International Journal of
**Technology and
Systems**
(IJTS)

Significant Advances in Application Resiliency: The Data Engineering Perspective on Network Performance Metrics

Significant Advances in Application Resiliency: The Data Engineering Perspective on Network Performance Metrics

Jayanna Hallur

Sr Lead Engineer in Observability Engineering and Data Architect, Richmond, VA

https://orcid.org/0009-0007-9789-2672

Accepted: 23rd Nov, 2024, Received in Revised Form: 4th Dec, 2024, Published: 17th Dec, 2024

Abstract

Purpose: The purpose of the article is to explore how network metrics like latency, packet loss, and throughput, combined with application logs, can help organizations improve the reliability and performance of their applications. It focuses on how these insights support Site Reliability Engineering (SRE) teams in proactively addressing issues to achieve better application resiliency, enhancing user experience and market trust.

Methodology: The article uses a combination of case studies and analysis to demonstrate how monitoring specific network metrics and application logs helps identify and resolve performance issues. It examines real-world scenarios where proactive adjustments based on these metrics improved application reliability and aligned with organizational objectives.

Findings: The findings show that by analyzing network metrics and application logs, organizations can pinpoint causes of transaction failures, such as high latency, packet loss, or misconfigured firewalls. Proactive resolutions based on these insights result in smoother application performance, reduced downtime, and increased user satisfaction.

Unique Contribution to theory, practice and policy: This article makes valuable contributions to theory, practice, and policy. For theory, it expands the understanding of how network metrics and application logs can work together to improve application resiliency, offering a framework for integrating Site Reliability Engineering (SRE) principles with network observability. For practice, it provides clear, actionable steps for SRE teams to identify and resolve performance issues, helping organizations enhance reliability and user satisfaction. For policy, it highlights the importance of proactive network monitoring and metric-driven decision-making, encouraging organizations to adopt policies that prioritize resiliency, ensure consistent performance, and meet service-level agreements (SLAs).

Keywords: *Data Engineering, Network Observability, Telemetry, Reliability, Resiliency, Scalability, FCAPS*

I. Introduction

In today's interconnected digital world, network devices such as switches, routers, and firewalls serve as the backbone of IT infrastructure, supporting the seamless flow of data across complex networks. The performance of these devices directly impacts the reliability and efficiency of applications, including those hosted in cloud environments and accessed through mobile platforms. As organizations increasingly depend on real-time and cloud-based services, maintaining optimal network performance has become critical to ensuring a high-quality user experience [1].

Network performance metrics have emerged as a key focus area in this context, enabling deeper visibility into both network conditions and application behaviors. By analyzing performance metrics like latency, throughput, jitter, and packet loss, network administrators can gain valuable insights into the health of their network infrastructure. These metrics are essential for identifying bottlenecks and optimizing data flow, thereby enhancing application performance. For example, managing latency and jitter is particularly crucial for real-time applications such as video conferencing and VoIP, where even small delays can degrade the user experience

Equally important is the role of fault management in maintaining network stability. Monitoring metrics such as interface status, power supply conditions, and device health allows administrators to detect hardware issues early and prevent unexpected failures. This proactive approach helps minimize downtime and maintain continuous service availability [2]. Additionally, integrating AI-based anomaly detection tools has further enhanced the ability to identify and respond to network anomalies in real time, providing an edge in maintaining robust network operations [1].

This paper explores the importance of these performance management metrics, emphasizing their role in achieving network observability and ensuring consistent application availability. It also highlights how a data-driven approach to network management can support organizational goals, such as operational efficiency and customer satisfaction, by providing timely insights into network conditions. As the demand for digital services continues to grow, understanding and leveraging these metrics is essential for organizations aiming to maintain a competitive edge in today's digital landscape.

II. Monitoring

Monitoring of applications is very crucial to ensure the functionalities of applications are providing seamless experience to the customers, ensuring high reliability, setting up automatic dynamic scaling, identify the opportunity for performance improvements, tracking the user experiences for all the geographical users, security and compliance, early detection of issues, sending notifications on detections of anomalies and many more [2]. Currently monitoring capabilities are enabled at individual applications and the respective teams to take the required actions for their applications for functionalities stated above. Observability enhances monitoring by providing a more comprehensive view of an application's behavior and state. Observability focuses on understanding the internal states of a system based on the data it generates, such as logs, metrics, and traces, and enabling proactive troubleshooting and optimizations. Network observability is a comprehensive approach to monitoring and understanding the state and behavior of the network infrastructure in real-time. It provides deeper insights and remediation capabilities than traditional network monitoring, going beyond metrics like uptime, latency, and packet loss, into how data flows through the network, the interactions between different network components like switches, APs, firewalls, and applications, and the overall health of the network environment [1].

