International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

Performance Optimization Techniques in Computer Networks

D. Priyadharshini[1], P.Yamuna[2]

[1,2]*Assistant Professor, Department of Computer Science, Mahalashmi Women's College of Arts & Science, Avadi, Chennai, Tamilnadu, India*

***Abstract*— Performance optimization in computer networks is vital for ensuring high throughput, low latency, reliability, and efficient resource utilization. As modern applications demand real-time responsiveness and large-scale data handling, the need for optimized network infrastructures has grown significantly. This paper provides an in-depth review of performance optimization techniques across various layers of the network architecture, including physical, data link, network, transport, and application layers. Approaches such as congestion control, traffic shaping, load balancing, caching, Quality of Service (QoS) provisioning, software-defined networking (SDN), and emerging AI-driven methods are discussed.**

***Keywords*— Computer Networks, Performance Optimization, Congestion Control, QoS, SDN, Load Balancing, Traffic Engineering.**

I. Introduction

Computer networks today form the backbone of almost every digital service, ranging from cloud computing and real-time communication to large-scale data analytics and Internet of Things (IoT) applications. As the volume of data traffic continues to grow and networked applications become increasingly latency-sensitive, ensuring high performance has become a critical requirement. Performance optimization in computer networks focuses on improving various operational characteristics such as throughput, latency, jitter, packet loss, resource utilization, and energy efficiency. These improvements are essential for supporting Quality of Service (QoS), guaranteeing user satisfaction, and meeting the demanding needs of emerging services such as 5G, autonomous systems, and edge computing.

Network performance is influenced by several factors, including protocol efficiency, congestion behavior, routing decisions, hardware limitations, and traffic patterns. Traditional optimization methods centered around static configurations and manual tuning are no longer sufficient because modern networks are highly dynamic, heterogeneous, and distributed.

As a result, new strategies—such as Software-Defined Networking (SDN), intelligent routing, machine learning-based controllers, advanced queuing disciplines, multipath protocols, and energy-aware communication—are increasingly adopted to enhance performance.

This paper examines key performance optimization techniques across different layers of the network architecture. It discusses classical and modern approaches, evaluates their effectiveness, and highlights open research challenges. By understanding and applying these optimization strategies, network designers and researchers can build more efficient, resilient, and scalable network infrastructures capable of meeting future demands.

Below is a focused, citation-backed **literature review** on *Performance Optimization Techniques in Computer Networks*. I organize it by major technique areas (congestion control, AQM/queuing, traffic engineering & SDN, caching/CDN, edge/offload, ML-driven approaches, and evaluation). I cite representative, high-impact works and surveys so you can follow up on primary sources.

II. Literature Review

1. Overview and scope

Performance optimization in networking is interdisciplinary: it spans protocol design (congestion control, reliability), device-level techniques (buffering, scheduling), control/management (traffic engineering, SDN/NFV), content placement (caching/CDNs), compute placement (edge/fog), and recent data-driven methods (machine learning). Surveys and reviews emphasize that no single technique suffices — practical systems combine multiple approaches and trade correctness, latency, throughput, fairness, and deployability.

2. Transport-layer congestion control

Congestion control remains foundational to network performance. Classic loss-based TCP variants (Reno, CUBIC) detect congestion via packet loss; while simple and widely deployed, they can cause persistent queuing (bufferbloat) or perform poorly on shallow-buffered paths.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

Delay-based and hybrid algorithms (*e.g.* Vegas, modern hybrids) use RTT/ECN signals to reduce queues and lower latency. Notable recent work includes DCTCP for data centers — which leverages ECN marks to keep queues small and improve latency for short, latency-sensitive flows — and BBR, which departs from loss-based heuristics by explicitly estimating bottleneck bandwidth and RTT to select pacing rates. These algorithms demonstrate important performance gains in specialized environments and highlight the need to consider deployment-interactions with legacy flows and middleboxes.

3. Queuing disciplines and Active Queue Management (AQM)

AQM designs and scheduling policies control queue occupancy, affecting delay, loss, and fairness. RED historically introduced early random drops to signal congestion, but required careful tuning. CoDel (Controlled Delay) proposed a “no-knobs” AQM that bounds queuing delay by measuring packet sojourn time, and has seen wide practical interest as a bufferbloat mitigation technique. Combined solutions (e.g., FQ-CoDel, DRR + CoDel) combine fairness and low latency; buffer sizing guidance (BDP vs. tiny buffers) remains an active practical consideration, especially across data centers vs. wide-area links.

