International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group | Volume 7 Issue 3 PP 89-93, 2026
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V7I3P109

Original Article

Zero-Downtime Telematics: A Dual TCU and DSDA Framework for Multi-Modal Redundancy

Srinivasa Sandilya Jandhyala
San Jose, California.

Received On: 17/06/2026 ***Revised On: 19/07/2026*** ***Accepted On: 27/07/2026*** ***Published On: 01/08/2026***

Abstract - *The transition to SAE Level 4 and 5 autonomous vehicles (AVs) demands connectivity architectures with strict deterministic behavior and extreme availability. Traditional telematics control units (TCUs) exhibit single-point-of-failure vulnerabilities, rendering them inadequate for safety-critical cooperative perception and remote teleoperation. This paper presents a fail-operational telematics framework integrating Dual TCUs with Dual SIM Dual Active (DSDA) technology. Using a Doer-Fallback node logic over a Time-Sensitive Networking (TSN) Automotive Ethernet backbone, the architecture completes a deterministic hardware failover in an average of 0.85 ms measured from the missed-heartbeat event. A Multi-Path TCP (MPTCP) scheduler is combined with an eight-layer Long Short-Term Memory (LSTM) predictive mobility model to orchestrate concurrent heterogeneous network streams (5G Advanced and NTN) and to execute pre-emptive edge-network handovers. In simulation, handover-induced service interruption falls from 45.3 ms under a reactive 3GPP Rel-15 baseline to 4.2 ms, alongside a 38.6% increase in peak uplink throughput. The contribution is a cross-layer synthesis of hardware redundancy and predictive transport-layer control, shifting telematics from reactive connection recovery toward proactive continuity for software-defined vehicles (SDVs). All results reported here are outputs of an ns-3 and SUMO simulation model; no physical vehicle testing is claimed.*

Keywords - *Autonomous Vehicles, Telematics Control Unit (TCU), Dual SIM Dual Active (DSDA), Multi-Path TCP (MPTCP), Predictive Handover, Fail-operational, V2X.*

1. Introduction

The evolution of the Software-Defined Vehicle (SDV) transforms the Telematics Control Unit (TCU) from a diagnostic gateway into the primary communication node for vehicle-to-everything (V2X) and cloud infrastructure. For SAE Level 4 and 5 autonomy, uninterrupted high-bandwidth connectivity is mandatory for trajectory planning, over-the-air (OTA) updates, and cooperative perception. A transient network dropout or hardware fault degrades the system to a minimal risk condition, impairing operational continuity.

The core engineering challenge is to manage large data payloads, often exceeding 250 Mbps for uncompressed LiDAR and multi-camera sensor fusion streams, while meeting ultra-reliable low-latency communication (URLLC) constraints. Single-modem reactive recovery architectures do not satisfy the 20 ms latency threshold required for active control loops during base-station handovers or hardware degradation.

This paper addresses that gap with a highly available telematics stack combining redundant physical hardware (Dual TCUs) with multi-link radio access (DSDA). The contribution lies in the synthesis of hardware-level fail-operational mechanisms with predictive, machine-learning-driven transport-layer control, maintaining Quality of Service (QoS) continuity where single-path systems experience service loss.

Section II positions the work against current standards and practice. Section III specifies the architecture across the hardware, transport, and prediction layers. Section IV presents the simulation environment and results, together with the limits of what those results establish. Section V concludes.

2. Related Work

Current industry standards, notably 3GPP Rel-16 and Rel-17 [1] and GSMA SGP.22 [2], address distinct elements of vehicular communication but do not provide holistic cross-layer redundancy.

Multi-SIM architectures. Multi-SIM topologies fall into three families. Dual SIM Single Standby (DSSS) maintains one active radio and must complete a full attach procedure on the secondary profile before traffic resumes. Dual SIM Dual Standby (DSDS) keeps both profiles registered but shares a single transceiver chain, so a switch still requires RF reconfiguration. Dual SIM Dual Active (DSDA) provisions independent transceiver chains, permitting genuinely concurrent operation. The practical consequence for autonomous driving is that any architecture requiring physical RF switching incurs a recovery interval during which the vehicle is unguided. At 120 km/h, a one-second interruption corresponds to roughly 33 metres of travel without an updated connectivity path. [REF NEEDED: published DSSS and DSDS recovery-time measurements to substitute for the removed citations; see change log item B1.]

