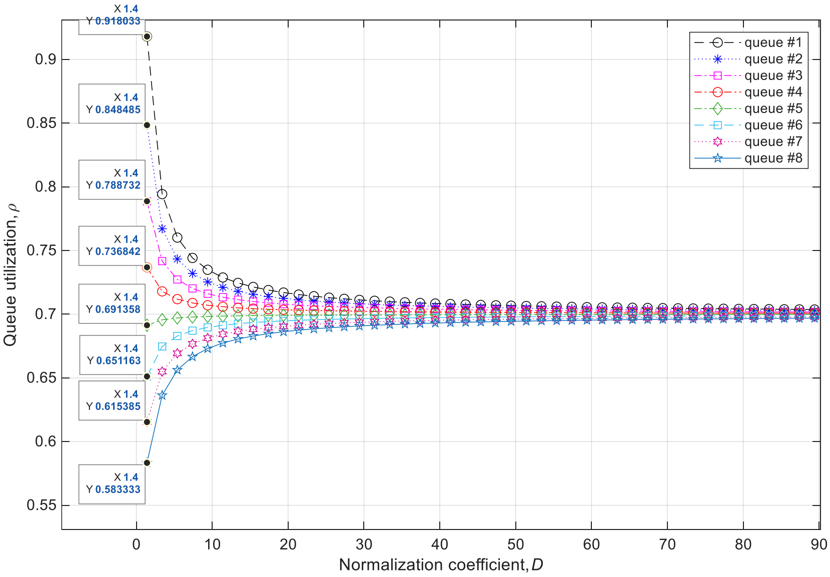

Figure 2: Variation of utilization coefficients across class queues for input data variant 15.

These options share the same traffic structure at the interface, with only queues 1 through 5 being loaded. In variant 12, five class queues were created, whereas in variant 9, eight queues were created; however, queues 6 through 8 remained empty, meaning no packets were assigned to them. From a practical perspective, the ECB-TEQ model produced the same interface bandwidth allocation among the active queues from first to fifth in each case. This resulted in identical values of ρ_{min} and ρ_{max}, but different values of D_{min}, since expression (6) explicitly accounts for the total number of queues organized on the interface (N).

Accordingly, in Table 1, when the number of queues varied from five to eight (variants 9-12), the value of D_{min} changed, while ρ_{min} and ρ_{max} remained constant. A similar situation is observed, for example, in input data variants 4–5 and 6–8. In variants 6–8, between six and eight queues were organized on the interface, but only six were actually utilized (Table 1).

Another factor affecting the determination of D_{min} was the structure of the traffic arriving at the interface. Even with the same number of organized and active queues and the same total interface load R, differences in flow intensity across individual class queues (variants 1-3) led to variations in D_{min}, although these variations were not substantial.

Therefore, within the CB-TEQ model, the heuristic determination of the minimum value of the normalization coefficient D in expression (6) is a nontrivial task. Its solution is influenced by numerous factors, including the interface utilization, the number of organized and active class queues, and the traffic structure both at the interface in general and within individual queues. Based on experimental results, specific recommendations for selecting the minimum value of D can be formulated, as demonstrated in [19], [20]. However, even minor errors in this process may result in significant changes in queue utilization coefficients (Fig. 1 and 2) and the loss of packet service differentiation.

For this reason, the ECB-TEQ model provides a more robust solution. The study conducted using the input data presented in Table 1 confirmed its effectiveness. The ECB-TEQ model consistently computed the optimal (minimum) values of D (Table 1), ensuring the highest level of packet service differentiation across class queues (Fig. 1 and 2). The trade-off for this extended functionality is the transition from solving a linear programming problem (as in CB-TEQ) to a nonlinear programming problem (ECB-TEQ).

5 CONCLUSIONS

This study examines the well-known Class-Based Traffic Engineering Queue (CB-TEQ) model (1–10), which enables the organization and management of class queues on a router interface while supporting differentiated and balanced packet servicing. In the CB-TEQ framework, the queue utilization coefficient, and consequently delay and packet loss, are directly determined by the queue class. The higher the class of a queue, the lower its utilization coefficient and the higher the level of QoS provided to the packets assigned to it.

