Bridging 5G and 6G for Multimedia IoT: Structural Limitations and Architectural Directions

Alex Piccioni*†, Claudia Rinaldi†, Carlo Centofanti*‡, Andrea Marotta*‡,
Piergiuseppe Di Marco*†, and Fabio Graziosi*†
**Department of Information Engineering, Computer Science and Mathematics*, *University of L'Aquila*, L'Aquila, ITALY
{alex.piccioni, carlo.centofanti, andrea.marotta, piergiuseppe.dimarco, fabio.graziosi}@univaq.it
†*CNIT*, *University of L'Aquila Research Unit*, L'Aquila, ITALY, claudia.rinaldi@univaq.it
‡*CNIT*, *FIBERS Lab.*, L'Aquila, ITALY

Abstract—**The convergence of Multimedia and Internet of Things (IoT) (M-IoT) is driving a new class of applications that combine interactive, high-throughput content with real-time sensing, inference, and actuation. While 5G systems have introduced key enablers such as edge computing, network slicing, and broadcast/multicast delivery, current deployments reveal structural limitations when exposed to the heterogeneous demands of immersive Extended Reality (XR), cooperative perception, and holographic services. This paper consolidates M-IoT requirements from a system-level perspective and analyzes how 5G architectures respond to these demands through the modeling of physical resources of HD and 4K streaming workloads under unicast, sliced, Non-Public Network (NPN), and 5G Multicast/Broadcast Service (5MBS) delivery configurations. These results reveal structural limitations related to throughput saturation, rigid resource partitioning, and unidirectional delivery constraints that cannot be resolved through incremental enhancements alone. In this background, we provide a structured mapping of identified limitations to targeted 6G architectural directions aligned with IMT-2030 objectives, considering that future networks must move beyond capacity upgrades toward integrated support for compute–communication co-design, scalable radio access, and intelligent orchestration. These features are essential to meet the evolving needs of M-IoT and enable resilient, low-latency, and scalable media services.**

Index Terms—**Multimedia Internet of Things (M-IoT), 5G, 6G, Mobile Networks, Multimedia.**

I. INTRODUCTION

Multimedia Internet of Things (M-IoT) captures the convergence of Internet of Things (IoT) sensing/actuation with high-bandwidth media streams (e.g., audio, video, and immersive content), where heterogeneous devices cooperate to enable interactive services such as multi-view Extended Reality (XR), industrial video analytics, smart-city monitoring, and real-time cyber-physical feedback loops [1], [2]. Compared to scalar telemetry, M-IoT pipelines expose the communication substrate to sustained high bitrates, stringent end-to-end latency, tight multi-modal synchronization, and strong requirements on reliability, security, and energy efficiency [3]. These characteristics make M-IoT a fundamentally cross-layer problem, requiring integrated optimization across physical infrastructure, transport protocols, and service orchestration.

Recent deployments in vertical domains such as remote assistance, collaborative robotics, and immersive learning have revealed the practical impact of these challenges, where performance degradations often stem from structural limitations in the underlying network stack. In particular, many applications must sustain heterogeneous uplink/downlink profiles, react to rapidly changing workloads, and operate across distributed topologies with limited local compute capabilities. Such constraints become even more pronounced in large-scale deployments that span enterprise sites, public venues, and mobile outdoor environments.

Fifth generation of mobile technologies (5G) systems have introduced a set of enabling mechanisms that are directly relevant to M-IoT, such as: virtualization and service-based cores, multi-access edge computing, network slicing for differentiated Quality of Service (QoS), Non-Public Networks (NPNs) for local determinism, and multicast/broadcast capabilities for one-to-many media delivery [4]. The evolution toward 5G-Advanced further strengthens these capabilities and expands the design space for multimedia services running over shared infrastructures [5]. Nevertheless, current deployments and studies often treat these mechanisms in isolation (e.g., slicing without joint edge/media orchestration, broadcast without adaptive unicast integration, or private deployments without systematic trade-off analysis), leaving open questions on end-to-end architectural governance across sensing, processing, transport, and application layers [1].