As organizations offer more and more services to their customers or expansion of their services, it requires managing the increasing complexity of the distribution systems. The traditional monitoring [2] provides insufficient capabilities for maintaining highest reliability required for the increasing modern business and customer satisfaction. Observability provides deeper insights into system behavior, allowing teams to detect, diagnose, and resolve issues more efficiently [3].

III. OBSERVABILITY

Monitoring helps to understand if something is wrong, while observability helps to understand why it is wrong. Together, they provide a robust strategy for managing and maintaining modern complex systems.

While monitoring involves tracking predefined metrics and alerting on specific thresholds, observability delves deeper by analyzing comprehensive data from the network to understand the underlying state and causes of issues. Observability can be seen as a superset of monitoring, encompassing the collection and correlation of metrics, logs, traces, and events. Refer the scholarly article [4] about enhancing system reliability and team productivity by adapting observability in the organization. Observability transforms the traditional monitoring by focusing on the internal state of each system and services. The observability enables teams to check beyond the predefined metrics. This shift is very crucial for handling the complexity of modern systems.

IV. DATA ENGINEERING FOR NETWORK PERFORMANCE METRICS FOR RESILIENCY

In today's connected world, network performance equates directly to business outcomes. Networks that are slow, unreliable, or insecure will lose revenue, reduce productivity, and ultimately damage the reputation of the company. Observability allows organizations to identify and respond to issues, anticipate and prevent them, while improving overall user experience and operational efficiency.

Network observability means an organization's ability to deeply understand the state of a network infrastructure and applications through continuous monitoring for performance, health, and security. In contrast to traditional monitoring, which notifies an administrator only in cases of trouble, observability extends the view to include proactive management and troubleshooting. The need for observability will be driven by increasingly complex networks created by the incorporation of cloud services, IoT devices, and remote workforces that demand high network uptime and performance.

A. Performance management of network devices

Network device management is a critical aspect of ensuring the reliability, efficiency, and security of a network. The FCAPS model, which stands for Fault, Configuration, Accounting, Performance, and Security management, is a well-established framework used to manage network devices. Among these, Performance Management and Fault Management play crucial roles in enhancing network observability to improve the resiliency of the enterprise applications, which, in turn, significantly improves application monitoring, customer experience, availability of critical business functionals, and overall network health.

Performance management of network devices focuses on measuring, analyzing, and optimizing the performance of network devices.

- Performance Monitoring: Continuous tracking of metrics like CPU usage, memory usage, latency, and throughput.
- Trend Analysis: Analyzing historical performance data to predict future behavior and capacity needs.
- Optimization: Adjusting configurations and resources to improve device performance and network efficiency.

Examples of Performance Management in Network Devices:

- Bandwidth Monitoring: Tracking utilization to ensure critical applications have sufficient resources.
- Latency Measurement: Monitoring delays to ensure real-time applications operate smoothly.
- Resource Utilization: Identifying devices that are over or underutilized, enabling better resource allocation.
- Network Traffic Metrics: Monitoring the total incoming, outgoing, packets dropped, and queuing in equipment.

The benefits of network device performance management include the following:

- Better Network Efficiency: The performance data ensures that continuous optimization is performed and the network resources are put to good use.
- Enhanced User Experience: Application performance is highly improved due to reduced latency and minimization of packet loss.
- Predictive Maintenance: Trend analysis helps in giving insight into performance issues that will pop up so that remedial actions may be taken in time.