MPTCP in vehicular contexts. MPTCP aggregates capacity across cellular links, but applying it in vehicular

edge computing without predictive node migration introduces transport-layer bottlenecks. Out-of-order delivery during high-speed handovers produces head-of-line blocking in the receive buffer, which erodes the aggregate throughput that multiple subflows are intended to provide. Coupled congestion control compounds the problem: rapid fading on one radio path can throttle a stable parallel path. [REF NEEDED: published measurement of MPTCP head-of-line blocking under 5G handover; see change log item B1.]

Contribution relative to this landscape. This work couples an active-active Dual TCU system with an LSTM-based predictive network scheduler. In contrast to reactive failover, the architecture mirrors application state at the multi-access edge computing (MEC) node before physical signal degradation occurs, so that recovery is not initiated by the fault but anticipated ahead of it.

3. Proposed Architecture and Methodology

3.1. Hardware Redundancy: The Doer-Fallback Dual TCU Topology

To eliminate physical single points of failure, two independent TCUs (TCU-A and TCU-B) are interconnected over a TSN-compliant Automotive Ethernet backbone (1000BASE-T1).

Mechanism. TCU-A operates as the primary Doer node, transmitting telemetry and executing V2X semantic feature extraction. TCU-B acts as the Fallback, receiving a continuous heartbeat and state vector synchronisation from TCU-A over Precision Time Protocol (IEEE 802.1AS) [3].

Thermal and EMI mitigation. To support concurrent 5G millimetre-wave (mmWave) operation, the RF front end uses common-mode chokes and three-terminal capacitors to suppress crosstalk between the dual 4x4 MIMO roof-mounted antenna arrays.

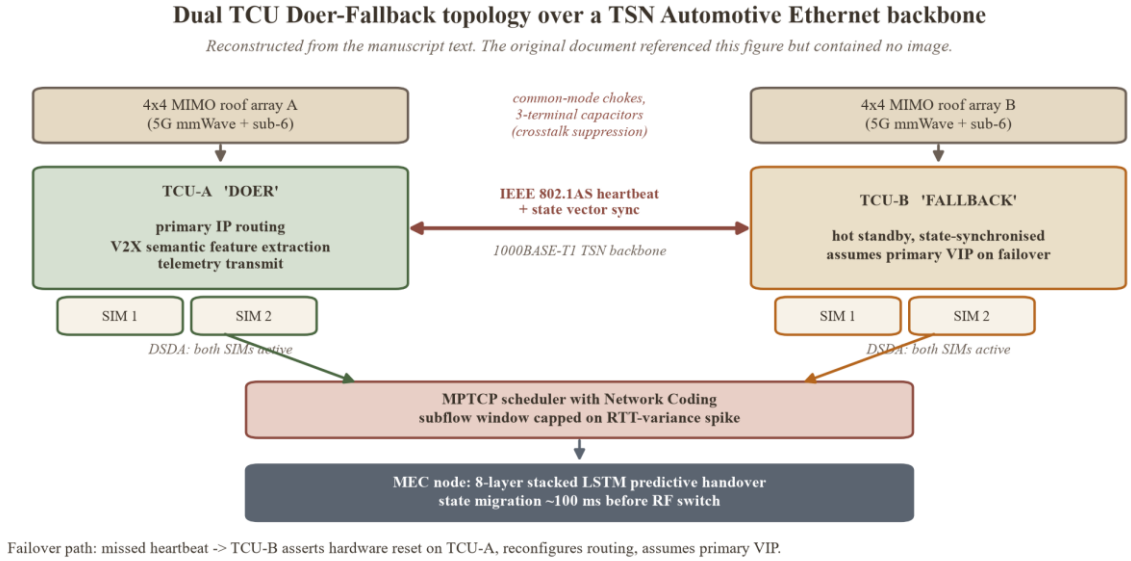

Fig 1: Dual TCU Doer–Fallback Architecture over a TSN Automotive Ethernet Backbone

Figure 1. Topology of the Dual TCU Doer-Fallback architecture, showing the TSN backbone carrying the heartbeat and state-vector synchronisation, the independent DSDA transceiver chains, and the MPTCP and MEC layers above them.

