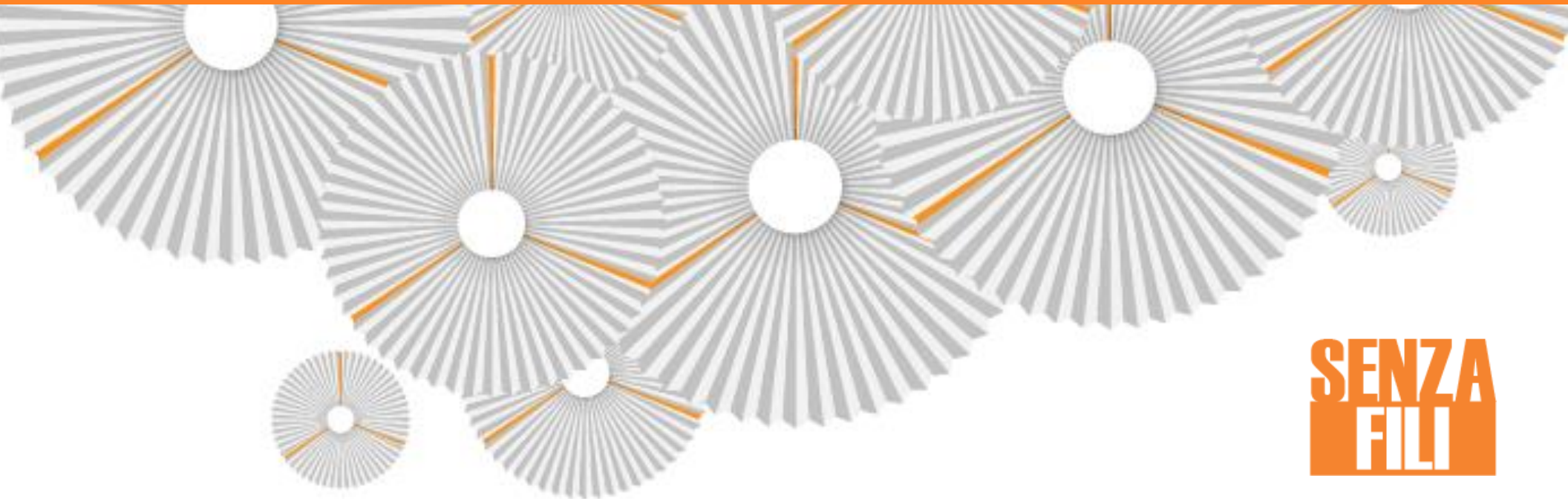
Getting the best QoE

Trends in traffic management and mobile core optimization

By Monica Paolini



SENZA FILI

We thank these companies for sponsoring this report:

ascom ALEPO Anritsu envision : ensure

VIAVI guavus FLASH networks

In collaboration with

RCRWireless INTELLIGENCE ON ALL THINGS WIRELESS

Table of contents

I. Market report

1. Introduction.

A Copernican shift to QoE-based network optimization

The way mobile operators optimize their networks has started to change. It has to change to keep up with the growth in usage and the evolution in traffic composition, and operators are aware of the urgency of the change and eager to embrace it. The end goal – give subscribers the best experience allowed by the efficient use of the network infrastructure – remains the same. The way to achieve it is what is changing today.

In the legacy environment that still dominates, operators primarily maximize performance at the element level. They make sure each network element performs at its best, with performance measured by industry-accepted KPIs, on the assumption that this gives subscribers the best experience the network can support. Operators try to push as many bits as they can, as fast as they can, through their network to increase network utilization. Optimization at the network-element level maximizes throughput; hence it works in environments with homogeneous traffic composition – the type of environment mobile operators faced until recently. However, the greater volume and variability in traffic have rendered this approach inadequate in optimizing the use of network resources and expanding revenues.

The new approach to optimization no longer assumes that what is good for the network is good for subscribers. Operators first establish what the experience is like for their subscribers, and then optimize the network based on the subscriber experience. This change does not amount to swapping existing KPIs with new ones. Instead, it is a Copernican shift from a network view to a subscriber view – a shift that requires a new perspective on how to assess network performance. Successful optimization depends on the operator having an end-to-end view of the network that is tied to QoE – one that is deep enough to relate QoE measurements (and performance issues) to specific network elements, and broad enough that it sees what is happening in the core and the RAN at the same time. And this must be done in real time.

Legacy: Element-based optimization	Evolution: QoE-based optimization
Maximize throughput, minimize latency across network and across applications	Measure and maximize QoE at the application level, using multiple data sources, such as geolocation or crowdsourcing
Test, monitor and optimize performance of individual network elements	Use end-to-end view of network performance to identify root causes of QoE performance issues
Static networks	Dynamic networks
RAN and core largely operate separately	Integrated RAN and core optimization
Historical and averaged KPIs	Real-time, location-aware and application-aware QoE data, complemented by network KPIs

2. Drivers and enablers for QoE-based optimization.

It is not just about network capacity and speed – it is about how they are used

Until recently, a well-behaved core and a fast RAN were the objectives of optimization. In this report we argue that this approach is no longer sufficient or cost effective, because it does not allow operators to use network resources efficiently and it constrains revenue protection and creation. We argue, as well, that the tools exist to move to a QoE-based approach to optimization.

The first half of the report explores six drivers we identified that make the legacy approach inadequate and that create the opportunity for a more effective way to manage traffic in a mobile network. The six drivers jointly create an environment in which mobile traffic is more complex and diversified. This makes optimization more difficult, but it also allows operators to extract more value from their networks.

Services on mobile networks today are priced and managed much like real estate would be if all houses were priced based solely on square footage, and if developers saw building new houses as their only way to expand revenues. QoE-based optimization allows operators to move to a model more similar to the actual real estate market: in the same way that two similarly sized houses may sell for a different price depending on location, construction quality or design, network resources are more selectively allocated to the services that create more value – e.g., a higher QoE.

New technologies and new solutions have recently become available to support this shift, giving mobile operators the tools they need to access, analyze, and use the performance data they need – and at the time resolution and granularity that is relevant. The context-aware, end-to-end network view that operators have makes it possible to understand and quantify QoE, and then use this information to manage traffic. The second part of the report looks at these enablers and how they help operators address the challenges they face as traffic keeps growing in volume and complexity.

Drivers: What matters
Time
Location
Devices
Applications
RAN conditions
Policy

Enablers: The way forward
Measuring QoE
Real-time traffic management
Big data, analytics
Virtualization
End-to-end integration
Mobile Edge Computing (MEC)

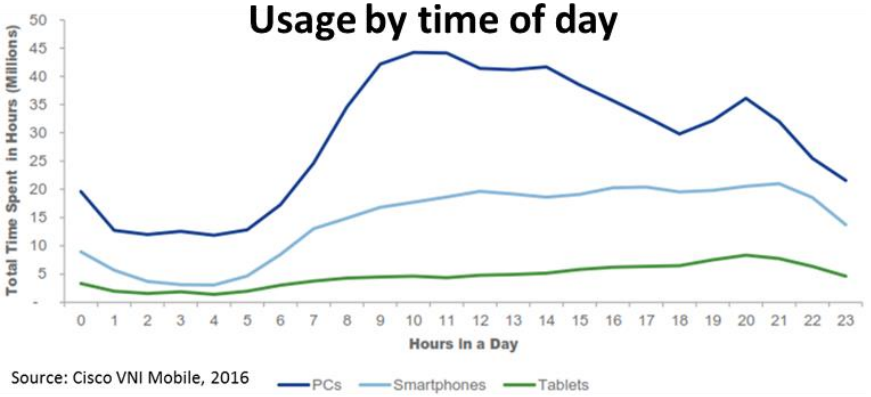
3. What matters: Time.

Unpredictable time fluctuations make a case for real-time optimization

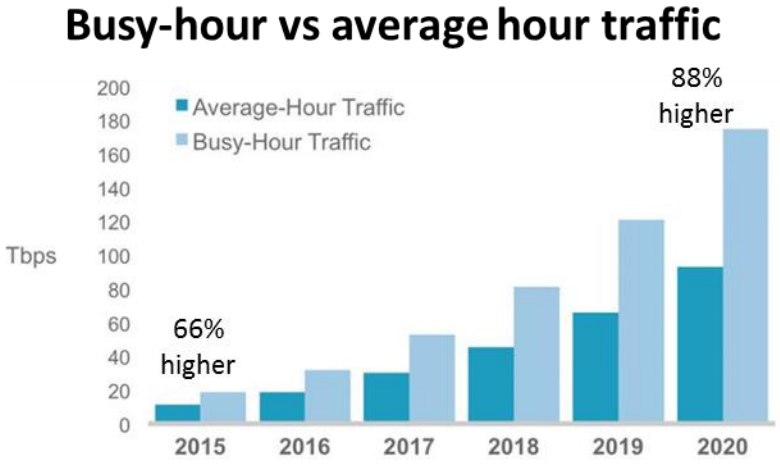
The first driver to QoE-based optimization is nothing new: usage depends on time, with traffic concentrated during peak hours. What is changing is that mobile operators have access to data on traffic fluctuation and the tools to manage it. For instance, mobile operators can optimize video traffic as a function of the network load at a given time – e.g., only when the network is at capacity or congested.

Time-of-day usage patterns interact with location, magnifying the impact of traffic distribution through time. Time-of-day curves like the one on the right show the average of usage across the network footprint, so they do not capture the extent of the extreme traffic concentration that we can observe at the single-cell level and that creates the need for real-time, location-specific optimization. Several different temporal dimensions affect usage and performance:

- **Predictable time-of-day changes.** Historical data can be used to predict and accommodate traffic patterns due to factors like commuting or school schedules. These account for most of the variability in the long term, and they are the easiest to manage, because they do not require real-time optimization.
- **Unpredictable changes and residual variability.** Traffic changes due to accidents or other unexpected events cannot be predicted. Mobile operators need to respond in real time to prevent disruption and, eventually, they should become able to anticipate usage peaks as they notice traffic growing.
- **Microbursts.** These extremely short-term fluctuations are due to the inherent variability in IP traffic at a high time resolution. As a result, even in a network running at capacity, there are time slices during which the network is underutilized unless the operator purposely optimizes traffic management to smooth out access – e.g., with application awareness in traffic management. This type of optimization requires the ability to access and analyze network data at a fine temporal resolution.



Source: Cisco VNI Mobile, 2016



Source: Cisco VNI Mobile, 2016

4. What matters: Location.

Real-time optimization shores up densification in the RAN

Along with time, location is the other fundamental dimension of traffic concentration. Locations such stadiums, public transportation hubs, malls, and city centers attract a disproportionate amount of mobile traffic. The interaction of time and location produces an extremely uneven distribution of traffic. Even in highly urbanized countries like Japan, areas with dense usage are few, and mostly have extremely high traffic loads only during a few hours each day. Yet these are the locations where availability of mobile services is highly valuable to subscribers. Events like the Super Bowl in the US – and the prominence that operators give to the event both in terms of investment and disclosure of usage data – remind us of how crucial it is for operators to retain good QoE at times of extremely high traffic load.

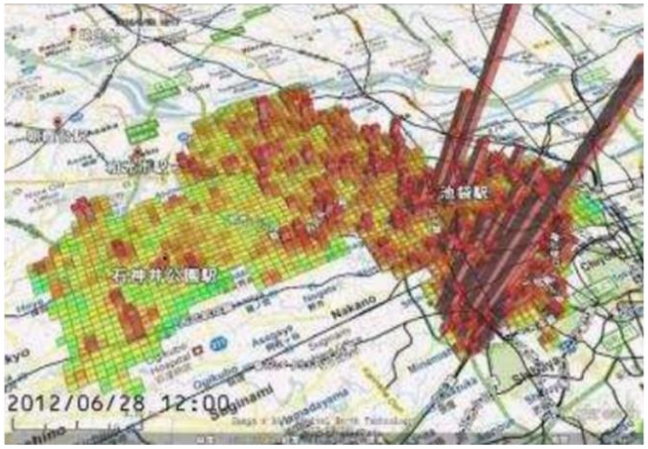
RAN densification efforts are crucial to addressing the concentration of usage, but optimization in the core is a highly valuable complement that helps operators manage both predicted and unpredicted usage peaks. Expanding capacity to meet demand is a very expensive proposition, and to some extent a self-defeating one. It has been shown many times that as soon as capacity increases, subscribers find a way to use it and average usage goes up. This is in part due to the fact that current network usage does not necessarily reflect demand – a congested network by definition does not meet demand. But in addition, more capacity and faster speed allow subscribers to use more services during the same amount of time.

As a result, we should think of mobile networks as being capacity constrained, even as technology evolves. Over the last decade, growth in usage has largely outstripped the growth in performance, and growth in usage is not matched by the growth in revenues. So it is financially unfeasible for operators to expand RAN capacity to meet demand. Real-time optimization is a necessary tool for managing traffic in a capacity-constrained environment, because it allows operators to extract the most value – in terms of QoE and in terms of profitability – from existing network assets.

Super Bowl 2016
Verizon stats from the stadium

- $70 million cost to update the infrastructure in the area.
- 7 TB of data.
- More than 35,000 estimated unique devices.
- Recorded download speeds reached 120 Mbps from inside the stadium.
- Video was 20% of traffic, social media 18%.
- Facebook was the preferred social media app, with 12% of total usage.
- Messaging and email accounted for less than 3% of data.

Extreme concentration of traffic in Tokyo



Source: KDDI

5. What matters: Devices.

Moving beyond smartphones with IoT

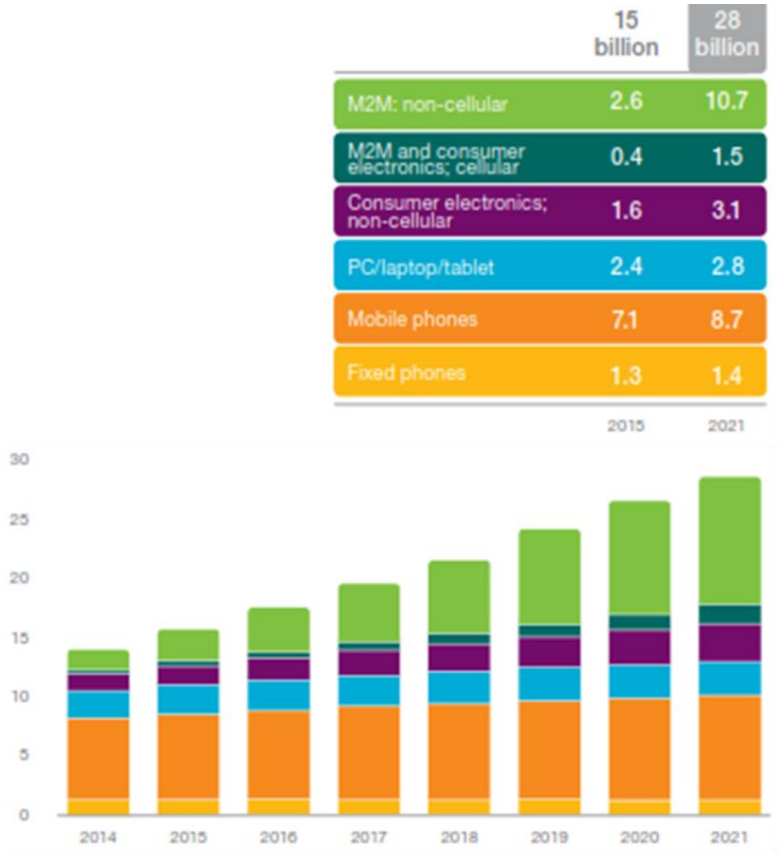
Mobile phone and smartphone adoption have both been amazing success stories. In 2015, 62% of the population worldwide had a mobile phone, and this percentage is expected to grow to 69% by 2020, according to Cisco VNI Mobile. By comparison, in 2015 only 37% of the population had a landline phone, and 41% had running water. As we reach saturation in some markets, growth is slowing down.

IoT is set to fill in the growth potential for connected devices. Its growth has been slow to date, because IoT requires a complex ecosystem, new business models and new use cases. But there are many encouraging signs, from multiple vertical sectors, that growth has started to accelerate.

The high volume and diversity of devices – along with their widely diverging requirements – is the most pressing challenge from IoT. Some IoT devices require high mobility and are linked to individual users (e.g., connected cars); others are fixed and stand-alone (e.g., utility smart meters). Some have strict performance and bandwidth requirements (e.g., security webcams), while others need to send only a few bits intermittently (e.g., sensors). Revenues per device will vary greatly too.

This diversity increases the complexity of traffic management for operators used to dealing with a limited number of form factors and a homogeneous set of service plans, compared to what we expect to see with IoT.

At the same time, IoT represents the opportunity to extract more value from the network by optimizing traffic depending on demand for different IoT services, and applications and devices. For instance, smart meter communications can be treated as low-priority traffic that is transmitted when there is unused capacity, and hence the marginal cost of supporting the service is very low; in many markets low cost is a requirement to attract utilities that cannot afford expensive contracts.



Traditional landline phones are included for legacy reasons.
Examples of M2M include connected cars, machines and utility meters, remote metering.
Note: A connected car is herein counted as one "thing" though it may have hundreds of sensors.
Examples of consumer electronics devices include: smart TVs, digital media. boxes, Blu-ray players, gaming consoles, audio video receivers.

Source: Ericsson Mobility Report, 2015

6. What matters: Applications.

QoE depends heavily on application and traffic type

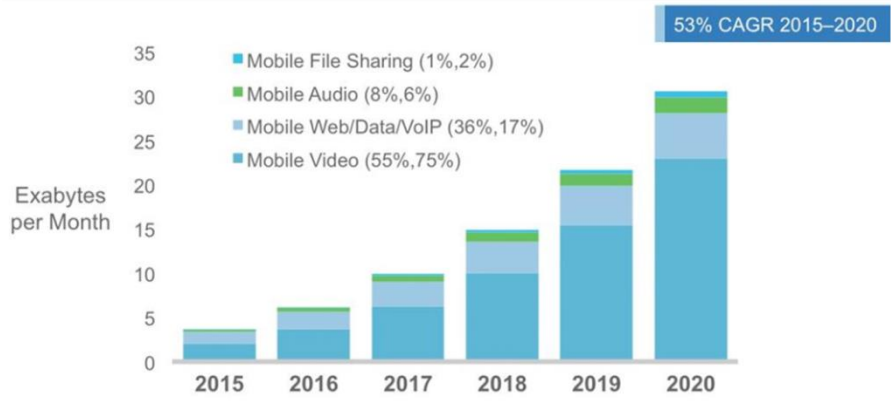
The growth in mobile usage tells us that we do more with our connected devices, more frequently and for a longer period. By the end of 2014, Nielsen estimates, subscribers spent 37 hours a month using mobile apps, up from 23 in 2012. According to Google, in 2015 smartphone users had an average of 36 apps installed. The change that is most significant, however, is the increase in video traffic: Cisco VNI Mobile estimates a 55% CAGR for video will occur in the 2015–2020 period, with video accounting for 75% of mobile traffic by 2020. The dominance of video traffic comes primarily from the bandwidth-intensive nature of video, but also from the increase in demand for video content and the networks' ability to carry better-quality video than in the past.

Sandvine data shows that the top application continues to be Facebook (19%), and the top downstream application is YouTube (21%). But below the top applications there is considerable fragmentation (Skype is 1% of upstream traffic, Google 3% of downstream traffic). Voice accounts for less than 5% of traffic, according to Ericsson Mobility Report. With IoT we should expect to see increasing variability in the types of traffic and applications.

QoE is extremely sensitive to what we do with our mobile devices. The same latency and throughput on a given device may translate into different QoE depending on whether the user is texting or watching a video. There are three key elements that affect QoE:

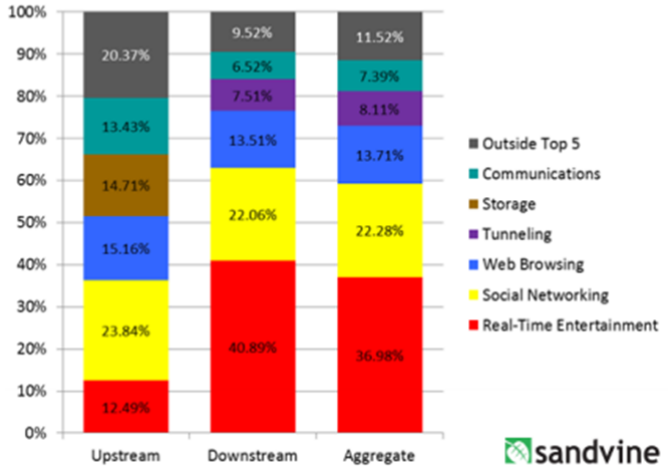
- Traffic type: voice, video, best-efforts data, each with its own set of performance requirements
- Service: video could be streaming (downstream), conversational (upstream and downstream), or broadcast; typically, conversational video requirements are more stringent

The evolution of traffic type



Figures in parentheses refer to 2015 and 2020 traffic share.
Source: Cisco VNI Mobile, 2016

Peak Period Traffic Composition (North America, Mobile Access)



- Application: OTT or managed by the operator (e.g., Skype vs. VoLTE), or supporting different services (YouTube and Netflix for entertainment video; Skype and Zoom for conversational video)

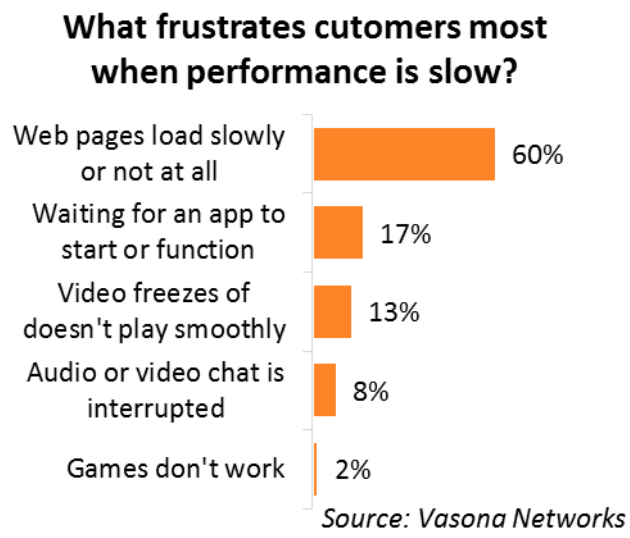
Not only does the QoE depend on the traffic type, service and application, the relevant QoE metrics and optimization strategies change.

