SANDVINE

Disruptive Mobile Analytics

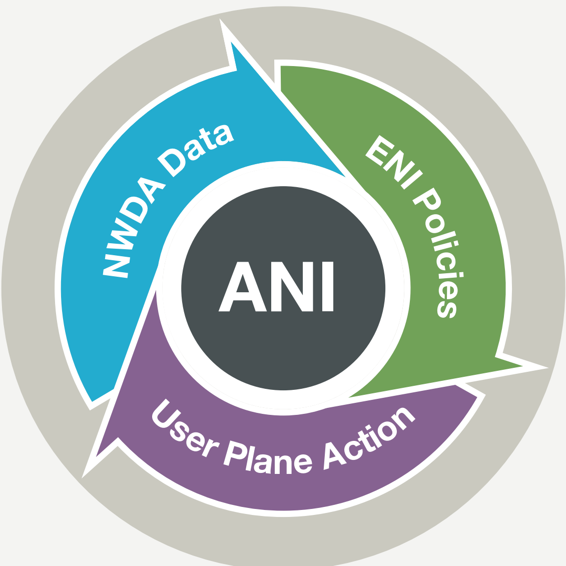
Experiential Network Intelligence

A critical foundation of experiential network intelligence is context – and situation – awareness

The 3GPP 5G standard introduced a Network Data Analytics (NWDA) function to mobile network architectures, and has defined a collection of initial use cases that specify what kind of data can be provided to feed 5G network decisions. ETSI has recognized the importance of the customer experience and launched an Experiential Network Intelligence (ENI) ISG that leverages artificial intelligence on network analytics to adjust services based on changes in user needs, environmental conditions, and business goals; ETSI has also defined a set of initial use cases that demonstrate how an ENI system would enhance networks. When these two powerful solutions are combined into a single network intelligence solution, they will create a disruptive advantage for a mobile operator – today and in the 5G future.

Figure 1

5G NWDA and ETSI ENI mapped to Sandvine's Active Network Intelligence



To a consumer, all that matters is if an application is working – right now. If a YouTube video is low resolution and stalling, Skype voice is skipping, or Facebook is loading images slowly, then the user is having a bad experience. If every other subscriber on the network is getting good service, the network has still failed that one subscriber. The challenge today is that network operators don't have the visibility to determine that this has occurred, much less to fix the problem.

Sandvine's Active Network Intelligence solutions are served by an unparalleled foundation of contextual data that meets and exceeds the data requirements specified for many of the NWDA and ENI use cases. Using artificial intelligence (AI) and machine learning (ML) techniques, Sandvine takes data collected in an NWDA role and processes it as an ENI engine would, in order to help solve network issues like those described in various NWDA and ENI use cases.

This paper is the first in a three part series that will explore the principles of NWDA and ENI, and map them to specific use cases that can be deployed today with Sandvine solutions. This whitepaper explores the network data (also known as telemetry) needed to measure experience and how it needs to be contextualized to make it actionable – comparable to an NWDA. The rest of the series will focus on processing and analyzing the data (as an ENI engine would) and turning that into intelligent policies to close the loop – the process that Sandvine calls Active Network Intelligence.

DEFINING SUBSCRIBE EXPERIENCE METRICS AND KPIS

Most operators would claim that they have plenty of data on their network performance. The question is if it is the right data. Let's take a look at the initial statement on the consumer's view of network experience: "If a YouTube video is low resolution and stalling, Skype voice is skipping, or Facebook is loading images slowly – the user is having a bad experience." Any measurements taken that do not contribute to the actual experience delivered to specific customers or portions of the network do not help answer this question.

Knowing that a network segment is congested, maybe by data collected from network elements in that segment or an active probing system, does not actually translate to the experience received by individual subscribers in that segment.

So what is the right data? As more and more of the data on a network is encrypted, through TLS/SSL, getting application and quality data from existing systems has been problematic. The required data determines that the consumer is having a bad experience, while also enabling the network operator to resolve this issue in real time – which in 5G would drive User Plane Function (UPF) selection based on load.

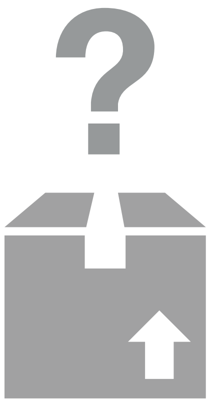
Each application's quality of experience (QoE) is dependent mainly on three specific factors:

Bandwidth

Bandwidth is an important measure of a subscriber's broadband experience, and most broadband plans are priced and marketed based on their peak bandwidth capability. Many applications, video streaming in particular, are bandwidth-hungry and rely on having access to a high bandwidth to perform well. Many broadband measurement "scorecards" are primarily focused on throughput – like the Netflix ISP Speed Index or the Ookla Speedtest. Bandwidth measurement is also an "instantaneous" metric – the interval that you measure bandwidth can have a big effect on that measurement. Measured across 250ms, throughput may be reported as several Mbps, but that same transaction across a minute time frame results in a measurement of Kbps.