B. Adapting Network Device Performance Metrics for applications resiliency

Building applications resiliency using health network devices relies on key performance indicators to provide actionable insights into the behavior, health, and performance of the network. These metrics are necessary for pinpointing bottlenecks, troubleshooting issues, and ensuring optimal service delivery. The following are the core metrics underpinning effective network observability:

Throughput

Throughput measures the amount of data successfully transmitted through a network device within a given time frame, typically in Mbps or Gbps. High throughput is essential for bandwidth-intensive applications such as video streaming or large file transfers[6].
Example: A router sustaining 1 Gbps throughput during peak hours indicates it is performing well. A sudden drop to 500 Mbps may highlight congestion, insufficient hardware capacity, or packet drops due to interface errors.
Indicator: Consistent high throughput signifies optimal performance, while significant deviations from expected throughput suggest a performance bottleneck or degraded network conditions.

Bandwidth Utilization

Bandwidth utilization represents the percentage of available bandwidth being consumed by the network device. Monitoring this metric helps in capacity planning and identifying congestion points.
Example: A switch consistently operating at 60-70% utilization is well within its capacity. If utilization climbs to 95% or higher, it may lead to queuing delays, packet drops, or jitter for real-time services.
Indicator: Moderate utilization reflects efficient operation, whereas sustained high utilization signals a risk of congestion and potential performance degradation.

Packet Processing Rate

The packet processing rate indicates how many packets a network device can handle per second. It reflects the efficiency and capacity of the device in managing traffic load[7].
Example: A firewall designed to process 1 million packets per second maintains performance under normal traffic. However, exceeding this limit during a DDoS attack results in delays and packet drops.
Indicator: A high and consistent packet processing rate shows good performance, whereas a declining rate under increasing loads suggests device stress or insufficient resources.

Latency

Latency measures the time a network device takes to process and forward packets. It directly impacts the performance of time-sensitive applications like VoIP and online gaming.
Example: A firewall with a latency of 5 ms under normal load performs as expected. If latency spikes to 50 ms due to high CPU utilization or complex rule sets, user experience degrades significantly.

Indicator: Low and consistent latency reflects optimal performance, while increasing latency indicates device overload or misconfiguration.

CPU Utilization

CPU utilization measures the percentage of processing power a device uses while performing tasks like routing, switching, or deep packet inspection[8]. **Example**: A router operating at 60% CPU utilization under normal traffic is healthy, but if usage spikes to 90% during a traffic surge, it may result in packet delays or processing failures. **Indicator**: Moderate CPU utilization indicates adequate performance, while consistently high utilization can lead to performance degradation and potential device failure.

Memory Utilization

Memory utilization tracks how much of the device's memory is being used for tasks like maintaining routing tables or NAT sessions. **Example**: A switch with 80% memory utilization is efficiently managing its resources. However, exceeding 95% may cause dropped packets or inability to establish new sessions. **Indicator**: Low to moderate memory usage signals healthy performance, while high usage risks overflows and operational inefficiencies.

Packet Loss

Packet loss is the percentage of data packets that fail to reach their destination. It impacts the reliability of applications like video conferencing or financial trading platforms[7]. **Example**: A router experiencing 1% packet loss under heavy traffic may still perform acceptably. However, 5% packet loss during a live trading session could disrupt critical operations. **Indicator**: Minimal or no packet loss indicates good performance, while increasing loss rates suggest congestion, hardware failure, or misconfigured interfaces.

Jitter

Jitter represents the variability in packet arrival times, which affects the quality of real-time applications like VoIP and video conferencing[10]. **Example**: A call center VoIP system with jitter below 20 ms ensures clear audio quality. Jitter exceeding 50 ms due to network congestion results in choppy or delayed communication. **Indicator**: Low jitter reflects stable performance, while high jitter signals potential issues with traffic prioritization or overloaded devices.

Interface Utilization

Interface utilization measures the bandwidth usage on individual device ports or interfaces. It helps in detecting overloaded or underutilized connections[9]. **Example**: A switch port operating at 80% utilization during peak hours is efficient, but sustained utilization above 95% risks congestion and queuing delays.

Indicator: Balanced utilization across interfaces indicates good performance, while high utilization on specific interfaces may highlight a need for traffic redistribution or capacity upgrades.

These metrics, collectively, provide a holistic view of network device performance, enabling proactive management and ensuring the network supports organizational objectives effectively.