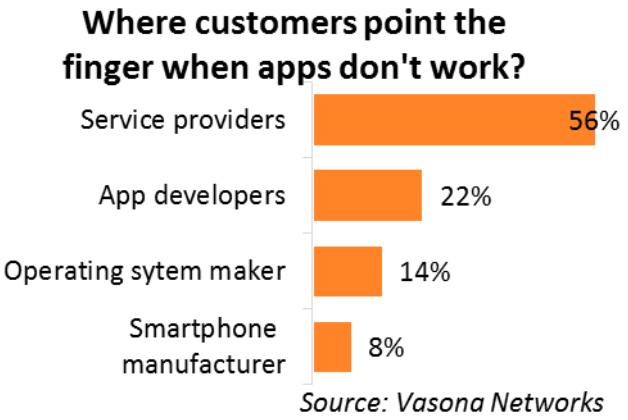
Usage models and subscriber expectations are still evolving, and mobile operators have only started to measure QoE quantitatively. So we still do not have established QoE metrics with industry-wide acceptance that are specific to traffic type, service and application. But it is clear that we need to establish a better approach to measuring QoE, and that mobile operators have to set the values they find acceptable for the metrics relevant to different types of data use.

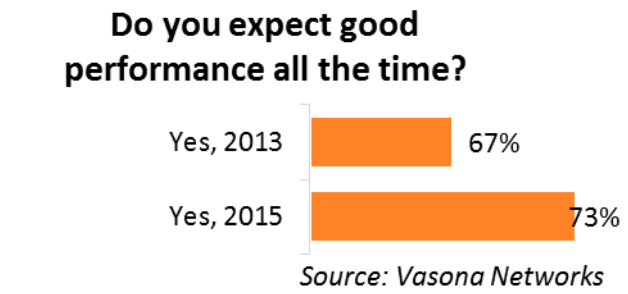
For instance, latency and stalling are relevant for video, but not for web browsing and texting. Dropped audio and video calls are still a major element in QoE, but less so for streaming video – it is easier to restart a video stream than to call someone back. At the other end, subscribers will hardly notice high latency on messaging or social applications, but reliability becomes crucial. Subscribers may accept that they cannot view a YouTube video, but they may be less tolerant if the operator fails to deliver their texts.

The dependency of QoE on the perceived quality of a service or application, rather than the averaged network KPIs or averaged QoE metrics, makes a strong case for optimizing network traffic at the traffic type, service and application levels. Operators are already doing this with VoLTE when they give it priority over other types of traffic. But more can be done by managing traffic in an application-aware and traffic-aware fashion.

Splitting traffic into network slices, each with specific requirements, is a path some operators have started to follow. It gives operators a framework within which to manage traffic by taking into account different requirements and constraints. Adoption of IoT will further increase the appeal and usefulness of approaches such as network slices.







7. What matters: RAN conditions.

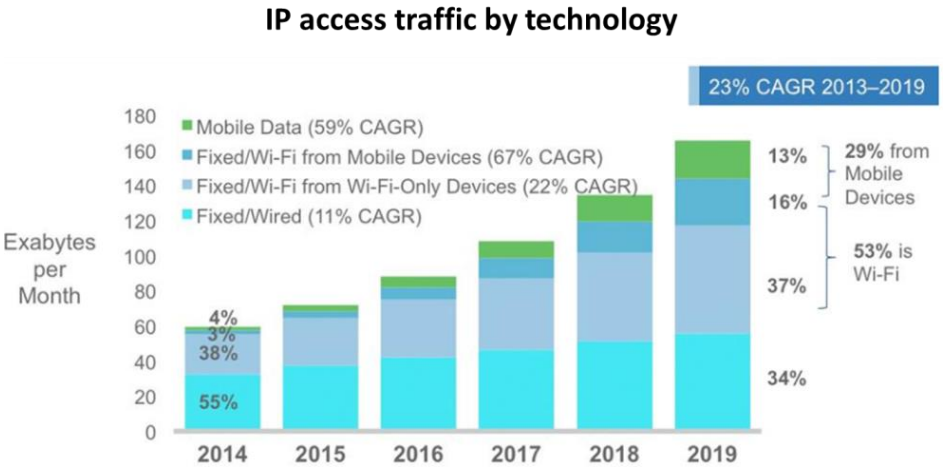
RAN and core performance are mutually strengthened by optimization

In legacy, element-based optimization, mobile operators optimize the RAN and the core as two largely separate units. In the RAN, the goal is to meet specific targets – e.g., for throughput, latency, packet loss, dropped calls. As the focus of optimization moves to QoE and end-to-end network optimization, the role of the RAN remains prominent, but is more closely tied to the core and has to be related to the QoE measurements. Virtualization and initiatives like MEC are already dismantling the border between RAN and core, by moving them physically closer to each other and making them functionally more interrelated. With HetNets and SON, the use of unlicensed spectrum, and C-RAN topologies, the RAN is becoming more dynamic; hence the opportunities for optimization are expanding.

More crucially, however, RAN and QoE are closely intertwined, although the link is not a deterministic one: a bad QoE may be caused by RAN congestion or performance issues, but other causes may explain it; bad performance in the RAN is likely to be reflected in the QoE, but subscribers may not notice it if they are only texting.

The challenge for operators is to understand when the RAN is responsible for a bad QoE and how to address the issue. Ideally this should be done in real time, as soon as QoE starts to degrade, so that the operator can take corrective or preventive action right away. This action may or may not involve the RAN, however. In many cases, core optimization may be more effective at addressing limitations in RAN performance – e.g., those due to congestion.

Awareness of RAN conditions in real time is also required to put QoE measurements in perspective. While operators strive to provide the best service at all times and places, there are environments, such as deep-indoor locations, where QoE may not be great, but connectivity is still valuable. In these environments, operators can use core optimization to improve QoE or to improve how they manage the traffic.



Source: Cisco VNI Mobile, 2016

8. What matters: Policy.

The crucial link between optimization and monetization

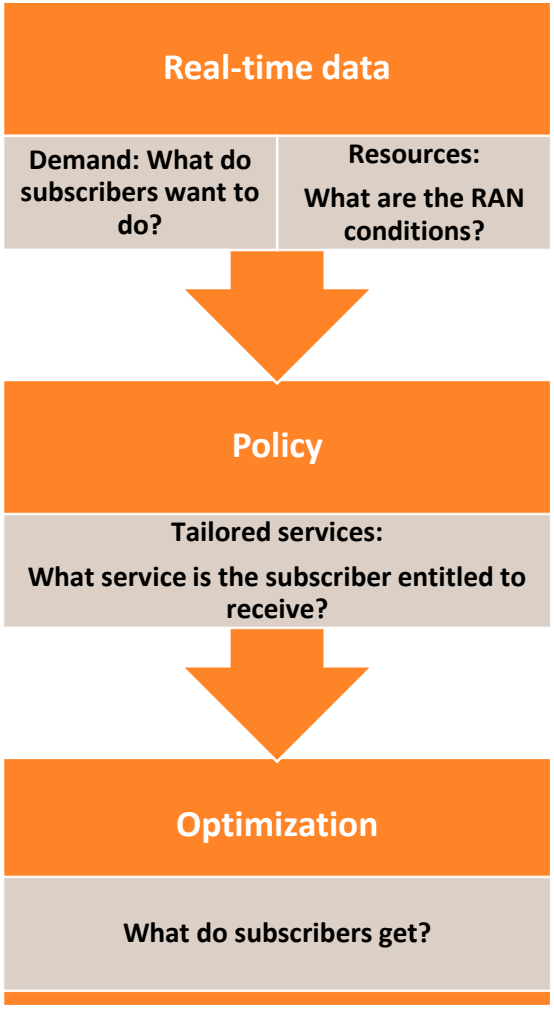
Operators do want to provide subscribers with a good experience. But as they try to be guided by QoE measurements to optimize their networks, operators will not necessarily be able – or desire – to provide all subscribers the same level of service.

As discussed before, mobile networks are capacity constrained, and when a network is running at capacity or has become congested, demand exceeds resources, and this limits what subscribers get. At the same time, mobile operators that provide customized, tiered or otherwise tailored services have to provide the level and type of service each subscriber has paid for. In both cases, operators set policy rules ahead of time to help them manage demand in relationship to the capacity and other performance capabilities in the network.

Policy can be thought of as a tool that fine tunes the level of service a subscriber gets. Today, policy is most commonly applied without taking into account real-time demand and resources. This is effective in implementing service definitions and contractual agreements with subscribers, but it does not allow the mobile operator to take full advantage of the dynamics of network utilization, nor does it allow the subscriber to extract all the benefits from the resources available.

Mobile operators can integrate policy and optimization by taking real-time demand and RAN conditions as inputs that shape the way policy is implemented in the live network. In doing so, operators gain two advantages:

- Operators maximize the value they extracts from the existing resources, giving priority to the data flows that are most valuable (while respecting the applicable regulations, such as net-neutrality requirements).
- They can offer new tailored services that give subscribers better service (e.g., giving subscribers on low-cost plans free access at nonpeak hours when there is unused network capacity), without adding to their own marginal costs.



9. What matters: Recap.

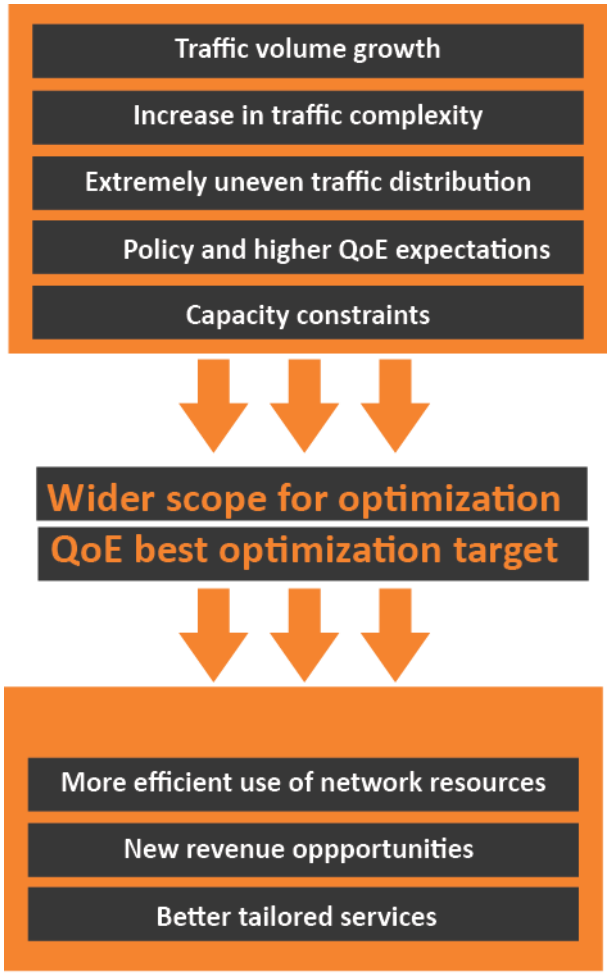
A wider scope for optimization drives the move to QoE focus

The first half of this report looked at what matters in defining the subscriber experience and identified six drivers, all of which contribute to defining or shaping QoE. In a legacy environment, traffic homogeneity limits the impact that these drivers have, because the scope for optimization is restricted in the absence of traffic differentiation. As traffic volume and complexity grow, operators can leverage traffic diversity through optimization. The six drivers then become more powerful as data sources to support optimization efforts in the mobile core, because they have a key role in defining QoE. Furthermore, these drivers interact with each other, creating new opportunities for mobile operators to use their network assets more efficiently and to extract more value from them.

The increased expectations that subscribers have for quality of service, and the ability of mobile operators to leverage the quality of service with policy, make the focus on QoE even more compelling, because it gives operators another dimension along which they can extract maximum value from their existing assets.

The advantages that operators can gain from QoE-based optimization in the core provide a strong motivation toward this new approach. The transition, however, is not easy, because the wireless ecosystem has much work in store to learn to measure QoE, track it in real time, and use it as an input to network performance optimization.

The main challenge is that, unlike the KPIs currently used, QoE is a holistic, subjective measure of subscriber perception that reflects the performance of the network, but only in a nonlinear, indirect way. Yet QoE is what matters to operators – it is the subscriber experience, not network performance, that generates revenues, retains subscribers and attracts new ones. In the new QoE-based approach to optimization, subscribers no longer have to adapt to what the network can offer; it is the operators that have fine tune the network to meet subscribers' expectations.



Source: Senza Fili

10. The way forward: Measuring QoE.

Testing and monitoring take a more active role in optimization

Operators – as well as vendors supporting them – have to learn what the relevant QoE measurements are, and how they relate to the performance of the end-to-end network and the elements or functions within it. The challenges listed on the right reflect the fact that operators have not previously used QoE explicitly as input for network optimization. To make these challenges manageable, operators need to move to a new way to test and monitor their networks.

To date, testing and monitoring are largely separate from operating – and optimizing – the network. They are mostly confined to the network deployment and upgrade phases, and, later, to fixing performance issues. With the transition to QoE-focused, real-time optimization in networks that are inherently more dynamic, testing and monitoring become ongoing tasks that tend to converge, making the demarcation between the two less sharp and less relevant. Measuring performance becomes a continuous activity in networks in which:

- New elements such as small cells or Wi-Fi access points are added as needed, as part of the organic growth of the network or on a temporary basis
- Network capacity may change, for instance when using unlicensed spectrum, which is shared with other operators or users
- SON changes the RAN configuration in response to traffic demand
- Real-time traffic management alters traffic flows to increase network utilization
- Because of virtualization, the hardware substrate supporting network functions continuously changes

Not only do testing and monitoring acquire a more prominent role in ensuring good network performance; they also become more closely linked to core optimization efforts, increasingly providing the inputs needed to measure QoE and relate it to the performance of network elements and of the network end-to-end.

The challenges of measuring QoE

QoE is subjective. QoE is measurable in a quantitative way by multiple metrics. However the relative importance of these metrics (e.g., in video, how does loading time compare to stalling?), and what subscribers rate as good or unacceptable, are largely subjective and vary from one subscriber to the next. Even for the same subscriber, QoE may vary on factors such as time, location, device or application. This makes it difficult for operators to define a target QoE, or how much QoE degradation is acceptable under which circumstances, for instance.

QoE reflects the subscriber's perception of what the device delivers. And yet it is the network, not the device, that operators can optimize. Bad QoE may be caused by device issues, but those have to be dealt with separately – e.g., through customer support. To gain insight from QoE measurements, mobile operators have to understand how QoE is linked to end-to-end network and element performance – and to the KPIs that measure their performance.

QoE degradation causes are difficult to trace. The same low QoE measurements (e.g., video stalling, or high latency in voice calls) may be due to different causes – and different elements or functions – in the network. Or they may be caused by issues from outside the network – e.g., on the internet or other operators' networks.

QoE depends on traffic type, service, and applications. The same latency, for instance, has a variable impact on QoE. Latency, for instance, is crucially important for voice and video, but much less so for social networking or messaging.

There are no standard QoE metrics for data applications. There are many ways to quantify QoE. The need to go beyond established KPIs is accepted, but there is still no agreement on how best to measure QoE or what the right metrics are.

11. The way forward: Real-time traffic management.

The gradual path toward real time in QoE-based optimization

Network optimization is still mostly done using historical data, under the assumption that networks are largely static and changes are to be planned in advance and not too frequent. There is a clear trend, however, toward running mobile networks in real time, because this enables operators to push optimization further. Moving toward real time enables operators to take into account network activity that historical data cannot predict, and to address performance issues in a timely manner. Because real-time network management brings additional complexity, operators are treading lightly to make sure the transition is safe and at a pace they are comfortable with.

While QoE historical data is useful in providing a high-level view of network performance, much of the QoE value stems from the real-time component that allows operators to address any QoE degradation quickly. Managing networks in real time carries multiple requirements:

- Ability to measure real-time performance of the network
- Ability to analyze large amounts of data from multiple sources at variable time resolution and depth of analysis, and to quickly identify appropriate corrective or preventive action
- Automation of the testing, monitoring and optimization processes, which is required by the need to continuously process large amounts of data
- Organizational and cultural change to accommodate the lower level of direct control of network operations that results from automation.

What is real time?

Is it milliseconds, minutes, hours or days? The answer depends on the task, and it is crucial to pick the appropriate temporal resolution for different tasks or functions. Too high a resolution leads to unnecessary overhead, or may lessen the benefits of optimization. Too low a resolution may bring only little changes. When dealing with RAN congestion or network disruption, a high temporal resolution is advantageous. When fine tuning a well-behaved network, the need for accuracy may warrant a slower pace.

Time resolution is getting shorter. The transition to real-time optimization and traffic management is gradual. Mobile operators have just started to move toward managing their networks in real time, but invariably they are still far from the targets they have set. The temporal resolution in the solutions available from vendors continues to shrink, though, and mobile operators are becoming more comfortable with the faster pace of change that real-time operations require.

Prediction is the ultimate goal. Real-time optimization is a clear improvement over using historical data alone, but to maximize performance and minimize disruption, operators ideally would predict demand surges and performance problems before they emerge. Monitoring network performance can be used to detect early signs of congestion or disruption, and enable operators to take preventive action.

Historical data is still important. With real-time data, operators can address unplanned and unpredicted changes in the network. Most of mobile usage and network operations, however, are predictable – and historical data effectively predict them. The adoption of real-time optimization does not remove the need for and value of historical data in providing mobile operators with the basic intelligence for the initial – and more substantial – network optimization.

12. The way forward: Big-data analytics.
Learning to harness complexity

The focus on QoE in network optimization requires real-time traffic management to be effective. In turn, running a mobile network in real time – or nearly real time – adds complexity. Measuring network performance, identifying events that may have or might soon cause service degradation or disruption, finding what caused those events, and, finally, resolving performance issues or inefficiencies in the network – this process requires a massive amount of work if done manually. Mobile operators could not afford it; they would not have the financial or human resources to do it. So they have largely kept using the legacy approach to optimization, which has limitations but is trusted, affordable and manageable.

The introduction of big-data analytics – aided by automation – is rapidly changing this, by making the evolved, QoE-based optimization approach feasible and cost effective. Dealing with the additional complexity is becoming worth the effort. With time, we expect big-data analytics to widen its scope by providing operators with a complete end-to-end view of network performance and the ability to drill down to single network functions when and where needed.

Access to data is not enough, though. Mobile operators have always had access to network performance data, for instance with probes. The challenge has always been in finding a way to collect and use this data in real time, without being overburdened by the complexity and sheer amount of processing needed, and then to be able to use it effectively. Analytics solutions do not change the availability of network performance data: they allow mobile operators to leverage it. This is not easy: processing power is needed, but what is crucial is the ability to identify what is relevant within the data collected, and learn how the data helps in assessing network performance. The answer depends on the specific task at hand – e.g., optimizing voice performance in a congested network, caching, or using SON in a HetNet – and it requires that mobile operators have the flexibility to access the same data at different levels of granularity in different contexts.

Big data and analytics: Divide and conquer through time
Before: Plan
Understand which data needs to be collected and analyzed for different tasks: relevant metrics, appropriate level of detail and temporal resolution. Decide how to analyze this data. Define trigger points or thresholds that warrant analysis or action.
During: Detect events throughout ongoing activity
Measure network performance, collect data. Detect events to be investigated.
Immediately after: Analyze event, resolve issue
Zoom in or zoom out of the data collected, to find the appropriate level of analysis and source of data. Identify performance events that lead to suboptimal use of resources, cause performance issues, or cause QoE degradation or disruption. Identify root causes for these events, and assign them when possible to specific network functions. Propose a solution, with corrective, protective or preventive action, if warranted.
After: Learn
Learn, from collected data and the analysis of it, how to define events that require investigation, how events relate to network functions, and which solutions worked and which did not.

13. The way forward: Virtualization.

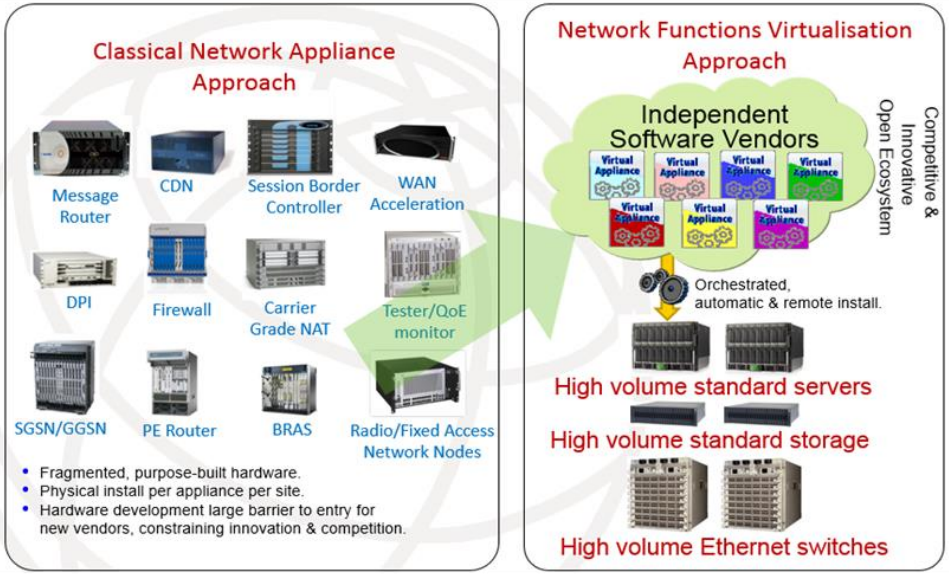
NFV and SDN create the dynamic core that real-time optimization needs

The path toward virtualization has been enthusiastically embraced by mobile operators across the world, even though each proceeds at its own pace, following its own strategy. In the US, AT&T, one of the most aggressive supporters of virtualization, is committed to virtualizing 75% of its functions by 2020, starting with mobile core functionality. "No part of our network will be unaffected," AT&T Senior EVP John Donovan has said. "We have catalogued the hundreds of network functions that we manage, and decided which will be relevant in the future and which are becoming obsolete. Of these 200 critical future functions, we plan to move 75% to this software-defined architecture in the next five years."