These gaps highlight a broader need for network architectures that go beyond throughput and latency improvements, embracing adaptability, content-awareness, and full-stack programmability. IMT-2030, which defines the requirements for next generation mobile networks (i.e., 6G) is expected to broaden the scope from connected things to intelligent, immersive, and resilient cyber-physical systems, motivating architectural shifts toward Artificial Intelligence (AI)-native control loops, tighter compute-communication co-design, and support for extreme media workloads such as high-fidelity volumetric streaming and immersive bidirectional interaction [6]–[8]. In this context, bridging 5G-Advanced and 6G for M-IoT calls for a principled analysis of which emerging requirements cannot be met by incremental upgrades alone, and which architectural directions can enable a smooth transition path while preserving deployability and interoperability.

This work consolidates the requirements of M-IoT systems

ISBN 978-3-903176-82-9 © 2026 IFIP

from a multimedia-centric perspective and repositions them within the broader evolution from 5G to 6G. Unlike prior works that address multimedia networking or IoT connectivity in isolation, or that treat 6G directions at a purely qualitative level, this paper combines a system-level requirement analysis with quantitative PRB-based modeling of representative streaming workloads under realistic 5G deployment configurations, providing a structured mapping of identified limitations to targeted 6G architectural directions grounded in IMT-2030 objectives.

II. M-IoT REQUIREMENTS AND APPLICATION SCENARIOS

The M-IoT landscape is shifting from the concept of media sensing to interactive, closed-loop multimedia services in which audio-visual streams are not only consumed, but also continuously interpreted and acted upon. This evolution amplifies the coupling between communication, computation, and media processing, and turns network support into a first-order design constraint rather than an implementation detail [1], [3]. This paradigm shift is reflected in a growing class of M-IoT applications that place unprecedented demands on the underlying infrastructure, spanning immersive XR, cooperative perception, and scalable group media services.

Immersive and wearable XR. Interactive XR experiences (e.g. multi-party immersive sessions and wearable Augmented Reality (AR)/Mixed Reality (MR) form factors) require sustained high-throughput delivery, stringent motion-to-photon latency, and stable jitter bounds to avoid perceptual discomfort and preserve interaction fidelity [9]. Moreover, XR workloads are inherently multi-modal (audio, video, spatial metadata, haptics) and can involve asymmetric uplink/downlink patterns due to viewpoint feedback, scene updates, and multi-view contributions, making them structurally different from conventional downlink-centric streaming [10].

Cooperative perception and distributed media intelligence. Smart spaces and connected robotics increasingly rely on cooperative perception, where multiple sensing agents share intermediate representations (raw streams, features, or semantic objects) to expand spatial coverage and robustness [11]. Such pipelines produce bursty traffic, frequent state updates, and strict freshness constraints, which push the network toward content-aware prioritization and compute-communication co-optimization at the edge [3], [11].

Scalable group media and hybrid events. Hybrid collaboration and live media experiences often combine personalized unicast (e.g., individualized viewpoints and adaptive bitrates) with one-to-many delivery (e.g., common scene anchors, shared audio objects, or event feeds), stressing the need for seamless integration of unicast, multicast/broadcast, and edge-based adaptation within a unified service architecture [1]. As the number of simultaneous consumers grows, unicast-based delivery becomes inherently inefficient, since identical content is replicated over separate radio bearers for each receiver, consuming spectral resources proportionally to the audience size. Broadcast and multicast mechanisms address this inefficiency by decoupling content transmission from the number of receivers, making them a structural requirement rather than an optional feature for large-scale M-IoT deployments. However, pure broadcast delivery lacks the flexibility to support personalized streams, uplink interaction, and adaptive quality selection, which are essential in immersive and collaborative scenarios. This tension between scalability and interactivity motivates the need for hybrid delivery architectures capable of dynamically combining broadcast efficiency with unicast adaptability, a challenge that current 5G deployments only partially address through mechanisms such as 5G Multicast/Broadcast Service (5MBS).

