Lyapunov Optimization Based Queue-Aware Traffic Shaping for 5G-TSN in Industrial Environments

Kouros Zanbouri
University College Cork
Cork, Ireland
k.zanbouri@cs.ucc.ie

Md Noor-A-Rahim
University College Cork
Cork, Ireland
m.rahim@cs.ucc.ie

Cormac J. Sreenan
University College Cork
Cork, Ireland
Cormac.Sreenan@ucc.ie

Dirk Pesch
University College Cork
Cork, Ireland
dirk.pesch@ucc.ie

***Abstract*—Manufacturing companies look increasingly at Private 5G networks to manage Automated Guided Vehicles (AGVs). While 5G promises Ultra-Reliable Low Latency Communication (URLLC), line-of-sight blockage in dense metallic environments degrades capacity faster than transport protocols and rate adaptation can react to, and the shortfall accumulates in the modem's transmission buffer as bufferbloat and latency spikes that violate safety margins. The backlog serves as a sufficient statistic for the impairment, capturing the capacity deficit without requiring explicit channel measurements or predictions. We address this with a cross-layer rate control algorithm based on Lyapunov Drift-plus-Penalty theory embedded at the 3GPP Release 16 Device-Side TSN Translator, which trades service utility against queue stability from the instantaneous buffer state alone. We validate the approach in a trace-driven simulation framework driven by 3GPP-mapped capacity data. Numerical results demonstrate that while baseline scheduling suffers catastrophic queue accumulation and excessive delays upon reconnection, the proposed Lyapunov controller eliminates bufferbloat, ensuring immediate low-latency operation as soon as the channel recovers.**

***Index Terms*—5G; TSN; Time-Sensitive Networking; URLLC; Lyapunov optimization**

I. Introduction

Industry 4.0 relies heavily on the deployment of dependable wireless networks, such as private 5G networks, to support mobile robots and Automated Guided Vehicles (AGVs). Unlike traditional factory automation, these mobile agents require a wireless network infrastructure that handles mixed-criticality traffic: Enhanced Mobile Broadband (eMBB) to support high-throughput video streaming, and URLLC to guarantee the responsiveness required for teleoperation and safety monitoring [1]. However, many industrial radio environments are challenging, with large metallic structures, moving cranes and high-density racking creating severe multipath fading and shadowing. In particular, networked AGVs are susceptible to frequent LOS blockage events, where the channel capacity can degrade to near-zero very quickly [2].

To integrate these wireless capabilities with deterministic industrial control systems, 3GPP Release 16 introduced the 5G-TSN framework, which logically models the entire 5G system as a single Time-Sensitive Networking (TSN) bridge [3]. At the 5G network boundaries, time-synchronization and protocol translation are handled by the Network-Side TSN Translator (NW-TT) at the core network and the Device-Side TSN Translator (DS-TT) co-located with the User Equipment (UE). In the standard architecture the DS-TT functions as a Layer 2 bridge that, while performing residence-time computation for gPTP synchronization, lacks any awareness of the wireless channel state or queue conditions within the 5G modem, and therefore cannot adapt its forwarding behaviour to the underlying radio dynamics.

A critical challenge arises when physical blockage events interact with this passive architecture. Standard video transmission mechanisms (e.g., UDP/RTP) and congestion control algorithms rely on packet loss or Round Trip Time signals to adapt their transmission rates [4], and in a blockage scenario these feedback loops are too slow. As the channel capacity drops, the application layer continues to push high-bitrate video frames into the modem's transmission buffer. Commercial 5G modems typically feature deep buffers (e.g., 4 to 8 MB) designed to absorb bursty traffic and maximize TCP throughput. For an 8 Mbps video stream, saturating such a buffer introduces multiple seconds of queuing latency, which can be catastrophic for real-time teleoperation safety margins [5]. In safety-critical scenarios, such latency renders remote machinery uncontrollable.