Efforts to move to a virtualized mobile network and to adopt real-time, QoE-based optimization are mutually reinforcing. Virtualization transforms the core network to become truly dynamic, giving mobile operators the ability to allocate off-the-shelf hardware resources to the functions that need them, in real time. In turn, real-time optimization provides operators the input they can leverage to determine which resources each function needs.

For instance, a RAN that is congested or approaching congestion can trigger a degradation of QoE. This can be averted by using content optimization – and, more specifically, video optimization, since video is the dominant traffic type. On the other hand, when content optimization is unnecessary, limits revenue, or becomes an overheard burden, it can be switched off. In a virtualized core, the hardware resources can be dynamically allocated to (and removed from) content optimization when needed. This in turn provides the basis for the capex and opex savings that are expected to come from virtualization.

Classical and virtualized approaches



Source: AT&T

14. The way forward: End-to-end integration.

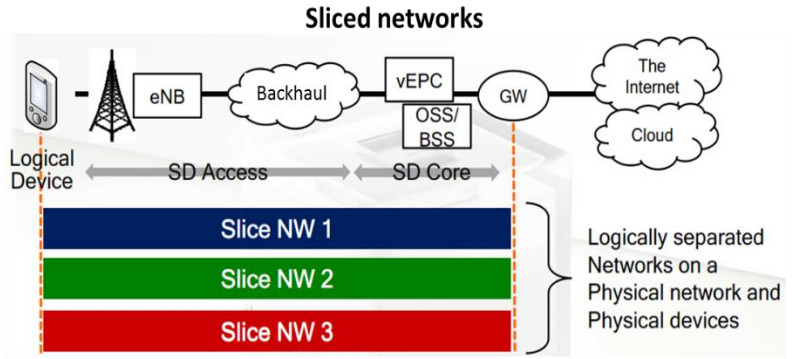
The border between core and RAN vanishes

The core and the RAN have been long treated as two largely separate entities, with the core in a supporting role relative to the RAN. The RAN captured most of the attention, investment and performance accolades. This separation is gradually fading, because it no longer serves a purpose, and instead hinders the ability of mobile operators to improve resource utilization in both the core and the RAN.

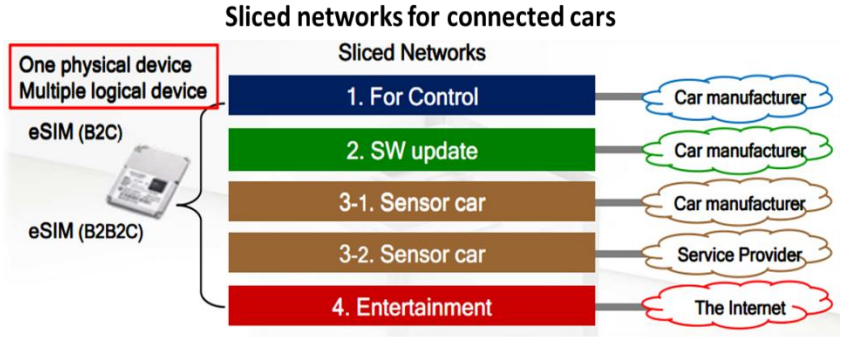
Specifically, when assessing end-to-end network performance to support a high QoE, the demarcation between RAN and core becomes irrelevant. To address an instance of QoE degradation, an operator needs a reliable end-to-end view into the network first; then it can identify and focus on the elements – whether they are in the core or in the RAN – that may be causing the problem. To do this efficiently, operators need tools that give them equal access to both RAN and core.

To successfully bring RAN and core closer to each other, mobile operators have to pursue a cultural and organizational change that can be challenging. The teams that work on the RAN and core need to collaborate more closely, and they may eventually merge. With the tighter integration of policy with end-to-end traffic management, the marketing and network operations teams also need to develop a stronger relationship.

Network slices have emerged as a way to conceptualize traffic optimization and management in terms of traffic type, service or any other features that the operator deems relevant, using a horizontal approach which cuts across the RAN versus core separation. Operators can split traffic into slices and then optimize and manage each slice separately from the others. This approach enables operators to respond to QoE degradation in a very precise way, by limiting their actions to those streams that require or benefit from fine tuning.



Source: KDDI



Source: KDDI

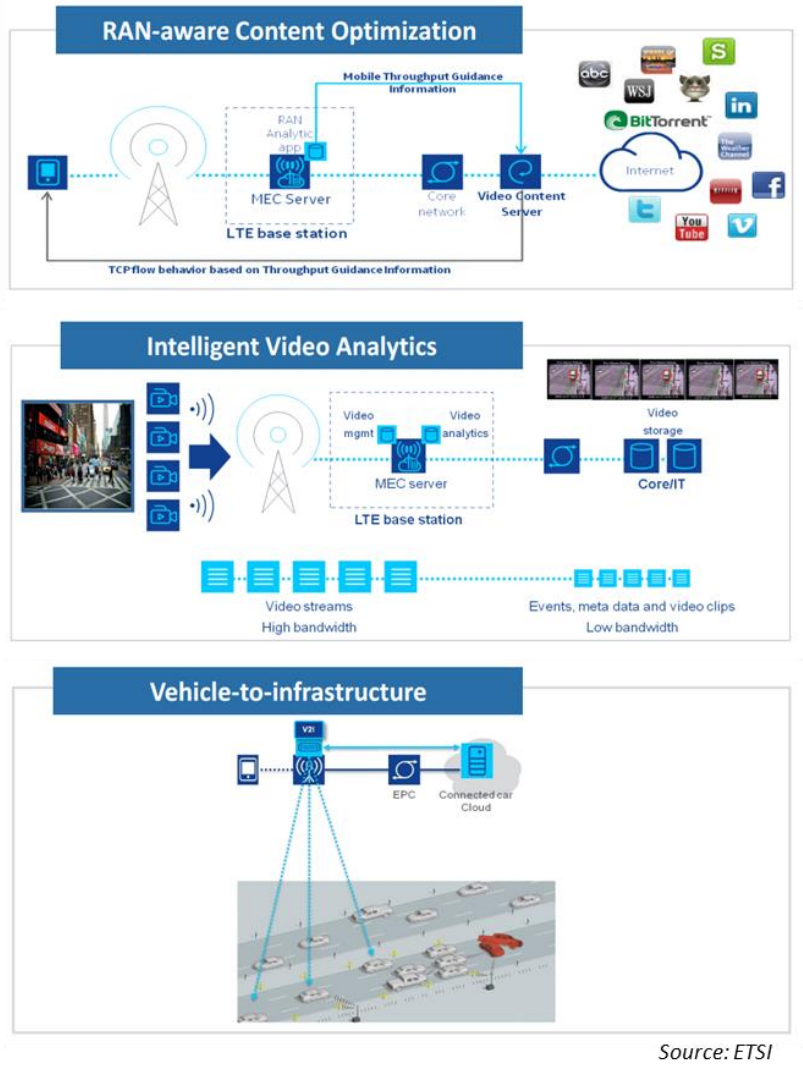
15. The way forward: Mobile Edge Computing (MEC).
Taking virtualization to the edge

Virtualization adds the possibility of closer interrelation between core and RAN. RAN topologies move the edge elements toward a more central location, and the virtualized core enables operators to locate core functions where they are most effective.

Network functions have become independent not only from hardware, but also from location. This does not mean that location no longer matters. On the contrary, location can be used to improve the value and performance of network assets. Functions with limited real-time requirements, such as policy, are likely to remain centralized, because this avoids duplication and complexity. Functions that are latency sensitive, such as content optimization, video caching, or video analytics, are more likely to benefit from being pushed to the edge – possibly merged with the C-RAN or vRAN infrastructure – because this can lower latency. In some cases, a distributed approach may improve resilience, scalability, automation and redundancy.

In this perspective, MEC, an ETSI initiative with strong operator support, provides an organic extension to virtualization that enables mobile operators to use their RAN and core resources more efficiently, by moving some of the core functionality that today is centralized to the edge.

MEC and, more generally, efforts to move functionality to the edge may enable a more efficient use of network resources – i.e., improve RAN performance and increase traffic served – and hence improve QoE. Because network functions are then closer to the subscriber, MEC further expands the scope for optimization, making it easier for operators to differentiate how they fine tune different parts of their networks.



Source: ETSI

16. Implications.

The path to a higher QoE and more efficient network resource utilization

Mobile operators are moving from a network-based to a real-time, QoE-based approach to optimization, in which they explicitly maximize QoE rather than network performance.

Network capacity is still important. But what is crucial is how operators use this capacity to meet subscribers' expectations.

Increased complexity and volume in traffic, along with traffic's uneven distribution, have made networks more dynamic but still capacity constrained. This leads to a prominent role for policy, and widens the scope for end-to-end optimization.

QoE-based optimization can yield a more efficient use of network resources, new revenue opportunities, and better-tailored services.

The transition to QoE-based optimization requires time and learning. It is challenging to measure QoE and to relate QoE metrics to network KPIs and performance, but operators and vendors are committed to meeting these challenges.

Together, a new approach to testing, the ability to collect and analyze real-time data with analytics solutions, and virtualization provide a solid base for QoE-based optimization to succeed.

As subscribers' usage models evolve, we can expect the concept of QoE – what subscribers consider a good QoE, what defines a good QoE, and what performance tradeoffs subscribers are willing to accept – to change as well. As operators and vendors try to find the appropriate metrics to define QoE, we need to stay flexible enough to keep updating our definition of QoE and how it can guide network optimization.

II. Vendor profiles and interviews

Alepo

Core optimization solutions

Since 2004, Alepo has provided software solutions for networks of service providers worldwide. It has extensive experience with both tier-one operators and smaller ones, and in both developed and emerging markets. Customers include mobile operators, wireless ISPs, MVNOs, Wi-Fi operators and wholesale operators. Alepo's solutions cover the LTE EPC (OCS, PCRF and HSS), AAA and policy, convergence gateways, OSS and BSS, mobile device management, and carrier Wi-Fi monetization and offload.

Alepo's PCRF is the product that bridges policy and QoE to drive network optimization. It is a policy control engine for Diameter EPC networks that is fully compliant with 3GPP standards and is interoperable with equipment from other vendors.

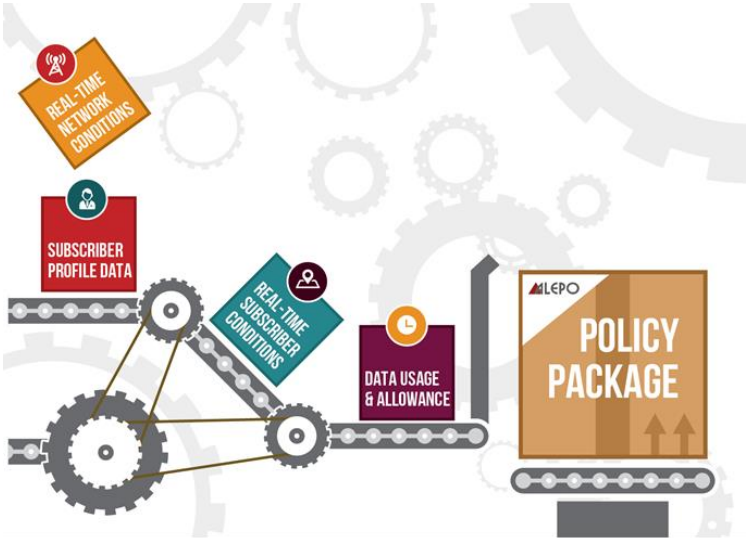
The PCRF provides operators with a wide range of service templates they can deploy quickly. Operators can choose use cases from the templates that their subscribers find compelling and that the network can support, and they use the templates as the basis for optimizing and managing traffic in real time. Examples include

- Bundled applications and content
- Roaming data pass
- Sponsored or toll-free applications
- Tiered services
- Location-based services
- Bandwidth on demand
- Shared data and data gifting
- Parental controls
- Data rollover
- Device tethering
- Fair use policies
- Data allowance then pay-as-you-go (overage)

Alepo's PCRF provides a growing, subscription-based library of templates that operators can adapt to define and launch new services, or modify existing ones, without having to use a scripting language to customize them. This approach allows operators to change their service offering rapidly to reflect market conditions, operator strategy, network capacity, or feedback from subscribers.

These services can be based on policy that is aware of network conditions, giving them a crucial role in optimizing network performance. For instance, policy may dictate how to manage traffic based on RAN load. The same content and the same application used by the same subscriber may be treated differently depending on time, location or real-time RAN conditions.

The flexibility and ease in launching new services or refining existing ones not only increases the operators' efficiency in using network resources; it also supports their efforts to monetize their services and to retain their subscribers, by offering services that subscribers find more attractive. Operators can do this by widening the options available and making them more transparent to subscribers, by using analytics to evaluate subscriber feedback and refine the offering, or by enabling subscribers to customize the services themselves, based on their own preferences and requirements.



Using real-time data to build context-aware policy packages *Source: Alepo*

Alepo

Optimizing network performance with a view to service creation and monetization

A conversation with Danielle Elaine Smith, Marketing Director

Monica Paolini: Welcome to our conversation with Danielle Elaine Smith, Marketing Director of Alepo. Today we will talk about how to get the subscriber viewpoint and perspective in optimizing the traffic management in mobile networks.

To get started, can you tell us what you do at Alepo, and what is it Alepo does in this area?

Danielle Elaine Smith: I am the director of marketing at Alepo. I've been with Alepo for the better part of five years, here in Austin, Texas. In my tenure with Alepo, I've worked pretty heavily in the area of policy and charging control. I was fortunate to start right as LTE as a network technology gained a lot of momentum in the marketplace, so I've seen it evolve pretty significantly in the past five years.

At Alepo we make software technology for the core network and service provider IT. We primarily focus on policy and charging control, as well as real-time convergent charging and billing, and BSS/OSS. We also work pretty heavily in carrier Wi-Fi, enabling Wi-Fi offload, and Wi-Fi hotspot monetization.

Our focus as a technology provider is to help mobile operators and Internet service providers to be more successful in next-generation IP-based data networks, on the fixed side with technologies like DSL and cable, as well as the mobile broadband side in LTE. We're working with operators to help them make the transition to an IP network environment gradually and gracefully. We also look to help them sell more data and sell more data services.

Monica: That is a crucial trend right now. We see operators trying to be more and more attuned to the subscriber experience but, at the same time, finding it difficult to find out what subscribers want. Can you tell us how you have seen this field evolving over the last few years?

Danielle: We've seen policy control, specifically, evolve pretty significantly over the past few years. Originally, policy control was used mostly as a traffic management tool to help operators better allocate network resources. Today, we're seeing policy and charging control as the genesis of an operator's data business.

We're seeing use cases that not only define how network resources should be used when there's congestion in the network, for example, but also how to deliver application-based data bundles, or how to deliver roaming data passes specific to certain areas or countries. We are seeing the use cases for policy control expand, and move from just being network-resource oriented to being customer oriented and revenue focused as well.

Monica: At the same time, there seems to be a trend to move away from a policy that is strict and has to be pushed very slowly. Operators used to – and many still do – plan a month or so in advance for a new package to be launched. Now operators are moving to something much closer to real time, where they can easily and quickly introduce new services. Are you seeing that as well?

Danielle: Yeah. The functional role of policy control, like a PCRF, or policy and charging rules function, in the network is to use real-time and highly contextual information from the network, from the subscriber, and from other IT systems in order to make real-time policy decisions.

The PCRF will consider whether there is network congestion right now. Is the subscriber roaming right now? Which devices are the subscribers on and connecting to? Which services are they trying to use? Which applications? All of that real-time information becomes the basis for building policy rules and then building policy-driven data packages and data offers.

Monica: Basically, what's happening is that you have a policy in place that is based on the operator's understanding of how its network operates and what the subscribers want. The operator then uses policy in conjunction with QoE data to automatically adjust to the network conditions. That can be done in real time, with policy working more closely with everything else in the core network. Is that what you are seeing?

Danielle: Absolutely. Not only that, but it's not a one-way street. It's not that the operator makes a business decision, builds a policy rule, a policy package, and sends it out into the universe and hopes for the best on the subscriber end. What

we're seeing is more of a continuous feedback loop wherein the operator makes those business decisions, sets the policy – the policy is determined by the real-time network information – then the operator sends it out to the subscriber.

Then, using real-time analytics and insights from the subscriber and from the network, we're able to feed that back into the operator's decision-making process and then feed that back into the types of packages that the operator is building and sending out. We really see it as a kind of constant feedback loop: put a product out, collect analytics on it, make decisions, and refine and refine.

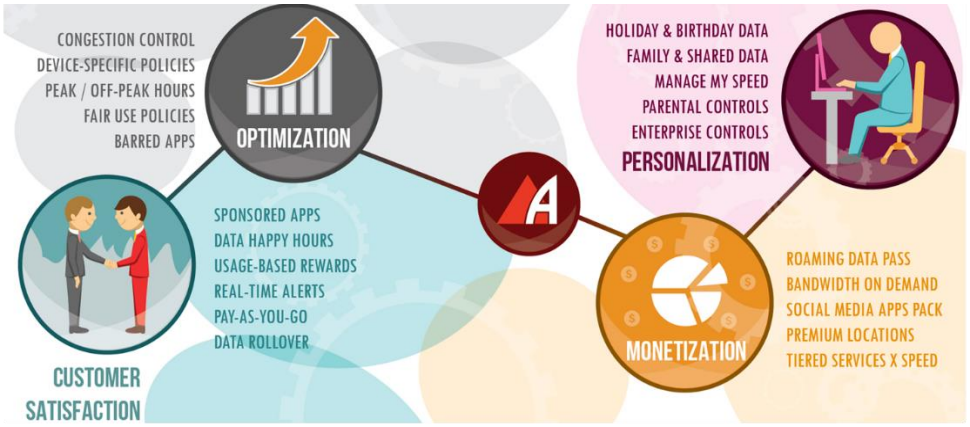
I think that the subscriber experience can be gleaned from the types of analytics that operators are able to collect through a PCRF in order to make better decisions for the subscriber but also for the network. A lot of times, when we talk about network capacity, we're talking about how we handle capacity crunch. But we can also expose opportunities where maybe the network is underutilized or where we can find points of value in the network and expose those to deliver better quality of experience and to better monetize services.

Monica: You're talking about a very important issue here, because, for an operator, it's very difficult to actually understand what QoE is, because it's so subjective.

One good way to learn about it is to try different services and see what happens. It's a huge learning opportunity there. The ability for the operator to change in response to external events, what their competitors do, or what the subscribers want is an opportunity to learn. So it's a continuous learning process.

What that means is that operators need to be able to change their product offering quickly. They cannot just wait for a month, because that period is too long.

What are you doing to compress the time it takes to change policy or to offer a new service?



Real-time policy control *Source: Alepo*

Danielle: We've done with a few things with our policy and charging control solutions to, as you say, compress that time between idea or business decision and product on the market.

The first is that we deliver our product with a pre-built library of policy assets. Each of those policy assets is really mapped to a business use case. In our library, you would find use cases for location-based services, for social-media bundles, for data happy hours, for sponsored apps. Our customers gain assurance that, as new use cases emerge in their markets, the policy assets are ready and available to them.

So that's one way that we are shortening the time frame. On the actual user experience side of the product - which I won't go into too heavily - we've gone to great lengths to make it very simple and to reduce the learning curve to get products out the door faster. You can drag and drop policy rules. They're mapped to your business use case. You can build really business-friendly policy packages, and push them out the door to your subscribers in a matter of minutes.

Monica: That's very useful; most operators worldwide deal with the same choices. In the past, they would have to customize their solution. So it was much more of an ad hoc type of a development of a new service. Now, they can choose. I guess it's more of a toolbox approach. Is that what you think?

Danielle: Absolutely. It's a toolbox approach of business use cases. Another thing that we're doing is we're making this library incredibly dynamic. We're constantly developing new use cases based on operator feedback, based on what we see in the market, and delivering those to our customers.

One important point worth noting, because it ties into this overall conversation about evolution from legacy services into a kind of IP-driven world, is that, for us, it's important that operators are able to deliver, to test, new data offers without impacting or jeopardizing their legacy mobile services.

So the way we're developing our products is to work alongside those legacy services, to plug into

your existing network and plug into your CRM and your legacy charging system, and to be able to test and deliver new types of data offers without impacting those legacy services.

Monica: Another thing that operators are increasingly doing is listening to what their subscribers say. So not just telling them, "This is what you can do. This is the offer," but also getting their input. From a policy point of view, there is also the possibility of giving subscribers some flexibility and some choice.

Danielle: I think it's important to include the subscriber in the mix of conditions in determining policy. We know that – as you said, because QoE is so highly subjective – in some cases, it's really the subscriber himself who is the best judge of what QoE he needs in that moment. So what we're seeing is a lot more peripheral tools that bring the subscriber into the mix.

