

Evaluating 5G Network Performance and Ultra-Low-Latency Communication: A Study of QoS and RedCap for Industrial IOT

Dr Saima Aleem¹, Dr Nasruddin²,

¹Department of Computer Science Engineering, Khwaja Moinuddin Chishti Language University, Lucknow U.P. India

² Regional Centre for Urban & Environmental Studies (RCUES)
Lucknow U.P. India
nasar.rcueslko@gmail.com

Abstract - The deployment of 5G technology has commenced in India and other regions, aiming to facilitate higher data rates, reduced latency, and enhanced network capacity. Nonetheless, the nationwide implementation remains inconsistent, with service quality exhibiting considerable variation among operators and urban locales. This review evaluates 5G performance, particularly emphasizing ultra-low-latency communication by assessing round-trip latency, throughput, and packet loss through empirical drive tests and application-level observations. Additionally, the study explores 5G Reduced Capability (RedCap) devices as a crucial facilitator for industrial IoT, where it is imperative to minimize cost, power consumption, and hardware complexity while maintaining reliable connectivity. The paper underscores existing Quality of Service (QoS) deficiencies in 5G deployments and discusses the technical and policy interventions required to enhance latency-sensitive services in India.

Key Words: 5G, ultra-low latency, URLLC, RedCap, QoS, industrial IoT, network performance

I. INTRODUCTION

Current 5G deployments in India by operators like Jio and Airtel support emerging use cases such as autonomous vehicles and industrial automation. However, achieving the sub-1ms latency required for mission-critical applications remains a challenge requiring further optimization. The deployment of 5G in India is recognized as a foundational technology for critical applications, including autonomous vehicles, remote surgery, and industrial automation [1][2] (Mistry & Yadav, n.d.) (Kaur & Singh, n.d.). These sectors rely on Ultra-Reliable Low-Latency Communication (URLLC), which is designed to enable mission-critical functions by supporting sub-1ms latency and 99.999% reliability [3] (Joshi, 2025). While 5G represents a significant advancement over previous generations for cyber-physical systems, current implementations are often viewed as a first generation that remains limited in terms of end-to-end predictability and scalability [4]

Operator landscape and performance

Major Indian telecommunications providers, specifically Reliance Jio and Bharti Airtel, have been at the forefront of 5G diffusion in the country [1] (Mistry & Yadav, n.d.). These operators are increasingly partnering with

stakeholders to implement 5G-enabled IoT and low-latency services [2] (Kaur & Singh, n.d.). For instance, Jio has utilized its network infrastructure, such as Jio Air Fiber, to support low-latency requirements for specialized projects like smart tractor vehicle-to-infrastructure (V2I) communication [5] (Ray et al., 2025). Furthermore, both Jio and Airtel have significantly expanded their 5G footprint, with their combined connections reaching 12 million in the context of fixed wireless access (FWA) services [6] (Bott, 2025).

Latency optimization and challenges

Achieving the performance benchmarks necessary for mission-critical URLLC requires rigorous network optimization. Research indicates that while 5G URLLC improves performance compared to 4G, specific technical configurations are required to meet sub-1ms targets [7][3] (Urbaniak et al., 2023) (Joshi, 2025) Table 1 summarizes three latency optimization strategies commonly considered in real-time and ultra-reliable communication systems. It highlights how each technique affects latency and what cost or limitation is associated with its use. The comparison helps show that reducing delay often comes with compromises in throughput, control overhead, or scheduling flexibility.

Table 1. Latency Optimization Strategies and Their Trade-offs

Optimization Strategy	Latency Impact	Trade-off
2-symbol mini-slots	52% lower latency	30% throughput reduction
Grant-free uplink	40% control plane reduction	N/A

TDD constraints	frame latency	+1.8ms latency	Requires dynamic adaptation
--------------------	------------------	-------------------	-----------------------------------

Processing heterogeneity in user equipment (UE) can introduce significant latency variations, specifically $\pm 0.8\text{ms}$ in FDD and $\pm 1.2\text{ms}$ in TDD, highlighting the need for standardized hardware capabilities [3] (Joshi, 2025). Additionally, the integration of security features, such as user plane integrity protection, introduces overhead that can affect end-device performance, prompting operators to weigh security benefits against potential latency impairments [8] (Heijligenberg et al., 2023). Ultimately, the transition to 6G and the evolution of edge computing are identified as necessary steps to bridge the gap between current network capabilities and the requirements for truly deterministic end-to-end communication [3][4] (Joshi, 2025)

II. ULTRA-LOW LATENCY IN 5G

Research from 2024-2025 highlights that while 5G achieves sub-1 ms latency via grant-free uplink and edge integration, commercial deployments face challenges. RedCap supports varied bandwidths across FR1 and FR2, though channel estimation remains a critical optimization bottleneck.

