

From Network Visibility to Network Intelligence: Modernizing Telecom for Real-Time Decision Making in a 5G and Edge Era

Shandilya Avadhanam Venkata Krishna Sastry

Y&L Consulting Inc., USA

ARTICLE INFO

ABSTRACT

Received: 08 Dec 2025

Revised: 19 Jan 2026

Accepted: 29 Jan 2026

The telecommunications industry is at an inflection point where outdated network visibility tools cannot effectively support the new 5G and edge computing architectures. The industry needs to shift from reactive systems that run on demand to automated smart systems that can make decisions in real time across a distributed core and edge infrastructure. Modern fifth-generation wireless technologies overlap with edge computing and introduce further challenges due to dynamic workload placement, super-low latency, and massive connectivity. Legacy systems cannot manage these workloads. This allows AI/ML-driven clever systems to enable operators to predict the deviation events, optimize resources, and cleverly orchestrate services proactively instead of reactively after they impact the customer. Cloud-native ingestion pipelines and distributed intelligence layers handle scales and speeds of telemetry data that humans, unaided, cannot analyze. They readily transform raw metrics into understanding within an operationally meaningful timeframe. Network slicing and automated orchestration allow operators to slice isolated, optimized virtual networks for different service needs from a shared physical infrastructure while exercising resource efficiency. Adopting the resilience engineering principles of stateless services, graceful degradation, and autonomous edge operations in a connectivity platform will reduce service degradation when components fail or connectivity is disrupted. Moving from infrastructure management to experience orchestration is a business imperative as well as a programmatic evolution; it will enable new revenue opportunities, cost efficiencies, and the transformation of operators from commodity connectivity providers into platforms of innovation in a digital services economy.

Keywords: 5G Network Intelligence, Edge Computing Orchestration, Real-Time Network Automation, Artificial Intelligence Telecommunications, Distributed Network Architecture

1. Introduction: The End of Predictable Networks

Telecommunication networks that have predictable patterns of traffic evolve in a slow, planned way and rely on predicted traffic patterns for capacity planning. Today, the network needs to simultaneously support real-time digital experiences, massive device density, localized processing, and very low tolerance for downtime. 5G network architecture is a progression from previous generations of mobile broadband networks toward support of a wider range of services and use cases, such as massive machine-type communications and ultra-reliable low-latency communications, thus requiring improved and more complex management than has previously been used in mobile communications networks [1]. Edge computing is a model that overcomes some of the limitations of centralized cloud computing by moving computation and data closer to the point of need, reducing latency and bandwidth utilization, and enabling new classes of applications that require real-time or near-real-time data processing [2].

Yet many telecom operators continue to rely on legacy architectures and happen to miss out on actual opportunities to leverage enormous amounts of network data telemetry, logs, alarms, analytics, and performance measurements. Operationally, fault detection after customer impact, capacity threat detection after a breach occurs, and root cause detection with different tools, people, and handovers are becoming increasingly impractical with the advent of network slicing and edge workloads. Information is available in big data, but it lacks the real-time intelligence and orchestration across core, edge, and radio to convert big data into intelligence in an operationally useful time frame.

2. The Fundamental Shift: Why Visibility Alone Fails in 5G Networks

Customary monitoring solutions had provided visibility into individual elements in the network, whereas with 5G, traffic is no longer centralized, and workloads move between core and edge. The architecture on which 5G networks are built is highly complex from a monitoring point of view, with many radio access technologies, cloud-native network functions, and software-defined networking principles across multiple domains, and it constitutes a clear generational shift compared to previous generations of cellular network technologies. 5G networks are based on a service-based architecture (SBA) and a different approach to network function chaining, in which functions are decoupled from specific hardware boxes and created dynamically, which increases the complexity beyond what traditional network monitoring methods can manage [1]. The network must support use cases with completely unique requirements, such as improved mobile broadband with high throughput and ultra-reliable low-latency communications with response times in the sub-millisecond range.

Latency-sensitive applications such as autonomous systems, industrial IoT applications, augmented and virtual realities, and smart cities need to make decisions on the order of milliseconds and thus are incompatible with centralized monitoring and control architectures. Mobile edge computing (MEC) extends cloud-computing capabilities to the edge of the mobile network, thereby bringing computation and storage resources closer to the mobile end-users and the mobile IoT devices, as it is designed to reduce end-to-end delay and backhaul bandwidth. [2] With such a distributed architecture, there are exponentially more data points to monitor than in a customary network. The correlations are beyond the capabilities of humans to analyze. Dashboards and centralized alerting fail because they have no context, correlation, or predictive capabilities for distributed topologies. Signals from radios, edge nodes, transport layers, and applications all need to be interpreted together in real time.