4. Traffic engineering, routing optimization, and SDN/NFV

Traffic engineering (TE) aims to place and route flows to avoid hotspots and meet SLAs. Traditional techniques (OSPF weight tuning, MPLS) give way to centralized and dynamic methods in the SDN era: controllers with global visibility can solve constrained routing/TE problems and program forwarding state quickly. Segment Routing (SR) and programmable data planes expand expression power for TE. Surveys of SDN-based TE discuss algorithmic formulations (multi-commodity flow, constraint-based routing), practical heuristics, controller placement, and scalability/rollback concerns. NFV enables flexible placement of middleboxes/functions to shape traffic and performance through service chaining.

5. Caching, CDNs, and content placement

Caching reduces backbone load and user-perceived latency by placing content nearer to demand.

Research spans cache-replacement algorithms (LRU, LFU, TinyLFU variants), cooperative caching, and optimization for cache placement under capacity and demand uncertainty. CDNs combine server selection, caching and adaptive bitrate streaming to jointly optimize start-up latency, rebuffering, and bandwidth costs. Placement and replacement are often posed as knapsack/coverage problems or solved using demand prediction and heuristic placement strategies.

6. Edge computing and computation offloading

Shifting compute/storage to the edge (edge/fog computing) reduces latency for interactive and IoT workloads. Optimization problems include task offloading decisions, joint routing-and-compute scheduling, placement under heterogeneity, and migration costs. Studies use integer programming, bin-packing heuristics, and online algorithms; practical deployments expose trade-offs between latency gain and management complexity. Recent surveys summarize models and open issues around heterogeneity, mobility, and resource scarcity.

7. Machine learning and data-driven optimization

ML has been applied across networking tasks: traffic prediction for TE, ML-assisted congestion control, anomaly detection, learned packet scheduling, and resource allocation. Seminal surveys show ML techniques (supervised, unsupervised, reinforcement learning) improve adaptability and forecasting but face practical challenges: distribution shift, requirement for labeled data, inference latency, explainability, and safety guarantees for closed-loop control. Promising directions include hybrid model-based + ML controllers (e.g., model predictive control informed by forecasts) and safe RL with fallback heuristics.

8. Energy-aware networking and coding techniques

Energy efficiency is increasingly considered alongside traditional metrics. Techniques include link rate adaptation, powering down idle components, energy-aware routing, and trade-offs that balance energy vs. latency. Additionally, network coding and FEC are used for reliability and multicast efficiency; they introduce CPU/latency trade-offs but can reduce retransmissions and improve goodput in error-prone links. Recent work explores joint energy–performance optimization in heterogeneous infrastructures.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

9. Evaluation practices and reproducibility

Robust evaluation is critical: researchers combine simulations (ns-3, OMNeT++), emulation (Mininet, NetEm), testbeds, and production traces. Best practices include using realistic mixes (mice vs. elephant flows), reporting CDFs of latencies and flow completion times (not just means), multiple runs with confidence intervals, and releasing artifacts (configs, scripts) to support reproducibility. Surveys and experimental papers emphasize the sensitivity of conclusions to traffic mixes and parameter settings (e.g., AQM thresholds, buffer sizes).

III. Gaps And Open Research Directions

Key open problems highlighted across the literature include:

- Robust, deployable ML controllers that safely adapt under distribution shift.
- Cross-domain (multi-AS) coordinated optimization without violating administrative policies.
- Explainable, auditable black-box controllers for operator trust.
- Better joint optimization across compute, storage, and network resources (edge + CDN + TE).
- Scalable centralized control that gracefully degrades to distributed modes for resilience and latency.

IV. Network Performance Metrics

1. Throughput

Throughput refers to the rate at which data is successfully transmitted across a network channel. Optimizing throughput involves minimizing retransmissions, reducing congestion, and improving bandwidth utilization.

2. Latency

Latency measures the delay experienced by packets while traveling from source to destination. Low latency is essential for real-time applications such as VoIP and online gaming.