Technical and commercial latency gaps

The gap between theoretical technical capabilities and commercial implementation is defined by the integration of specialized network architectures. While 5G networks are architected to achieve latency under 1 ms through mechanisms such as grant-free uplink transmission and preconfigured resources, practical industrial deployments often struggle to maintain these thresholds outside of controlled environments. Research indicates that while dedicated Ultra-Reliable Low-Latency Communication (URLLC) slices can successfully facilitate safety-critical control by reducing latency, the reliance on centralized cloud infrastructure frequently introduces delays that necessitate a shift toward localized edge computing. The synergy of 5G and edge AI is specifically intended to bridge this divide by minimizing centralized processing, thereby enabling real-time decision-making for applications like autonomous vehicles and industrial automation.



Figure 1: 5G's Sub-1 ms Latency

Effectiveness of grant-free uplink and edge computing
Grant-free uplink transmission is identified as a primary enabler for meeting sub-millisecond latency requirements in high-density IoT environments. By bypassing traditional scheduling

requests, this method significantly reduces the control-plane overhead that typically hinders time-sensitive applications. When combined with edge computing, which shifts localized processing closer to the data source, the overall system response time is substantially improved compared to standard 5G configurations. Furthermore, techniques such as mini-slot allocation and grant-free uplink are essential for maintaining synchronization in 6G-enabled digital twin architectures, proving that these features are foundational to both current 5G-Advanced and future network standards.

5g Nr Redcap Channel Estimation Advances

Enhanced battery lifetime, cost, and chip-size reduction in 5G New Radio (NR) RedCap devices necessitate single-antenna reception, making efficient channel estimation critical; [10] Jendraszyk et al. demonstrate that a Time-Domain-Enhanced (TDE) MMSE algorithm outperforms standard MMSE by 0.9 dB in multipath conditions with only 9% silicon-area overhead, offering superior robustness without prior channel knowledge.

RedCap performance in FR1 and FR2 bands

RedCap (Reduced Capability) technology is designed to balance performance requirements across diverse industrial IoT applications by operating within both Frequency Range 1 (FR1) and Frequency Range 2 (FR2) bands. In FR1, RedCap supports flexible bandwidth configurations, with capabilities typically specified at 20 MHz [9] (Veedu et al., 2022) and scaling up to 100 MHz [11] (Pan, 2025). Conversely, the use of the FR2 spectrum facilitates the deployment of more sophisticated, high-capacity architectures that support higher throughput requirements [10] (Pan, 2025). These configurations often leverage advanced MIMO antenna setups, such as 4Rx, to maintain connectivity and performance stability despite the reduced complexity of the user equipment.

RedCap channel estimation and throughput bottlenecks

The performance of RedCap is constrained by fixed radio parameters and the inherent complexities of channel estimation in dynamic industrial environments. While RedCap aims to provide a middle ground between low-power wide-area networks and high-end 5G, its reliance on standardized sub-carrier spacing and fixed configurations can create throughput bottlenecks when network conditions fluctuate. Effective channel estimation is critical to mitigating these limitations, particularly when operating in the higher-frequency FR2 bands where signal propagation is more susceptible to interference [10] (Pan, 2025). Consequently, research emphasizes the need for advanced control strategies to manage these parameters and ensure consistent performance during transitions between 4G and 5G network layers. Channel estimation and performance optimization are crucial for the effective deployment of 5G New Radio (NR), especially for Reduced Capability (RedCap) devices with limited hardware resources. These devices often face challenges such as multipath fading and reduced antenna configurations, which can degrade performance. Advanced channel estimation techniques, such as Time-Domain-Enhanced (TDE) MMSE methods, offer improved robustness and efficiency, making them suitable for RedCap and industrial IoT applications. These methods, along with flexible estimation frameworks, help optimize performance while maintaining manageable implementation overhead.

Time-Domain-Enhanced Mmse Methods

TDE MMSE methods outperform standard MMSE algorithms by 0.9 dB under strong multipath conditions, without requiring additional channel knowledge, thus enhancing robustness[11](Jendraszyk & Korb, 2024). The TDE approach includes unallocated spectrum extrapolation and threshold-based significant tap selection, which are crucial for RedCap devices with single-antenna reception[11](Jendraszyk & Korb, 2024).

Flexible Estimation Frameworks

Flexible channel estimation frameworks, such as those implemented on vector digital signal processors (vDSP), allow adaptation to various IoT standards and workloads, achieving a 15.9x speed-up compared to non-vectorized approaches[12](Damjanecvic et al., 2021). These frameworks are essential for handling the high application variability and weak radio signal conditions typical in IoT deployments[12](Damjanecvic et al., 2021).