Latency

Latency is an important metric for interactive applications. Few consumer plans offer latency as a service level agreement (SLA), but it is sometimes included as part of a managed service offering for business connections. Anyone that has ever played in an online first-person shooter game can tell you that latency "literally" kills! Voice connections that experience high latency exhibit this by the two speakers talking over each other, creating a very frustrating conversation. Excessive buffering, congestion, or the simple physics of transmission across long distances (i.e., overseas or satellite connections) on the network can all be the source of network latency. Some latency, such as transmission latency, cannot always be fixed in the network.



Packet Loss

Packet loss can also cause customer dissatisfaction with their broadband connection due to the impact it can have on some applications. Packet loss can result in increased buffering and stalls in video streaming, slow web page load times and jittery voice applications, reducing the goodput (as opposed to throughput) on the network. Packet loss wastes bandwidth on the network as packets are retransmitted and, depending on the application types, can create havoc with the subscriber experience. For a Web shopping session in a browser, packet loss of even just 1% can result in double page-loading times, significantly impacting the QoE.

These three factors – throughput, latency, and packet loss – are commonly identified by telecom regulatory authorities (TRAs) as the key metrics for a broadband operator to either measure and self-report on, or the TRAs themselves will increasingly deploy measurement solutions themselves as an independent authority on network performance; however, the question that immediately follows the collection of these key performance indicators (KPIs) is "Are those good numbers?" The answer is "It depends on the application."

WEB SURFING

STREAMING VIDEO

GAMING

SOCIAL MEDIA

UPLOADING

DOWNLOADING

VOICE APPLICATIONS

Different applications have different network requirements in order to provide good quality of experience

END USER PERSPECTIVE: IT'S ALL ABOUT THE APPLICATION

With the three critical quality measurements available, any exploration of the subscriber experience needs to factor in the expectations of the applications that a subscriber might be using. The table below highlights the expectations different applications have from the network and how each of application type reacts to throughput, latency, and loss changes on the network.

Application Type	Throughput	Latency	Loss
Web	Needs short bursts of download performance	High latency leads to slow page load times	Packet Loss can lead to slow page load times
Video	Sustained throughput delivers good quality	Not usually a concern except for initial loading of video	Less sensitive to loss unless it affects throughput
Social Media	Needs short bursts of download/upload performance	High latency can slow interactive sharing experience	Packet Loss can slow interactive sharing experience
Gaming	Most games do not require high bandwidth	High latency leads to lag in real-time games	Packet Loss leads to lag in real-time games
Upload	Sustained bursts of upload performance	N/A	Less sensitive to loss unless it affects throughput
Download	Sustained bursts of download performance	N/A	Less sensitive to loss unless it affects throughput
Voice	Low throughput requirements	High latency leads to poor voice experience	Some loss can be tolerated, high loss leads to perceived latency

As you can see from the table above, simply delivering a good average throughput on the network will not necessarily result in a good subscriber experience with applications like Web Surfing, Voice, or Gaming. A good average score does not translate to a consistently good experience eithe,r as conditions may vary throughout the day, week, or month. However, if the operator has collected these key metrics, they can create their own network experience scorecard that they can use to improve the experience they deliver to their subscribers, which Sandvine represents as a letter grade just like if you were in school – from A to F.

MAKING DATA ACTIONABLE: CONTEXT AND SITUATION AWARENESS

The early ETSI work on Experiential Network Intelligence (ENI) highlights that a critical component of experiential network intelligence is context- and situation-awareness. ENI defines context as: "The context of an Entity is a collection of measured and inferred knowledge that describes the state and environment in which an Entity exists or has existed." Analyzing that definition leads to a crucial difference between ENI and traditional mobile analytics – the combination of known facts that can be measured with inferred data (or analytics gleaned from ML/AI actions).

Environment is also extremely important to the definition, as the network environment is changing in real-time, and decisions made without real-time visibility and feedback are more likely to be ineffective. The definition also hints at historical knowledge or past decisions that were made that may influence future decisions.