Customary software and hardware architectures are not designed to support real-time ingestion, correlation, and analysis of data streams coming from any number of sources. Latency and the necessity to make decisions at the edge of the network, including the data center (centralized edge), more regional Data centers (regional edge) and micro-edge deployments at the cell site pose additional challenges. The downside of this solution is that operators are left with manual triaging, centralized remediation, and network operations center personnel end up spending most of their time correlating and prioritizing alerts and alarms rather than resolving them.

Traditional Monitoring	Intelligent 5G Networks
Centralized dashboards and alerts	Distributed intelligence at the core and edge
Reactive fault detection after customer impact	Proactive anomaly prediction before service degradation
Manual alert correlation and triage	AI-driven automated correlation across network layers

Limited context across network domains	Holistic context spanning radio, transport, core, and edge
Processing thousands of events per second	Processing millions of events per second in real-time
Minutes to hours for issue detection	Sub-second to millisecond decision latency
Single-network-tier architecture	Multi-tier distributed processing architecture
Static capacity planning based on historical trends	Dynamic resource allocation based on predictive models

Table 1: Evolution from Visibility to Intelligence in 5G Networks [3, 4]

3. Platform-Centric Architecture: Building Intelligence into the Network Fabric

For the full value of 5G and edge computing, telecommunications operators will require an intelligence architecture that spans core and edge environments. The platform must provide cloud-native ingestion pipelines that can integrate high-frequency telemetry from intermittent communication and latency-sensitive environments. Edge-aware ingestion can apply first-pass analytics at the edge for high-value data and transmit less critical data back to centralized analytics engines. The convergence of mobile edge computing and 5G creates new opportunities for service delivery and new business models by using distributed computing resources to support latency-sensitive applications while optimizing resources across multiple network layers [3]. The modern architecture of a network must include data processing capabilities at different layers, from real-time stream processing in the edge to resource-hungry batch analytics in the core, thus forming a hierarchy of intelligence that balances latency requirements, processing complexity, and available computing resources.

Sitting above this ingestion layer is the intelligence layer powered by artificial intelligence, where machine learning models can be applied to time, location, slices, and services to identify congestion and degradation of the radio and perform anomaly detection before affecting service level agreements. The application of artificial intelligence/machine learning network operations has transitioned from reactive to predictive, where predictive models can analyze network conditions and stimulate proactive actions to prevent service degradation. [4] More advanced versions combine models, such as time-series forecasting (to define capacity forecast), graph neural networks (to analyze network topology), and deep learning (to determine correlation patterns across multiple modalities from multiple inputs). Intelligence can be centralized or distributed: some decisions are taken centrally for global optimization and capacity planning purposes, while others are made at the edge to steer traffic and quality of service locally.

The hybrid intelligence architecture can reduce the decision-making latency compared to centralized architectures by supporting localized decision-making at the edge nodes, while still enabling the central Control planes optimize edge decisions through a coordinated architecture. The platform can support continual learning, supporting model training on historical data sets in central facilities and deploying the trained models onto the network edge for inference, while enabling model updates and auditing to account for changes in network conditions. By embedding AI within all aspects of the network life cycle, from planning through optimization to troubleshooting, a network operator can handle complexity at scale, drive down operations costs, and improve service quality. The smart layer also helps explain decisions so a network operator can see how a choice was made, trust the automation, and know when human involvement is needed for important network tasks.

Architecture Layer	Key Capabilities	Operational Benefits
Cloud-Native Ingestion	High-frequency telemetry processing, edge-aware filtering, intermittent connectivity tolerance	Bandwidth reduction while maintaining signal fidelity
AI-Driven Intelligence	Time-series forecasting, graph neural networks, and multi-modal correlation	Congestion prediction, radio degradation identification, anomaly detection
Distributed Decision-Making	Central optimization for network-wide planning, edge autonomy for local adjustments	Reduced decision latency, localized traffic steering
Model Management	Centralized training, edge deployment, continuous learning	Accuracy maintenance as conditions evolve
Explainability Features	Decision transparency, human oversight capabilities	Trust building in automated systems
Integration Layer	OSS/BSS connectivity, orchestration framework integration	Seamless information flow across domains

Table 2: Platform Architecture Components for Real-Time Network Intelligence [5, 6]

4. Orchestration: The Defining Capability of Modern Telecom Operations

Orchestration is the primary technology that enables 5G, as orchestration engines can be used to coordinate slices, edge workloads, and transport paths. Traffic steering, edge resource scalability based on localized demand, and the ability to adapt, prioritize, and throttle network slices dynamically at runtime according to service requirements are made possible. Cloud Radio Access Network (Cloud RAN) deployment enables centralized baseband processing functions, allowing coordination and dynamic sharing of resources between multiple cell sites in near-real-time to support network features such as coordinated multipoint (CoMP) and interference management [5]. It has enabled clever orchestration with optimization of radio resources over a large geographical area, smart traffic management across cells, and tuning of the transmission parameters to maximize the user experience and spectral efficiency. Today's orchestration platforms often include multiple domains, such as the radio access network (RAN) configuration, transport network path computation, core network functions lifecycle management, and edge computing resource allocation, across the entire network.