Graph Neural Networks for Channel Estimation

Lightweight Graph Neural Networks (GNNs) like GraphNet offer a scalable solution for channel estimation, matching or exceeding the performance of traditional methods in dynamic environments[13](Norouzi et al., 2025). GraphNet's low computational footprint makes it suitable for real-time deployment on edge devices, enhancing robustness in challenging channel conditions[13](Norouzi et al., 2025).

Statistical Modelling and Optimization

Novel statistical models using LS and MMSE algorithms improve channel estimation accuracy by incorporating precise measurements of channel coefficients, enhancing performance in various fading scenarios[14](Sabuncu & Bilgehan, 2025).

These models are particularly effective in non-line-of-sight (NLOS) and multipath environments, offering a nuanced approach compared to pilot-based estimation[14](Sabuncu & Bilgehan, 2025).

While these advanced methods significantly enhance channel estimation and performance optimization, challenges remain in real-world 5G deployments, such as those in India. Variability in service quality across regions and operators highlights the gap between theoretical capabilities and user experience, emphasizing the need for continued innovation and adaptation in channel estimation techniques[15](Karimi & Pan, 2009). The integration of Time-Sensitive Networking (TSN), Deterministic Networking (DetNet), and network slicing within 5G networks significantly enhances their capability to support mission-critical applications. TSN synchronizes devices and manages packet timing, while DetNet ensures deterministic behavior across IP networks. Network slicing allows for the separation of various service types, such as Ultra-Reliable Low-Latency Communication (URLLC), enhancing performance and reliability. The following sections elaborate on these mechanisms.

TSN and DetNet Integration

TSN provides precise timing and synchronization, crucial for industrial automation, ensuring minimal packet loss and robust communication under challenging conditions[16](Abuibaid et al., 2023). DetNet extends these capabilities across IP networks, guaranteeing end-to-end latency and packet loss assurances for real-time applications[17](Deshpande et al., 2024).

Network Slicing Benefits

Network slicing enables the creation of virtual networks tailored to specific service requirements, such as low latency for URLLC and higher reliability for industrial applications [18] (Vaideghy et al., 2025). Simulations indicate that network slices can maintain latencies as low as 6 ms for vehicular traffic, even under peak load conditions [18](Vaideghy et al., 2025).

Practical Implementations

The integration of TSN with network slicing allows for orchestrated management of resources, enhancing the performance of critical applications like factory automation [19](Bhattacharjee et al., 2021). Beyond 5G (B5G) networks are anticipated to further improve URLLC capabilities, leveraging mature transport networks for enhanced reliability[20] (Abbou et al., 2023).

While the integration of these technologies presents significant advantages, challenges remain in ensuring that non-critical services do not adversely affect the performance of critical applications, highlighting the need for careful management and resource allocation.

III. CHALLENGES AND FUTURE SCOPE

Despite its promise, 5G still faces several challenges in delivering consistent ultra-low latency. These include radio congestion, limited spectrum availability, backhaul delays, uneven infrastructure rollout, and operator-specific optimization differences. Future work should focus on edge intelligence, AI-based scheduling, RedCap optimization, and joint radio-core-network tuning. In addition, policy support and infrastructure planning are needed to improve QoS in high-density urban areas. As 5G matures, it is likely to play a central role in real-time communications and industrial automation.

REFERENCES

- [1] M. Mistry and P. Yadav, "The Diffusion of 5G in India: An Analysis of Unprecedented Pace and Systemic Speed Limits", [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5639530
- [2] K. Kaur and J. Singh, "GROWTH OF INDIAN TELECOM SECTOR IN THE CHANGING POLICY PARADIGM", [Online]. Available: https://www.nrjbe.in/images/paper_pdf/GRO-6a3276d5d4463.pdf
- [3] A. Joshi, "Optimization of Latency in 5G Networks," Indian Scientific Journal Of Research In Engineering And Management, vol. 09, no. 04, pp. 1–9, Apr. 2025, doi: 10.55041/ijsrem46438.
- [4] "Towards Deterministic Communications in 6G Networks: State of the Art, Open Challenges and the Way Forward," Apr. 2023, doi: 10.48550/arxiv.2304.01299.
- [5] J. Ray, B. Sharma, S. Khatua, and A. Biswas, "5G URLLC test bed design and smart tractor V2I implementation at GangaSagar Island exploring 5G new radio and AI", [Online]. Available: <https://link.springer.com/article/10.1007/s41870-025-02882-6>
- [6] A. Bott, "Differentiated connectivity services in 5G standalone architecture", [Online]. Available: <https://media-publisher.eu/wp-content/uploads/2026/01/5-6-2025.pdf>
- [7] D. Urbaniak, J. Rosell, R. Suárez, and M. Suppa, "Computación frontera: influencia de latencias en la precisión de robots," pp. 779–784, Aug. 2023, doi:10.17979/spudc.9788497498609.779.

