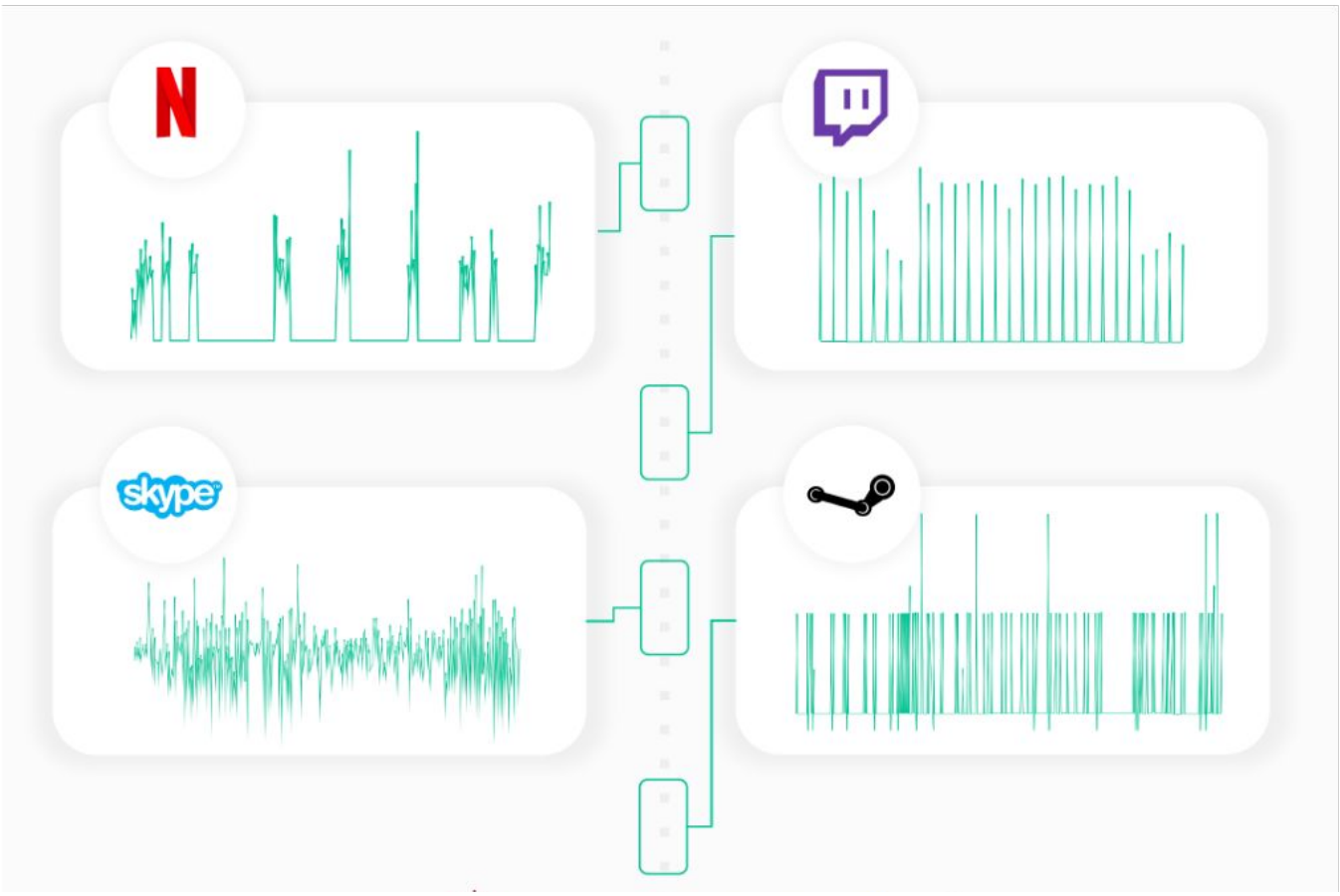
White Paper | Aug 2024

Powering the Network AI Ecosystem with Super-Netflow



CANOPUS NETWORKS

Contents.

Executive Summary.

NetFlow (also known as IPFIX), by which Internet routers export records of the traffic conversations they transport, has been the de facto standard for over 25 years to extract information from the network on flow endpoints, volumes, and timestamps. This standardization has fostered a multi-billion ecosystem of analytics tools that take NetFlow as a source of network data, for uses ranging from network troubleshooting and performance optimization to service assurance and security analytics. The dramatic growth of AI in recent years brings opportunities to enhance these areas, as well as open up new analytics use-cases in areas relating to automated operations, customer support, and personalised marketing, which are inadequately served by NetFlow as it stands today.