To address this, we propose a queue-aware rate control algorithm embedded in the DS-TT, a *Smart DS-TT*, which couples application-layer video adaptation to link-layer buffer state. Because it acts as a UE-side ingress shaper it requires no base-station modification and no cooperation from the peer endpoint, and it therefore applies to unresponsive UDP video, which transport-layer congestion control cannot shape. We evaluate our approach in a trace-driven simulation across three industrial blockage scenarios, under a finite modem buffer, and against a non-adaptive UDP baseline.

II. Related Work

Standard 3GPP mechanisms like Logical Channel Prioritization (LCP) [6] enforce QoS at the link layer but cannot throttle application source rates during sudden channel degradations. Conversely, end-to-end Adaptive Bitrate frameworks including WebRTC and existing Lyapunov-based video streaming algorithms [7] rely on Receiver Reports to estimate network conditions, introducing an unavoidable RTT delay; during sudden 5G RF blockages this delayed feedback can cause massive bufferbloat before the congestion signal ever reaches the sender [4]. To manage bufferbloat closer to the radio edge,

979-8-3195-0167-7/26/$31.00 ©2026 IEEE

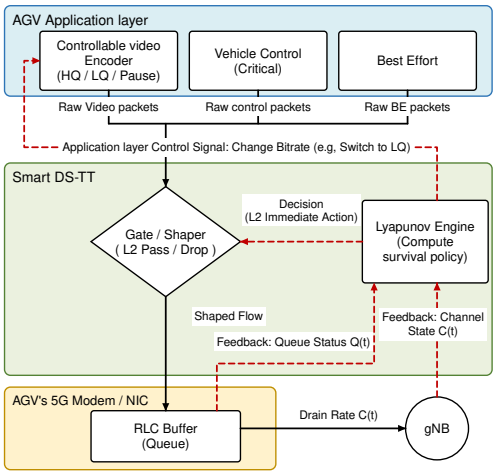


Fig. 1: Smart DS-TT architecture and queue-state feedback loop.

Active Queue Management (AQM) mechanisms such as CoDel have been evaluated within the 5G SDAP layer, but CoDel relies on static timing heuristics that fail to adapt to abrupt 5G bandwidth shifts, and its congestion control depends on TCP's window reduction response to packet loss. For UDP-based applications, dropped packets do not trigger any rate adaptation at the source, rendering the queue management ineffective as a throughput control tool [8]. Recent cross-layer architectures rely on network-assisted proxies or base-station coordination [9], at the cost of infrastructure modifications, or track SINR end-to-end to suppress spurious retransmission timeouts [10], requiring protocol modifications at both the mobile station and the remote server and remain inapplicable to UDP-based real-time flows. Our work differentiates itself by using a lightweight Lyapunov control law that requires only local buffer state information, avoiding the rigid parameter tuning of AQM and the signaling overhead of network-assisted scheduling.

III. Proposed Smart DS-TT Architecture

Fig. 1 shows the proposed architecture. The Smart DS-TT sits between the application layer and the 5G modem and closes a feedback loop between exactly two layers: it reads the link-layer buffer occupancy and sets the application-layer encoding rate. The AGV generates three traffic classes, which the Smart DS-TT maps to standard 5G QoS flows and serves under a strict priority discipline: critical control traffic (URLLC, 5QI 82), adaptive video streaming (eMBB, 5QI 2), and best-effort background data (5QI 9). The URLLC queue is maintained as a single-packet *Freshness-First* buffer in which new state updates overwrite older ones, so that stale telemetry is discarded at source and the modem always transmits the freshest state upon channel recovery. Taking the modem's transmission buffer occupancy as the state variable, the system dynamics follow the standard discrete-time queue evolution equation [11]:

$$Q(t+1) = \max\big[\, Q(t) + A(t) - C(t),\ 0 \,\big] \qquad (1)$$

where $Q(t)$ is the video backlog at slot t, $C(t)$ the number of bits the radio can carry during the slot, and $A(t) = R_{video}(t)\Delta t$ is the volume admitted by the shaper. The video source is controllable across a set of four discrete profiles ($\mathcal{P} = \{1080p, 720p, 480p, \text{Pause}\}$, generating 8, 4, 1 and 0 Mbps and carrying utility weights w_p of 8.0, 5.0, 2.0 and 0.0 respectively), and the instantaneous utility is $U(t) = w_p$ for the selected profile.