- [8] T. Heijligenberg, C. Bohm, D. Rupprecht, and K. Kohls, "BigMac: Performance Overhead of User Plane Integrity Protection in 5G Networks," May 2023, doi: 10.1145/3558482.3581777.
- [9] S. N. K. Veedu et al., "Toward Smaller and Lower-Cost 5G Devices with Longer Battery Life: An Overview of 3GPP Release 17 RedCap," in *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 84-90, September 2022, doi: 10.1109/MCOMSTD.0001.2200029.
- [10] Jendraszyk, D., & Korb, M. (2024). Hardware-Efficient Time-Domain-Enhanced Channel Estimation for 5G New Radio RedCap. 359–364.
<https://doi.org/10.1109/meditcom61057.2024.10621363>
- [11] X. Pan, "Detection and Compliance in Financial Management via 6G-Enabled Standardized AI-Powered Digital Twins," in *IEEE Communications Standards Magazine*, doi: 10.1109/MCOMSTD.2025.3639269.
- [12] S. A. Damjancevic, S. A. Dasgupta, E. Matus, D. Utyansky, P. van der Wolf, and G. Fettweis, "Flexible Channel Estimation for 3GPP 5G IoT on a Vector Digital Signal Processor," pp. 12–17, Oct. 2021, doi: 10.1109/SIPS52927.2021.0001
- [13] S. Norouzi, M. Rahmani, Y. Chu, T. Braun, K. R. Chowdhury, and A. Burr, "Lightweight Graph Neural Networks for Enhanced 5G NR Channel Estimation," pp. 1–7, July 2025, doi: 10.1109/pimrc62392.2025.11274706.
- [14] Ö. Sabuncu and B. Bilgehan, "Novel Statistical Modelling and Optimization Techniques of Fading Channel Coefficients for 5G Network Performance," *Journal of Network and Systems Management*, vol. 33, pp. 42–42, Feb. 2025, doi: 10.1007/s10922-025-09905-4.
- [15] Karimi, M., & Pan, D. (2009). Challenges for Quality of Service (QoS) in Mobile Ad-hoc Networks (MANETs). 60, 1–5. <https://doi.org/10.1109/wamicon.2009.5207262>
- [16] Abuibaid, M. A., Ghorab, A., Saruhan, A. A., St-Hilaire, M., Parsons, G., Farkas, J., Varga, B., Moldován, I., Mate, M., & Naqvi, S. H. R. (2023). Integration of DetNet/TSN Reliability Functions in 5G Systems: A Case Study and Measurements. 369–375.
<https://doi.org/10.1109/cscn60443.2023.10453205>
- [17] Deshpande, Y., Diederich, P., Luthfi, M., Becker, L., Fontalvo-Hernández, J., & Kellerer, W. (2024). Integrating Deterministic Networking with 5G. *arXiv.Org*, abs/2409.13400. <https://doi.org/10.48550/arxiv.2409.13400>
- [18] Vaideghy, A., Abbas, H. M., & Shreef, H. (2025). Network slice management in mobile internet for critical applications. *Journal of Internet Services and Information Security*, 15(3), 394–407.
<https://doi.org/10.58346/IJISIS.2025.I3.027>
- [19] Bhattacharjee, S., Katsalis, K., Arouk, O., Pellegrini, F., Slear, R., Nandakumar, R., & Sabella, D. (2021). Network slicing for TSN-based transport networks. *IEEE Access*, 9, 62788–62809. <https://doi.org/10.1109/ACCESS.2021.3074802>
- [20] A. N. Abbou, K. Samdanis and J. Manner, "URLLC in B5G Networks: Use Cases, TSN/DetNet Extension, and Pending Issues," 2023 33rd International Telecommunication Networks and Applications Conference, Melbourne, Australia, 2023, pp. 25-30, doi: 10.1109/ITNAC59571.2023.10368531.

3. CONCLUSIONS

The online version of the volume will be available in LNCS Online. Members of institutes subscribing to the Lecture Notes in Computer Science series have access to all the pdfs of all the online publications. Non-subscribers can only read as far as the abstracts. If they try to go beyond this point, they are automatically asked, whether they would like to order the pdf, and are given instructions as to how to do so.

ACKNOWLEDGEMENT

The heading should be treated as a 3rd level heading and should not be assigned a number.

REFERENCES

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. *Int. J. Digit. Libr.* 1 (1997) 108–121
2. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): *Theoretical Aspects of Computer Software. Lecture Notes in Computer Science*, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415–438
3. van Leeuwen, J. (ed.): *Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science*, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
4. Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)