Additionally, the complexity of the setup, management, and optimization of 5G networks is increased by network slicing, which allows the configuration of multiple virtual networks to meet the needs of different services using the same physical infrastructure. A new architectural concept supported by the 5G system, is "network slicing," through which network resources are partitioned into isolated virtual networks, each Each network slice is dedicated to a specific use case and offers the corresponding service capabilities and quality of service guarantees [6]. The orchestration system should also ensure the lifecycle management of the network slices, including instantiation, auto-scaling in accordance with traffic load, and decommissioning of network slice instances. Resource management also includes resource optimization and isolation, ensuring resource allocation and interference control between slices. Service level agreements can be enforced using continuous monitoring, dynamic resource allocation, and closed-loop control strategies that react to deviations from performance targets by automatically adapting the deployed resources.

This allows operators to move from a reactive, manual method of detecting problems to proactive policy-based automation, enabling the orchestration of a service spanning multiple network elements to meet and maintain service level agreements. Advanced orchestration of networks enables zero-touch provisioning and closed-loop automation of network services based on real-time information about customer experience and can be performed without human involvement. The orchestration platform needs to communicate with the operational and business support systems and expose an application programming interface (API) that third-party applications could use to request network services, ensuring performance and supporting new business models.

Orchestration Domain	Functions	Transformation Impact
Network Slicing	Slice instantiation, resource isolation, lifecycle management	Multiple logical networks on shared infrastructure
Radio Access Network	Configuration management, coordinated multipoint transmission, and interference management	Spectral efficiency optimization across cell sites
Transport Network	Dynamic path computation, traffic rerouting, and quality of service enforcement	Sub-second convergence for traffic redirection
Core Network Functions	Virtualized function lifecycle, scaling operations, policy enforcement	Minutes instead of weeks for service provisioning
Edge Computing Resources	Container orchestration, localized workload management, and demand-based scaling	Rapid deployment cycles for edge applications
Service-Level Agreements	Continuous monitoring, dynamic resource adjustment, closed-loop control	Proactive compliance instead of reactive breach response
Zero-Touch Provisioning	Automated instantiation, self-optimization, autonomous remediation	Operational complexity reduction with improved consistency

Table 3: Orchestration Capabilities Across Network Domains [7, 8]

5. Engineering for Reality: Scale, Resilience, and Ultra-Low Latency

5G and edge computing create new needs for on-demand capacity provisioning to handle traffic spikes, localized outages, and temporary services that are always starting and stopping. These demands necessitate architectural approaches that can adapt to ongoing changes. The need of stateless cloud-native technologies for asynchronous messaging, distributed data and processing, and design-based resilience. Modern platforms should also adopt microservices architectures and containerization technologies to ease the deployment and scaling of network functions and support the capability of networked intelligence. Orchestration systems should be able to instantiate a new service instance within seconds [7]. The failure rate of infrastructure components is also expected to be higher than customary data centers due to distributed deployment, heterogeneity of environments, and the geographical location in which these components are deployed. Such an outcome requires the use of complex fault tolerance mechanisms to ensure service continues in the event of component failure.

It should be possible to design systems for partial failure and graceful degradation. For example, an edge platform should be able to make decisions and conduct missions if disconnected from a central control system. Data pipelines should use priority queuing, adaptive sampling rates, and smart data

retention policies that keep important operational data while dropping less important data when the network is congested or unavailable [8]. Modern resilience engineering is adding redundancy throughout the stack, from compute to multi-path networking to replicated data stores that can be restored quickly. Operations at edge platforms require them to rely on local policy engines, as well as a locally cached state of configuration and management data, and distributed consensus management protocols to maintain continuity of service, even for extended periods, when disconnected from centralized systems.

Finally, integration becomes vitally important when network intelligence platforms use standard application programming interfaces (APIs) to connect to operational support systems (OSS), business support systems (BSS), edge orchestration systems, and partner ecosystems. Integration capabilities need to support high transactional and low-latency processing of data to ensure timely orchestration of diverse network functions and external systems. Without these integrations, insights remain trapped within their silos, and operational automation is not possible. Operators cannot reap the full benefit of clever network operations. On the other hand, organizations with integrated platforms arrive to market faster and operate more efficiently because data can flow across organizational and technology boundaries smoothly. As such, it must support both synchronous request-response for interactive operations and asynchronous event-driven architectures for monitoring and analytics, as well as any other integration patterns that may be required in the network ecosystem.