Canopus is offering **Super-NetFlow** – an extension that enriches NetFlow with application-level information identifying the type & provider of each traffic flow (e.g. Netflix streaming video, CS:GO gaming, Zoom conferencing, etc.), as well as its real-time quality-of-experience (QoE) measures such as video resolution, buffer stalls, gaming lag spikes, and conference stutters. In this whitepaper we show that Super-NetFlow enables new categories of AI applications that can provide better customer care using behavioral data; better capital allocation that can directly maximize user experience; and personalized marketing and advertising that can employ behavioral user segmentation. We describe how the multimedia telemetry data is represented and exported in a standardized way so it can be fed to a wide array of AI analytics platforms. We outline the schema, and illustrate it in a live deployment with representative examples of video streaming and gaming. Super-NetFlow will be the fuel that powers predictive and generative AI applications much needed by network operators to maintain and enhance profitability in the decade to come.

Better AI Demands Better Data.

Artificial Intelligence is revolutionizing the Telecommunications industry in many ways. Network planning, maintenance, and troubleshooting are increasingly leveraging AI to realize the dream of a self optimizing network, while customer care is being augmented with genAI-based chatbots. AI however needs high-quality network data as the fuel to run on. Unfortunately NetFlow, which has been the de facto standard for the export of flow data from network elements for over 25 years, is very limited – it only provides basic information on the flow endpoints, volumes, and durations, which limits the benefits AI can bring to the business.

We introduce Super-NetFlow, an augmentation to network data that provides timely and accurate application information, including classification of the type and provider corresponding to each traffic flow (such as Netflix streaming video, Twitch live video, Call-of-Duty gaming, Zoom conferencing, etc.), along with high-fidelity measures of the user's real quality-of-experience (aka QoE), such as video resolution and buffer health; game latency and lag spikes; conference stutter and dropouts, etc. Super-NetFlow is the fuel that powers the new ecosystem comprising predictive and generative AI applications, as illustrated in **Fig. 1** below.

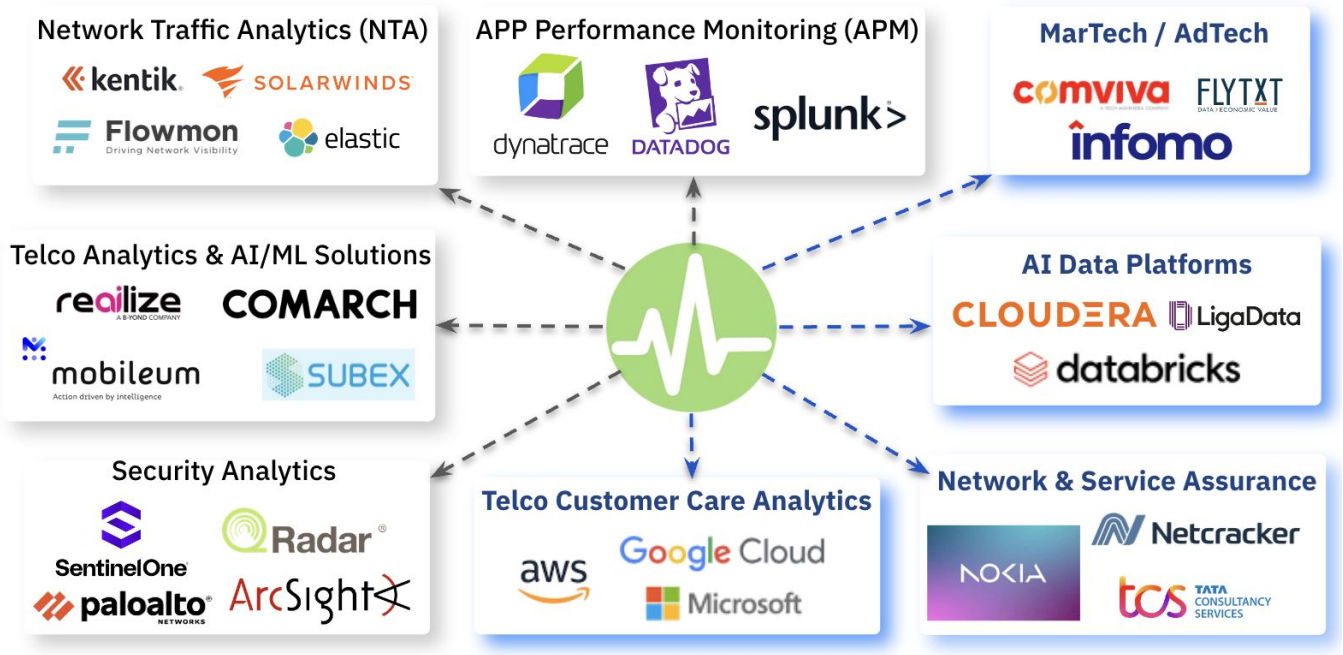


Fig. 1 The rich AI ecosystem enabled by Super-NetFlow.