Let me give you a concrete example. We've developed a companion mobile application for end users so they can get bandwidth on demand, buy a turbocharge package. Let's say if they need to download a heavy file or to watch an HD movie, they can purchase maybe a temporary bandwidth boost.

In other cases, we've seen subscribers who want to make their data package stretch to the end of the month. They self-select to reduce their bandwidth speeds – reduce their QoS maybe for certain applications or maybe for all applications – in order to make their data services last according to their own needs.

So I think what we're understanding is that overall in policy control, in QoE, it's not about always delivering the highest quality of service to subscribers, but rather to deliver the quality of service that is ideal for the subscriber, for the network and for the operator's business goals at the same time.

Monica: It's more than just up selling. It's more than just offering the best service. It is also a very good way to leverage the network resources, for an operator to not always provide the best, but provide what is most valuable to the subscriber, even though it's not top quality. The subscribers may have constraints in terms of how much they're willing to spend that may make them choose a lower performance level for their service.

You spend a lot of time talking to operators. What is it you see that right now they need the most? What are their major pain points that they're trying to address?

Real-time analytics can be used to optimize policy control *Source: Alepo*

Danielle: I think what's been a theme of this conversation, and a lot of my conversations with operators, is that operators are still looking for how to best monetize data services. We know that, in most markets, mobile voice revenues are either plateauing or declining. Operators are really looking to data services as a strategy to fill that revenue gap.

Getting services to market rapidly is always a pain point. So, giving operators the tools to be highly responsive and highly agile is really important to us. It constantly comes up.

From a use-case perspective, from a PCRF use-case perspective, I'm always surprised that fair-use policy is always in the top three use cases. It seems such a basic use case, but it's always an important one for operators.

From there, we're seeing a lot of divergence in the types of data offers that operators are looking to implement. We tend to work in emerging markets and in highly competitive places, in Asia, in Africa, and Latin America, where operators really need to be able to differentiate.

Being able to offer data rollover, emergency top-up, temporary data passes, free Facebook after 10 pm – we see all sorts of different use cases that really vary and run the gamut. But the common denominator is really that the use cases require advanced real-time policy control to bring it together and make it work.

Monica: Well, this is actually quite interesting, because we look at the big operators in developed markets as innovation leaders, but often there is more innovation amongst smaller operators in emerging countries. Why do you think it's so? Why are they more willing to try new services?

Danielle: That is a great question. We work with a lot of greenfield operators that might just have more of that startup mentality and say, "Let's innovate and test things, because we're starting out." Because we work in highly competitive emerging markets, the status quo in a lot of places hasn't been developed yet, so it's really fun for us to get to work in those environments and test new things.

For us, again, it's important to be able to allow or to help our customers to test new services, test them quickly, but also test them with a small segment and collect that data before they open them up to their broader customer base. One thing I'm excited about – just to give you a heads up on a new product feature we're working on – is creating more test simulation within our product itself. Looking ahead, we're looking at how can we help operators to not just guess how the policy is going to work when they get it out the door, but to simulate and to run a test bed with users before they open it up more broadly.

Monica: What is special about Alepo? What sets you apart in this field?

Danielle: I would say a few things. First, as an organization, we tend to be much more lean and agile and responsive to our customers. We don't have long, complex development cycles. We're a product-oriented company, so we put a lot of momentum into getting new features and new products out the door quickly to our customers.

We're also very business oriented when it comes to data. Like I said in the beginning of our conversation, our focus right now is on how we help our customers sell more data services. The way we build and design our product is not just for the CTO but also for the CMO. We're taking into consideration not just building technical products but building products that work for business, as well as for the network itself.

Monica: You're actually hitting on a very important issue, which is that your audience within the operator's organization is also widening. So you want to make sure, by having an easy-to-use interface that different groups can all contribute. It's a much more accessible product.

Have you seen that to be a change in the way operators use policy, in the sense that it's not confined to just a few individuals, but it becomes more of an operator-wide effort in which larger groups of people are involved?

Danielle: Yeah, absolutely. The stakeholders in policy control have evolved alongside the functional role of policy. As we move from it being a broadband traffic manager into being a data-monetization ecosystem, we've definitely seen the users and the stakeholders and the decision makers change and broaden, absolutely.

Monica: In closing, this is an area that's really changing very fast. What do you expect to see over, say, the next five years? What type of new features, products or directions are you working on at Alepo?

Danielle: One thing that I do see starting to happen is that a lot more of these use cases that traditionally have been limited to mobile broadband networks – 3G and LTE – now, we see the same use cases on carrier Wi-Fi networks.

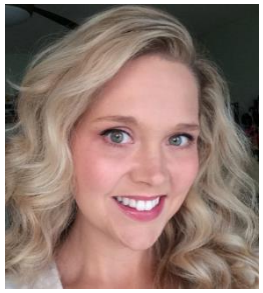
I think we'll see much more advanced policy control in Wi-Fi networks, to enable the same types of services and the same types of offerings that operators are offering over LTE networks.

Obviously, we'll see much more convergence of these different network technologies, and operators being able to provide more consistency and convergence not only in how they package these types of services and offers, but how customers move fluidly from LTE to Wi-Fi and back and forth.

About Alepo

Alepo makes carrier-grade software solutions that enable global communications service providers to accelerate revenue growth, market share, and customer experience on fixed and mobile broadband networks. Alepo's innovation spans advanced policy and charging control, convergent charging and billing, BSS / OSS, Wi-Fi hotspot monetization, Wi-Fi offload, AAA infrastructure, and managed services. Established in 2004, Alepo is a mature company, present in all regions of the world, with offices in Argentina, France, India, and corporate headquarters in Austin, Texas. Alepo has served leading CSPs, including Orange, Saudi Telecom, and Digicel. For more information, please visit www.alepo.com.

About Danielle Elaine Smith



Danielle Elaine Smith is the Director of Marketing at Alepo. She joined the global technology provider in 2010 and has since worked to connect emerging technologies to the business value they deliver to global communications service providers. Danielle earned a B.S. in Life Sciences Communication from the University of Wisconsin-Madison. She can be reached at Danielle.Smith@alepo.com and on LinkedIn at www.linkedin.com/in/hellodanielle.

Anritsu
Profile

Anritsu is a leading test and measurement company that offers solutions for all the mobile network components – from the RAN, to the core, to the mobile devices –for both legacy and new technology. Mobile operators use Anritsu's solutions to test and measure components in the RAN, wireless backhaul, transport, core, IMS and mobile devices.

Testing and measuring equipment includes base station analyzers, cable and antenna analyzers, conformance test systems, handset testers, spectrum analyzers to identify sources of interference, PIM analyzers, power meters and sensors, signal generators, signaling testers, signal and spectrum analyzers, and trace management tools.

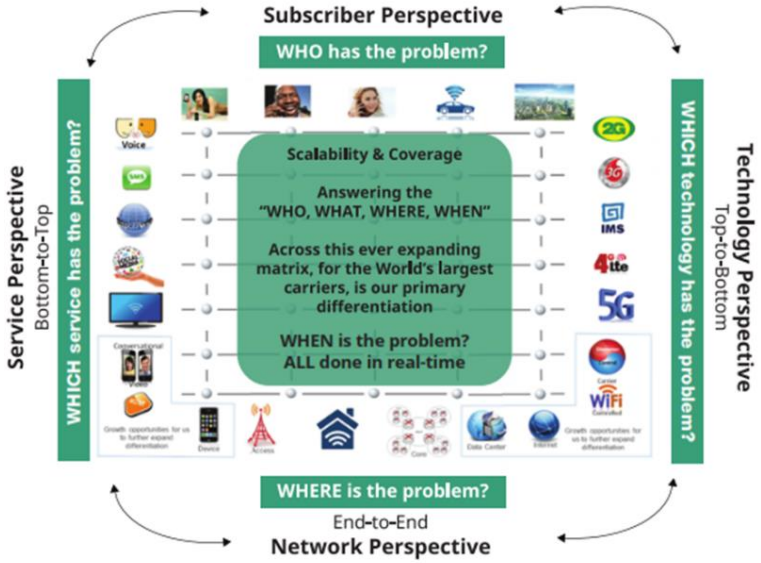
Complementing the test and measurement equipment, Anritsu has recently launched a portfolio of analytics solutions that focuses on capturing information on the QoE of individual subscribers and in specific locations, device types, and applications. The solution helps operators in multiple service assurance areas, including:

- Fault management
- Root-cause analysis
- Performance monitoring and analysis
- Traffic trending
- Congestion control
- A better understanding of QoE in real time can also be a valuable tool for monetization and customer retention.

The eoMind, a real-time streaming analytics platform, enables operators to use the testing and monitoring data they've collected to get a sharper understanding of QoE and overall network conditions in real time. eoMind uses big-data algorithms based on machine learning techniques to detect performance anomalies, identify their root causes, and suggest steps to address the problems. It does so in real time, analyzing data as it is collected, without the need to store it in a database for postprocessing.

A complement to eoMind, eoSight is a visualization tool that covers subscribers, devices, networks and applications, and shows the results of the data analysis from eoMind. It provides real-time analysis of root causes of network performance; it also provides predictive analysis, using a multidimensional data analysis approach that enables operators to get answers to queries that span multiple data sources or levels of data granularity.

Finally, operators can use the eoSight dashboard to view and optimize network performance and resource utilization, improve subscriber satisfaction, and reduce MMTR. It tracks KPI and other network metrics that the operator collects within a unified tool. With eoSight, operators can look at overall network performance or drill down until they reach the individual subscriber. As is true of other analytics tools, eoSight works in real time to identify and address performance issues, but it also alerts operators to new issues as they emerge, before they disrupt service and impact QoE. Operators can set the granularity and extent of the ongoing monitoring of their network to meet their requirements.



Addressing performance issues in real time *Source: Anritsu*

Anritsu

Tracking down QoE at the subscriber and network levels

A conversation with Neil McKinlay, Director of Product Management

Monica Paolini: For our conversations on core optimization and traffic management, today I have with me Neil McKinlay, Director of Product Management of Anritsu.

Neil, thanks for being with us today. Anritsu has been working on network testing, monitoring and assurance for a long time. What do you focus on? And what do you personally work on?

Neil: Anritsu is a very large company. We have, as you say, a long history in the test and measurement area. We also have a long history in network monitoring and service assurance. We have a set of products that can monitor networks end to end, across all technologies and all different networks, whether they are mobile or fixed, and whether it's voice or data.

I lead the product management team. We define the strategy and direction of our multi-dimensional assurance, analytics and big data solutions.

Monica: You've been able to see the way the networks have evolved with the introduction of LTE. Now we're starting with virtualization, different usage patterns, and higher usage volumes. How is that changing the way we test, monitor and operate mobile networks?

Neil: The key thing coming through is personalization. It's a strong buzzword in the industry at the moment, but when we talk about personalization, we're actually talking about understanding the quality of experience of each and every consumer of data on the network.

No matter where the subscribers are, no matter what device they're using, no matter what service they're using, the key thing is being able to understand the experience of each and every device – human controlled or machine, and also, critically, understand why some have a good experience and some perhaps don't have a good experience, even if they're located in the same location using the same device, even the same application.

It's not necessarily the case that those two different devices will get the same experience. The key thing is really understanding and identifying the reasons why an experience is different for different devices when, on the surface, everything should be the same.

Monica: What has changed that has made personalization a key element in mobile operators' strategy?

Neil: It's the technology that's available now around big-data solutions and streaming analytics that enables operators to get to that actionable intelligence, that actionable information, to be able to understand.

In the past, the technology forced people down certain routes of bucketizing individuals' experience. In the past, if you looked at QoE, it was around location or around particular devices, because the technology wasn't really there to deliver the right insights at the right time so that operators could use them to improve consumers' QoE.

Whereas now, the big-data technologies – your Hadoops, your streaming analytics – enable operators to gather this data from many sources, not just from a monitoring system such as Anritsu and other players provide, but from any source or sensor in the network – including network elements and applications.

By merging all that together, you get a high-definition view of QoE. That's really the difference: being able to manage and interpret that data, and also, critically, doing so at a cost that is acceptable. It was possible in the past, but the cost was prohibitive. Technology's moving very quickly to enable that.

Monica: How can an operator quantify QoE, given that it's so subjective?

Neil: There's a lot of knowledge around in the industry. For example, the TM Forum published a set of guidelines to define what customer experience is, what QoE is.

Typically, that's your starting point. If you have a certain application that customers are using – let's say they are streaming video with Netflix – there

are bandwidth and latency limitations that the application can deal with.

Your first step is to look at that in a passive way without getting data from the application itself, and to say, "For these users in these locations, the network delivered that amount of bandwidth, that amount of latency."

If it was all within specific boundaries to enable the quality of that application, we can say, "Yes, everything is in place to make sure that the content is within the limits and can be delivered."

Of course, there will always be exceptions to that. But, in general, if you take that as your starting point, then it makes sense to say, "This was good" or "This was not so good," and start from that.

There are many different measures out there. If you think about voice, for example, mean opinion score, or MOS, has been around for a very, very long time. Almost every solution will be able to deliver the MOS score. For VoLTE, for example, you can have a MOS score for every call, for every user, for example. That's a good starting point.

Monica: This enables you to go beyond averaging, because you can get the MOS for the whole network, but as you drill down, you can get the MOS and other metrics for each call, for each user, for each location. You get much more granularity in your data.

Neil: Exactly right. It's about the granularity. It's that high-definition view. You're not saying, "We've got two million subscribers, and they all had this MOS score." You can say, "for every single subscriber who was running VoLTE, we measured the MOS per second, and we store that." The technology is available to make that happen, at an acceptable cost point.

Monica: You can then still see the average MOS score, but you can also see the distribution, so you understand where it good and where it is bad.

Neil: Absolutely, this is all about that high definition view. You can look for patterns in our data and uncover insights that would not be visible without that granularity. You can say, "All of these people had this particular experience." What is the common denominator? It may be that the application they're using is causing the problems, rather than the network. The operator can rapidly identify the root cause of any quality issue. It's actually getting value from those large amounts of data, but it's also enabling operators to pinpoint exactly why that particular instance of poor experience is happening.

Anritsu's eoLive dashboards *Source: Anritsu*

Monica: We're talking about VoLTE in this case, but generally the performance depends on the application. In an LTE network, it's all IP packets really. But, if you just look at the performance of your IP packets themselves, you're not going to be able to capture QoE, because it depends on whether it's voice, video, text messaging, or downloads. You need to define QoE at the application level.

Neil: Yes. You can't get away from it. In the past, understanding QoE was relatively straightforward, because you sent text messages and spoke. Whereas now, you may use have VoLTE or any of the OTT apps, such as Viber, WhatsApp, Skype, and people are speaking, but it is all packet based. To understand what the experience is, you need to understand the apps the people are using. Can the network deliver the performance required for those applications? Solutions such as ours deliver the granularity, the visibility, and not only from a network viewpoint, but also from a subscriber viewpoint.

That's the big change, if you troubleshoot for a specific subscriber. Let's say you identified the 50,000 subscribers who have this particular app. In the past, it was almost like the needle in the haystack with a magnet to find it.

Big-data technology and customer-experience analytics enable a much more effective way to identify issues, starting from the symptom rather than microscope, for example: "This is a common denominator for all of these subscribers. Let's drill in and do the troubleshooting that way."

This way is a much more holistic, top-down view, and looking at the applications is a critical component of that. You can't separate the apps that people are using from their experience as they are integral to that experience.

Monica: Analytics is clearly a big enabler in terms of crunching all this data and extracting the valuable bits.

Let's step back to how we collect the data. There's been a lot of use of probes, of call traces. Now we're starting to get more into crowd sourcing, asking the subscribers what they think, and looking at the performance in the device. What are the most valuable sources, and how do you combine them?

Neil: I think you touched on most of them, but I think operators will always need an independent view of what's going on in their network. There's huge value in data from probes, whether they are passive probes that sit alongside the network or, as we're seeing now, moving to virtualized environments in NFV/SDN networks. Probes have become part of the fabric. A probing solution that cannot deal with a virtual EPC or a virtual IMS core that's spun up will have little value in future.

Independent data that's specifically focused on the subscriber experience will continue to have a lot of value. Then, of course, there is lots of information from the network itself. It is also becoming possible to add in data from the apps themselves, if they generate information, and from the customers. Customers can give highly valuable feedback.

Also we're seeing a big trend with network operators monitoring social media. For example, operators are very sensitive if they're seeing lots of tweets coming through that a particular location is suffering problems.

Pulling all of this data together is absolutely critical to understanding both the subscriber and the network viewpoints. When you start tying all that together and do it in real time, then we start to move as quickly as the networks need to move.

I think we'll see a change from capturing, storing, enriching and then analyzing data to analyzing data as soon as it becomes available. The process will become much quicker and more automated.

Monica: At the core of all this is you need to have a good engine that is able to pull this information together. That's where analytics comes in.

What is it that you do at Anritsu, and what are the challenges you are facing as you try to develop a solution?

Neil: I think the key things is being able to focus on the strengths of your organization, in the past many companies tried to build everything themselves, even if outside their core competencies, such as visualization tools, databases as just two examples. That meant it was much harder to focus maximum resources on what actually adds value. In Anritsu we recognize this so we integrate best in class components that help us get to market and deliver the greatest insights for our customers.

What I think you touched on in your question is the understanding of, one, what the data is; two, the value of that data and what it contains; and three, how it can be interpreted and presented, so that we enable users of our solutions to make the right decisions to improve quality where it has the highest impact. For example, we have a large team that spends a lot of time with the network operators, understanding how the networks live and breathe.

We take that knowledge and productize it, so that the data is interpreted correctly. I like to use the expression "drinking from the fire hose of information." It doesn't necessarily give you knowledge, it just means that you drank.

It's really about identifying the crucial pieces of information from the deluge and saying, "This is what we can use. These are the key components. This is the key, root cause of why we're seeing this problem in the network, and this is what needs to happen to resolve it." It's distilling that knowledge into a set of products that people can use.

What we're seeing is that many of our customers are saying, "Our teams cannot cope with all of this deluge of data. We have lots of data. We don't know what to do with it." That's where companies like Anritsu come in. We distill out, "What are the key things? What is the key knowledge?" and present them in a usable format.

Monica: This has to be done in real time or close to real time. Data that is one month old is good for historical purposes, but it's no longer enough for optimizing the network or solving problems.

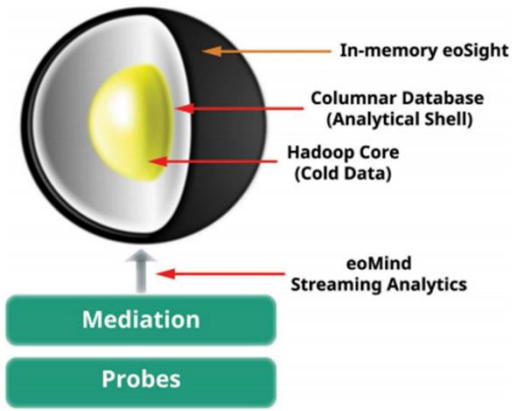
Neil: Our solutions are positioned that way. We have a real-time customer experience dashboard that we call eoLive. That's for people in operations. If you see something happening, you need to take action right away. You can't wait even an hour before taking steps to resolve the issue.

We also have our eoSight, which gives the long-term trends and analysis. That's for understanding, "Is there anything changing, fundamentally, in the network that we need to do?"

Last, we're due to launch fairly soon a brand new product that is as close to real time as possible, with its streaming analytics from very many different sources.

It's really matching up with what we're hearing from lots of our customers, who are saying, "We've got virtualization coming along. We're moving very strongly to NFV and SDN infrastructure to cut our costs, to make the network more dynamic."

As the network changes to be real time, any solution also has to change to be real time, because even minutes is too long a time to actually take the right steps to change things. Of course, as networks become more autonomous, any solution that is feeding information back is part of our feedback and decision loop. If it's not real time, then it's not going to succeed in the future.



Anritsu's eoSight architecture *Source: Anritsu*

Monica: With the combination of real time and multiple data sources, you need automation. What's the role that you're playing on the automation side of things?

Neil: It's the interpretation of the data and automating the thought process that a high-functioning expert would go through to say, "This is the problem. This is how it's occurring. This is the root cause. This is why it's happening. This is who it's impacting."

It doesn't help you to have hundreds or thousands of alarms going off all the time to say, "You've got a problem." That just adds to the burden that people are under. It's really interpreting what's going on and saying, "You have X number of customers who are impacted by a problem, this is the root cause, and this is what you can do to resolve it. These are the next best steps."

Many customers are asking for this, and it's not that their organizations don't have great skill. It's simply there's too much data and not enough hours in the day, and the networks have become too complex for a team of humans to respond quickly enough to resolve the issues. That's really what the automation is all about that we're working on.

Monica: Virtualization is also another element that might not necessarily add complexity, but it definitely adds dynamicity in the networks.

When you're measuring something, you no longer measure a box, because the same function can be instantiated in different ways, and that might have an effect on performance. In terms of monitoring and also testing, this makes things more complex. You're testing a moving target.

Neil: Absolutely, and it's something that we've been working on for a while. Without going into much detail, any probing solution has to be part of the orchestration layer. You can imagine if the network determines, "I need more capacity over here. I need to have an additional serving gateway," for example, "and a virtualized EPC."

If you want to monitor that, if you want to understand what the experience of customers is in that area, you have connections through that particular device. If you're not part of the orchestration, if you're not part of the fabric of the network, then you have to have an individual person saying, "I need to type some code in. I need to click on some buttons."

It's too late. That part of the capacity is already spooled up. It's already transferring data, and you don't know what's going on. The virtualization makes it a real challenge to increase and decrease resources, and also deliver the right information to any system or solution that's monitoring the customer's QoE.

Many of the systems and the technology in the past were force-fitted to a very static environment. That's paradigm can't last for too long because of how quickly the technology is changing.