The working group highlights the NWDA specification within 5G as one potential use case for ENI – specifically within network slices. In a 5G environment, it is possible that the behavior of one slice can effect other slices, especially during times of network stress (outages, natural disasters, or peak hours). In these scenarios, understanding the context of slices will be another aspect of context that will need to be factored. But slice-awareness is more of a future requirement rather than a today one. What context awareness is needed now?

A critical foundation of experiential network intelligence is context – and situation – awareness

PUTTING IT ALL TOGETHER: EXPERIENTIAL NETWORK INTELLIGENCE

Just measuring network-level throughput, latency, and packet loss does not deliver experiential network intelligence. As shown above, understanding the application being used by the subscriber is a bare minimum for determining if the network is delivering the required network performance to the subscriber.

With more and more traffic being encrypted on networks, a strong application identification engine (which already leverages machine learning to recognize applications) is mandatory. Location, time of day, device type, and the consumer's service plan (among many attributes) can all impact the measured QoE. Different OTT providers (Skype, YouTube, Netflix, etc.) use multiple protocols or streaming codecs with differing resolution levels and may originate from Content Delivery Networks (CDNs) with different levels of robustness or congestion. Service providers need a more granular understanding of all of these metrics combined together to provide the context- and situation-awareness they need to feed an experiential network intelligence system. Associating the KPIs in context with the subscriber and their attributes enables the data to be actionable in real-time.

With this type of context, the service provider can quickly determine if their network is delivering a good experience for applications that drive subscriber usage. The data should be able to directly lead to the root cause of the network impairment and enable network engineers to improve the network experience. Investigation of a degraded experience might reveal a specific location is experiencing systemic congestion and over-utilization; splitting the node would resolve the issue. Combining the technical issues with service degradation ensures that the service provider can make the right business decisions on where to invest in their network to improve their quality and have the best return on investment.

When adding context to quality metrics, subscriber identification and application are two of the most important. Several others are worth exploring because they can provide valuable insight into how the ENI engine will determine what the root cause of an issue is and how to rectify the problem with an intelligent policy. The contexts that deliver the most value are:

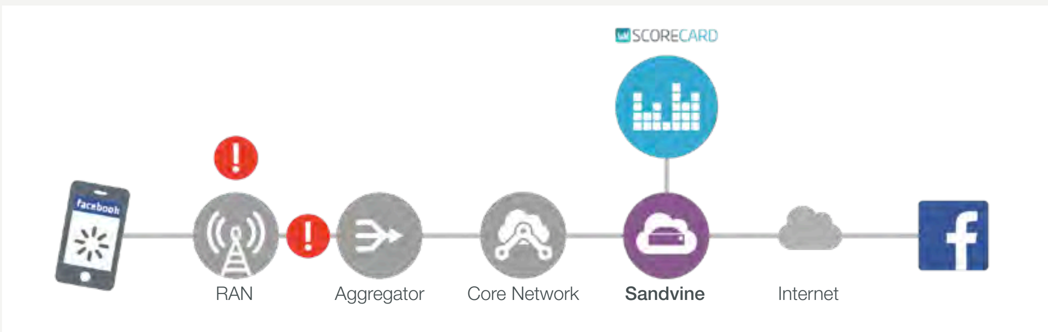
- Location (Cell ID, City)
- Virtual Network Function (VNF), network slice, or topology
- Device (Handset model or brand)
- Subscriber Tier

Location

Measurements in a mobile core network need to be enriched with information about their source location in order to become truly valuable. When measurements are enriched with location information, it is possible to group them and understand the impact the location's cell is making on the quality of the sessions carried by it. A breakdown of the performance per location is showing the KPIs of all the sessions that originated or terminated in a particular Cell ID.

Figure 2

Metrics identify problem RAN from high latency measurement



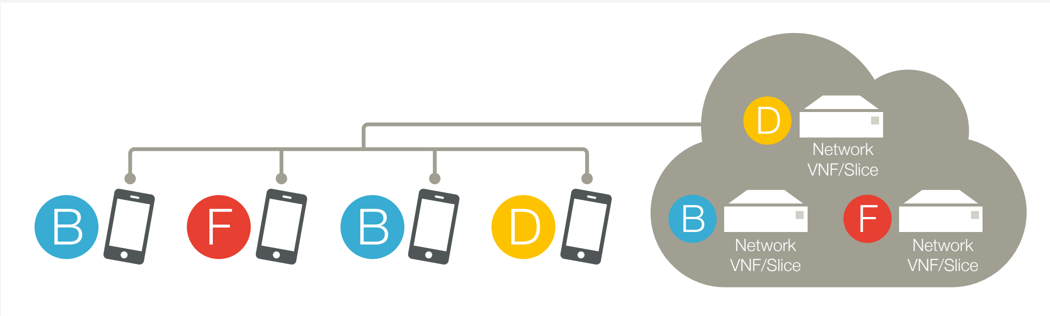
While each individual session could have been impacted by many different factors (including device type, OTT service quality etc.), looking at all sessions with the same location in common over time shows if there are commonalities and that the root cause of the issue is directly related to the environment in that cell – which can directly lead to the policy decision on how to manage the network to rectify quality issues in the cell area.