Defining the quadratic Lyapunov function $L(t) = \frac{1}{2}Q(t)^2$ and the one-step drift $\Delta(t) = \mathbb{E}[L(t+1) - L(t) \mid Q(t)]$, applying $\max(x,0)^2 \le x^2$ to Eq. (1) bounds the drift by $B + Q(t)\mathbb{E}[A(t) - C(t) \mid Q(t)]$ for a constant B. Following the drift-plus-penalty framework [12] and incorporating the utility, and minimizing the resulting bound at each slot is equivalent to maximizing

$$\underset{p \in \mathcal{P}}{\text{maximize}} \quad V w_p \ - \ Q(t)\, R_p\, \Delta t \qquad (2)$$

where $V \ge 0$ dictates the strictness of the trade-off: a larger V prioritizes utility, a smaller V prioritizes a small backlog. In this policy $Q(t)$ acts as a dynamic penalty weight, so the controller throttles the source before congestion can accumulate, and it does so without observing $C(t)$ at all. With Q in Mb, R_p in Mbps and the constant Δt absorbed into V, as in our implementation, equating the metrics of adjacent profiles yields deterministic switching thresholds in closed form: $Q = 0.75V$ from 1080p to 720p, $Q = V$ from 720p to 480p and $Q = 2V$ from 480p to Pause. The decision engine requires only $O(|\mathcal{P}|)$ arithmetic operations per TTI and stores three scalar queue counters, making it suitable for resource-constrained industrial gateways or embedded modems.

IV. Trace-Driven Simulation

We developed a discrete-time, trace-driven simulation framework that models the interaction between wireless channel dynamics and the AGV's internal queuing system. Uplink channel-quality data was generated with the **Simu5G** framework in OMNeT++ for a 3GPP Indoor Factory scenario with a single AGV moving at 2 m/s, numerology $\mu = 1$ and a 51-PRB allocation; CQI is mapped to spectral efficiency using 3GPP TS 38.214 Table 5.2.2.1-2 and multiplied by the granted bandwidth to obtain a per-slot bit budget on a 0.5 ms grid over a 60 s window. The simulation compares the proposed Smart DS-TT against a Baseline UDP agent that transmits 1080p regardless of channel conditions, mimicking current non-adaptive deployments. Background traffic comprises an 800-bit URLLC payload every 20 ms and a 2 Mbps best-effort stream.

Three impairment windows are introduced into the capacity series, representative of different industrial obstacle types: a *transient deep blockage* of 2 s at 0 Mbps, modelling a passing forklift; an *extended deep blockage* of 5 s at 0 Mbps, modelling the AGV navigating behind large static machinery; and a *severe partial degradation* of 3 s at 2 Mbps, modelling a non-metallic obstacle. A complete 0 Mbps outage is realistic in dense Indoor Factory deployments, where NLOS shadowing drives the SINR below the threshold required for signal decoding at the lowest modulation and coding scheme, and it serves as a strict worst-case boundary test.