III. M-IoT REQUIREMENTS AND TRAFFIC DEMANDS

A. Architectural Viewpoint and Scope

Following established architectural models for multimedia-aware IoT systems [1], we adopt a streamlined layered abstraction that separates device-level operations, media processing, distributed intelligence, and communication infrastructure. The lower layer includes sensing, actuation, and user-facing devices that interact with the physical environment and generate media-rich data. This content is acquired, encoded, and rendered through application-specific pipelines that must adapt to heterogeneous modalities and device capabilities. Data is then processed and orchestrated across edge and cloud resources, which manage inference, storage, and content adaptation. All these elements are connected through a programmable network substrate that spans access, transport, and core segments, enabling flexible service composition and quality assurance. While each layer introduces specific requirements and constraints, this paper focuses on the limitations arising at the network level (Fig. 1). The increasing complexity of multimedia workloads, characterized by high-throughput flows, ultra-low latency demands, and strict coordination between distributed components, exposes structural boundaries in current 5G deployments. These challenges reveal not just performance bottlenecks but architectural tensions that cannot be fully addressed by incremental enhancements, and which call for a more integrated, media-aware network design as envisioned in the evolution toward 6G.

B. Quantitative Requirements from XR, Haptics, and Holography

The evolution of M-IoT is increasingly shaped by interactive and immersive applications, where sensing, inference, and rendering are orchestrated in tightly coupled loops spanning devices, edge infrastructure, and network substrate. These closed feedback cycles impose compound demands that extend beyond conventional throughput or latency metrics, and require service continuity under heterogeneous and dynamic conditions.

Among the most prominent modalities, XR introduces a particularly complex blend of asymmetric traffic patterns. Unlike traditional downlink-centric media services, XR couples high-rate visual and audio delivery with uplink streams conveying sensor data, head and body pose, and real-time interaction feedback. These flows are further intensified by edge-based

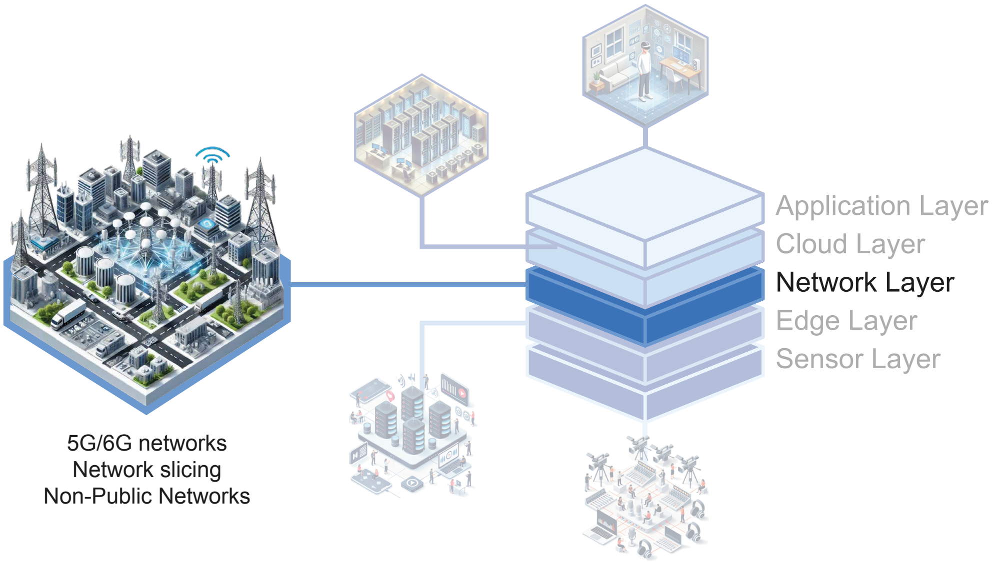


Fig. 1: Multimedia architecture.

rendering functions and perception offloading mechanisms, leading to dynamic and asymmetric traffic profiles. According to 3GPP technical reports, representative XR workloads range from $0.1 - 1$ Gbps for 360° video up to 10 Gbps for ultra-high fidelity interactive scenes, depending on scene complexity and user interaction frequency [10]. These workloads are also highly latency-sensitive: interaction loops must be completed within 5 to 20 ms to preserve responsiveness and perceptual coherence.