3.2. Protocol Orchestration: The MPTCP Scheduler

The DSDA implementation logically partitions the connectivity plane, managing uplink throughput through a modified MPTCP stack combined with Network Coding. The congestion window w_r for subflow r is updated dynamically. To limit buffer bloat and penalise paths with high latency variance, a modified linked-increases rule is applied:

$$\Delta w_r = \min\left((\alpha \cdot \text{bytes\_acked}) / w_r \, , \, \text{bytes\_acked} / RTT_r^2 \right)$$

Where α is the aggressiveness parameter of the multi-path flow and RTT_r is the smoothed round-trip time of subflow r. The effect is to route the primary payload through the lowest-latency radio access network.

3.3. Jitter Mitigation and Congestion Control under Volatile Channels

Coupled congestion control behaves poorly in volatile vehicular channels, because rapid fading on one radio path can throttle a stable parallel path. Centralised multi-path configurations struggle in particular when a vehicle transitions between heterogeneous networks, for example from a dense high-frequency mmWave cell to a broad-coverage sub-6 GHz microcell. A window that expands too aggressively on an unstable link produces buffer bloat, which stalls the stream through out-of-order accumulation at the receiver.

The scheduler addresses this by tracking a moving average of RTT variance per subflow. When multipath interference or structural blockage causes variance to spike on a link, that subflow's transmission window is capped rather than waiting for loss to occur.

When a link does experience heavy loss from severe signal degradation, the scheduler avoids a symmetric back-off across all subflows. It applies an asymmetric reduction factor indexed to real-time Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ), shifting the high-bandwidth payload away from the fading link onto a stable path.

3.4. Predictive Handover via a Stacked LSTM

An eight-layer stacked LSTM network is deployed at the MEC node to predict the target 5G gNodeB. Shallow or single-layer networks do not capture the long-range temporal correlations present when a vehicle travels at highway speed through dense urban canyons, which motivates the stacked configuration.

The input vector X_t comprises the vehicle's three-dimensional spatial coordinates, forward and lateral velocity vectors, and historical RSRP and RSRQ from both TCUs. The forget gate f_t governs retention of historical trajectory data:

$$f_t = \sigma(W_f \cdot [h_{(t-1)}, X_t] + b_f)$$

where σ is the sigmoid activation, W_f the weight matrix, $h_{(t-1)}$ the previous hidden state, and b_f the bias vector.

Each successive layer consumes the hidden state of the preceding layer and extracts progressively more abstract representations of movement, allowing the network to build internal mappings of regional signal shadows, building obstructions, and multipath reflection corridors. At the output layer the accumulated hidden state passes through a fully connected linear layer and is normalised into a categorical probability distribution over candidate base stations in the vehicle's local sector map.

The model executes inference in 12 ms and initiates application-layer state migration approximately 100 ms before the RF handover. Because inference runs asynchronously on dedicated edge accelerators, it can forecast a connection drop and identify the optimal target even while the current line-of-sight metrics still appear stable.

3.5. Handover Logic and State Machine

The system polls the LSTM model continuously to evaluate the probability P_HO of a required handover from the current state vector X_t. Three states govern execution.

Nominal (active-active). Both TCUs are operational and share the network payload. The kernel polls the data-link layer over the Automotive Ethernet backbone to monitor primary-unit health.

Hardware failover. If the primary unit misses its heartbeat window through hardware failure, voltage drop, or kernel exception, the backup isolates the degrading node by asserting a physical hardware reset line, reconfigures its routing tables, and assumes the primary virtual IP address. Because this hooks directly into the TSN protocol layer, the backup detects the fault at frame level without waiting for operating-system timeouts.