Engineering Principle	Implementation Approach	Operational Outcome
Cloud-Native Architecture	Microservices, containerization, stateless services	Rapid deployment and scaling capabilities
Fault Tolerance	Compute redundancy, multi-path connectivity, and distributed replication	Service continuity during component failures
Graceful Degradation	Priority-based queuing, adaptive sampling, and intelligent retention	Critical data preservation under resource constraints
Autonomous Edge Operation	Local policy engines, cached configuration, distributed consensus	Extended service continuity without central connectivity
Asynchronous Communication	Event-driven architectures, message queuing, decoupled services	Reduced coupling and improved system resilience
High Availability Design	Redundancy at multiple layers, rapid recovery mechanisms	Minimal downtime and quick restoration
Integration Resilience	Standardized APIs, multiple communication patterns, flexible protocols	Seamless coordination across diverse systems
Adaptive Resource Management	Dynamic allocation, demand-responsive scaling, load balancing	Optimal utilization under variable conditions

Table 4: Resilience and Scale Engineering Principles [9, 10]

Conclusion

This report presents a major model shift for the telecommunications operators, from maintaining a passive network business focused on commodities to being a systems integrator of clever and adaptive systems that deliver real-time digital experiences. Fifth-generation wireless networks, natively distributed by design, along with an edge computing framework, present new opportunities and challenges that existing visibility and monitoring models cannot address. Artificial intelligence at each network stack level initiates remediation before service becomes perceptibly degraded, reducing

mean-time-to-repair and avoiding unnecessary manual intervention for operational efficiencies, including predictive maintenance. Network slicing enables known, repeatable service outcomes, with high levels of compliance with service level agreements (SLAs) globally, making customers confident in their network capabilities. Operators creating smart technology systems are now seen as sources of innovation, helping to support fast, critical applications and a large number of connected devices, along with business services that meet promised service levels. Companies that start using these new features are already seeing significant increases in revenue, not just from new services that 5G allows, but also from becoming more competitive in business markets, which lets them charge more for better performance. Businesses need to move from centralized, reactive, and often manual operations to smart, self-managing systems, as this is essential for their success, not just a technical upgrade, to help operators thrive as networks become vital for the digital transformation of the entire economy. These companies will realize the true potential of 5G and edge computing as innovation and reliability platforms for sustainable growth in competitive telecommunications markets with rising customer expectations.

References

- [1] Mansoor Shafi et al., "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," IEEE Journal on Selected Areas in Communications, 2017. Available: <https://ieeexplore.ieee.org/document/7894280>
- [2] Weisong Shi et al., "Edge Computing: Vision and Challenges," IEEE Internet of Things Journal, 2016. Available: <https://ieeexplore.ieee.org/document/7488250>
- [3] Petar Popovski et al., "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View," IEEE Access, 2018. Available: <https://ieeexplore.ieee.org/document/8476595>
- [4] Yuyi Mao et al., "A Survey on Mobile Edge Computing: The Communication Perspective," IEEE Communications Surveys & Tutorials, 2017. Available: <https://www.researchgate.net/publication/319299183>
- [5] Aleksandra Checko et al., "Cloud RAN for Mobile Networks—A Technology Overview," IEEE Communications Surveys & Tutorials, 2015. Available: <https://www.researchgate.net/publication/273706693>
- [6] Xenofon Foukas et al., "Network Slicing in 5G: Survey and Challenges," IEEE Communications Magazine, 2017. Available: https://www.researchgate.net/publication/316903300_Network_Slicing_in_5G_Survey_and_Challenges
- [7] Shuo Wang et al., "A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications," IEEE Access, 2017. Available: <https://ieeexplore.ieee.org/document/7883826>
- [8] Mobasshir Mahbub and Raed M. Shubair, "Contemporary advances in multi-access edge computing: A survey of fundamentals, architecture, technologies, deployment cases, security, challenges, and directions," Journal of Network and Computer Applications, 2023. Available: <https://www.sciencedirect.com/science/article/abs/pii/S1084804523001455>
- [9] Rongpeng Li et al., "Intelligent 5G: When Cellular Networks Meet Artificial Intelligence," IEEE Wireless Communications, 2017. doi: 10.1109/ICWIN.2015.7057932. Available: <https://ieeexplore.ieee.org/document/7886994>
- [10] Nutanix, "What is Edge Computing?" 2025. Available: <https://www.nutanix.com/info/cloud-computing/edge-computing>