Monica: One application that is still special – and complex – is voice. Is there anything special that you're doing with VoLTE?

Neil: We already have solutions for VoLTE that use our dashboards and analytics solution. We have a valuable capability that helps network operators understand the QoE of each and every VoLTE subscriber, in real time.

Not only does it give an aggregated view, it actually gives a high-definition view of every call and the MOS score on a per-second basis, on a per-cell basis, and a per-device or per-application basis. That's something that we've had out for about 12 months now. We're seeing a strong acceptance in the market.

Many of the customers we talk to say they're not going to launch VoLTE if they can't adequately monitor it, if they can't ensure the quality. Obviously, VoLTE is a way to compete against the free, over-the-top apps – WhatsApp, Viber – which have no guaranteed bandwidth.

VoLTE offers subscribers a higher quality experience via a guaranteed bandwidth, a guaranteed quality, and allow operators to differentiate on quality.

If you're using Skype at home and the quality gets too bad, you resort back to your old fixed line, because you know that the quality's going to be there. With VoLTE, it's the same thing. For business, high-quality speech is critical. Obviously, if you're differentiating on quality and you can't guarantee that quality, that's a difficult thing to sell.

Monica: What should we expect to see in the next few years? We're learning more and more about our networks. What is changing, and what are you working on at Anritsu?

Neil: I think what you'll see, certainly from our perspective and based on what we're hearing from quite a few customers, is the completion of taking data from any source and any device.

Whether that's structured data that you typically get from probes, semi-structured data that you get from another vendor's probes or from test equipment, or unstructured data from social media feeds, what is changing is the availability of data and the ability to take in data from many sources. Over the next few years, we will see a growth in putting all of that together and making sense of it.

Also, network virtualization will be complete in the next few years. A lot of the vendors – your Intels, your HPs, your Ciscos – are all working very hard to make that a reality, and also to deliver the performance that this requires.

We'll continue to see the amount of bandwidth-hungry applications grow and grow. We'll see a massive explosion in the connected devices. The ability to manage massive amounts of data from massive numbers of connected devices, and still deliver that visibility, still deliver that insight, is going to be the key thing that we'll see going forward.

We'll see much more automation, much more machine learning. The networks, the services are not going to get less complex. In fact, they are going to get much more complex, and it will be an even bigger change.

About Anritsu

Anritsu Company is the United States subsidiary of Anritsu Corporation, a global provider of innovative communications test and measurement solutions for 120 years. Anritsu's "2020 VISION" philosophy engages customers as true partners to develop wireless, optical, microwave/RF, and digital solutions for R&D, manufacturing, installation, and maintenance applications, as well as multidimensional service assurance solutions for network monitoring and optimization. Anritsu also provides precision microwave/RF components, optical devices, and high-speed electrical devices for communication products and systems. The company develops advanced solutions for 5G, M2M, IoT, as well as other emerging and legacy wireline and wireless communication markets. With offices throughout the world, Anritsu has approximately 4,000 employees in over 90 countries.

About Neil McKinlay

Neil McKinlay is the Director of Product Management within Anritsu's Network Infrastructure Business Unit and is responsible for driving Anritsu's strategy for Multi-Dimensional Assurance, Big Data and Business Intelligence solutions. He has spent his whole career in telecoms all the way from 2G through 4G and 5G starting out in Hewlett-Packard with ground-breaking solutions such as AcceSS7 which helped create the passive monitoring market. Neil is passionate about understanding markets and defining and developing solutions that truly make a difference to Anritsu's customers'. Neil lives in Scotland with his family and his interests include motorcycling and, predictably, is a member of The Scotch Malt Whisky society, describing himself as "a keen amateur in understanding whisky."

Ascom Network Testing

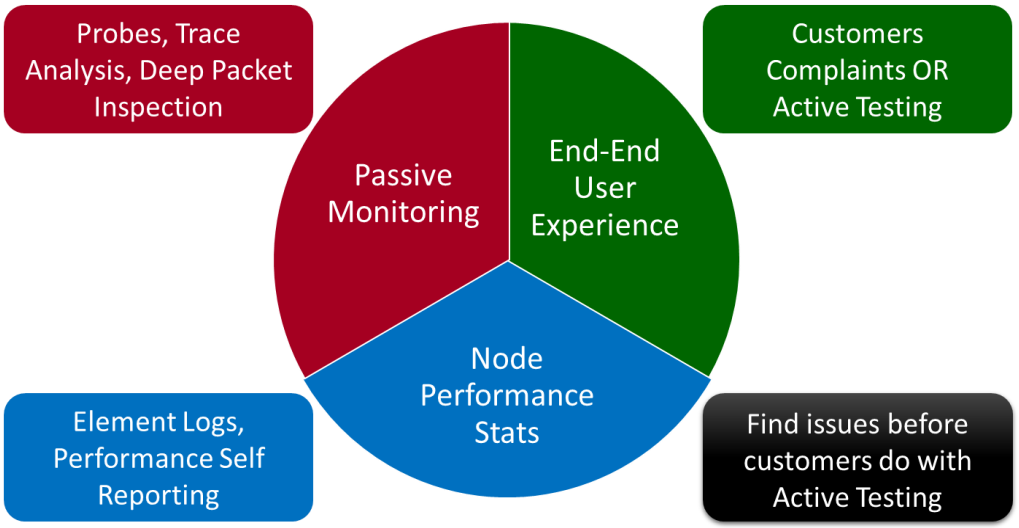
Profile

Ascom Network Testing is a division of the international technology company Ascom. It provides a testing, monitoring and optimization platform for wireless communications and helps mobile operators worldwide improve the efficiency and performance of their networks.

The TEMS Portfolio covers the mobile network infrastructure end to end, from the RAN edge to the core. It includes multiple solutions for drive testing, benchmarking, monitoring and analyzing network performance. Operators use TEMS Solutions when deploying networks or services, and to maintain and optimize them.

The modules of the TEMS Portfolio are:

- TEMS Automatic, a tool for automated data collection for testing, monitoring, benchmarking, audit and verification, via probes in mobile devices that can be remotely controlled from a central location.
- TEMS Capacity Manager, for planning and managing network capacity, which gives mobile operators visibility into network resource utilization and end-user experience. It identifies bottlenecks, congestion areas and other performance issues that affect subscriber experience, and analyzes how they relate to different traffic types (voice, video and other data applications).
- TEMS Discovery, for analytics and reporting, which combines and analyzes data from drive tests, mobile devices, core network elements, crowdsourcing data, and application data. It supports modules for VoLTE, subscriber experience, indoor performance, HetNet traffic management, SON, handset analysis, subscriber KPIs and benchmarking.
- TEMS Investigation, for drive testing of the air interface and service quality.
- TEMS MobileInsight, to capture QoE from mobile devices by collecting KPIs and through direct subscriber feedback. The data is collected by software agents installed in the devices. This module can be used for crowdsourcing, SLA monitoring, and location-aware network monitoring and optimization.
- TEMS Monitor Master, to test and monitor network performance using simulated traffic at the application layer for different devices. Mobile operators can use it for voice services, OTT services, revenue assurance and billing, and roaming. For voice services (IMS and VoLTE), TEMS Monitor Master measures and analyzes performance, availability, capacity and QoE. For OTT services, it can collect performance, availability and other metrics at the application and traffic stream levels, and simulate application use and interactions.
- TEMS Pocket, enabling indoor and pedestrian-area data collection from handheld mobile devices.
- TEMS Symphony, for benchmarking.



Testing and monitoring mobile networks *Source: Ascom Network Testing*

Ascom Network Testing

Testing and monitoring wireless networks to capture and optimize QoE

A conversation with Tim Brooks, Account Director

Monica Paolini: Our conversation today on core optimization is with Tim Brooks, account director at Ascom Network Testing.

Tim, thanks for being with us today. First, I would like to ask you, what do you work on at Ascom?

Tim: I'm a sales director at Ascom responsible for solution sales – helping wireless carriers do active testing on their networks to find problems that other testing doesn't find, basically.

As the networks move to being all IP, it becomes a lot more complex. It's quite difficult for carriers, sometimes, to figure out that something has gone wrong, because it can be quite subtle. When things have gone wrong, network operators need to do a lot of work to find out where it's gone wrong along the chain.

They have a lot of tools for digging deep into the IP packets, but that doesn't always catch everything. Our job is to help find things that their existing systems don't catch and to draw their attention to that.

Monica: As you mentioned, the networks have been changing with the move to IP. Does it mean that we have to look at how wireless works and network performance in a different way now?

Tim: I think we do. Wireless networks are a lot more complex. There are a lot of boxes in a row all interacting with each other, working end to end to create a service. Those boxes are often coming from different vendors using standard interfaces to communicate to each other. But at the end of the day, it's really each vendor's interpretation of the standards.

Things don't always work when they're plugged in. Each vendor, as a matter of course, is typically doing a lot of firmware and software upgrades. Each time that happens, things change. Things that were working yesterday don't work today. A problem was fixed in one area, a problem was created in another, and that becomes difficult for operators.

Monica: Yes. There is this complexity in the network. We're trying to get the network to do more. Also, isn't it true that with the move to IP, the way you actually measure performance is also changing? You have more complexity, and the metrics are different.

Tim: The metrics are different, and the metrics are more complex. The KPIs are more difficult to measure. There are minute things happening in a SIP message stream or an IMS registration or something like that – all things that are measured in milliseconds.

Traditionally, when you monitored services, it was reasonably easy to tell when they broke. I call it "red light, green light." Things worked, or they didn't work. With IP, there are a lot of opportunities for things just not to work as well as they should.

The IP network reroutes around things that don't work, so the service continues to get delivered, but it isn't delivered with the quality that it should have. Things like latency can creep in. And when errors occur, latency can increase and quality can go down, but the service doesn't break. It's a little bit more subtle, but more difficult to detect, when things are wrong, and that's where we come in.

Monica: Voice is a very good example, where in a circuit-switched subcore environment, you have a dropped call, and that is a failure. As you said, with VoLTE, you might not have a failure, but the latency is high, perhaps unacceptably high. How do you deal with this situation?

Tim: I think it is a couple of things. One is we're measuring KPIs, and we're measuring them across time. If a KPI changes, like latency gets worse, we can detect that change. There are recognized thresholds, so if the change takes you across a threshold, you go from having OK quality to questionable quality to not-OK quality, and we can track those things.

The other thing we can do is track individual elements of the core performance. Does call setup succeed? In particular with VoLTE, each audio direction is a separate data stream, so you can have one-way audio.

We may see a case in which the audio sets up fine, works fine for a minute or two, and then drops off in one direction. Things like that are new types of problems that you never traditionally faced with a voice call. Either it worked or it didn't, or it dropped.

Monica: This gives the operator a way to differentiate from other operators, because not all of them have the same thresholds for good performance.

Tim: Operators all choose their own thresholds, which is interesting for us. We provide recommendations, obviously, but we do notice that different carriers are approaching the same challenges in different ways and have different ideas about what they need to do to make them successful, and we can support that.

Monica: Do operators have different thresholds depending on where the voice call takes place? If it's a place where coverage is an issue, subscribers may be willing to accept somewhat of a lower quality, because it's better than having no call at all.

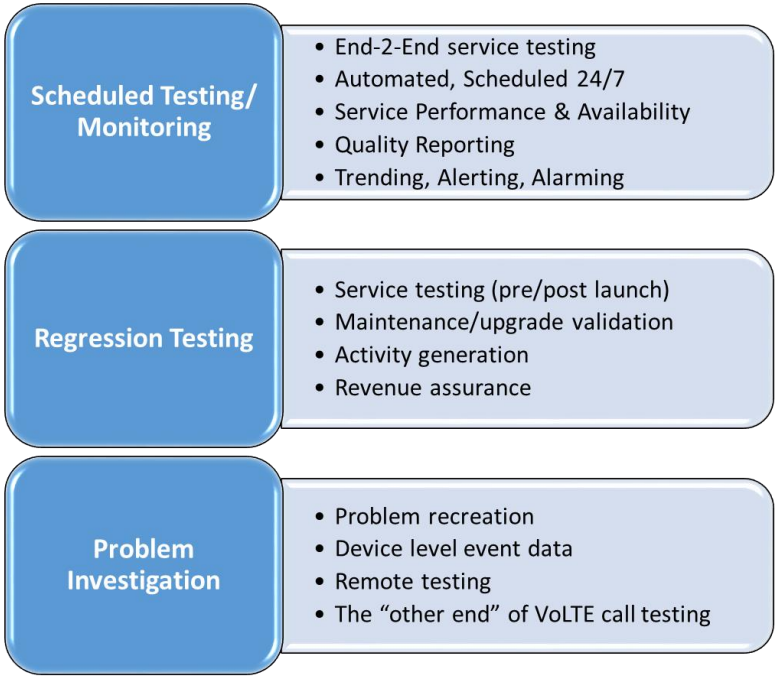
Tim: We're not seeing that so much. Usually operators set a standard and adopt it everywhere. As video becomes more and more pervasive, there might be opportunities and challenges that operators face as they try and optimize their capex, deciding on whether they do want to provide the same video quality everywhere, and that's an opportunity for them.

Monica: Operators are taking a long time to deploy VoLTE, and it seems like it's more complicated. So testing becomes crucial in ensuring that everything works as planned and continues to work as you monitor the system. What are the specific challenges that you see in testing IMS for VoLTE?

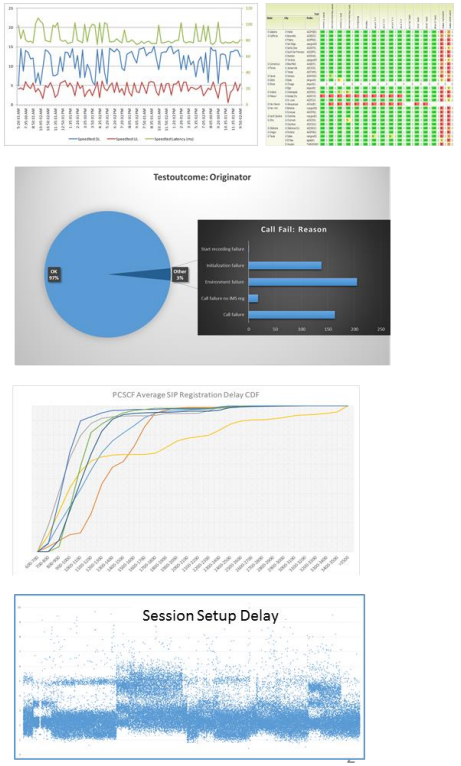
Tim: As operators set up their IMS networks, as you pointed out, they're taking a long time to come to market with VoLTE, primarily because they can. They can choose to, because they've got legacy voice that's doing the job for them now. They only get one chance at making a first impression with VoLTE, so they want to get it right before they go live.

The other side of it is that VoLTE is very complex, and so carriers need to spend quite a bit of time getting all of their different building blocks in place testing them, testing that they work together, working with the different vendors, and getting software loads.

Each carrier configures VoLTE and operates VoLTE in a slightly different way. They're not all deploying a vanilla standard, so there are lots of variances and things to nut through before they decide to go to launch.



Active testing



Source: Ascom Network Testing

We can help carriers in three areas. One is getting ready for launch. The second is, once they're operating, help them with regression testing every time they make a change to the network. Also, we can help them with continuous monitoring so they can see that the service that was working yesterday is working today, is working tomorrow.

Carriers can identify if they get any changes in their KPIs, and they can identify when any of those changes occur, like degradation. If you've had degradation at 3:00 am yesterday, then the first thing you can do is go and look at all the logs for what changed on the network at 3:00 am. That gives you a reasonably good indication as to what might have been the cause.

Monica: You have developed an active testing and monitoring solution. What is different about it? What kind of pain points does it address?

Tim: First of all, it's different because it is active testing. There are a lot of testing solutions out there that are passive. Operators are monitoring the traffic that is going past, and they're collecting huge amounts of information. Theoretically, they can tell you everything that's going on all the time, but they really have too much information to be able to sort through to do that. We also find that that doesn't quite catch everything.

We do active testing. We're emulating or doing what the customer would normally do, and we're reporting how it turns out. We can do that using actual smartphones that the customers use, testing end to end over the RAN and over the core.

We can also emulate the phones using a proprietary IMS client. What that enables us to do is to quickly cycle through and test lots of different types of phones, so that we can identify if there are any types of phones – since they all behave slightly differently from each other – that are not working well with a network configuration and might cause issues.

Monica: There is a growing emphasis on QoE. Operators want to capture performance at the subscriber edge, rather than trying to measure the network elements – which obviously is still crucial, but in and of itself might not give you the whole story.

Tim: Agreed, agreed. A call consists of the words from your mouth that go through the phone, through the RAN, through the core, and back out the other side. It's all an end-to-end call through multiple components, one of which is the handset.

The network operator has very little control over the handset and over what individuals do – whether they dropped it or put an unusual combination of apps on there, or whether there is something that really changes the state of it. It's good to be able to test the network independently of the devices using it, so that we can help identify whether it is a network problem, a handset problem, a problem with a particular type of handset or version of firmware, or those sorts of things.

Monica: Is it possible that you might have bad voice quality, but all the core network's elements might be working just fine? If you were to look just from the network point of view, you'd have the false impression that everything is fine, except that the call is not good.

Tim: That is possible. If that is happening, you want to try and identify if it is the phone, so you don't spend too much time trying to triage your network on something that you really can't fix.

Monica: From your point of view, testing using emulated data gives operators the ability to go in depth without having to deal with huge amounts of data. This is a real issue when talking to operators, because they feel overwhelmed with the amount of data they have to cope with.

Can they just look at simulated data? How can you figure out the right amount of detail in the performance data?

Tim: I don't know if there's one main way to go. I think there are lots of complementary things carriers need to do in parallel. So we see the active testing as complementary to the passive monitoring that they do.

Different types of monitoring and testing can detect different types of issues, so the real challenge for them is to detect if something has gone wrong. Then if something is wrong, to quantify it, to repeat the problem, then to get enough information so they can investigate it.

The way I describe it is that carriers need to assign the trouble ticket. Is it a RAN issue? Is it a core issue? Is it a handset issue? Is it an IMS, a node issue? These are the sorts of things they deal with.

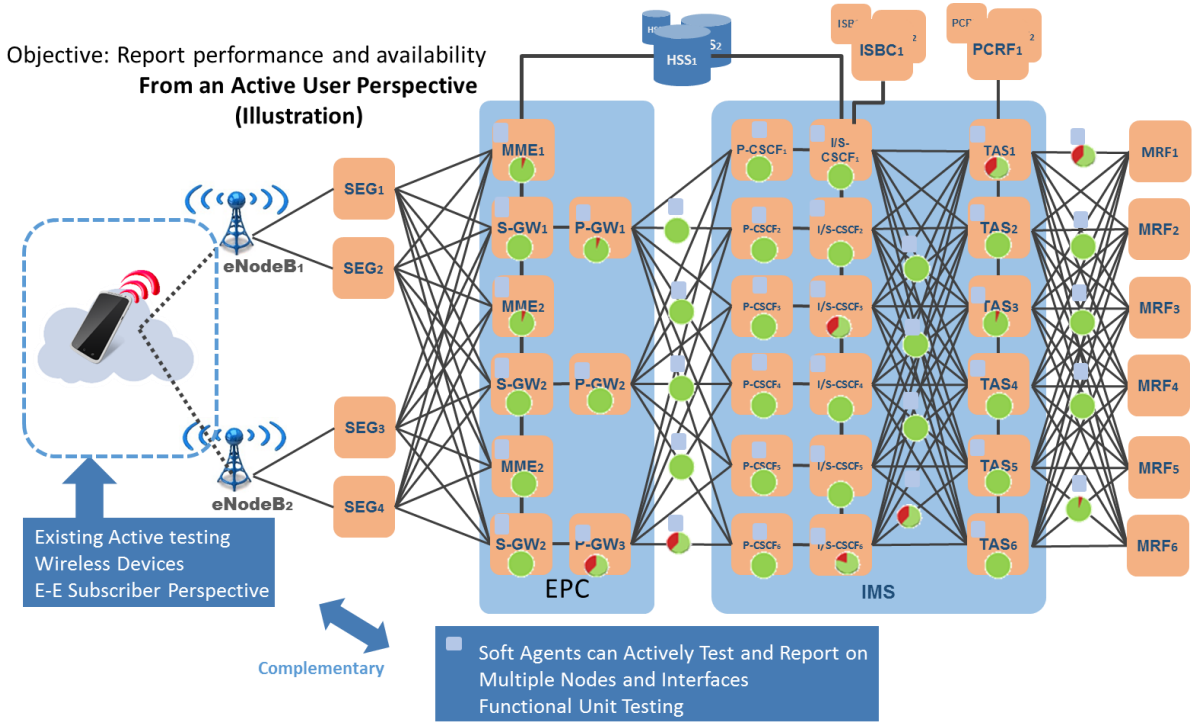
We contribute with other types of monitoring and testing to get the carriers that information. To do that is a challenge for them. It's all new. The way the network is configured now crosses a lot of the traditional RAN switching transmission organizational boundaries.

Carriers almost need to start thinking about whether they want to rearrange their organizations to manage some of this, and I know some of them are.

Monica: And it seems that testing becomes, in many ways, more important, in the sense that it's a continuous type of activity because the networks keep changing. You need to keep testing them, and you need to keep monitoring them. So having more dynamic networks means that, for operators, testing becomes a more integral part of running their business.

Tim: It does. And actually, we've seen quite an uptick in active testing because of it. When you sit in one place, and you make call after call, you could be using completely different routes through the network. Each of the nodes that you hit does some work for you, like the MME or the IMS core. There are often different instances of those that you might hit, depending on load and depending on how your individual call is routed. And within each instance, there are usually multiple servers operating in parallel or doing the same job, and you might hit a different instance of those.