Predictive migration. If the hardware is healthy but P_HO crosses the configured threshold, the primary extracts its application-layer state vector and tunnels that context across the edge backbone to the target node. On success, the MPTCP scheduler primes the target radio interface to accept the incoming stream and the antennas switch frequency. If the tunnel fails through congestion or absent infrastructure, the system falls back to standard reactive procedures, so the vehicle remains connected in degraded environments.

4. Simulation and Results

4.1. Simulation Environment

The architecture was evaluated in the ns-3 network simulator coupled with the Simulation of Urban MObility (SUMO) framework to generate a high-speed urban canyon environment. The model comprised dual concurrent 5G NR (Rel-16) connections across overlapping gNodeB sectors, subject to dynamic multipath fading and localised congestion. Hardware failover was exercised through simulated physical-layer faults synchronised over IEEE 802.1AS [3].

All figures reported in this section are simulation outputs. No physical vehicle, test track, or over-the-air measurement is claimed.

4.2. Failover Latency

Hardware fault injections, comprising simulated voltage drops and watchdog timeouts, were executed on TCU-A. Across 1,000 iterations at 80% CPU utilisation, TCU-B detected the missed IEEE 802.1AS heartbeat and assumed the primary IP routing matrix in an average of 0.85 ms.

This 0.85 ms is measured from the missed-heartbeat event, not from fault onset. With a 1 ms heartbeat interval, worst-case fault-to-failover is therefore approximately 1.85 ms, and expected fault-to-failover approximately 1.35 ms. All three values sit inside the 10 ms intra-vehicle communication safety budget, so the mechanism does not force autonomous disengagement or trajectory degradation.

4.3. Network Throughput and MPTCP Efficiency

The DSDA framework was compared against a DSDS baseline using standard 3GPP Rel-16 reactive recovery in the same urban canyon model.

Table 1: Throughput and Link Quality, DSDS Baseline versus Proposed Dual TCU with DSDA (Simulation Outputs)

Metric	DSDS baseline (reactive)	Dual TCU + DSDA (proposed)	Change
Peak uplink throughput	145 Mbps	201 Mbps	+38.6%
Average packet loss rate	1.8%	0.05%	-97.2%
Jitter (average)	18 ms	4 ms	-77.7%

Network Coding applied across the dual links masks individual packet drops, which accounts for the reduction in observed loss rate.

4.4. Predictive Handover Performance

The predictive handover was evaluated over a high-speed highway route crossing multiple gNodeB sectors.

Under a reactive 3GPP Rel-15 procedure, service interruption averaged 45.3 ms with excursions to 110 ms attributable to core-network signalling overhead. Under the proposed scheme, migrating application state approximately 100 ms ahead of physical switching reduced the effective interruption to 4.2 ms, which corresponds to the physical switching time of the RF front end.