VNF, network slice, or network topology

As operators move to a virtualized infrastructure, measuring the performance delivered by each virtualized network function (VNF) is even more important. Segmentation can also be applied (per 5G definitions) to specific network slices to meet the high level NWDA requirements of network load and environment for each active network slice. This can also be generalized to network topology context, which can reveal if particular infrastructure nodes are not performing up to standards. Poor QoE metrics that are isolated to a specific node or slice can be an indication that there is a misconfiguration of the nodes.

Figure 3

See score per VNF or network slice instance

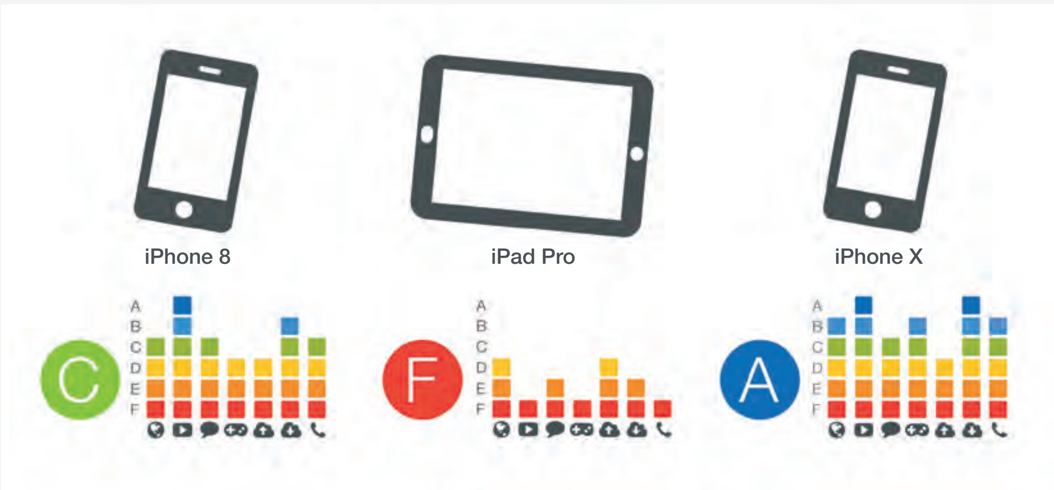


Device

Devices can have a huge impact on QoE delivered. Delivering good quality video on a small mobile screen is different than delivering that on an iPad Pro screen. Many subscriber problems on the network can be isolated to specific device's behavior or versions of the OS that may have specific bugs. With real-time data enrichment of the measurement reports with device types, issues can be isolated if they are related to the devices.

Figure 4

See scores based on device type



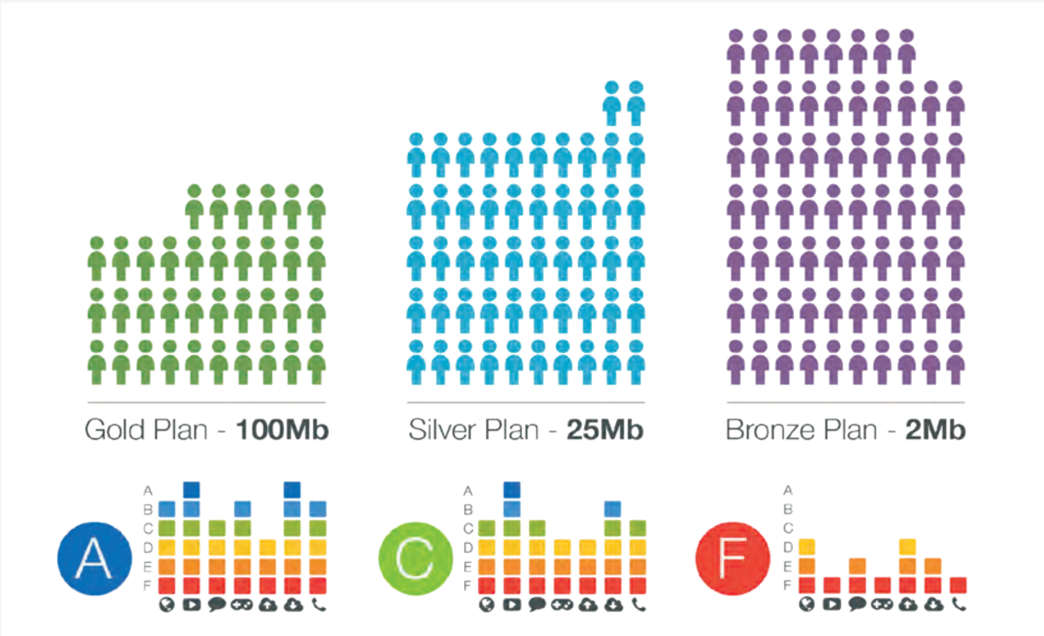
Subscriber expectations for their service plan is a key factor in perception of QoE metrics and ENI