We're finding that by making lots of calls and by looking at some of the IP addresses and other elements, you can see in the trace which network node, which network element within the node you were using; you can start to detect whether you've got a bad instance of a server or a node.



Active testing in a complex network *Source: Ascom Network Testing*

I call it peer review. You're looking at multiple things that should be behaving the same, and finding one that doesn't. It may not be broken, but it's not performing properly. That helps the carrier to see that something needs some attention.

Monica: You need to understand the interactions between all the different elements and all the different metrics you are collecting.

In terms of use cases, what do you see when working with operators?

Tim: The first one that we focus on is regression testing. It's whenever operators make a change, they can automate what is becoming a very long list of things that they want to test afterward.

There are primary things like making a call, sending a message, watching a video. But there are also secondary things like turning on call forwarding, things that people need to be able to do from their phone but don't do very often. You still need to test that those work; so you have quite a long list.

You do your test of the network. You push the button. It runs through the list, and you get a green light to go live with that change.

That replaces a lot of manual activity that people are undertaking today. By automating it and

systemizing it, it gets tests done each time, it gets it done faster, and it gets it done cheaper. Those are all things that are interesting to carriers.

The second use case is continuous monitoring, usually a subset of the regression tests. You don't have to test everything all the time. You test the important things. You can test them every 15 minutes or every 30 minutes.

That's helpful for spotting if something breaks. It's also helpful for spotting degradation, so you can monitor performance, monitor KPIs, see trends, see sudden shifts, and see when those sudden shifts occur.

Lastly, what we're seeing every now and then is that something happens to an operator. I'll give you an example. One of our customers has special services for some private networks and major customers. They were identifying that some of those customers were having poor performance, and they saw that none of their existing monitoring was catching that.

They talked with us, and we did a short-term, focused piece of monitoring on that specific service for those specific customers to help them drill down and investigate what was going on. Also, to give them visibility into things, that their other solutions couldn't provide them.

So the third use case is a problem investigation triage, focusing on something specific that's giving a problem, and to help the people fix that problem.

Monica: The constant theme here is the need to get the high-level picture, see that there is something not right, and then be able to go down and find what's causing the problem. And this is where having the average latency or other average KPIs for the whole network doesn't really help you, because the averaging may mask the problem.

Do you have a different type of approach to look at the issues, even to visualize what is going on?

Tim: Definitely. And there are lots of tools available that provide lots of views into that. The active testing we do end to end provides the view from the customer. That's an ultimate view. Also, the traces that we provide from the device – there is information in there that doesn't appear in the traces from the network, so we really are capturing everything from the user's perspective.

Monica: To understand the interaction across different elements, you may want to look at the distribution of performance metrics, rather than their average.

Tim: Yeah. We can capture performance many, many times rather than doing an average, so you end up with more of a scattergram approach to what the outcomes were, what the experience was.

Also, with our configurable IMS client, we can recreate interactions with the network that might not be intended. If you have a phone that is actually doing something wrong, and that's suspected as being the cause of a problem in the network, then we can configure our IMS client to simulate that bad behavior and help the network operator understand how that's impacting the network, and how it's flowing through.

Operators are not just interested in catching problems. They're interested in replicating problems and understanding how to address them.

Monica: By replicating the problem, an operator knows whether it has found a solution that is working or not. It's a way to validate the changes made.

Tim: Really, it is quite complicated. There's not a buzz sentence to describe what we do, because we do lots of different things at different times for different people.

Monica: One thing that is changing in the networks is the move towards virtualization. How is that going to change the way you test the network?

Tim: It's not really going to change the way we test the network, because you still want to test it end to end, but we are ourselves virtualizing some of our testing equipment. Instead of putting as much hardware out in the field to do the testing, we're putting less hardware in. We're virtualizing some of the components that do command and control on the hardware, and putting that in the cloud.

There are two main advantages. The first one is that it reduces costs and deployment time.

The other advantage is that it dovetails with the strategies that the carriers have. If you're dealing with a carrier who has the objective of virtualizing as much of their toolset as possible in the next five years, we can contribute to that.

Monica: How do you see testing and monitoring to optimize the mobile network and traffic management evolving over the next five years?

Tim: I see it becoming more complex, so things that we're doing a little bit of now, we'll be doing a lot more. Especially as more and more services are layered onto the same data pipe and they are starting to share the IMS core, and those carriers are starting to have their IMS core talking directly with other IMS cores – so the VoLTE-to-VoLTE calls between carriers, and roaming.

Then the complexity goes up. The traffic all looks the same, because it's all data, and so you really need to have lots of ways of testing and looking at things to understand: Is it still working? Is it working as well as it was yesterday? Has any customer issue started to creep in?

About Ascom Testing Solutions

Ascom Network Testing offers expertise and solutions that enable wireless operators to expand network capacity, improve operational efficiency and deliver a premier customer experience. Ascom develops its own line of mobile network testing, monitoring and post-processing solutions that have been trusted by mobile operators for decades under the TEMS™ brand. Today, those solutions enable field testing (drive, in-building, autonomous) of mobile networks, automated post processing of data collected via these – and other third-party – probes, OSS-based network troubleshooting and optimization, and application testing and monitoring.

About Tim Brooks

Tim has more than 25 years' experience working in wireless telecommunications – including cellular, wireless data and satellite systems and applications. His career has focused on ensuring that customers with critical communications requirements receive successful, efficient and optimized services, technologies and work flows.

Tim joined Ascom in 2010. He is responsible for consultative sales for Ascom's TEMS Monitor Master Active Testing and TEMS Capacity Manager solutions in the Americas, which has included nationwide Monitor Master deployments at all tier one US carriers for testing LTE, VoLTE and other value-added services, among them Voice over Wi-Fi.

Flash Networks

Profile

Flash Networks was founded 19 years ago to optimize data traffic in the mobile core network, enabling operators to use and monetize their network assets efficiently. Flash Networks helps operators accelerate data transmission, support a higher number of devices, and generate new revenues.

Flash Networks solutions are based on the company's Harmony optimization platform for 3G and 4G networks. The Harmony Mobile Internet Services Gateway enhances QoE by managing traffic flows from multiple applications.

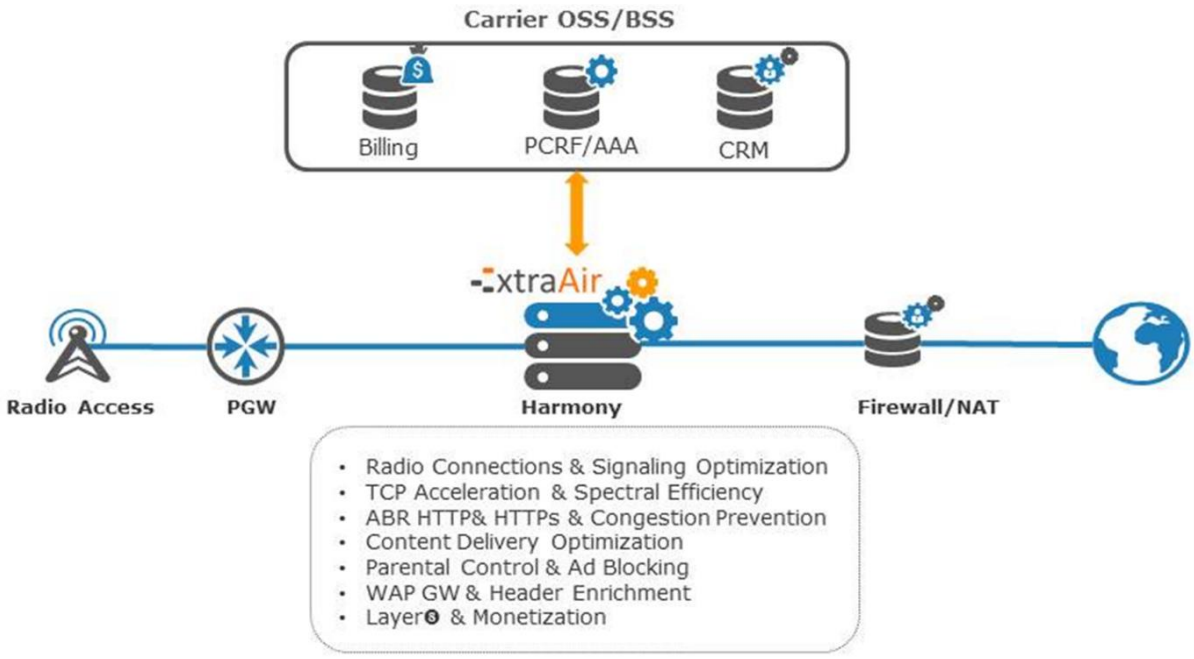
On the monetization side, it supports Layer8 and MoreFor.me, two solutions that operators can use to develop new services for subscribers, and make new content available to them.

On the optimization side, Flash Networks has recently introduced xtraAir as the latest, fully virtualized solution that runs on the Harmony platform. xtraAir is designed to improve spectral efficiency in networks that have high signaling traffic levels and need to support a larger number of connections. With xtraAir, operators can increase the number of connections by 35%, according to Flash Networks.

There are four key components in xtraAir:

- Signaling and radio connection optimization. xtraAir manages signaling traffic directly from the core – without needing to install a client in the device – by sending control messages in batches. As a result, devices change from idle to active state less often. This frees up radio resources – an important benefit, because signaling traffic in mobile networks is growing faster than user-plane traffic, as the number of devices and the number of active applications per device increase. IoT devices and services will lead to a further increase in signaling traffic.
- Acceleration. xtraAir uses its proprietary TCP-4TE algorithm to optimize TCP transmission on the basis of real-time RAN conditions. xtraAir works on all types of traffic, including encrypted data and messages, and with all mobile air interfaces.
- Optimization of streaming video. When congestion is detected or expected, xtraAir uses adaptive bit-rate management of video traffic to reduce video volume by up 40%, according to Flash Networks. Encrypted video traffic is optimized, too. xtraAir relies on real-time RAN load conditions to determine the bitrate that maximizes video quality, while minimizing stall incidence.
- Web and download optimization. Flash Networks estimates that its solution reduces traffic in the RAN and backhaul by 50%.

xtraAir can also take policy information into account to determine how to manage traffic in a congested environment.



Flash Networks xtraAir running on the Harmony platform *Source: Flash Networks*

Flash Networks

Increasing spectral efficiency in the RAN by optimizing transmission in the core

A conversation with Ofer Gottfried, CTO

Monica Paolini: Welcome to our conversation with Flash Networks. I'm talking here today with Ofer Gottfried, the CTO at Flash Networks.

Ofer, what is it that Flash Networks does in this area?

Ofer: Flash Networks helps operators to do more with less – enabling the operator to utilize their assets and radio infrastructure to serve more users with a better quality.

Today we are focusing more on radio data spectral efficiency. From the core, we are able to streamline the data for better utilization of the radio. By doing that, we enable the radio to serve more users, and more radio connections, and more throughput to the operator's audience.

Monica: This is an area where there have been a lot of changes over the last few years. What are the main changes that you see when you talk to operators? And what are the major pain points that you are trying to address?

Ofer: Today we see more and more radio load – because of both higher throughput and higher video resolution – and more and more HTTPS video from one end. From the other end, we see more connections and more radio connectivity from the internet of things – or internet of everything – and connected cars, as well as more smartphones and other devices.

These two factors are loading the radio not just from the throughput or load or user consumption perspectives, but also from a radio connectivity and radio connection perspective.

Those are the things that we are today able to optimize. We can give better service for the users on the radio, even while serving up to 35% more users on the same radio infrastructure.

Monica: You mentioned the internet of things. Is IoT just generating more traffic, or is there more to it? Because if there is more traffic, you just need to add capacity.

Ofer: Yes, it's not that simple, and it is costly. Throughput is not the only KPI that matters, or the main bottleneck of the radio. Sometimes there is enough bandwidth, but as there are more connections on the radio -- more users or more devices – that are going on and off from the network, they create radio signaling and core signaling, in addition to data traffic.

Signaling is sometimes the input to the radio that is a bottleneck, preventing the network from serving more throughput or serving more users on the same infrastructure.

In order to get more from your radio network, you need a solution that handles not only the throughput, with traditional optimization or new optimization of HTTPS video, but also the connectivity and radio connection optimization.

That also includes handling the signaling load associated with every device or element that is requesting a radio connection or radio service, and then asking to be disconnected from the radio service. We are able to optimize the radio from both those angles in order to get higher spectral efficiency, or spectral utilization.

Monica: You are working on to improve the resource utilization in the RAN, but you are not necessarily just working in the RAN, you're working end to end. You are optimizing the data flow before it gets to the RAN.

Ofer: Yes, you are absolutely correct. Looking at the data and understanding the data. We've been doing it for the last 19 years. We have a lot of experience in managing mobile data.

We are able to find those scenarios that are not optimized or not designed for the mobile radio. For example, if a server wants to send a control packet to the radio or to the device, it just sends it. It doesn't care whether the device now is in idle state, or about the connect state in the radio.

With our experience we are able to detect and to know what is the state of the device, and then to optimize, or to better utilize, the state of the device with the traffic from the internet or from the servers. We are able to reduce the number of state changes of the user devices – from idle to connect, and from connect to idle – and by doing that to reduce the load on the radio from the core.

Monica: Does this also work in a HetNet environment, where you have multiple elements that cover the same area?

Ofer: It's applied, really, in the same manner. Whatever radio access serves the user, we are able to see and to monitor all the traffic from this user.

By seeing, monitoring and streamlining all this user traffic from the TCP layer to higher layers, we are able to prevent unnecessary state changes or unnecessary load on any type of radio, regardless of the type of radio. As a result, it takes less radio to serve more users.

Monica: How do you improve the spectral efficiency in the RAN?

Ofer: Basically we have four main features that we are applying today.

The first one is the radio connection and signaling optimization. We have the ability to detect TCP control messages that are being sent from the server to the device or element when it is in the idle state; this is a control message that is not necessary for the application to behave properly at that time. We can queue this message and send it later on, when the device is in a connect state. By doing that, we eliminate the connect and disconnect state changes, and the signaling associated with them. This way we can save a lot of radio connections.

This feature also synchronizes keep-alive or repetitive data. If there are a few servers connected to the same device, that are running applications that have some keep-alive data that may come every few minutes, we can synchronize them to a single phase so they are sent together once a minute. That again saves state changes in the radio, and so it saves radio connections.

The second feature is about TCP acceleration. TCP was not designed for mobile and does not work well in the dynamic mobile environment.

For the last 19 years we have been improving our TCP acceleration, a feature that enables the cell site to transmit higher throughput. With our TCP acceleration, we can gain 10% in cell throughput while improving, also, the quality for the end user.

The third element in our design is about congestion prevention from a throughput perspective. We are able to detect a congestion situation when the user is getting less than x throughput, or x megabits per second – x is configurable, of course.

When we detect congestion, we can control the load. For example, we can limit the maximum resolution that a YouTube video will now load onto the network. We enable the operator to control its load, preventing the content vendors from overloading its network. The operator can control this overload condition autonomously.

The fourth element is about traditional web and download optimization where we can reduce the throughput load from transcoding, image compression and caching.

Both from a throughput perspective and from a radio connection perspective, we are able to reduce the load and prevent congestion from reaching the radio.

Monica: As you move closer to real time, you can manage the network and traffic load depending on the RAN conditions and QoE metrics. A lot of the time you want to prevent congestion from happening, so you need to start acting before the network is in a congested state.

How can you predict the onset of congestion before it starts? Because there is a lot of spikiness in the data, the fact that there is a high utilization at this very moment doesn't necessarily mean it is going to be protracted and result in congestion. How do you decide where you're about to get into trouble with congestion?

Ofer: We have a few layers of detection. One of them is user congestion. We are able to identify the situation when the user is getting less than, for example, 2 Mbps from the network – from the radio.

This can be congestion level one, so a low congestion level. With this congestion level you can start optimizing or start controlling the load at a low level – for example, by reducing YouTube to 720p, and not allowing YouTube 4K.

Once the user can get only 1 Mbps, you can squeeze tighter: further control the load from YouTube, for example, to limit it to 360p or 480p. Still good quality, but you allow your network to serve more users with more bandwidth, while controlling the load from others. You do this only when you need it, only when you detect this congestion. This is the congestion on the user level.

What we are able to do, also, is to get the cell ID and identify the congestion level for the users who are connected to the same cell. By doing that we can prevent further congestion when we detect that 10% of the users are getting congestion: "Let's

start to do some proactive controlling of the load." With this algorithm we're able to prevent congestion in the cell.

Monica: In this case you use geolocation to help in detecting congestion, right?

Ofer: Yes, and this is something that we can get from various resources, either from radios, or through the Gx interface, or we have proprietary implementations that we are doing with some partners of ours, radio partners that can send this data on the TCP header. It's in real time, without any need for loading the other elements or other areas in the core.

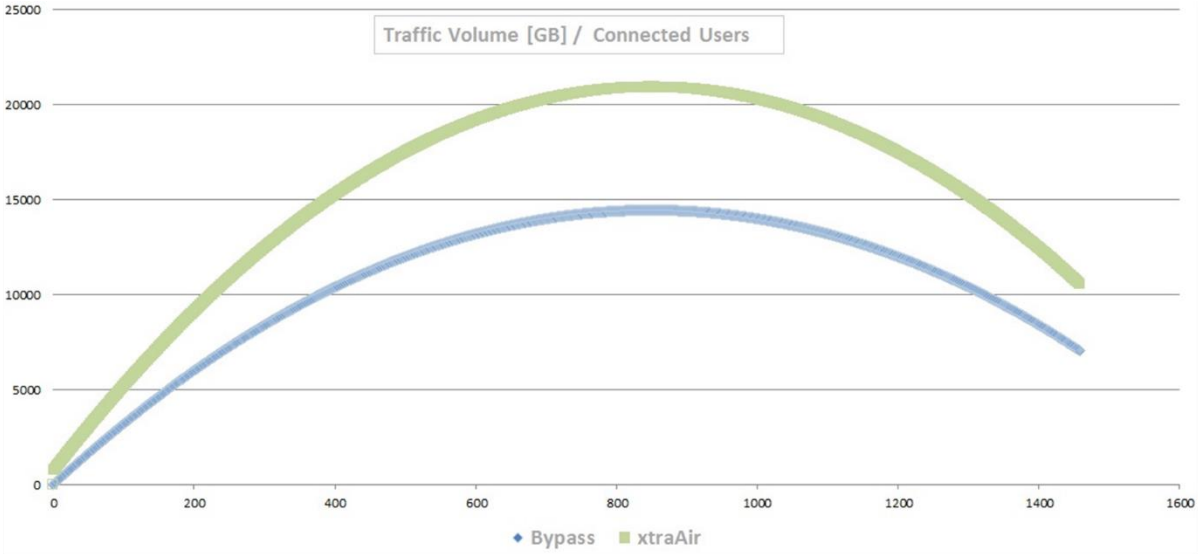
We also have the ability both to enable this by using congestion detection alone, which we have autonomously. For this, we don't need any external information, and we can enhance it with external information about the cell ID or geolocation.

Monica: Basically, your solution works regardless of who the RAN vendor is, in the sense that you can get inputs from them, but you're not dependent on them?

Ofer: Exactly.

Monica: We talked mostly about streaming video, but there is also conversational video. There, the quality requirements, as well as the expectations from the end users, are different.

Is there anything different that operators would do to manage conversational video? Does it make a difference that conversational video is symmetrical, and not just downstream traffic?



Traffic volume versus connected users with and without xtraAir *Source: Flash Networks*

Ofer: Yes, we have a few solutions for that. We do acceleration for uploads too, which are also crucial in conversational video. Sometimes the upload is the bottleneck, especially in bidirectional conversations.

Most of those services today also are getting either into VoLTE or other guaranteed quality of experience, and getting some high marks on quality of experience. Operators now have the ability to give subscribers these special resources that they need in order to get the right quality. We are able, also, to control the load on those services once they support ABR videos, because they also have the ability to go up and down in the bit rate that they are consuming, based on network conditions.

We can limit or mimic a network condition that will allow them or that will require them to reduce the bit rate slightly in order not to overload the radio, because the radio resources are needed, \also, for other users.

Monica: You can basically do that on a dynamic basis, based on the network condition and on policy?

Ofer: Exactly. We can get, also, a user plan or a user ID for different users to apply different optimization techniques or different features in real time.

Monica: What are the major concerns that you're trying to address that you hear from operators these days?

Ofer: Today the main problem of the network planners is about radio connections and radio connectivity. A few years ago, or last year, it was about throughput, and about the load of videos. Today it's more and more about radio connections

and how to maximize the number of network radio connections.

This is also the main benefit of our latest solution for radio spectral efficiency. We already are doing four trials on this new solution that we launched late last year.

I once met with an operator's VP of finance, and explained to him what we are doing and what our solution is about. Even the VP of finance knew about the radio connection problem in the network – and he's not the technical guy.

This is something that is definitely a problem. For every operator it's a problem, in a different phase of its network evolution.

Sometimes it's an operator that has just launched LTE, but still a majority of his customers are running on 3G, so he needs to better utilize his current 3G assets without upgrading them to LTE. The operator utilizes our solution to get more users on its 3G network until his 4G network is ready to be loaded more heavily and to have more users moving to LTE.

There are other operators that are far ahead in their LTE launch, and the LTE network is already loaded with more and more connections and more and more users. Here, there is a need for acceleration, data control, congestion prevention, and, of course, radio and signaling optimization for the radio itself.

Monica: Different operators have very different needs depending on what their infrastructure is, what their subscribers are doing, what their policy is. You need to be flexible in offering a solution that meets the requirements of all those different environments, right?

Ofer: Yes, definitely. We still have, also, our old and very powerful legacy optimization from the transcoding and transrating days, which supports caching, header enrichment, and things like that, that operators still need.