Super-NetFlow can enhance existing network traffic analysis, security analytics, and application performance monitoring platforms by providing application-level information. It can further fuels AI applications in many ways:

Customer Care: Telecoms operators are partnering with cloud providers like Microsoft Azure [1] and GCP [2] for customer care chatbots – these can be augmented with context on what the user was doing (e.g. gaming) and what experience issues they had (e.g. latency spikes), so the chatbots can provide meaningful and contextual help; indeed, Super-NetFlow enables proactive AI-based detection and remediation even before a support call is made.

AI Analytics and Automation: AI is increasingly being incorporated into areas such as revenue assurance [3], fraud detection [4], Customer Lifetime Value (CLV) optimization [5], and automated capacity management [6]. Platforms such as Mobileum, Subex, Reailize, and Comarch, as well as internal Telecom analytics teams that use data platforms such as Cloudera and DataBricks, can greatly increase the power of their AI with fine-grained data on application usage and experience.

Marketing and Advertising: By leveraging data from the network on the consumption of applications by their users, Telecoms operators can increase revenues and loyalty by creating personalized bundles and offers [7] (powered by analytics platforms like Comviva and Flytxt), as well as operate advertising brokerage platforms (e.g. Infomo).

Super-NetFlow Data Schema.

Super-NetFlow extends NetFlow by adding three new blocks of data (see **Fig. 2** and refer to [10] for more details):

- **Application & User Platform:** This block associates an application session with the application type (e.g. video, gaming), provider (YouTube, Call of Duty), client OS, and software agent (e.g., Chrome browser or native app).
- **Network QoS:** This block includes network quality of service (QoS) statistics for this application, e.g. volume, throughput, latency, jitter, and packet loss, aggregated across the various flows of the application session.
- **Application QoE:** This block contains the quality of experience (QoE) metrics for each multimedia application type, as listed in the table in Fig. 3(a) below and described next.

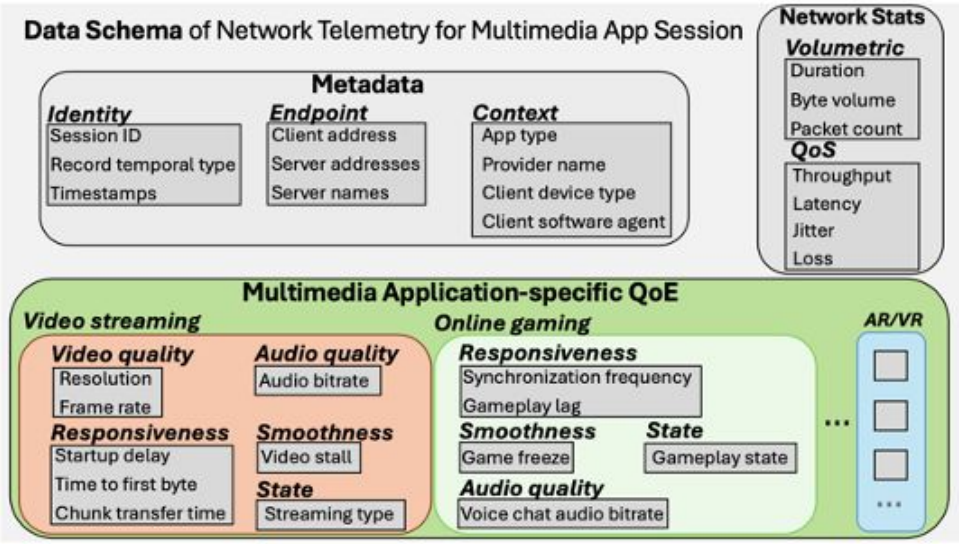


Fig. 2 Telemetry data schema for multimedia applications

Application type	Popular provider	Example QoE metric
On-demand video streaming	YouTube, Netflix, Hulu	Resolution, frame rate
Live video streaming	Twitch, YouTube Live	Startup delay, video stall
Online gaming	Call of Duty, Fortnite	Gameplay lag, game freeze
Voice/video calling	WhatsApp, Zoom, Teams	Voice mute, frame drop
Music streaming	Spotify, Apple Music	Audio bitrate, loading delay
Cloud gaming	GeForce NOW, xCloud	Resolution, frame rate
AR/VR	VRChat, RecRoom	Activity lag, motion freeze
Social networking	X, Facebook, Instagram	Page load time

Fig. 3(a) QoE metrics for eight example application types.