Subscriber Tier

Through integration with the BSS/OSS system, measurement reports can be enriched with what rate plan the subscriber belonged to at the time of recording the metrics. Grouping this data provides the means for an operator to see the quality delivered to gold, silver, and bronze subscribers or separate QoE per reseller channel. Subscribers can be in several groups at the same time, enabling simultaneous monitoring of demographics, channels, and rate plans. Two subscribers that receive the exact same quality measurements may have very different opinions on their perceived QoE if one is on a gold plan and one is on a best effort plan. If the subscriber is receiving the QoE that matches his SLA (even if the performance is low), no action may be required from an ENI engine, emphasizing the importance of this specific context.

Figure 5

Discover scores based on subscriber tier



DEPLOYMENT SCENARIOS

From an end-to-end perspective, the measurements should take place as close to the subscriber as possible, but can be taken anywhere in the network. Measuring at the OTT applications' location is providing one app's end-to-end view, but does not represent the scoring of the service provider's networks in question; it also leaves out relevant metadata so that it does not allow the results to become actionable.

At the same time, measuring at the subscriber's location would result in too many measurement points becoming costly and difficult to manage. The ideal location is close to the interconnect point between the access network and the peering partners, at a location where all traffic can be captured.

The biggest advantage for an operator in today's mobile infrastructure is the ability to gather this data with an orchestrated and managed virtual probe solution. Sandvine's eVolution virtual experience probe can be deployed anywhere in the mobile network, anywhere from the edge of the RAN trough the packet core, and even to the peering point. With the right level of integration and enrichment to the data as has been described in this paper, similar levels of visibility can be attained regardless of the location of the probe.

Building to 5G networks requires network intelligence even before the 5G network goes live. Combining the NWDA and ENI functionality into a single entity will deliver the highest ROI for mobile operators

EVOLVING TO 5G MOBILE BROADBAND NETWORKS

As operators move toward building out their 5G infrastructures and deploying their new packet cores, understanding how subscribers use the existing networks with applications is a key factor. It is also critical to understand that many early 5G adopters will be the power users – in other words, those who both use the network more than normal subscribers and have much higher expectations from the network.

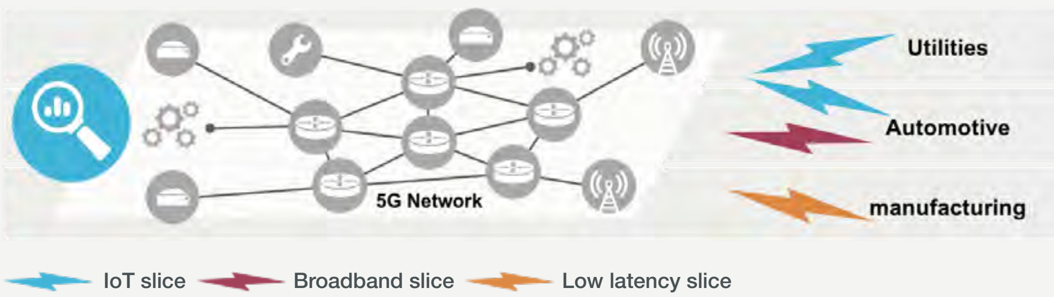
These power users are also willing to pay more ARPU than most subscribers as the early adopters as well. In fixed mobile substitution use cases, many mobile operators will target replacing fixed line broadband (like DSL) with 5G mobile connections. The usage on these links will be radically different than traditional mobile usage, so gaining visibility into this usage pattern from the initial deployments will help the operator ensure that their network quality does not degrade as users adopt 5G for this use case. It will also help guide the building of monetization models and service plans for 5G power users – providing options like location-based pricing (different quota usage at home than while mobile), time of day allowances (more video bandwidth during primetime viewing hours), device-based plans, and even closely tracked performance SLAs for small business plans using mobile hot spots.