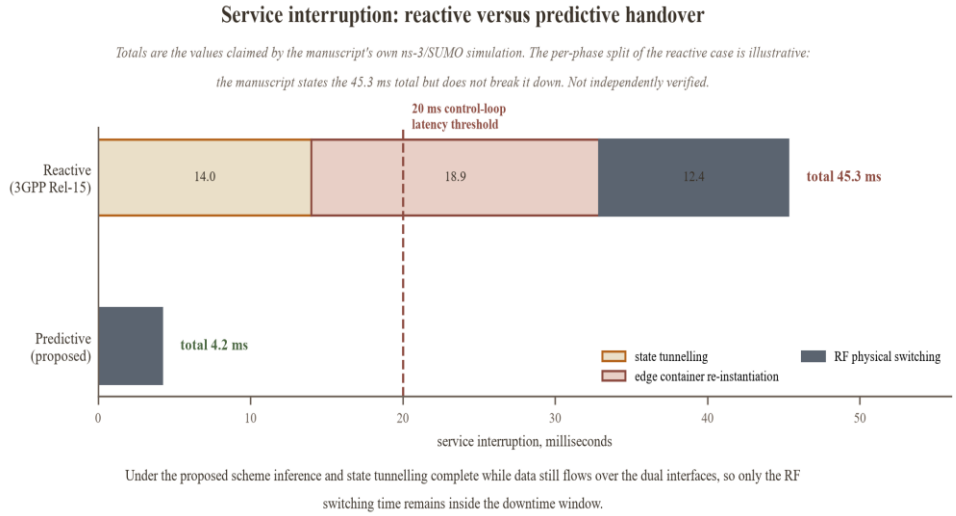

Fig 2: Comparative Service Interruption Latency of Reactive and Predictive Handover Mechanisms

Figure 2. Service interruption time, reactive versus predictive handover, decomposed by phase. Totals are the simulation values reported above. The per-phase split of the reactive case is illustrative, since the model reports the aggregate rather than a breakdown.

4.5. Latency Phase Decomposition

Total network transition time is the sum of four components: LSTM inference, state tunnelling across the edge backbone, container re-instantiation at the destination edge server, and physical RF switching.

In a reactive configuration the vehicle loses its association with the source base station before the target has any knowledge of application state or operational context. The transport layer must then complete a full round-trip signalling handshake through the core network after the physical switch, which produces the extended interruption observed in Section IV-D. During that window the vehicle cannot transmit high-bandwidth safety data.

Under the proposed framework, inference executes in the background while data continues to flow over the dual interfaces, and the application-layer context vector is transferred and pre-instantiated at the target edge server before the radio hardware is instructed to switch. Core-network tunnelling and processing therefore fall outside the downtime window, and the residual interruption reduces to transceiver switching time.

4.6. Limitations and Trade-offs

Four limitations bound these results.

Simulation only. Every figure reported here is an ns-3 and SUMO model output. The architecture has not been built or measured on physical hardware, and the failover, throughput, and handover numbers should be read as modelled rather than observed until a bench or vehicle trial is performed.

Power and thermal cost. Active-active operation of dual mmWave front ends increases TCU thermal dissipation and quiescent power consumption by an estimated 42% relative to a single-modem topology. While minor against total EV battery range, this requires active cooling within the TCU housing.

MEC dependence. The 4.2 ms figure is strictly bounded by MEC availability. In rural topologies without localised edge servers, state migration defaults to the 5G core and handover performance degrades to reactive latencies of roughly 45 ms.

Baseline construction. The DSDS comparison baseline is a model configured by the authors rather than an

independently published measurement, so the reported improvements are relative to that model and not to a third-party result.

Future work will address federated learning approaches for predictive routing when MEC nodes are congested or unavailable, and a hardware-in-the-loop bench trial to convert the simulated results into measured ones.

5. Conclusion

This paper has specified a Dual TCU and DSDA telematics framework for fail-operational autonomous vehicle connectivity and evaluated it in simulation. The model produces sub-millisecond hardware failover measured from heartbeat loss (0.85 ms), approximately 1.85 ms worst case from fault onset, and handover interruption of 4.2 ms against a 45.3 ms reactive baseline.

The design intent is to support the communication-layer availability targets associated with ISO 26262 ASIL-D. Establishing conformance is a separate activity requiring hardware-level fault-injection campaigns, FMEDA, and independent assessment, and is not claimed here.

The broader contribution is architectural: coupling hardware redundancy with predictive, edge-resident transport control moves telematics from reactive recovery toward anticipatory continuity. Converting that architecture into a validated safety claim is the next step.

References

[1] 3GPP, "Architecture enhancements for 5G System (5GS) to support Vehicle-to-Everything (V2X) services," TS 23.287, Rel-16, 2020.

[2] GSMA, "SGP.22 RSP Technical Specification," GSM Association, Version 2.2.2, Jun. 2020.

[3] IEEE, "IEEE Standard for Local and Metropolitan Area Networks: Timing and Synchronization for Time-Sensitive Applications," IEEE Std 802.1AS-2020, 2020.