A Benchmarking Framework for Video Streaming Protocols

Sangeetha Abdu Jyothi

Assistant Professor, Computer Science, University of California, Irvine
Amazon Scholar, Prime Video & Studios CoreTech
sangeetha.aj@uci.edu

We will design an extensible benchmarking framework to evaluate video streaming protocols across diverse network environments. The architecture will consist of four interconnected modules, each responsible for a distinct aspect of the evaluation pipeline.

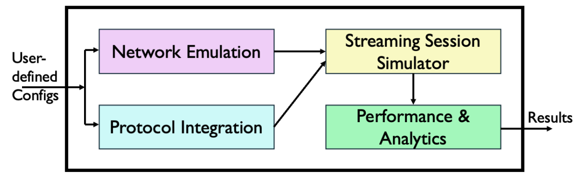

Figure 1: Overview of the proposed video streaming benchmarking framework.

Network Emulation Module. This module will integrate a wide array of real-world traces from both traditional terrestrial networks and LEO satellite networks. For terrestrial networks, we will leverage existing datasets (e.g., Puffer [1]) and collect new traces to fill coverage gaps. For LEO networks, we will conduct extensive measurements using a Starlink terminal across different times of day and weather conditions. To ensure comprehensive coverage, we will also develop a parametric trace generator that can synthesize realistic network conditions based on empirical models. This will enable controlled, repeatable experiments that isolate the impact of specific network phenomena on protocol performance.

Protocol Integration Module. This module will integrate a wide variety of streaming formats and ABR algorithms to ensure broad applicability. This module will provide built-in support for industry-standard protocols such as DASH (Dynamic Adaptive Streaming over HTTP) [4], its low-latency variant LL-DASH [2], Media over QUIC [3], and WebRTC [5]. The framework will handle the underlying mechanics of each protocol, allowing researchers to focus solely on performance evaluation. The module will feature a standardized API, creating a "plug-and-play" environment where any ABR algorithm can be implemented and tested. This design makes it straightforward to benchmark a single ABR algorithm across typical application scenarios such as video streaming and online video gaming or, conversely, to compare different ABR algorithms within a single video streaming protocol like DASH, all without altering the core simulation engine.

Streaming Session Simulator. This core component will model a client-side video player, managing the playback buffer, processing video segment requests, and simulating the user's viewing experience. It will include a Scenario Generator that will allow researchers to construct complex, reproducible evaluation scenarios by combining specific traces, introducing controlled stressors (e.g., background traffic), and defining key parameters such as initial buffer state and available video bitrates, ensuring that all published results are fully auditable and replicable. This module will interact with the Protocol Integration Module to obtain bitrate decisions and with the Network Emulation Module to simulate the download of video segments under a specified set of network conditions.

Performance and Analytics Module. This module will be responsible for capturing and analyzing the results of each simulation. It will log a rich set of low-level data, including segment download times, buffer occupancy levels, selected bitrates, and packet-level statistics. From this raw data, the module will compute typical Quality of Experience (QoE) metrics. The framework will also compute statistical confidence intervals to mitigate measurement noise. Additionally, the module will generate detailed reports and visualizations to facilitate systematic comparisons, helping to uncover the fundamental limitations of existing algorithms and guide the principled development of new ones.

We will design the benchmarking framework with extensibility and reproducibility as first-class concerns. We will implement all components as modular packages with clearly defined interfaces. Experimental configurations, including network traces, video selections, and algorithm parameters, will be specified using configuration files that can be version-controlled and shared alongside published results. The framework will automatically generate detailed logs and provenance information for every experiment, enabling other researchers to reproduce results exactly or to understand how specific methodological choices influenced outcomes. We will release the framework as open-source software, accompanied by comprehensive documentation and a curated collection of baseline experiments that establish reference performance levels for common protocols under standard conditions. Beyond serving as an evaluation tool, the benchmarking framework will function as a discovery platform for identifying fundamental principles of ABR design. By systematically varying network conditions, video characteristics, and algorithmic parameters, we will conduct large-scale sensitivity analyses that reveal which design choices matter most under different scenarios.

References

[1] Puffer. https://puffer.stanford.edu/results/.

[2] DASH-IF/DVB. DASH-IF/DVB Report on Low-Latency Live Service with DASH. https://dashif.org/docs/Report%20on%20Low%20Latency%20DASH.pdf, 2025.

[3] IETF. Media Over QUIC (moq). https://datatracker.ietf.org/wg/moq/about/, 2025.

[4] T. Stockhammer. Dynamic adaptive streaming over HTTP– Standards and Design Principles. In *Proceedings of the second annual ACM conference on Multimedia systems*, pages 133–144, 2011.

[5] WebRTC-Initiative. WebRTC. https://webrtc.googlesource.com/src/, 2025.