Sandvine strongly believes that the future of mobile data analytics lies in the combination of the NWDA with an Experiential Network Intelligence Engine. Even more important is the ability to turn the output of the analytics system into direct action to resolve issues in real time – what we call Active Network Intelligence. We will explore this path from insights to action in the follow-on papers in this series, but the best way to turn this theory into reality is to start with an example of an ENI use case that can drastically enhance the user experience.

The ETSI ENI working group has an early draft of a use case document GR ENI 001 – V1.1.1 – Experiential Networked Intelligence (ENI); ENI use cases (1). One of the use cases in the document is Intelligent Network Slicing Management, where the operator needs to manage conflicts in capacity between slices as usage changes. As described in the use case document, the slices are created and configured, and the ENI system (through AI and ML capabilities) learns the configuration of the applied slices and as well the traffic patterns used by each of these slices. The ENI system also measures the utilization of the network and other relevant parameters that define the satisfactorily operation of each slice and that needs to conform to the service and network KPIs requested by the operators.

Figure 6

5G Networks subdivided into virtual networks, each optimized for one business case



- **Intelligence:** Data needs to be contextual and application QoE-centric
- **Policy:** Policy must be aware of all slices to maintain network quality

In this model, a single system needs to provide a comprehensive view of all network slice activity and the quality of each slice, so that the ENI system can make the right recommendations on creating new slices and managing QoE across the entire infrastructure by mediating resource consumption between slices.

If one or more of the slices begins to experience performance degradation (performance, latency, or packet loss), the ENI system will detect the issue and can determine the cause of the problem. In the 5G specifications, the NWDA needs to collect the right data to detect degradation of QoE – especially if a slice is specific to a particular application type or function (like a video slice, a VoIP/VoLTE slice, or even an industrial IoT slice). Adding experience intelligence to the NWDA will enable the operator to make far more intelligent decisions on how to detect and resolve issues on these slices and integrate with the UPF and Policy Control Function (PCF).

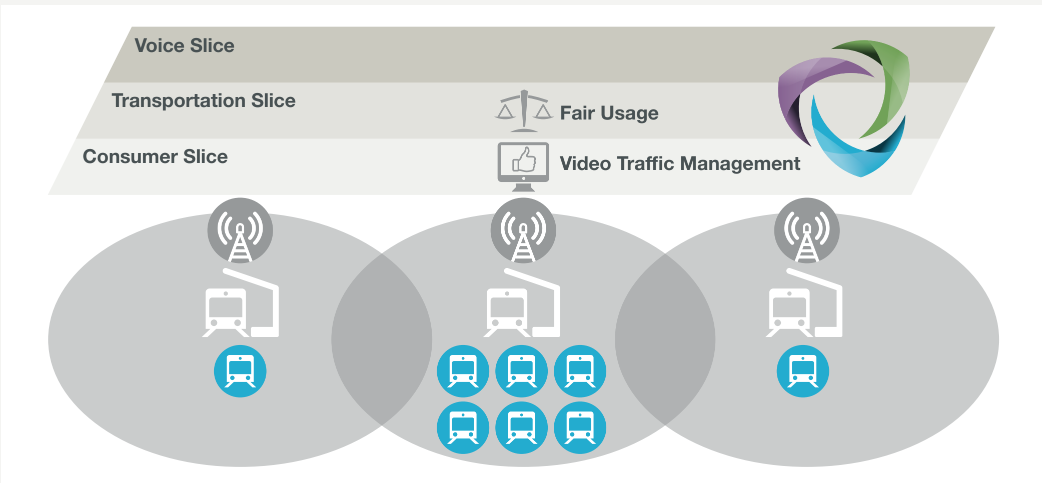
In this use case, by combining all of the parameters (discussed earlier in this paper), an ENI-powered NWDA would determine the specific issues causing the degradation and would propose the most surgical action that would resolve them across the slices the problem is affecting. If the root cause behavior is affecting all slices, the ENI engine would propose whatever policy actions needed to be taken across all slices to solve the issue – a common example here would be a DDoS attack that has a broad reach on both infected devices as well as attacked hosts.

In this example, the ENI engine may discover that there are two issues affecting the performance of the slices. The first issue is that video traffic on the consumer slice has spiked from a particular video content provider in a concentrated area, and the subscribers are experiencing poor network quality. The second issue is that simultaneously, a slice dedicated to transportation networks (i.e., trains, subways, and automobiles) is also congested due to the volume of users on the network in that same area. The first issue is solved by implementing a policy to manage video traffic down to a lower resolution for that specific video content type (since different video providers have different bandwidth needs for different resolutions). This immediately solves the problem for the consumer slice. The second issue is solved by implementing a fair usage policy among all of the devices attached to the transportation network to ensure that every user gets some bandwidth, but not targeting any specific application or user – an ineffective action, since the problem is related to the volume of users attached to the site. An alternate policy would be to allocate more resources to the transportation slice by "borrowing" capacity from other slices that were not using their allocation – through scale-out of the resources dedicated to that slice.