3. Jitter

Jitter is the variation in packet delay time, often affecting audio/video quality.

4. Packet Loss

Packet loss occurs when data fails to reach its destination, typically due to congestion, errors, or dropped packets.

5. Bandwidth Utilization

Efficient bandwidth usage ensures that network resources are used optimally without overloading links.

V. Performance Optimization Techniques

1. Congestion Control

Congestion control mechanisms maintain optimal traffic flow and prevent network overload. Common techniques include: TCP Congestion Control Algorithms: Tahoe, Reno, New Reno, Cubic. Queue Management (AQM): Random Early Detection (RED), Controlled Delay (CoDel)

2. Traffic Shaping and Policing

Traffic shaping regulates the data transmission rate to control congestion and meet QoS requirements. Techniques include: Token Bucket and Leaky Bucket algorithms

3. Load Balancing

Load balancing distributes traffic across multiple network paths or servers to reduce bottlenecks. Methods include: Round Robin and Weighted Round Robin, Consistent Hashing for distributed systems.

4. Caching and Content Delivery Networks (CDNs)

Caching frequently accessed data reduces latency and relieves network load. CDNs replicate data across geographically distributed servers to improve content delivery speed.

5. QoS Provisioning

Quality of Service mechanisms classify, prioritize, and manage network traffic. Common QoS methods include: Differentiated Services (DiffServ), Integrated Services (IntServ)

6. Routing Optimization

Optimized routing ensures efficient packet forwarding. Techniques include:, Link-state and distance-vector routing improvements

7. Network Virtualization

Virtualization improves flexibility and resource utilization. Software-Defined Networking (SDN) separates the control and data planes, enabling centralized network control.

8. Energy-Efficient Networking

Energy-efficient approaches minimize the energy consumption of network devices. Adaptive link rate (ALR), Energy-aware routing protocols

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

9. AI and Machine Learning-Based Optimization

AI-driven techniques can predict congestion, optimize routing decisions, and detect anomalies.

- Reinforcement learning for dynamic routing
- Predictive congestion management
- ML-based traffic classification

Table 1:
Key Performance Metrics in Computer Networks

Metric	Description	Purpose
Latency	Time taken for a packet to travel from source to destination	Measures responsiveness
Throughput	Amount of data transmitted per unit time	Measures transmission capacity
Bandwidth	Maximum data rate supported by the network	Determines potential speed
Packet Loss	Percentage of packets lost during transmission	Indicates network reliability
Jitter	Variation in packet delay	Affects real-time communication (VoIP, video)
Error Rate	Number of corrupted bits or packets	Impacts data accuracy
Availability	Percentage of time network is operational	Measures network reliability

VI. CONCLUSION

Network performance optimization is essential for meeting the demands of modern digital systems. Various techniques across all layers of the OSI model contribute to improving throughput, latency, security, and reliability. Emerging technologies, particularly AI and network virtualization, promise significant advancements in building adaptive and self-optimizing networks.

REFERENCES

[1] Tanenbaum, A. S., "Computer Networks".
[2] Kurose, J., Ross, K., "Computer Networking: A Top-Down Approach".
[3] RFC Standards: TCP Congestion Control, QoS, MPLS.
[4] Recent IEEE and ACM papers on SDN, NFV, and AI-driven networking.
[5] Liu, Y., et al. "Flow optimization strategies in data center networks." *Journal of Network and Computer Applications*, 2024.
[6] León, J. P. A., et al. "A machine learning based Distributed Congestion Control …" *Journal (ScienceDirect)*, 2023.
[7] Dang, X. T., et al. "A Survey on Energy-Efficient Design for Federated …" *Energies / MDPI*, 2024.
[8] Poornima, M. R., et al. "Holistic survey on energy aware routing techniques for IoT." *Journal (ScienceDirect)*, 2023.
[9] Guo, F., et al. "The application and performance optimization of multi …" *ScienceDirect*, 2025.
[10] "Performance Evaluation and Optimization of Software-Defined Networking in Large-Scale Computer Networks." *International Journal of Communication Networks and Information Security (IJCNIS)*, Nov 2024.
[11] Alahmari, S., et al. "A Comprehensive Survey on Energy-Efficient and Privacy …" *ScienceDirect*, 2025.