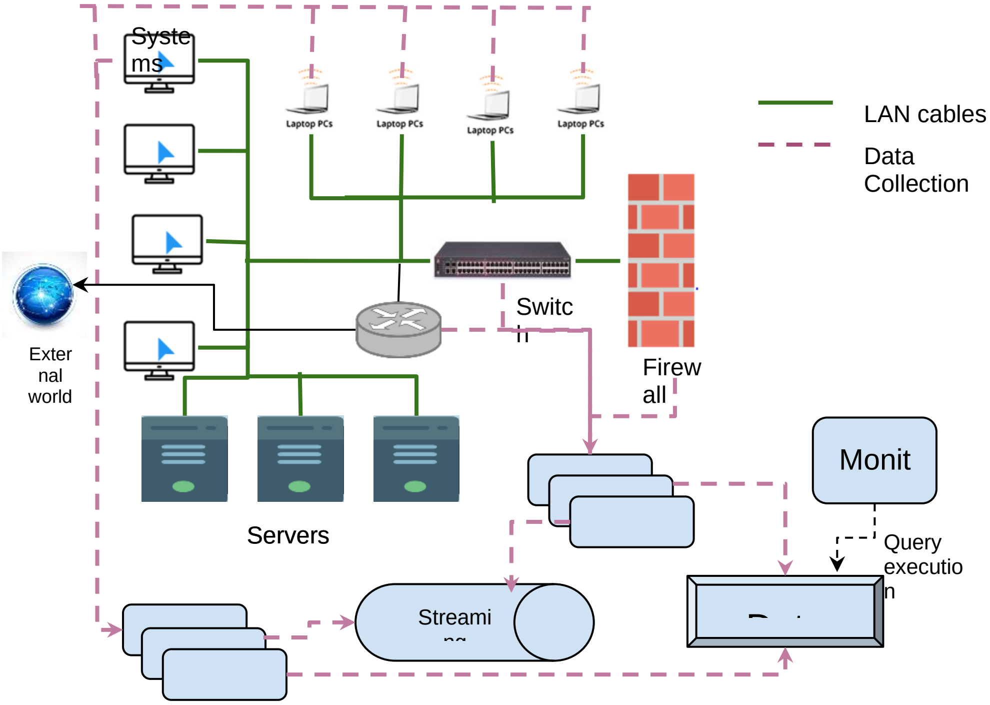

Figure 1: Example of data collection from network devices, servers and systems

As shown in the above picture, the metrics from all the network devices in the organization infrastructure, systems, applications, servers are collected using the supported agents on them and send the data to data collectors receivers. The data collectors can be configured to send the data to either streaming platform or directly to the data storage platform depending on the usecases identified as part of the organization's objectives. The data collectors shall be configured to send the data to the streaming platform for real time stream analytics, batch processing, data standardization, data transformations, etc. The data storage platform shall be configured to store the raw and structured data for long term as per the enterprise objectives and shall be used for running alerts, dashboards, adhoc searches for troubleshooting the issues, running AI/ML algorithm, etc. Once all the applications logs, metrics and network device metrics are ingested in the data storage platform, all these datasets shall be correctors to understand the reason for any latency or failure whether those are due to the performance metrics of network infrastructure or servers where the application is hosted.

Comparative Summary

Aspect	**With Metrics**	**Without Metrics**
Latency, CPU and Memory utilization	Proactively detected and optimized, reducing delays in user interactions and increases the resiliency of the applications.	Delays accumulate unnoticed, leading to slow application responses.
Jitter	Managed through QoS adjustments, ensuring clear communication in real-time apps.	Causes disruptions in voice/video calls, impacting leadership meetings and satisfaction.
Throughput	Allocated efficiently, prioritizing critical applications during peak times.	Bandwidth issues remain unidentified, causing congestion, delays and timeout of the requests.
Packet Loss	Reduced through rerouting and hardware replacement, ensuring smooth data flow.	Data loss leads to retransmissions, affecting app performance.
Interface Utilization	It helps detect overloaded or underutilized connections and proactively optimize the network infrastructure bandwidth utilization to ensure the applications get the right amount of bandwidth required for resilience and save the cost if the bandwidth is underutilized.	Delays in identifying the rootcause for applications performances when network bandwidth is overutilized and waste of money when underutilized.
Packet Processing Rate	Proactively detect the issues in the network infrastructure by monitoring this metric to ensure application resiliency and application interfactions.	Delays in identifying the rootcause for applications performances and interactions. May required interactions and meetings with the network engineering team to understand the actual rootcause.