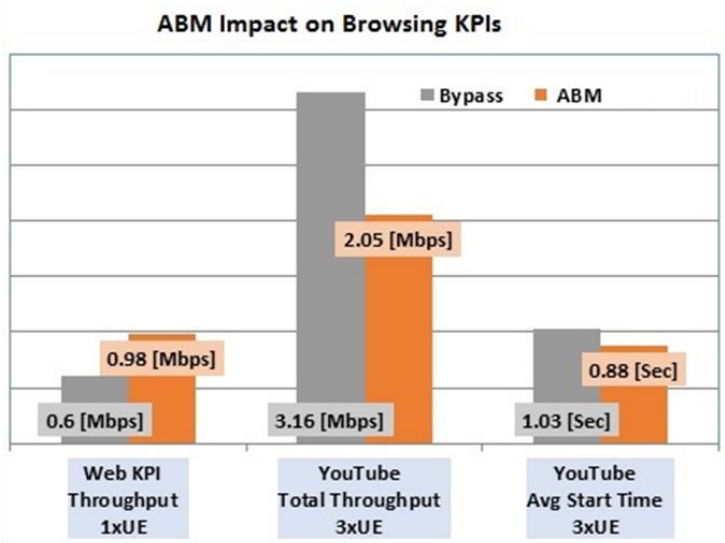
The traffic composition varies across the world. There are areas where most of the video now is HTTPS and ABR, and there are areas where progressive download, caching, and even buffer tuning or pacing are needed.

To support this, we have buffer tuning specifically for radio utilization. We call it "burst-mode" buffer tuning, because there you need to not overload the radio while doing pacing. We have a patent on it.

All those techniques are still with us and are still working in some parts of the world. We have the ability to configure or manage the system based on the requirements of the operator, who has unique conditions.

Monica: What are the features that set Flash Networks apart from everybody else?

Ofer: We have two main differentiators. One of them is our TCP acceleration. It is the best TCP acceleration out there. It's our own TCP control algorithm and congestion management algorithm. This is something that we are continually improving on and investing in all the time.



The impact of ABM on throughput *Source: Flash Networks*

The second one is the radio connection and signaling optimization. This is something unique, because we do this from the core, without the need to have anything downloaded to the device. We have no need for any cooperation with any content vendor or any device manufacturer. To optimize the radio connection and the signaling from the core is something unique that we are very proud to have launched in the market.

Monica: How is Mobile Edge Computing going to change your solution?

Ofer: Mobile Edge Computing is something that we are also working on. It is something, definitely, that we are monitoring. Today it's much easier to launch our solution on the core, or even outside the network. We have, also, the ability to have this solution on the cloud for the operator.

Monica: And what will be the impact of virtualization?

Ofer: Once the technology evolves, and it is evolving, applications on the radio side will be more virtualized and easier to deploy on the edge, and we will be able to utilize that. This will give us better granularity of real-time optimization on demand and the ability to use it only when it is needed, because we'll get more inputs from the radio. This will only enhance our solution that we have today.

Monica: In closing, what do you see happening in the future, over the next few years? What are you working on? What will the industry have to deal with?

Ofer: Definitely the internet of everything is now booming. One of the predictions is that every person will have 15 connected devices. This will be definitely something that will change the mobile environment, along with the continuous demand for more throughput and higher QoE for the user.

At CES we saw the 8K TV. Once 8K streaming is on the mobile, it will definitely demand a huge data and huge capacity increase in mobile networks.

Continuous improvement will be needed in the utilization of the radio infrastructure, which will become more and more scarce, and more and more expensive, as it has to load all this demand. This is something that we'll definitely see in our future, in our roadmap.

There is also another area we are looking at. Where there is more and more HTTPS traffic, there is also more and more risk that it may be coming from content that is not safe. We have the ability to sit in the middle of the network and to give users, via the operators, a smart, secure pipe, even on HTTPS, by controlling the traffic and monitoring the HTTP traffic. This is something that we believe will be needed, both for enterprises and consumers: a trusted pipe, from the operator, to gain more control and more security for the users.

About Flash Networks

Flash Networks is a leading provider of optimization solutions that enable operators to improve spectral efficiency, boost network speed, accelerate video and web traffic, and generate over-the-top revenues from the mobile Internet. Managing data traffic from the cloud or inside the core network, Flash Networks' xtraAir solution increases the number of connected users that can be served by 35%, and increases cell throughput by 10–15%, while accelerating network speeds for a superior user experience. In addition, Flash Networks' Layer 8 monetization solution enables operators to engage with their subscribers by offering an enhanced browsing experience while generating revenues from search, over-the-top content, and targeted advertising.

About Ofer Gottfried

Ofer Gottfried joined Flash Networks in 2007 after serving as General Manager and VP of Research and Development at NeuStar NGM (formerly Followap), a provider of instant messaging and presence products. Prior to joining NeuStar NGM, Ofer was VP of R&D at the Internet security company V-Secure, after being appointed CTO of Excel Switching, a supplier of development platforms for telecom applications and solutions. Ofer also served as General Manager and VP of R&D at Airslide prior to its merger with Excel Switching. Previously, Ofer held several senior positions at ECI Telecom and related companies managing the development of voice compression and VoIP products for the telecom market. Ofer holds a B.Sc. in Electronics Engineering from the Holon Academic Institute of Technology.

Guavus

Profile

Guavus is a big-data analytics company that works with a wide range of service providers – fixed and mobile operators, cable operators, and IP backbone carriers. It addresses the challenges that come with the need to collect and combine large amounts of data from multiple sources and to extract information from it that helps Guavus customers address performance and optimization issues. Guavus products combine scalability, real-time granularity, anomaly detection based on distributed machine learning and data science, and automated processing. Service providers use Guavus products for service assurance, customer care, and service personalization.

To mobile operators, Guavus offers products for marketing (Market Reflex), CEM (Care Reflex), operations (Service Reflex), and mediation and reporting (NetReflex for RAN and IP; Content Reflex for the core) that are all based on the analytics Reflex Platform.

The Reflex Platform uses an analyze-first, store-later approach that allows operators to collect large amounts of data in real time, analyze it, extract actionable insight, and save only the relevant information for future use. Operators can use the data they collect to:

- Explore and discover
- Ingest and normalize
- Contextualize and summarize
- Recognize patterns
- Target action

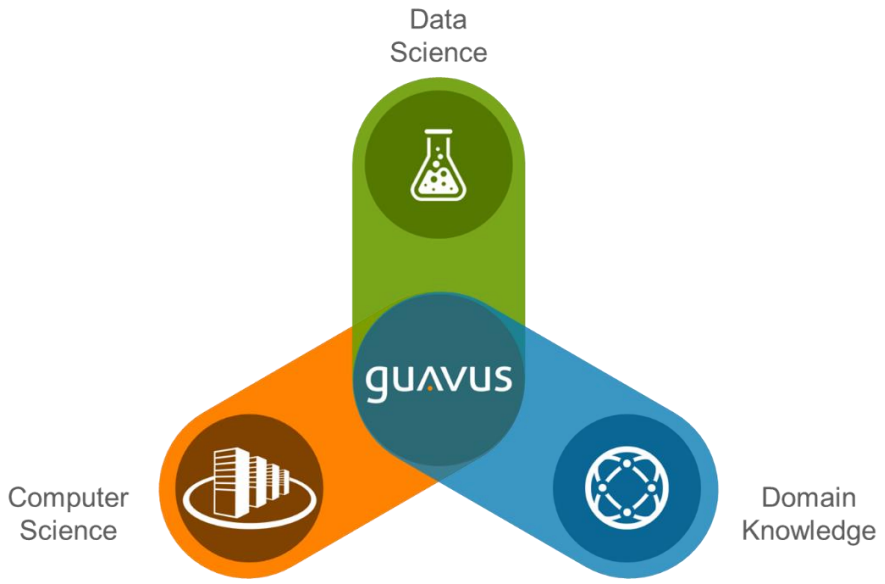
The platform is fully virtualized, but it can also work with legacy infrastructure and solutions utilizing only the specific platform components required for flexibility and extensibility.

In the context of core optimization and traffic management, Guavus analytics products help operators spot performance issues in real time, identify their causes, and find a way to address them.

Service Reflex leverages the Reflex Platform to optimize the subscriber experience. It uses anomaly detection and root cause analysis to identify network performance problems as they emerge, in order to prevent disruption or contain it. With Service Reflex operators can assess the impact that network events have on QoE and identify which subscribers they affect.

NetReflex is a correlation engine with streaming analytics that collects KPI and QCI data at the network element and subscriber levels and makes it available to operators within 5 seconds to one month, as appropriate. It collects data from network and non-network sources, and can integrate data from legacy databases.

Content Reflex is designed to understand the impact of different traffic types on network elements, helping customers improve network resource utilization, understand capacity requirements, and plan for capacity expansion. Mobile operators can use Content Reflex for encrypted content classification, trending and forecasting, optimization of video QoE, and traffic categorization.



Guavus overview

Source: Guavus

Guavus

Extracting and correlating relevant QoE and KPI data, using analytics

A conversation with Anil Singh, Director of Product Management, RAN

Monica Paolini: Welcome to this conversation with Anil Singh, the director of product management for the RAN at Guavus.

Anil, can you tell us what you do at Guavus on core optimization and traffic management?

Anil: Guavus has been around for about ten years now. We are primarily on the telecommunications domain, providing big data analytics products that deliver operational intelligence for wireless and cable service providers, as well as the regular classic wireline providers.

I am responsible for end-to-end product life cycle management of Guavus analytics products that cater to RAN and wireless access technologies in the mobile network.

Our analytics products are designed to focus on primarily two things. The first is the end-user usability, which means the ease to use, to deploy, to manage and to integrate. The second is investment conservation. Any product that we bring to the table has to have an incremental value. Any investment that an operator makes with us is future ready, and covers a broad set of use cases.

Monica: At Guavus, you're trying to optimize performance both for fixed networks and mobile networks. What's special about mobile networks that makes you do things differently?

Anil: The mobile network – and especially the wireless part of the wireless network between the handset and the base station, or the serving node – is actually the beast which creates the majority of the difference between these two sets of service providers.

Over time, mobility has evolved from being mobility for just voice into mobility for data in various types of networks, whether it was the classic cell site, to now consumers moving in from the wireless network under a small cell, to DAS networks, to Wi-Fi, and back into the macro-cell network. All of this adds a huge level of complexity in network management.

Monica: With the complexity comes the fact that you have large amounts of data, many things to keep track of in order to optimize performance, so operators often feel overwhelmed. How do we deal with that?

Anil: You're right. As the networks grow and as the types of services that are being offered on these networks continue to evolve – and faster than ever before – the huge volume and variety of data, what is called "big data," becomes a big issue.

Over time, this definition has changed quite a bit, from just being large volumes of data to also being different types of data, structured and unstructured, different velocities of data and different data sets that arrive at different time points. These sorts of complexities add to the pain that an operator would have when they start looking at combining all these data sets together and extracting actual actionable insights out of them.

Guavus was born with the understanding that analytics on big data brings new challenges and requires a new data-fabric approach. Our philosophy has been that we analyze data first and then store any other data that's needed. What that helps us do is that, whenever data is created or at data birth, we have the ability to "smart filter" it and analyze the useful data based on the business problem.

With smart filtering, we can identify whether the data is relevant or irrelevant. By irrelevant, I mean that 90% of the data today that is generally used for reporting purposes only. The data used for reporting has limited use in tactical operations management, and so you don't need to store that data.

For those occasions when there are anomalies in the network – the remaining 10% of the time – there is something wrong with the network. (I hope that this number is much less than that, but let's say for the sake of argument that this is the case.) Then the system must adapt and learn. The system needs to know that something went wrong, and now it needs to start storing more granular data.

But when nothing's wrong, the system only needs to store aggregated data. Guavus brings that machine learning and neural intelligence into the network at data birth.

Monica: Does this mean that you need to work in real time?

Anil: The concept of "real time" can be misleading, because by the time you take action for an event that may have occurred, the event has already passed. The best you could do is minimize your reaction time to certain things.

If you want a system that is truly real time, you would almost always have to predict an event's occurrence, within just enough time for that prediction to compensate for the lag in processing, reporting, transmission, and any action that needs to take place.

I don't believe any such system exists today. Anything which is feedback based, meaning an event occurs, you take feedback, the feedback executes, an action comes back. All of this takes time. It takes time for you to do, and by then, you lose the definition of real time. What we do is we try to make things possible in near-real time for whatever action that needs to be taken.

Now, if you were to think about NASA's definition of real time, how would NASA define it? NASA defines real time as the time it would take to save a human life, which means if an astronaut is light years away and for them to transmit data back to you, and you had to take a corrective action, they should have enough time for that astronaut in mission-critical conditions to save a life. How would you value anything that you are doing to the value of life?

If you mean customer experience must be conserved, or any service which is affecting a broader user needs to be conserved, this timeline for reaction is typically under 15 minutes. It would take you some time even to dispatch a guy to go out there and make some changes, or to do something remotely. Even for a remote electrical tilt to happen, it will take you a minimum of a few seconds to do that.

Guavus' products enable you to work and think in real time. A lot of our products cater to these near-real-time use cases. We have response times under 500 ms, if need be. Oftentimes, we find that most of the network controllers do not have that short a reaction time that, once we serve them that intelligence type, they can react on time for things to happen. Hopefully, that gap closes out in the near future.

Monica: Your solution allows mobile operators to drill down only when needed, so they don't get into too much complexity. Can you give me an example?

Anil: For an operator to move into real time versus not-real time requires a self-organizing network scenario if it's a highly mobile environment, There are checks and balances built into the network to try and conserve a call.

The first and foremost thing is the ability to save a call, whether it is a data call or a voice call. If you lose that call, do you have the ability to come back online quickly?

For the end user, this means to not feel that sort of delay in any application that he or she may be working through. Let's say, if you are looking at a video conference or a VoIP application, you have tolerances built in that can last you anywhere up to a second, which means that you can miss some frames and then all of a sudden you can blast these frames back to the customer.

You will often find that, when you're talking on the phone, suddenly there is a speed-up when the end user on the other side is trying to speak. These are all actions to conserve the call itself.

In terms of things that do not require real-time capability, you could be looking at things which are more on the customer care side of things.

For example, you have a user who has his call dropped or whose call has been repeatedly dropped. You tried your best to conserve the call – which is the real time aspect of management – but the non-real-time aspect of management also means that the customer still needs to be talked to, still needs to be kept informed.

You need to make sure that you are making the first contact with the customer, in letting them know, "Hey, we knew that the call was dropped and it's been happening for the past 30 minutes or for the past one hour, and here's what we're doing to help you."

Some of these systems can be automated, where you send a text message out to a customer, let him know that such and such is happening, which may or may not be real time, meaning the event has happened in the past, you just sent it 10 minutes down the road. The user does see that yes, somebody cares about him or her.

Monica: Traditionally, the type of data that mobile operators look at to optimize and to understand what is going on in the network is network KPIs averaged across the whole network or specific to an element.

There is an increasing awareness that that's not really enough. We need to look at QoE, we need to

look at new metrics, and that obviously complicates matters. How are you doing in dealing with different data sources?

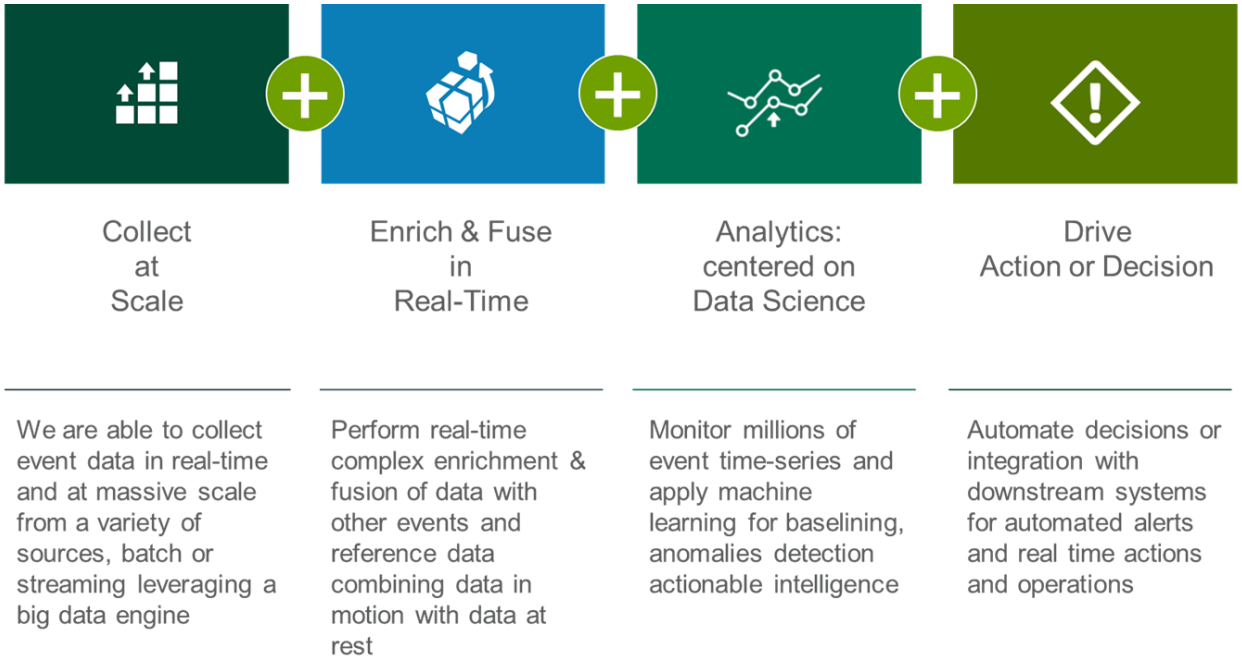
Anil: That's exactly where Guavus shines: the use of KPIs as the first beacon of actuality, meaning something just went wrong in the network and this is how I found out: in 15-minute aggregation, I found out that there was a 2% dropped call rate, or my PRB utilization just went through the roof.

While we use the KPIs as the initial indicator, there are a lot of things surrounding that event – either precursors to it, or at that same time – that if aligned will give you more intelligence.

It's not just that I know something went wrong, but how to respond back to it. What do I need to correct that? Do I have that intelligence of the five w – why it occurred, where it occurred, when it occurred, who was affected, and for what reason did that happen? If I'm to find the root cause, I do need multiple types of data sets.

Data sets that Guavus uses on a daily basis, along with performance measurement, are things like fault measurement, or FM data. We are also using trace data sets. We are using CDR data from the building site. We are also working with ingesting data from the handset nodes.

Nowadays, there is a lot of stress on moving away from simulated data, and moving towards actual user-provided information, which can be geotagged and has a lot more granularity. We do capture all of these data sets, align them on a time sequence, and then extract two types of information.



Core differentiation for Guavus applications

Source: Guavus

One is the information that is required for tactical management, meaning what should I be doing immediately? What is the root cause? Why should I send somebody to a site, or what parameters do I need to be fixing?

Secondly and most importantly, we manage a historical database, which, if a problem occurred, shows the precursor events.

In the future, we are moving towards fault isolation and preventative maintenance. You have a knowledge base that helps you address things before they start affecting the user, and sometimes that could be a big differentiator for operators.

Monica: That's quite important, because basically you use past historical data in addition to real-time data. You look at historical data to learn about the future to predict a problem before it arises. How do you add the learning into the system?

Anil: One of the key things that Guavus brings is a union of technologies that we adopt from the open-source community, but also we harden this technology for scale.

Open-source technology is not prime-time ready; it is meant for a plethora of use cases. How it caters to telecommunications data is something that Guavus specializes in.

Most importantly, we sprinkle it with a heavy influence of data science. There are a number of techniques we use when we are looking at predicting certain types of events, or when we are looking at identifying whether something is out of normal range.

I'll give you a classic example. One common misconception is that you require a humongous amount of historical data to make a prediction. There are now techniques available that do not require that. I won't discard historical data outright, but if you think about it, historical data has diminishing value, which means the older data may not be as relevant today.

The newer data set that just arrived is far more important and more heavily weighted than the earlier data sets.

We are building all of our products to have analytical capabilities, like predictive capabilities, the ability to detect anomalies, ability to correlate not just one or two or multiple but *n* number of data sets together with each other. This truly identifies at what scale we work. There are operators where we were ingesting more than 4 to 5 petabytes of data on a daily basis on a variety of data sets.

Monica: Let's look at VoLTE as an example. It's not as easy as people thought it would be.

Anil: You touch on a great point, which is a fantastic segue into what I was talking about earlier. A classic pain point that we are hearing from our customers is the fact that for every technology that an operator has deployed over time, a typical service provider's network has more than 10 different tools.