Haptic communication and teleoperation services further expose the limitations of current mobile infrastructures by emphasizing the discrepancy between radio-interface latency targets and the end-to-end requirements of closed-loop control. These applications operate under extremely tight temporal constraints, where round-trip delays must often remain below 1 ms. Use cases such as telesurgery, remote manipulation, or precision feedback in industrial settings additionally impose strict reliability requirements, depending on the operational risk and criticality. Sustaining such performance levels becomes especially challenging in dynamic environments affected by user mobility, radio interference, and resource contention across multiple service slices. These difficulties are exacerbated when high-rate XR media streams coexist with latency-sensitive haptic flows, creating contending demands on the same physical and logical network layers.

IV. 5G STRUCTURAL LIMITATIONS FOR M-IOT

A. *Structural Stress Points*

The traffic demands of emerging M-IoT services differ from conventional workloads not only in scale, but in their structural characteristics. As summarized in Table I, these applications combine sustained downlink streaming, bursty and high-volume uplink flows, and tight feedback loops involving latency-sensitive interaction between edge and device. Scenarios involving multiple concurrent XR sessions, for example, can quickly saturate the throughput provisioned for dense-urban environments, where the expected user-experienced data rates are limited to 100 Mbps (downlink) and 50 Mbps (uplink) [10], [12].

The coexistence of latency-critical and high-throughput modalities also creates a scheduling dilemma. Haptic or cooperative perception traffic must meet tight latency and reliability targets, but competes with bulk media delivery under shared infrastructure constraints. Although the 5G radio interface sets aggressive theoretical targets, such as 1 ms latency and $10^{-4} - 10^{-5}$ reliability for URLLC traffic [12], [13], practical deployments are limited by queuing delays, transport-layer variability, and cross-domain orchestration overheads. The cumulative impact of these effects across access, transport, and compute layers undermines the end-to-end determinism required by closed-loop control applications. Moreover, even if radio access mechanisms were enhanced to meet such demands, the long-term scaling trajectory implied by volumetric media places additional pressure on spectrum availability, edge computing density, and joint media-network optimization. Bridging these gaps requires a shift from isolated performance optimizations to integrated architectural solutions, which is a core objective of ongoing 6G research directions.

B. *Quantitative Evaluation of 5G Limits*

We adopt a comparative modeling approach consistent with established methodologies for evaluating network-layer performance under multimedia traffic workloads [1]. Specifically, we simulate a dense-urban deployment scenario based on a single-cell Urban Macro (UMa) configuration with an Inter-Site Distance (ISD) of 500 m and ideal radio conditions, focusing exclusively on Physical Resource Block (PRB)-based resource allocation strategies in Orthogonal Frequency-Division Multiple Access (OFDMA). Two representative multimedia

TABLE I: Indicative KPI ranges as 5G stress points for emerging M-IoT.

Modality	**Throughput / Bitrate**	**Latency / Delay**	**Notes**
IMT-2020 (5G targets)	20/10 Gbps peak; 100/50 Mbps user-experienced	1 ms (URLLC, unloaded); 4 ms (eMBB)	Baseline evaluation targets [12]
URLLC (RAN studies)	Small packets (reliability-defined)	0.5 ms UL, 0.5 ms DL; $1-10^{-5}$ for 32 B within 1 ms	Radio-interface viewpoint [13]
XR (3GPP)	$[0.1-1]$ Gbps (360° video); up to 10 Gbps (XR classes)	5–20 ms class (mode-dependent)	Interaction loop dependent [10]
VR (survey)	≥ 2.35 Gbps (tiling still high)	Application-dependent	Highlights rate inflation vs HD video
Holography (survey)	$[0.3-3]$ Gbps (point clouds); $[0.1-2]$ Tbps (LFV)	< 100 ms overall delay	Includes capture, compute, and render

TABLE II: Adopted parameters for network analysis.