Thus, network metrics and application logs could be used by the SRE team to identify which of the following might cause a transaction to fail: high latency, packet loss, misconfigured firewall rules, resource constraints at the server level. Using these data sets and proactive adjustments promotes application reliability as per the organization's resiliency objective. These approaches help to increase customer satisfaction and bring marketplace trust.

V. Future Trends in Network Observability Using FCAPS

With increasing adaptation of network observability, the FCAPS framework becomes so interwoven that leveraging network performance metrics for application resiliency becomes somewhat structured. The latest trend focuses on AI-driven fault detection, while predictive analytics for fault management is in place to engage networks in an automated mode to identify and mitigate possible problems before they may affect applications. For instance, AI-powered observability tools can automatically track latency spikes or drops in throughput in real time, isolating the fault domain in a router or switch. By correlating such metrics with application performance, an organization can take proactive steps such as rerouting traffic or adding additional resources to maintain application availability and performance[12].

In configuration management, intent-based networking (IBN) and infrastructure as code (IaC) are becoming pivotal. These are furthered in ensuring configurations of the network align with the demands of an application dynamically to reduce common human errors that lead to degraded performance. For instance, it is possible for an IBN system to reconfigure switches and routers for bandwidth policy in the peak periods of an application's use, thereby delivering consistent application performance. Security management today also brings together observability and performance metrics, enabling such network devices as firewalls to identify and adjust dynamically to changing attack patterns while preserving critical throughput and latency SLAs for application resiliency. These are but a few of the innovations marking a move toward turning autonomous, adaptive networks in ways that leverage FCAPS principles for monitoring, optimization, and protection of application performance[13].

A. Adapting Intent Based Networking (IBN) for application resiliency

IBN is a revolutionary approach of managing networks, leveraging automation, AI, and ML to translate high-level business requirements/intents directly into actionable network configurations. Whereas traditional methods, the professionals would have to author detailed manual step-by-step configuration inputs in order to achieve a specific desired configuration, but the IBN approach allows them to define their intent in plain business terms—such as "prioritize video traffic for remote users" or "segment IoT devices from critical infrastructure." The system automatically translates these into specific policies and configurations, then pushes them across the network while ongoing performance monitoring is conducted to ensure those intents are realized. This simplifies managing the network, reduces human error, and aligns network operations to organizational goals.

IBN systems operate in a closed loop, and continuous monitoring and verification ensure that the behavior of the network is aligned with what the intent had specified. This means that in case of discrepancy, such as increased latency or packet loss for prioritized traffic, the system automatically readjusts the network resources to previous performances. With the integration of AI and ML, IBN can predict impending issues, suggest optimizations, and adapt to changing network conditions ahead of time. For example, for an intent like "ensure zero latency for telemedicine sessions," a system will go into traffic prioritization, reroute data in cases of congestion, and keep the performance intact even when it works at peak usages. This capability enhances network resiliency and scalability while enabling complex environments with new demands: modern, cloud-based applications, or large-scale enterprise networks.

B. Adapting Infrastructure as Code (IaC) for application resiliency

Infrastructure as Code will be the key in ensuring application resiliency, integrated with AI, ML, and automation. IaC gives an organization the capability to define, provision, and manage infrastructure with code. This makes the deployment fast and consistent in diverse environments. By automating these processes, IaC minimizes human error, accelerates recovery from failures, and ensures infrastructure configurations remain consistent with application demands. In that direction, IaC combined with AI and ML can also find out configuration drifts on its own and make appropriate optimizations in advance to improve deployment, hence application uptime and performance based on both historic and real-time data.

For example, an ML-driven IaC system in a cloud environment can automatically scale resources during peak traffic to maintain service availability without human intervention. In the future, the integration of AI and ML with IaC will further transform application resiliency by enabling adaptive, self-healing infrastructure. AI-powered IaC systems will go a step further to monitor not only infrastructure health but also predict failures and trigger corrective actions automatically, including deploying replacement instances or correcting network configurations. Automation pipelines reconfigure infrastructure dynamically to best align with changing application workloads for peak performance in shifting conditions.