In reality, what may have happen in this example is that a number of trains pulled into a subway station at the same time, and all the users' connections transitioned into these two slices in a specific cell area. A few minutes from the detection of the issues, the users have dispersed and the trains have continued to their next destination. We will explore this same example more in the second part of this series, as it is likely that this is a daily recurring event when these trains pull into the station, qualifying this as a potential "planned event" that ENI can recognize as it begins to develop and implement the policies before congestion occurs.

Figure 7

Congestion Management example use case across multiple slices of a 5G network



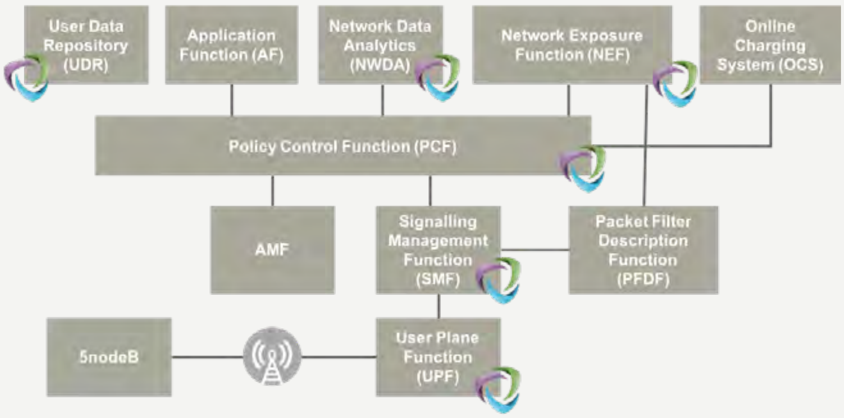
TRANSITIONING TO ACTIVE NETWORK INTELLIGENCE

Active Network Intelligence is Sandvine's strategy to enable network operators to leverage automated, closed loop use cases to meet the always increasing quality expectations of users while maintaining business profitability. Building world-class networks is hard, and the financial pressure to automate the management of networks is growing as networks virtualize and break the old deployment paradigms. Without the right data as a foundation for decision making, operators will struggle to meet these challenges.

Sandvine brings a unique value to mobile operators looking to make a financially successful evolution to 5G with our strong foundation of contextual data. As described in this whitepaper, the Sandvine eVolution probe and our Insights products deliver the user plane data required to gain a granular view of both overall network quality as well as individual subscriber quality, enabling an Experiential Network Intelligence Engine to make the right decisions on actions that need to be taken to resolve whatever issues or opportunities on the mobile network. As shown below, Sandvine can be deployed as a key part of an overall 5G architecture to deliver multiple components of a 5G architecture.

Figure 8

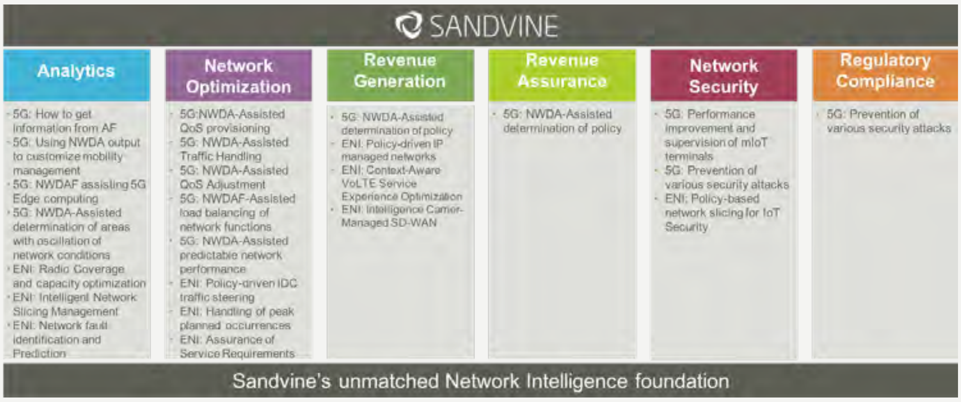
Sandvine applicability in the 5G architecture



Sandvine's Active Network Intelligence solution areas and use cases map directly to many of the use cases described by the 5G NWDA and ETSI ENI documents. In the Analytics area, both NWDA and ENI use cases map directly to the data that Sandvine can provide, especially in the area of QoE metrics. The use cases in the diagram below for analytics are focused from a Sandvine perspective on the use cases that Sandvine would not likely be part of the "action" taken to rectify the problem. The use cases in the Network Optimization, Revenue Generation, Network Security, and Regulatory Compliance areas all leverage the Sandvine analytics data to directly feed our implementation of an ENI engine and policy enforcement to deliver closed loop automation. The next paper in this series will examine several of these use cases in more detail, exploring what data would be needed and how different ENI engine decisions might be made depending on the real-time analysis of the data.