	Numerology	**Bandwidth (MHz)**	**SPS (kHz)**
5G Public Network	0	$[20, 50]$	15
Network Slicing	1	50	30
NPN	1	$[20, 50, 80, 100]$	30
5MBS		5	15 e.p.

services are considered, High Definition (HD) and Ultra High Definition (4K) video streaming, each associated with specific target per-user data rates [14]. The User Equipments (UEs) are incrementally added to the cell and uniformly distributed, and their peak data rate is computed by assigning each UE the minimum number of PRBs required to meet the service-specific bitrate, assuming the maximum spectral efficiency supported by the 5G Modulation and Coding Scheme (MCS) table [15].

The analysis compares four distinct network configurations commonly associated with 5G deployments. In the public network case, the available spectrum is fully shared among all UEs through unicast delivery and proportional-fair scheduling. The sliced public network introduces logical separation between groups of UEs, enforcing resource partitions without dynamic reallocation. The NPN scenario models an isolated local deployment, where a reduced number of UEs competes over a dedicated slice with full resource control. Lastly, the 5MBS configuration evaluates a one-to-many content distribution model, based on Orthogonal Frequency-Division Multiplexing (OFDM) broadcast waveforms and assuming a fixed 5 MHz bandwidth with consistent per-user efficiency. The parameters adopted for the analysis are reported in Table II.

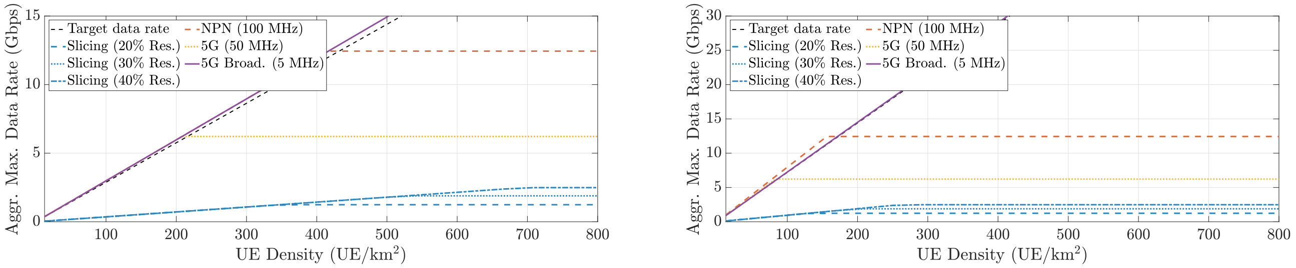
The resulting trends, shown in Fig. 2, highlight a structural mismatch between current 5G resource provisioning and the scaling needs of M-IoT video services. For both HD and 4K scenarios, the aggregate cell data rate saturates rapidly as the number of active UEs grows, while the average data rate per user sharply degrades once PRB resources are exhausted. Among the evaluated configurations, public 5G networks with unicast delivery and no slicing exhibit the steepest decline, followed by slicing-enabled setups with fixed resource partitions (i.e., 20%, 30%, 40% of the resource pool), which offer limited scalability due to rigid resource isolation. NPNs provide slightly improved performance under controlled loads, but still suffer from resource saturation as user density grows.

Notably, this degradation occurs well before reaching the simultaneous-user densities expected in real-world deployments, and is particularly severe for 4K workloads, where the higher bitrate requirements translate into a significantly larger radio resource demand per UE. These results underline the inability of unicast-based 5G configurations to support scalable multimedia delivery under high-density conditions, reinforcing the need for architectural enhancements in 6G, including broadcast-capable integration, spectrum elasticity, and media-aware traffic management. Notably, only the 5MBS-based configuration maintains stable data rate across all user densities, owing to its inherent one-to-many delivery model and fixed allocation strategy. However, even 5MBS shows structural limitations, such as the fixed bandwidth (5 MHz in this evaluation) imposes a hard ceiling on the total capacity, and its unidirectional delivery model lacks support for uplink interaction, limiting its applicability to purely broadcast-centric services.