Fig. 3(a) above shows the application types that are currently supported by Canopus Super-NetFlow, including the popular providers and typical QoE metrics. For instance, on-demand video (Youtube, Netflix) and live video (Youtube Live, Twitch) experience are best expressed in terms of resolution, startup delay, and freeze [8], while online gaming (Call of Duty, Fortnite) experience is represented by latency and lag spikes [9]. In addition to prevalent application types such as voice/video calling, social networking, and music streaming, Super-NetFlow is also ready for emerging applications like cloud gaming (Nvidia GeForceNow, Microsoft xCloud, Amazon Luna) and Augmented/Virtual reality applications on Oculus and Apple headsets. A detailed explanation of the data schema blocks shown in Fig. 2 above, and the various supported application types, can be found in [10].

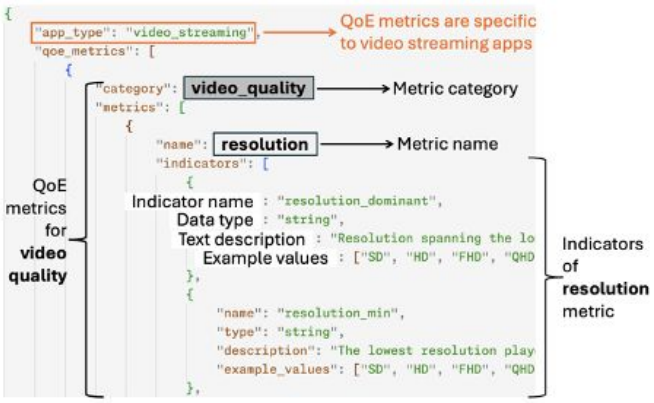


Fig. 3(b) Sample JSON snippet (see [11])

The Canopus probe exports Super-NetFlow records in JSON and YAML formats, which can be accessed from our github site [11]. **Fig. 3(b)** above shows a JSON snippet for a video streaming flow, depicting the name of each metric, its data type, a description, and reference values. This standardised format can be ingested by any of the hundreds of analytics platforms in the market today to leverage application information. We next show how this is applied to the specific case of a video streaming session and an online gaming session.

Example Use-Cases: Streaming & Gaming

We show below two illustrative examples of Super-NetFlow's ability to export application usage and experience data in the NetFlow/IPFIX format as described on the previous page. These examples are taken from a live deployment, and are corroborated by the users engaging in these applications.

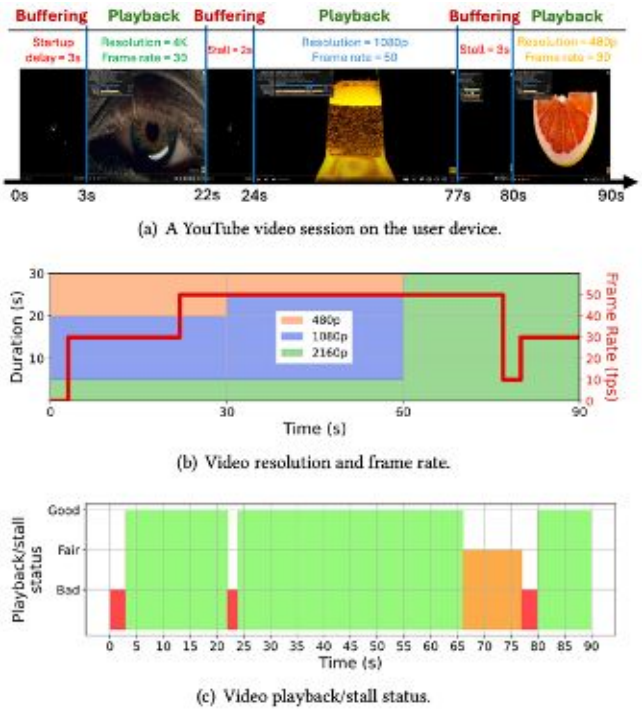


Fig. 4(a) QoE export of a YouTube video stream.