Figure 9

5G NWDA and ETSI ENI mapped to Sandvine's Active Network Intelligence

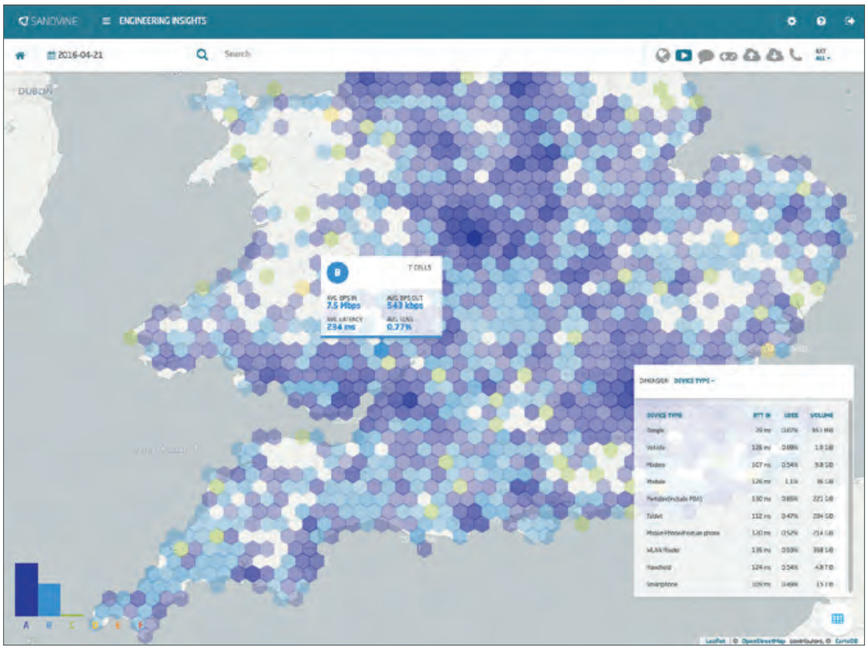


NEXT STEPS: ENI AND ACTIVE NETWORK INTELLIGENCE

This is part one in a three part series of papers on how Sandvine's Active Network Intelligence is closely aligned with the industry direction for next generation networking. This paper has focused on the data needed to power Experiential Network Intelligence and build out a 5G network with an NWDA to feed operational policies.

The next paper in the series will focus on analyzing the data to determine the correct actions for a number of different use cases leveraging artificial intelligence and machine learning in the ENI engine (as described by the ETSI documents). The final paper will focus on how to enforce the policies that are suggested by the ENI engine and how the enforcement granularity needs to match the data granularity to be truly effective.

If the data that was discussed in this paper is available, a network operator can very quickly visualize their network performance to quickly diagnose hotspots, outliers, and overall network QoE as shown in the example below. Without the right data, a mobile operator will not be able to meet the application QoE expectations of subscribers and even worse, will not know if their subscribers are having bad application experiences. We will delve into more detail in the second part of this whitepaper series.



v20180705

ABOUT SANDVINE

Sandvine helps organizations run world-class networks with Active Network Intelligence, leveraging machine learning analytics and closed-loop automation to identify and adapt to network behavior in real-time. With Sandvine, organizations have the power of a highly automated platform from a single vendor that delivers a deep understanding of their network data to drive faster, better decisions. For more information, visit **sandvine.com** or follow Sandvine on Twitter at **@Sandvine.**

SANDVINE

USA
47448 Fremont Blvd,
Fremont,
CA 94538,
USA
T. +1 510.230.2777

EUROPE
Birger Svenssons
Väg 28D
432 40 Varberg,
Sweden
T. +46 340.48 38 00

CANADA
408 Albert Street,
Waterloo,
Ontario N2L 3V3,
Canada
T. +1 519.880.2600

ASIA
RMZ Ecoworld,
Building-1, Ground Floor,
East Wing Devarabeesanahalli,
Bellandur, Outer Ring Road,
Bangalore 560103, India
T. +91 80677.43333