V. TOWARD 6G ARCHITECTURES

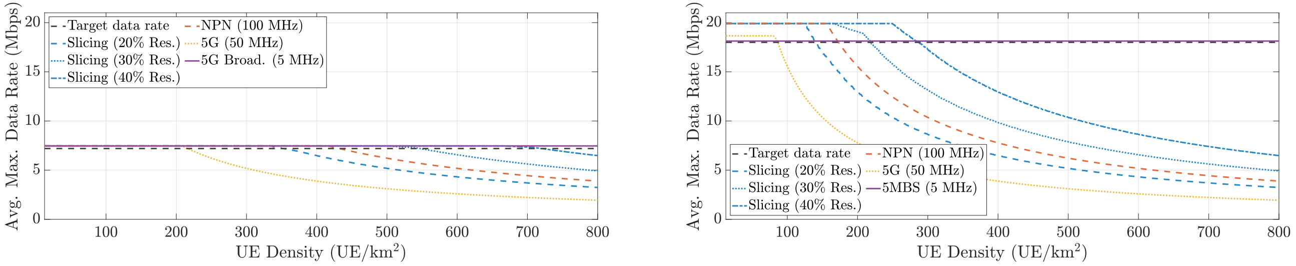
The results presented in Section III quantitatively confirm three structural stress points that current 5G deployments cannot resolve through incremental enhancements alone. First, unicast-based resource allocation saturates rapidly under growing user density, degrading per-user data rate well before reaching the simultaneous-user targets expected in dense M-IoT deployments. Second, fixed resource partitioning in network slicing provides limited scalability, as rigid isolation prevents dynamic reallocation in response to heterogeneous and bursty traffic patterns. Third, while 5MBS maintains stable aggregate data rate across all user densities, its fixed bandwidth ceiling and unidirectional delivery model make it inadequate for the asymmetric, interaction-driven traffic profiles of XR and cooperative perception services. These findings motivate a deeper architectural shift: rather than treating communication, computation, and media intelligence as separate concerns, 6G frameworks must integrate them into a unified, content-aware substrate. The remainder of this section outlines the key architectural directions through which this integration can be realized, mapping each direction to the limitations identified above.

The data rate saturation observed under unicast delivery stems partly from the inability of current 5G architectures to jointly optimize radio resource allocation and edge com-



(a) Aggregated data rate (HD)

(b) Aggregated data rate (4K)



(c) Average data rate per user (HD)

(d) Average data rate per user (4K)

Fig. 2: Simulated maximum data rate performance under 5G-like unicast and broadcast delivery for HD and 4K adaptive streaming.

putation placement. A defining principle of the 6G paradigm is therefore the convergence of communication and computing into an integrated infrastructure, enabling in-network decision-making, predictive scheduling, and content-aware optimization. This vision supports distributed intelligence directly within the network, and is especially crucial for M-IoT use cases where edge-rendering, cooperative perception, and real-time media adaptation must occur within latency budgets that traditional cloud processing cannot sustain. To this end, emerging paradigms such as joint communication-computation resource allocation, distributed training and inference over wireless links, and dynamic network function placement are expected to play a central role in enabling such tight integration [16].

The hard capacity ceiling exposed by the 5MBS configuration motivates a fundamental rethinking of spectrum usage and waveform design. At the radio access level, 6G will introduce more flexible waveform designs and spectrum usage models, potentially including sub-THz and visible light communications, thus enabling multi-hundred Gbps per-user data rates. These high-capacity channels are essential to support future workloads such as holographic streaming and ultra-multiview XR with spatialized feedback, where data rate requirements may exceed 100 Gbps per session. Furthermore, native integration of intelligent surfaces and reconfigurable radio environments (e.g., Reconfigurable Intelligent Surface (RIS)-assisted propagation) allows dynamic control of signal paths to improve coverage and reduce latency in dense deployments. Combined with enhanced Multiple-Input Multiple-Output (MIMO) schemes and full-duplex transmission capabilities, these innovations collectively extend the scalability envelope well beyond current 5G constraints [17].