V. Simulation Results & Discussion

A. Impact of Lyapunov Parameter V

Fig. 2 illustrates the peak video queue accumulated over the full three-impairment trace as V is swept logarithmically across 25 points from 0.1 to 100, revealing three distinct operating regions. Q_{peak} is the maximum over the whole 60 s run, not over a single window. For $V \leq 4.32$ the peak queue follows the linear relationship $Q_{peak} \approx 2V$ (Mbits), reproducing the $Q = 2V$ switching threshold derived in Section III rather than independently testing the asymptotic $O(V)$ result [11]. At $V = 0.5$ this yields $Q_{peak} = 1.00$ Mbit, corresponding to a worst-case video queuing delay of $D_{max} = 1.00\,\text{Mb}/78.5\,\text{Mbps} \approx 12.8$ ms at the mean unimpaired capacity of the trace. Both bounds on V are then explicit rather than empirical: the buffer requires $V \leq (Q_{budget} - R_{min}\Delta t)/2$, and reaching Pause within T_{react} requires $V \leq T_{react}/1.156$, the coefficients following from the profile set; the smaller binds. For $4.32 < V < 56.2$ the backlog can no longer reach $2V$ within the outage, since climbing the profile ladder takes $1.156\,V$ seconds; the peak is then set by T_{block} and $2V$ stays an unattained upper bound. A 4 Mbit modem buffer binds sooner, at $V > 2$. Beyond $V = 56.2$ the peak queue saturates at the Baseline value, the utility term dominating so completely that the controller never downshifts.

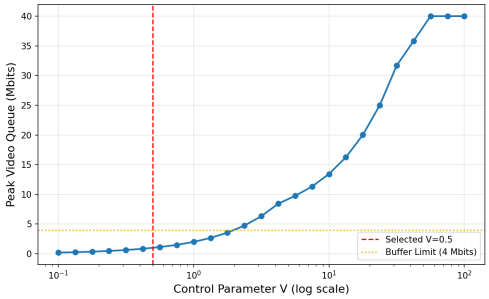


Fig. 2: Peak video queue against the Lyapunov parameter V, showing the linear control, transition and saturation regions.

B. Resilience to Variable Blockage Events

Fig. 3 shows the peak video queue accumulated by each agent across the three impairment scenarios, both agents being evaluated on the identical trace. Under the *transient 2 s blockage* the Baseline continues transmitting at 8 Mbps throughout, accumulating a peak queue of 16.00 Mb, exactly matching the theoretical deficit $R_{max} \times T_{block} = 16$ Mb. Every frame generated after reconnection queues behind that backlog, and the resulting delay is quantified in Section V-D. The Lyapunov controller instead begins downshifting 47 ms after onset, when the backlog first crosses $Q = 0.75V$, and steps down through 720p and 480p to reach Pause after 577 ms, capping the peak queue at 1.00 Mb, a **16**× reduction, and resumes low-latency operation instantly, with no backlog to drain.

Under the *extended 5 s blockage*, the Baseline accumulates 40.00 Mb, again matching the theoretical maximum, so the backlog grows linearly with the outage duration and is bounded only by the buffer; Section V-C shows what happens when that bound is reached. The Lyapunov controller maintains the same hard ceiling of **1.00 Mb** as in the 2 s scenario, confirming duration-independent stability: across a 2.5× longer blockage, the ceiling is determined by V, not by T_{block}.

Under the *3 s partial degradation* to 2 Mbps the Baseline, unaware of the capacity reduction, produces a shortfall of 6 Mbps and an 18.12 Mb backlog. The proposed controller applies a graduated downshift that matches the generation rate to the available capacity, capping the peak queue at 0.50 Mb, a **36**× reduction. The ceiling is lower than under total blockage because 480p at 1 Mbps is already covered by the residual 2 Mbps, so the climb halts at $Q = V$ rather than the $Q = 2V$ a full outage forces. Both behaviours emerge from the same drift-plus-penalty metric without any scenario-specific tuning.

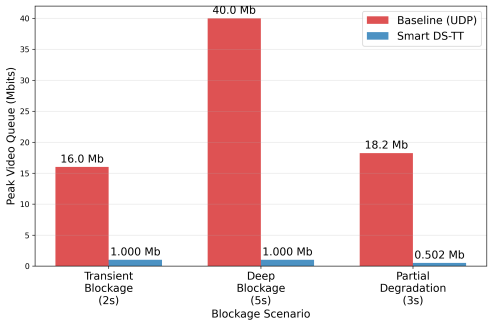


Fig. 3: Peak video queue comparison across three realistic industrial blockage scenarios.