For example, in an e-commerce platform, an AI-enhanced IaC framework could identify patterns of high demand during seasonal sales and preemptively deploy additional instances, configure load balancers, and optimize storage. The convergence of IaC, AI, ML, and automation will drive unparalleled levels of resiliency, enabling applications to operate seamlessly in the event of unexpected disruptions or rapid changes in demand[11].

VI. CONCLUSION

Network performance metrics play a crucial role in enhancing observability by providing detailed insights into the flow of data across networks, which directly affects application performance and user experience. Metrics such as latency, throughput, packet loss, and jitter help identify network bottlenecks and distinguish between network-related and application-related issues. This level of detail allows teams to perform more precise root cause analysis, especially in complex, distributed systems like microservices, where requests traverse multiple network paths.

Incorporating network metrics into observability enables teams to optimize resource allocation and improve reliability and resilience by detecting patterns that indicate network instability or potential security threats. By understanding how network conditions impact application performance, teams can make informed decisions about load balancing, traffic routing, and other configurations, leading to enhanced system performance and cost efficiency.

Network performance metrics also provide end-to-end visibility, crucial for applications serving users across various geographies. By correlating network metrics with application performance data, observability tools can offer a holistic view of how different components interact, facilitating proactive detection of issues and supporting adherence to Service Level Agreements (SLAs).

Overall, network performance metrics are essential for a comprehensive observability strategy, as they help ensure that applications run smoothly, securely, and efficiently. This integration of network insights enables teams to proactively manage and optimize the entire system, ultimately improving user satisfaction and operational effectiveness.

References

[1] https://joindigital.com/naas360/network-observability

[2] Jayanna Hallur, "The Future of SRE: Trends, Tools, and Techniques for the Next Decade", International Journal of Science and Research (IJSR), Volume 13 Issue 9, September 2024, pp. 1688-1698, URL: https://www.ijsr.net/getabstract.php?paperid=SR24927125336, DOI: https://www.doi.org/10.21275/SR24927125336

[3] What Is Observability? Key Components and Best Practices https://www.honeycomb.io/what-is-observability/

[4] Jayanna Hallur, "From Monitoring to Observability: Enhancing System Reliability and Team Productivity", International Journal of Science and Research (IJSR), Volume 13 Issue 10, October 2024, pp. 602-606, URL: https://www.ijsr.net/getabstract.php?paperid=SR241004083612, DOI: https://www.doi.org/10.21275/SR241004083612

[5] Sasongko, H., & Hadiwandra, T. Y. (2021). Cloud-Based NAS (Network Attached Storage) Analysis as an Infrastructure as A Service (IAAS) Using Open Source NAS4FREE and Owncloud. https://doi.org/10.25299/itjrd.2022.5712

[6] Xu, L., Jiang, W., Zhang, Q., & Wang, X. (2020). A comprehensive survey on bandwidth estimation techniques in high-speed networks. *IEEE Communications Surveys & Tutorials, 22*(3), 2038–2070. https://doi.org/10.1109/COMST.2020.3001035

[7] Mehani, O., Boreli, R., & Henderson, T. (2015). Latency and loss measurements in mobile networks: Empirical evaluation of the state of the art. *ACM SIGCOMM Computer Communication Review, 45*(4), 25–30. https://doi.org/10.1145/2831347.2831351

[8] Nagarajan, R., & Selvi, V. (2021). Resource utilization optimization in cloud-based network devices using deep learning. *Journal of Network and Computer Applications, 186,* 103095. https://doi.org/10.1016/j.jnca.2021.103095

[9] Johnson, D., Kumar, R., & Nguyen, T. (2018). Dynamic load balancing and interface utilization in software-defined networking. *Journal of Network and Systems Management, 26*(4), 987–1002. https://doi.org/10.1007/s10922-018-9469-5

[10] Ferrus, R., Sallent, O., & Agusti, R. (2014). QoS and jitter optimization in mobile multimedia networks. *IEEE Transactions on Multimedia, 16*(3), 759–768. https://doi.org/10.1109/TMM.2014.2299356

[11] Scalability and Flexibility https://softechds.com/it-solutions/cloud-solutions/

[12] IEEE Standards Association, "IEEE Standard for Information Technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements," IEEE Std 802.11-2016, 2016

[13] A. Brown, "Performance Monitoring in Modern Networks," Proceedings of the International Conference on Network Management, pp. 112-118, 2019.