The rigidity of fixed resource partitioning in network slicing calls for a fundamentally different approach to network control and orchestration. 6G is expected to embrace intent-based management and end-to-end QoS enforcement through AI-native service planes. Unlike policy-based slicing in 5G, the future control plane will rely on continuous feedback and self-optimizing loops, adapting resource allocation and routing decisions to real-time application semantics and user experience metrics. This is particularly impactful for M-IoT scenarios, where traffic patterns are inherently variable and tightly coupled with the state of sensing and actuation cycles. A distinctive innovation in 6G will be the pervasive embedding of AI and Machine Learning (ML) not only for infrastructure management, but also for content-aware adaptation and media orchestration, implying the use of learning models for dynamic bitrate adaptation, perceptual quality optimization, multimodal synchronization, and proactive flow scheduling. For instance, generative models may be employed to reconstruct immersive content from partial streams at the edge, reducing fronthaul load while preserving quality [18]. Reinforcement learning can guide resource allocation based on real-time Quality of Experience (QoE) estimates, adapting to varying network conditions and user interactions.

Interoperability is another foundational concern for 6G, especially as mobile systems expand into aerial, underwater, and space-based domains. The heterogeneity in frequency bands, addressing schemes, protocol stacks, and device types will require modular interfaces, AI-assisted translation layers, and hybrid protocol architectures [17]. Mechanisms such as Software-Defined Networking (SDN), Network Function Virtualization (NFV), and blockchain-based trust systems are expected to provide the foundation for flexible integration across vertical domains, including healthcare, mobility, and industrial media applications. Additionally, the concept of *network-as-a-sensor*, where the infrastructure continuously collects context data, opens new avenues for situation-aware streaming, adaptive security, and predictive media delivery. Finally, 6G architectures are expected to deliver natively synchronized, ultra-reliable services by embedding support for Time Sensitive Networking (TSN) and precision timing across all layers, a feature crucial for enabling tightly coordinated multi-modal media flows such as those required in distributed XR rendering or remote haptic interactions.

The 6G transformation represents more than a capacity boost. It redefines the network's role as an active, intelligent mediator of media-rich, feedback-driven services. For the M-IoT domain, this implies not only the resolution of current data rate and latency limitations, but also the emergence of a new service paradigm where the infrastructure actively collaborates with media processing pipelines to deliver scalable, immersive, and responsive experiences.

VI. CONCLUSIONS AND FUTURE DIRECTIONS

This paper has examined the evolving landscape of M-IoT services, highlighting how emerging applications such as immersive XR, cooperative perception, and real-time holography push mobile infrastructures far beyond traditional throughput and latency boundaries. Through a layered architectural perspective and a quantitative analysis of representative streaming workloads, we have shown that current 5G deployments, despite advanced features such as slicing, Multi-access Edge Computing (MEC), and multicast, are subject to structural limitations when exposed to the heterogeneous, asymmetric, and tightly coupled traffic patterns typical of M-IoT. These limitations manifest as scalability bottlenecks, latency degradation, and suboptimal resource utilization, particularly under dense user loads and real-time constraints.

Looking ahead, 6G represents a necessary architectural shift to address these challenges. Beyond raw capacity improvements, future networks must embrace native support for compute–communication co-design, semantic-aware traffic prioritization, and end-to-end service-level orchestration. This evolution will be crucial for enabling scalable delivery of multi-gigabit volumetric media, millisecond-level haptic feedback, and context-adaptive XR experiences. Furthermore, the role of AI-driven network functions and programmable data planes will become central to dynamically adapt to the service demands of M-IoT environments. These directions mark the transition from infrastructure-aware applications to application-aware infrastructures, positioning 6G not only as a faster network, but as an enabling fabric for next-generation media-centric IoT systems.