C. Finite Buffer & Packet Loss Analysis

In real-world industrial 5G modems buffer memory is strictly constrained, typically 2–8 MB. We introduced a hard buffer ceiling of 4 Mbits and measured packet loss during the 5-second deep blockage, initialising the queues empty so that all reported losses are attributable to the blockage alone. The results reveal a catastrophic failure mode in the Baseline. Transmitting continuously at 8 Mbps with no drain, it saturates the 4 Mbit buffer within **0.498 s** of blockage onset, after which the modem applies tail drop: of the 40 Mb generated during the blockage, **36 Mb (90%)** is irretrievably lost, yielding a data integrity score of **10.0%**. Tail drop is particularly destructive for compressed video, as it corrupts the GOP dependency chain and causes cascading decoding failures at the receiver. The Smart DS-TT eliminates packet loss entirely by throttling the video generation rate before packets enter the queue: the peak queue remains bounded at 1.00 Mb, 75% below the hardware limit, giving **zero packet loss** and **100% data integrity**, though over a far smaller volume: the Baseline generates 40 Mb and loses 36, the Smart DS-TT generates 1.00 Mb and loses none. Shaping chooses what to give up; tail drop does not. The same distinction bounds what AQM can achieve here: a dwell-time-based scheme holds its delay target by discarding arriving packets, but an unresponsive UDP source does not reduce its rate on loss, so the queue stays small at the cost of video integrity rather than by shaping the source.

D. Latency & Jitter

Fig. 4 compares the queueing delay $Q(t)/C(t)$ over the recovery phase; it is undefined while $C(t) = 0$, so slots inside the outage are excluded. Statistics are further conditioned on a non-empty queue, which the Baseline has in 1,066 of the 12,000 slots and the controller in 25. The Baseline exhibits a median latency of 240.7 ms with 95th-percentile and peak delays of 457.2 ms and 481.2 ms; a delay ranging from 240 ms to nearly 500 ms renders the system unusable for dynamic teleoperation, exceeding the human reaction threshold of < 100 ms by a wide margin. The proposed controller maintains a median latency of 5.6 ms with a 95th percentile of 11.0 ms, a $\mathbf{43}\times$ reduction on that conditional comparison and a 43-fold lower queue occupancy, because coupling the generation rate to the queue drift prevents the modem buffer from accumulating a backlog in the first place.

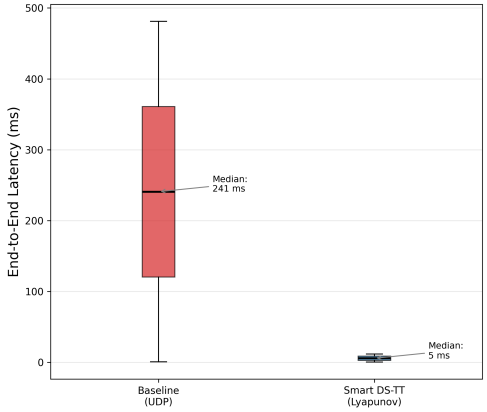


Fig. 4: Queueing delay over the recovery phase.

Jitter, defined as $|D(t) - D(t-1)|$ for sequentially received packets, tells a similar story. Both systems exhibit a similar median (≈ 0.45 ms), but the Baseline's low median is an artefact of bufferbloat: delays grow at a steady linear rate during the outage, keeping frame-to-frame differences artificially small. The failure is exposed upon recovery, when the saturated buffer flushes its stale content and the Baseline suffers a jitter spike reaching **481.20 ms**, violating the 10 ms threshold mandated for URLLC. The Smart DS-TT bounds its worst-case jitter to **11.56 ms**. Finally, URLLC traffic is insulated by the strict-priority discipline and the Freshness-First buffer rather than by the rate controller: with separate class queues the control payload is drained ahead of video in every slot, so both schemes serve it identically.