A key advantage of the CB-TEQ model is that the optimal allocation of interface bandwidth across class queues is achieved by solving a linear programming problem, which has a favorable impact on the computational complexity of its technological implementation. However, a certain disadvantage is the requirement for heuristic tuning of the normalization coefficient D. If D is set too high, packet service differentiation among queues is diminished. If it is set too low, the optimization problem with criterion (10) and constraints (1), (2), (7), and (9) may have no feasible solution. As the study showed (Table 1), selecting an appropriate value for D is complicated by its sensitivity to the interface state, including utilization level, the number of organized and active class queues, and the structure of incoming traffic.

Therefore, this work proposes an improvement of the CB-TEQ model by including the normalization coefficient D in the set of control variables in addition to β and b_i $(i = \overline{1, N})$. This modification required a revised optimality criterion (13) and an updated set of constraints, namely (1), (2), (9), (11), and (12). The computation of the optimal (minimum) value of the normalization coefficient within the ECB-TEQ model became possible by solving a nonlinear optimization problem, since constraint (11) is nonlinear with respect to the extended set of control variables. The study results presented in Table 1 confirmed both the feasibility of the proposed ECB-TEQ model and its effectiveness in automatically selecting the maximum level of differentiation in packet servicing across class queues.

Future research directions connected with the use of AI techniques to improve scalability and further reduce the computation time required to solve the ECB-TEQ problem. Extending the model to hierarchical queue management is also demanded. In addition, the proposed approach will be prepared for integration into programmable networking environments and adapted for deployment on corresponding network devices.

REFERENCES

[1] S. Troia, L. Borgianni, G. Sguotti, S. Giordano and G. Maier, "A Comprehensive Survey on Software-Defined Wide Area Network," in IEEE Communications Surveys & Tutorials, doi: https://doi.org/10.1109/COMST.2025.3594678.

[2] K. Zanbouri, M. Noor-A-Rahim, J. John, C. J. Sreenan, H. Vincent Poor and D. Pesch, "A Comprehensive Survey of Wireless Time-Sensitive Networking (TSN): Architecture, Technologies, Applications, and Open Issues," in IEEE Communications Surveys & Tutorials, vol. 27, no. 4, pp. 2129-2155, Aug. 2025, doi: https://doi.org/10.1109/COMST.2024.3486618.

[3] T. Vitalii, B. Anna, H. Kateryna and D. Hrebeniuk, "Method of Building Dynamic Multi-Hop VPN Chains for Ensuring Security of Terminal Access Systems," 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2020, pp. 613-618, doi: https://doi.org/10.1109/PICST51311.2020.9467953.

[4] J. Relington, QoS in IP Networks: Prioritization, Classification, and Traffic Shaping. Kindle Edition, 2025, 225 p.

[5] Cisco Systems, Inc., QoS: Congestion Management Configuration Guide, Cisco IOS XE 17. San Jose, CA, USA, 2019.

[6] M. Barreiros and P. Lundqvist, QoS-Enabled Networks: Tools and Foundations. Hoboken, NJ, USA: John Wiley & Sons, 2016, 256 p.

[7] H. Chahed, A. Kassler, "TSN Network Scheduling – Challenges and Approaches," in Network, 3, pp. 585-624, 2023, doi: https://doi.org/10.3390/network3040026.

[8] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, and E. Felstaine, "RFC 2998: A Framework for Integrated Services Operation over Diffserv Networks," 2000, doi: https://doi.org/10.17487/RFC2998.

[9] D. Black and P. Jones, "RFC 7657: Differentiated services (DiffServ) and real-time communication," 2015, doi: https://doi.org/10.17487/RFC7657.

[10] Q. Yu, J. Meng and J. J. Xu, "SW-EDF: A Single-Iteration Algorithm for Combined Input- and Output-Queued Switching," 2025 IEEE 26th International Conference on High Performance Switching and Routing (HPSR), Suita, Osaka, Japan, 2025, pp. 1-7, doi: https://doi.org/10.1109/HPSR64165.2025.11038878.