Fig. 4(a) depicts a 90-second YouTube video watched on a macOS laptop via the Chrome browser, with the experience phases depicted visually in the top plot. The Canopus platform ejects a Super-NetFlow message at the start of this stream, indicating the flow endpoints, application classification (YouTube), and user platform (macOS Chrome). It then ejects a Super-NetFlow message every 30 seconds, reflecting the video resolution (4k, 1080p, or 480p) and frame rate (10, 30, or 50), as shown in the middle plot, while the bottom plot shows the buffer health (stable, depleting, and buffering).

Key Takeaway

The ability for Telecoms operators to gain business value from predictive AI and generative AI is very dependent on high-quality data from the network. NetFlow as it exists today is insufficient; Canopus is offering Super-NetFlow that simply and elegantly incorporates information on application type/provider, user platform, and application QoE measures. Super-NetFlow is the fuel that can power the emerging AI ecosystem and help Telecom providers maintain and enhance their profits.

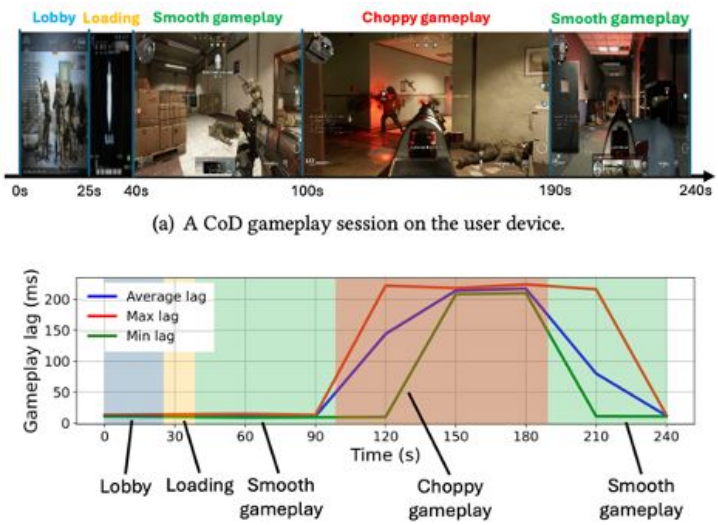


Fig. 4(b) QoE export of a CoD gaming session.

Fig. 4(b) depicts a 4-minute session of Call of Duty: Modern Warfare, played on a Windows PC, with the stages (lobby, loading, gameplay) annotated in the top plot. The Canopus platform ejects a Super-NetFlow message at the beginning identifying the gaming session and the platform. It then sends periodic messages every 30 seconds indicating regions of smooth versus choppy gameplay experience, along with the average, max, and min gameplay latencies as shown in the lower plot.

References

[1] "Azure for Operators", https://azure.microsoft.com/en-au/solutions/telecommunications.
[2] "Using AI to win the customer experience battle in telecommunications", Google Cloud, https://cloud.google.com/resources/winning-cx-in-telco-with-ai-whitepaper
[3] "What is revenue assurance? The ultimate revenue assurance guide for Telecom operators", Subex, https://www.subex.com/article/what-is-revenue-assurance-the-ultimate-revenue-assurance-guide-for-telecom-operators/
[4] "Fraud detection and prevention in the Telecoms industry", Infosys, https://www.infosysbpm.com/blogs/bpm-analytics/fraud-tactics-to-detect-and-prevent-in-the-telecommunications-industry.html
[5] "Winning in Telecom CX", McKinsey, https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/winning-in-telecom-cx
[6] AI/ML driven network efficiency", Comarch, https://www.comarch.com/telecommunications/ai-ml-driven-network-efficiency/
[7] "Unlocking the potential of real-time personalization", Comviva, https://www.comviva.com/blog/unlocking-the-potential-of-real-time-personalization-boosting-revenue-and-loyalty/
[8] S. Madanapalli, A. Mathai, H. Habibi Gharakheili and V. Sivaraman, "ReCLive: Real-Time Classification and QoE Inference of Live Video Streaming Services", IEEE IWQoS, Tokyo, Japan, Jun 2021. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9521288
[9] S. Madanapalli, H. Habibi Gharakheili, V. Sivaraman, "Know Thy Lag: In-Network Game Detection and Latency Measurement", Passive and Active Measurement (PAM) Conference, Mar 2022, https://dl.acm.org/doi/abs/10.1007/978-3-030-98785-5_17
[10] Y. Wang, M. Lyu, V. Sivaraman, "Standardizing Multimedia QoE Telemetry from Telecommunications Networks for Open Analytics", Proc. ACM Sigcomm Workshop on Emerging Multimedia Systems (EMS), Sydney, Australia, Aug 2024.
[11] Super-NetFlow data schema in JSON and YAML. https://github.com/WYF99/QoE-telemetry-standardized