VI. Conclusion

This paper presented a cross-layer Smart DS-TT framework for the robust teleoperation of industrial AGVs in blockage-prone 5G-TSN environments, embedding a Lyapunov-based decision engine that optimizes the trade-off between video utility and queue stability. Trace-driven simulation shows the advantage of this proactive approach over unoptimized UDP transmission. The controller eliminates bufferbloat, achieving a $43\times$ reduction in the median queueing-delay estimate; during deep blockages, it eradicates packet loss entirely, raising data integrity from 10% to 100%, whereas the baseline suffers queue saturation and video corruption; and it maintains a strictly bounded queue ceiling regardless of outage length. Future work will generate the impairments inside the propagation model rather than the capacity series, and evaluate the controller against tuned threshold and capacity-aware baselines; an extended treatment is available in [13].

Acknowledgment

This publication has emanated from research conducted with the financial support of Taighde ireann Research Ireland under Grant number 13/RC/2077 P2 (CONNECT: the Research Ireland Centre for Future Networks). For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] D. Segura, E. J. Khatib, J. Munilla, and R. Barco, "5G numerologies assessment for URLLC in industrial communications," *Sensors*, vol. 21, no. 7, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/7/2489

[2] F. Burmeister, N. Schwarzenberg, T. Hößler, and G. Fettweis, "Measuring time-varying industrial radio channels for D2D communications on AGVs," in *Proceedings of the 2021 IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1–7.

[3] 3rd Generation Partnership Project (3GPP), "System Architecture for the 5G System (5GS)," 3GPP, Technical Specification TS 23.501, March 2026, release 19.

[4] T. Kanaya, N. Tabata, and S. Yamaguchi, "A study on performance of CUBIC TCP and TCP BBR in 5G environment," in *Proceedings of the 2020 IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 508–513.

[5] L. Prause and M. Akselrod, "Sender-based approach to reducing the bufferbloat in heterogeneous 5G networks," in *Proceedings of the 2025 36th International Teletraffic Congress (ITC-36)*, 2025, pp. 1–5.

[6] 3GPP, "NR; Medium Access Control (MAC) protocol specification," 3rd Generation Partnership Project, Technical Specification 3GPP TS 38.321, 2025, release 19.

[7] J. Yang, B. Yang, S. Chen, Y. Zhang, Y. Zhang, and L. Hanzo, "Dynamic resource allocation for streaming scalable videos in SDN-aided dense small-cell networks," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2114–2129, 2019.

[8] M. Irazabal, E. Lopez-Aguilera, and I. Demirkol, "Active queue management as quality of service enabler for 5G networks," in *Proceedings of the 2019 European Conference on Networks and Communications (EuCNC)*, 2019, pp. 421–426.

[9] P. Netalkar, J. Chen, and D. Raychaudhuri, "mmCPTP: A cross-layer pull based transport protocol for 5G mmWave networks," in *Proceedings of the NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, 2023, pp. 1–9.

[10] M. Takahashi and A. Nakao, "Cross-layer congestion control algorithm and architecture for communication performance optimization in millimeter-wave networks," *IEEE Access*, vol. 14, pp. 35 823–35 837, 2026.

[11] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, 1st ed., ser. Synthesis Lectures on Learning, Networks, and Algorithms. Cham, Switzerland: Springer, 2010.

[12] L. Bracciale and P. Loreti, "Lyapunov drift-plus-penalty optimization for queues with finite capacity," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2555–2558, 2020.

[13] K. Zanbouri, M. Noor-A-Rahim, C. J. Sreenan, and D. Pesch, "Lyapunov Optimization Based Queue-Aware Traffic Shaping for 5G-TSN in Industrial Environments," https://arxiv.org/abs/2606.25823, 2026, extended version.