[11] W. Yang, H. Luo, S. Luo, Z. Zhang, X. Wang and T. Liu, "Efficient Scheduling Function for IETF 6TiSCH Networks Based on Multiweight Evaluation and Improved Q-Learning," in IEEE Internet of Things Journal, vol. 12, no. 17, pp. 35731-35743, 1 Sept.1, 2025, doi: https://doi.org/10.1109/JIOT.2025.3579390.

[12] H. Lin, H. Wang, N. Wang and D. Luo, "Research on Dynamic Traffic Scheduling Algorithm Based on Huawei Network Equipment," 2024 IEEE 4th International Conference on Data Science and Computer Application (ICDSCA), Dalian, China, 2024, pp. 578-584, doi: https://doi.org/10.1109/ICDSCA63855.2024.10859565.

[13] F. Alfredsson, P. Hurtig, A. Brunstrom, T. Høiland-Jørgensen and J. D. Brouer, "XDQ: Enhancing XDP with Queuing and Packet Scheduling," 2024 27th Conference on Innovation in Clouds, Internet and Networks (ICIN), Paris, France, 2024, pp. 52-56, doi: https://doi.org/10.1109/ICIN60470.2024.10494444.

[14] X. Yu, W. Chen and Y. Tian, "OWFQ: Reducing Packet Drops for Approximate Weighted Fair Queueing with Calendar Queues," 2023 9th International Conference on Computer and Communications (ICCC), Chengdu, China, 2023, pp. 540-544, doi: https://doi.org/10.1109/ICCC59590.2023.10507624.

[15] Y. Zhu, "Quality of Service Optimization for Satellite-Borne Router With Multi-Priority Scheduling Queues," in IEEE Communications Letters, vol. 27, no. 11, pp. 3003-3007, Nov. 2023, doi: https://doi.org/10.1109/LCOMM.2023.3314633.

[16] C. You, Y. Zhao, G. Feng, T. Q. S. Quek and L. Li, "Hierarchical Multiresource Fair Queueing for Packet Processing," in IEEE Transactions on Network and Service Management, vol. 20, no. 1, pp. 726-740, March 2023, doi: https://doi.org/10.1109/TNSM.2022.3197747.

[17] J. Pan, G. Chen, H. Wu, X. Peng and L. Xia, "Deep Reinforcement Learning-based Dynamic Bandwidth Allocation in Weighted Fair Queues of Routers," 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE), Mexico City, Mexico, 2022, pp. 1580-1587, doi: https://doi.org/10.1109/CASE49997.2022.9926628.

[18] G. Wu, "An Innovative Priority Queueing Strategy for Mitigating Traffic Congestion in Complex Networks," in Mathematics, 13, 495, 2025, doi: https://doi.org/10.3390/math13030495.

[19] O. Lemeshko, A. Persikov, O. Yeremenko, M. Yevdokymenko, "Method of Hierarchical Queue Management on Network Routers Based on the Goal Coordination Principle," in Advanced Smart Information and Communication Technology and Systems. MCT 2024, Lecture Notes in Networks and Systems, Springer: Cham, Switzerland, 2025, Volume 1470, pp. 200–214, doi: https://doi.org/10.1007/978-3-031-94799-5_11.

[20] O. Lemeshko, O. Yeremenko, L. Titarenko, A. Barkalov, "Hierarchical Queue Management Priority and Balancing Based Method under the Interaction Prediction Principle," in Electronics, 12, 675, 2023, doi: https://doi.org/10.3390/electronics12030675.

[21] O. Lemeshko, T. Lebedenko, M. Holoveshko, "Development and Research of Active Queue Management Method on Interfaces of Telecommunication Networks Routers," in Data-Centric Business and Applications, Lecture Notes on Data Engineering and Communications Technologies, Springer: Cham, Switzerland, 2021; Volume 69, pp. 1-20, doi: https://doi.org/10.1007/978-3-030-71892-3_1.