ACKNOWLEDGMENT

This work has been partially supported by the European Union through SNS Joint Undertaken SEASON project (GA: 101096120), and by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, through National Innovation Ecosystem (ECS00000041 - VITALITY - CUP E13C22001060006); and Mission 4, Component 2, Investment 1.3, CUP E83C22004640001, CUP D93C22000910001, CUP D43C22003080001, and CUP F83C22001690001 partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART").

REFERENCES

[1] C. Rinaldi, A. Piccioni, C. Centofanti, A. Marotta, P. Di Marco, and F. Graziosi, "Architectural Challenges and Solutions in Multimedia IoT Over 5G Networks," *IEEE Transactions on Broadcasting*, vol. 72, no. 1, pp. 46–61, 2026.

[2] G. Yang, M. A. Jan, A. U. Rehman, M. Babar, M. M. Aimal, and S. Verma, "Interoperability and Data Storage in Internet of Multimedia Things: Investigating Current Trends, Research Challenges and Future Directions," *IEEE Access*, vol. 8, pp. 124 382–124 401, 2020.

[3] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[4] X. Lin, "An Overview of 5G Advanced Evolution in 3GPP Release 18," *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 77–83, 2022.

[5] 3GPP, "Release 18 Description; Summary of Rel-18 Work Items," 3rd Generation Partnership Project (3GPP), Technical Report TR 21.918, 2025, v18.0.0 (2025-04).

[6] ITU-R, "Framework and Overall Objectives of the Future Development of IMT for 2030 and Beyond," International Telecommunication Union (ITU), Recommendation ITU-R M.2160-0, 2023. [Online]. Available: https://www.itu.int/rec/R-REC-M.2160-0-202311-I

[7] ——, "Future Technology Trends of Terrestrial International Mobile Telecommunications Systems Towards 2030 and Beyond," International Telecommunication Union (ITU), ITU-R Report M.2516-0, 2022. [Online]. Available: https://www.itu.int/pub/R-REP-M.2516-2022

[8] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[9] 3GPP, "Support of 5G Glass-Type Augmented Reality / Mixed Reality (AR/MR) Devices," 3rd Generation Partnership Project (3GPP), Technical Report TR 26.998, 2022, release 17 (e.g., V17.0.0).

[10] ——, "Extended Reality (XR) in 5G," 3rd Generation Partnership Project (3GPP), Technical Report TR 26.928, 2021, release 16 (e.g., V16.1.0).

[11] Z. Bai *et al.*, "A Survey and Framework of Cooperative Perception: From Heterogeneous Singleton to Hierarchical Cooperation," *IEEE Transactions on Intelligent Transportation Systems*, 2024.

[12] ITU-R, "Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s)," International Telecommunication Union, Tech. Rep. Rep. ITU-R M.2410-0, Nov. 2017, accessed: 2025-12-23.

[13] 3GPP, "Study on Scenarios and Requirements for Next Generation Access Technologies," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 38.913, Rel-17, V17.0.0, May 2022, published by ETSI as TR 138 913; Accessed: 2025-12-23.

[14] Cisco, "Cisco annual internet report (2018–2023)," Cisco, Tech. Rep., 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[15] 3GPP, "Physical Layer Procedures for Data (Release 18)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.214, 1 2024, Version 18.0.0.

[16] M. Alsader, A. A. Barakabitze, and I.-H. Mkwawa, "QoE-Driven Adaptive Video Streaming: Architectures, Techniques, and Future Research Challenges Toward 6G Networks," *IEEE Access*, vol. 13, pp. 157 408–157 441, 2025.

[17] S. Kharche and P. Dere, "Interoperability Issues and Challenges in 6G Networks," *Journal of Mobile Multimedia*, vol. 18, no. 5, pp. 1445–1470, 2022.

[18] M. Sheraz, T. C. Chuah, W. U. K. Tareen, A. Al-Habashna, S. I. Saeed, M. Ahmed, I. E. Lee, and M. Guizani, "A comprehensive survey on genai-enabled 6g: Technologies, challenges, and future research avenues," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 4563–4590, 2025.