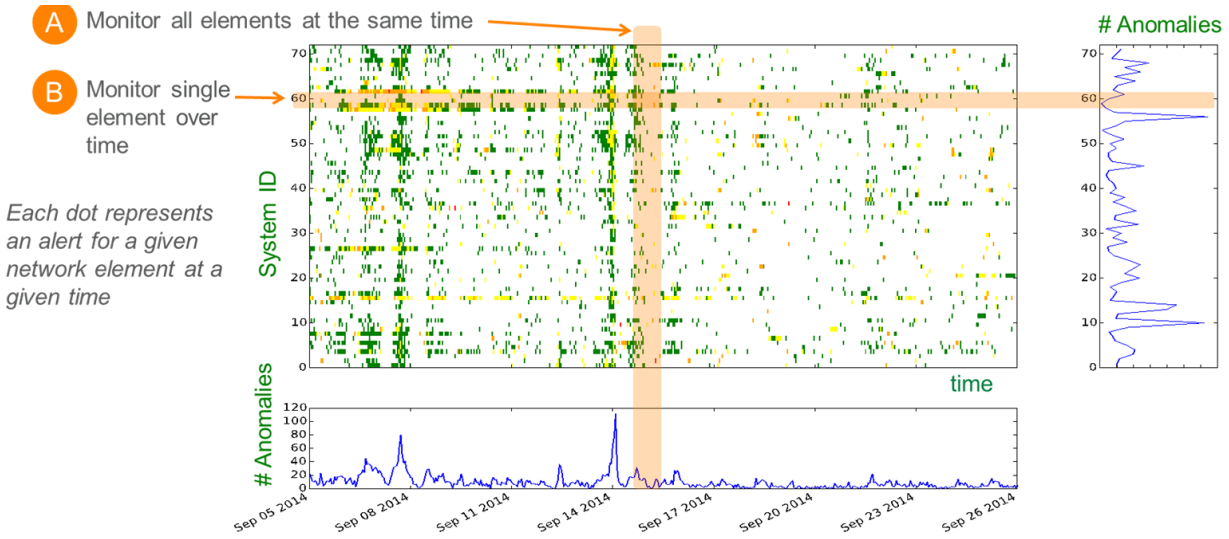
There is a tool to do capacity planning. There is a tool to do coverage planning. There is a tool that monitors the network. There are five types of probes. There are three types of equipment vendors, et cetera. At the end of the day, these tools are consumers of data. They're not creators of data, unless and until you're truly an equipment vendor, where you cater to the call itself, so you will be creating O&M traffic.

However, a lot of these tools, when you deploy them as a stack, want yet another tap of that existing data set. What happens is that there is an enormous overhead just to support individual tools.

How does that relate to VoLTE? I'll give you an example where one of our customers was looking at VoLTE calls. They had a specific tool which was meant to do real-time monitoring of VoLTE calls alone. But they found out that the tool spent most of its time deciphering whether the data was VoLTE or non-VoLTE – when at launch, let's say VoLTE accounts 5% or 7% of all your LTE traffic.

LTE traffic is huge. If only 5% of these calls are VoLTE calls, then why would you spend your assets, which cater to troubleshooting, just in identifying whether it's a VoLTE call or a non-VoLTE call? If it is a voice call, is it a troublesome VoLTE call or is it a non-VoLTE call?

These sorts of questions we've tried to address with our RAN mediation product. With that, we try to reduce the overheads that are currently out there in the market, from either transport of duplicate data packets, or just creation of duplicate packets or errors that arise because various tools aggregate data differently and so it does not match.



Network-wide view of anomalies *Source: Guavus*

Monica: There's more interest in looking at trace call data, but operators use probes as well. How do you see that changing over time in terms of the balance between these different types of information?

Anil: A lot of people think probe and trace data are the same thing, which in our view they are not. Trace data sets and probe data sets each have their own value proposition. Probe data is a lot more granular, and a lot of times it's the best granularity of data that's possible.

You ask yourself the question, "Do I always need the granularity of the data?" Let's be honest. Probe investment is pretty sizable. If an operator has to go deploy probes to cover 100% of its traffic all the time, 24/7, it's an overkill. At what times do they really need to capture this data? Additionally, it becomes a much more complex question, especially when you think about the impact of virtualization.

If the network is virtualized, where will the probe reside? Do the probe vendors have virtual probes that can replace the functionality that happens today? A lot of this is still speculative at the moment.

However, you think about the trace data sets: they're readily available; it's something that the equipment provider will give you. It's similar to your CDR calls on your BSS system. This data is available, albeit it cannot be as granular as the probe data set.

Guavus works with both. This means that we have the capability to ingest this velocity, variety and volume of data across a number of equipment vendors, whoever they may be. We provide a common platform for them to funnel the data set.

On top of it, because with Guavus you now have that capability, the service provider has Guavus as a partner in the long run. If they decide to shift away from one probe vendor and bring in another probe vendor, nothing changes on their downstream systems that rely on that data set, as long as you have an alternative.

We've often been challenged by our customers to create analytics, which combine trace data, to help minimize the dependency on the probe data sets.

Monica: There is an increasing use of geolocation, to relate QoE to a specific location. What are you doing in that area?

Anil: Guavus is not a geolocation company. We do have the knowledge base and we have the capability to identify geolocation data, but we do not market ourselves as a geolocation company. As an analytics company, geolocation information for us becomes just another, yet an important, data source.

To align geolocation data with actual call events is not trivial. We've seen that operators tend to use multiple vendors for geolocation or only use part of the network for geolocation. You don't need to geolocate 100% of your network.

Guavus offers operators that ability to pick and choose whatever vendor they want. We often tell our customers to evaluate geolocation partners from an accuracy perspective.

Monica: Virtualization is going to change the way operators think about their network and look at the performance. How is that changing things for you?

Anil: Every operator that we work with has some virtualization strategy identified. Either they are far down the path, or they have made significant investments in executing over the course of the next year or a couple of years.

Virtualization creates a more dynamic environment than what it is today. What makes it really complex is that it's an ecosystem play, in which the traditional classical players have their virtualization offerings, but you also have a whole class of innovative companies, newer players, which are bringing much value to the table, but who don't have an established relation with the operator.

All Guavus products are 100% virtualized – from capturing of data to the UI, it can all be done in the cloud or in an enterprise cloud within your firewall itself. We definitely see virtualization as being one of the key game changers in network transformation in the years to come. There is no avoiding it. It will happen, sooner or later.

Monica: What are the next big challenges that you're addressing at Guavus, or that the industry is addressing right now?

Anil: What we look forward to, and we are seeing a lot of traction for in the industry, is how are operators monetizing their bandwidth for non-telecommunications data sets? IoT becomes a huge thing.

When telemetric data is transmitted on that same network that also carries customer-sensitive data,

or let's say first responder data, how do you manage the network as a whole?

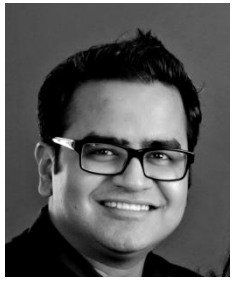
How does IoT become the game changer? At what point in time is it OK to transmit IoT data over that same infrastructure, or do you want to leverage something like unlicensed LTE?

We see that IoT is one of the key areas which can become a game changer, can be a new revenue source for the operators and definitely something that we are trialing in various parts of the world as we're proving how analytics can help operators monetize this opportunity.

About Guavus

Guavus solves the world's most complex data problems. Proven across Fortune 500 enterprises, Guavus provides a new generation of analytically powered big data applications to address specific business problems for next-generation service assurance, next-generation customer experience management and the Internet of Things. The company uniquely breaks down the barriers between Operational Support Systems and Business Support Systems to enable customers to more efficiently plan network capacity, improve service operations and deliver a better customer experience. Guavus' operational intelligence applications correlate and analyze massive amounts of streaming and stored business, operational and sensor data from multiple, disparate source systems in real time. Guavus products currently process more than two trillion transactions per day.

About Anil Singh



Anil Singh is Sr. Director – Product Management (Mobility) at Guavus and responsible for product life cycle management for Guavus products that cater to wireless access networks including RAN. Anil is an industry veteran with 15+ years of experience engineering products and developing solutions for next generation RAN. Anil also defines and executes mobility product strategy supporting a global sales team while managing partner relationships for our RAN analytics suite of solutions. Anil holds over 15 US patents covering wireless network capacity and performance management. Prior to his role at Guavus, Anil has held engineering and sales leadership roles at Samsung, Sprint, Nexius, Celcite (acquired by Amdocs) and MSI-Marconi. Anil holds a Bachelors degree in Instrumentation and Controls engineering from University of Mumbai, and a Masters of Science in Electrical Engineering specializing in Wireless Telecommunications Network from the University of Houston.

Viavi Solutions

Profile

Viavi Solutions (previously, JDSU, Arieso and Network Instruments) enables network operators and enterprises to gain visibility into wireless, wireline, enterprise and cloud networks and to monitor and optimize them. In the wireless area, Viavi covers the end-to-end network infrastructure throughout its life cycle: design and planning, installation, commissioning and verification, network optimization, acceptance, monitoring, and service assurance.

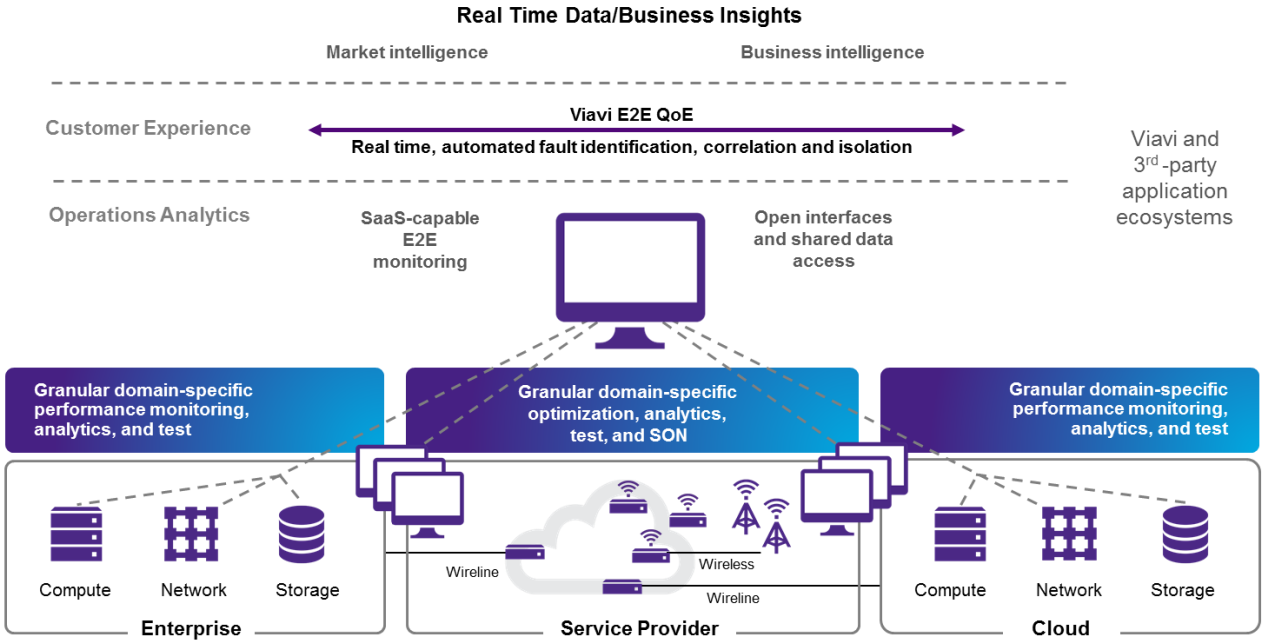
In addition to its RAN, backhaul and core optimization and assurance products, Viavi focuses on these areas to optimize core functionality and traffic management:

- Customer experience and service assurance, both to support customer retention and customer satisfaction, and to capture QoE in real time. Mobile operators can use QoE data to optimize their networks, isolate performance issues, and identify the network elements that cause these issues. Customer assurance solutions enable operators to correlate QoE metrics to network KPIs, and to gain end-to-end visibility in their networks.
- Testing, measurement and assurance of legacy and virtualized networks, helping operators to optimize their core networks as they transition towards NFV.
- VoLTE service assurance solutions specifically targeted at voice quality. Operators can use them to track R-factor MOS, PESQ scores, and IMS negotiation parameters, and relate them to location and RAN performance.

At the core of Viavi solutions for mobile operators is xSIGHT, PacketPortal and ariesoGEO. xSIGHT is a mediation, correlation and policy management platform that operates in real time to capture both network KPIs and QoE metrics. It delivers a multidimensional analysis of network performance that operators can use not only for core optimization and traffic management, but also to support customer services, reduce churn, and monetize new services.

PacketPortal is an additional solution that Viavi offers mobile operators, for collecting fine-grained data on subscribers, services and applications. While it is primarily targeted at improving customer experience, it also supports network optimization that is content, application, and service aware.

ariesoGEO locates, stores and analyzes data from mobile connection events, giving operators a source of intelligence to improve performance and user experience. Operators can use this intelligence to improve network resource utilization, refine their monetization strategy, and to deepen their understanding of QoE.



Capturing end-to-end QoE, with testing, assurance and analytics *Source: Viavi Solutions*

Viavi Solutions

An end-to-end, real-time approach to using QoE in network optimization

A conversation with Paul Gowans, Wireless & RAN Solutions Marketing Manager, and Ronnie Neil, Customer Experience Assurance Solutions Marketing Manager

Monica Paolini: Welcome to this interview that is part of our report on core optimization and traffic management.

Today, I have with me Paul Gowans, Wireless & RAN Solutions Marketing Manager, and Ronnie Neil, Customer Experience Assurance Solutions Marketing Manager, at Viavi Solutions.

Paul, Ronnie, thanks for being with me today. To get started, could you give us an introduction to what Viavi Solutions is doing in this area? Has there been a change in direction since you changed your name from JDSU?

Paul: Many people may not even be familiar yet with the Viavi Solutions brand. What we did was we took our Arieso product line, our network instruments product line, our JDSU product line, and we brought those together under one new brand image, Viavi Solutions.

We separated all those from our optical components group. It allowed us to bring a singular focus to our enterprise customers and our service provider customers to turn up, manage and optimize their networks. We help them deliver an end-to-end customer experience integrating both our instruments solutions and our systems solutions under one common platform.

Monica: You've been working on optimizing the end-to-end network – not just the core – for a long time. How is the evolution in technology and usage patterns changing the way we evaluate network performance and optimize networks to provide a better experience to the subscribers?

Ronnie: It's made a huge difference, Monica. If you think back a few years, we had one dominant application, and that was voice. The network was engineered from a performance standpoint to cope with that and deliver a high quality of experience for the voice service.

Now we have many types of applications. We have video, audio streaming and downloads – and we still have voice. These applications place different demands on the network to deliver a high quality of experience, or QoE.

To determine what QoE is being received, we need to know what application the end user is using at that time. We still make network performance measurements, but we need to also identify what the application is. Only by knowing what the application is can we relate those measurements to QoE; which is still the critical factor.

Monica: Looking at QoE at the application level increases the complexity of optimizing network performance.

Ronnie: It makes QoE more complex to measure.

Traditionally, solutions have used DPI to do that. DPI looks inside the packets to identify what the user does. But there are parts of the world where DPI is not allowed, because of data privacy concerns. Also, DPI does not work when you have an encrypted session, because it cannot identify the application.

We need to go beyond DPI and come out with a method that identifies the application in a way that is compliant with data privacy regulations and works with encrypted sessions. That's not easy.

Monica: When there is a QoE issue, how can you find out where the causes are?

Ronnie: That situation has gotten more complicated. As well as having different applications, now we have smart devices, we have content servers, and the intelligence is much more distributed across the network.

To identify issues, you have to look into multiple dimensions. What's the device? What's the application? Where's the location? You need to be able to analyze quality of experience by these dimensions.

A performance fault may be affecting only a particular device. It might be a device that recently has had an operating system upgrade, and suddenly it is causing lots of issues. Or it may be a network issue. You need that ability to not only measure QoE, but also to analyze that QoE by

dimension, and be able to identify the cause of the fault in real time.

Many traditional solutions could give you that dimensional analysis, but only in historical reports. We need to be able to do that in real time to quickly identify issues, prioritize them, and ultimately, resolve them. It's much more complex than it used to be.

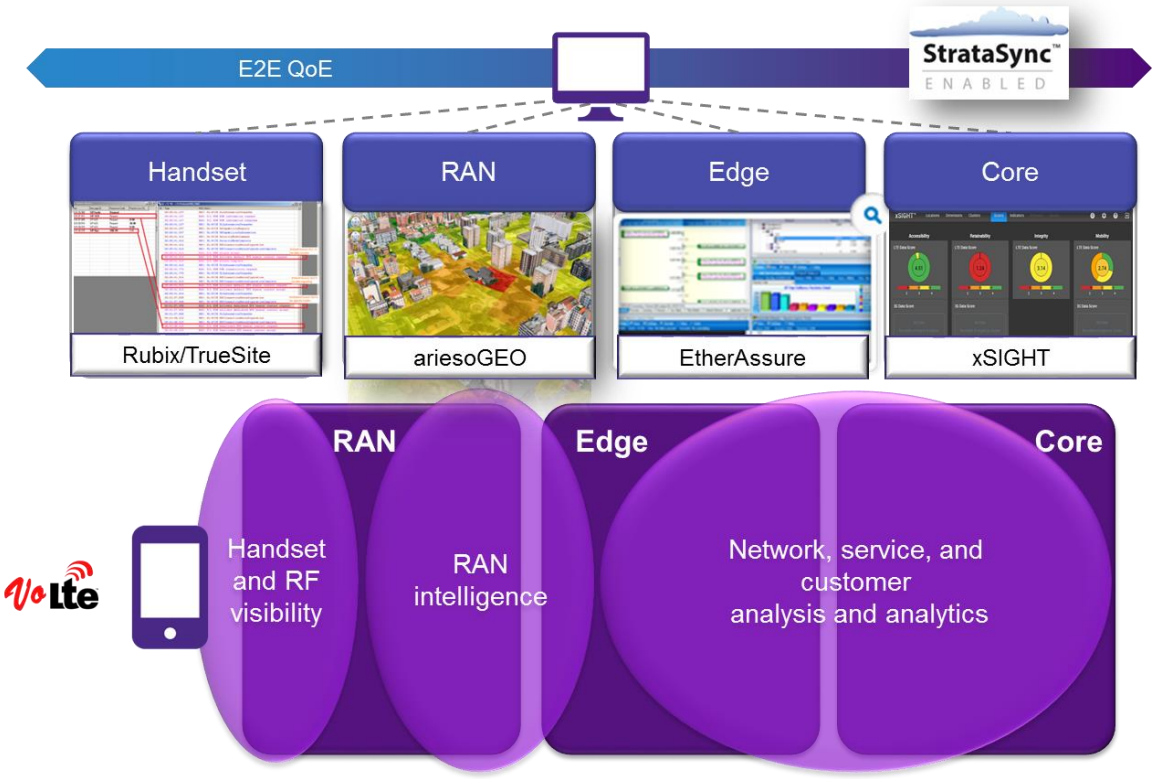
Monica: To add to this complexity, the way we use mobile networks is highly uneven. We tend to congregate in a small set of places at the same times. As a result, there is more pressure in some parts of the network at some times of the day. How can operators deal with that?

Paul: The extreme non-uniformity in mobile networks is a key challenge for our customers. There are four components. There's time, there's application, there's subscriber, and there's location.

If you look at utilization variation at a cell site, you can see a 30% difference in utilization in just a 50-minute period. That's a massive difference and a stress on the network.

Coupled with application usage, there may be specific areas where there's a lot of download of video. Maybe specific areas where there's a lot more voice. Maybe in some areas, there's more BitTorrent going on. Maybe in some areas there's more upload to other applications. All those applications can change at the different times of day.

Then you add in the subscriber usage. At one of the operators, we've found that 50% of the data was being consumed by one percent of the users.



VoLTE end-to-end visibility *Source: Viavi Solutions*

If we add location into that, we found that 50% of the data was being consumed in 0.35% of locations – 0.35%. That's a massive concentration.

The extreme variation in the network can have a major effect on operators trying to optimize the network for the best QoE to the customer.

Monica: This is really important, because you need to capture and leverage that variability to your advantage, and to protect the network, and to improve the experience. You need to really have an end-to-end look at QoE – capturing QoE and mapping it to the different elements in the network.

Just knowing what QoE is doesn't really help you unless you're able to do something with that information.

Paul: That's a great point. The visibility can be into the network, but then you're building on that visibility to do analytics, and to get to the root cause of things.

As an example, VoLTE is a service that is handset to handset typically, and that touches many parts of

the network, requiring particular bearers to be set up all the way throughout the network.

From a signaling and a user plane perspective, those bearers go from the handset, go through the RAN, hit the backhaul and then get to the core. You need to be able to get a view, end to end, right through all those components.

I need to know what the voice quality is in different parts of the network so I can determine what part of the network has an effect on the QoE. Getting that end-to-end view becomes fundamental.

Of course, that's not just for VoLTE. It could be for any kind of service you implement. It could be for video or video streaming, for example.

Monica: Traditionally, mobile operators have looked at KPIs, measuring the average network performance. That's no longer enough. What is it that you can help operators measure, beyond KPIs, to understand QoE?

Ronnie: Let's start with QoE. You need to identify what application is being used, so that you can relate the performance measurements to an actual, true QoE.

DPI was the traditional tool used to do that identification. As I mentioned before, there are some issues with that.

At Viavi, we worked with a tier one operator in Germany. They were using our solution, with the DPI, to identify the end user application. Germany has strict data privacy laws. The operator was asked to turn off the DPI and our solution. This meant that they couldn't do the QoE determination that we just talked about. At Viavi, we looked at how we could then identify that end-user application.

We've come up with an alternative method, which we have patented and which we use to identify the application being used without DPI. It complies with data privacy laws and also works with encrypted sessions – to the extent that, for that German tier-one operator, we were able to switch that functionality back on to identify the application being used.

Indeed, it allows us to identify the application used in over 90% of the sessions.

Monica: What are you doing on the other dimensions that contribute to QoE?

Ronnie: We've architected our assurance solutions to do that multidimensional analysis of QoE, or indeed of the network performance, in real time. That's a great step forward, because traditionally you could only get that in historical reports, perhaps produced once a week.

Now we are able to do that multidimensional analysis in real time, so you can measure QoE for a specific device, for a specific location, or for an individual over-the-top application. Also, you can do combinations. For instance, it may be that the issue affects devices only as they're using a particular application.

In this way, we're able to identify those issues that affect only a single dimension or perhaps a combination of dimensions. Not only are we able to give you that analysis, we then allow you to drill down on only those dimensions.

Let's say that you've identified an issue, and it does affect only a single device. We then allow you to drill down into the detail of that issue by looking at results associated with only that dimension, that device, or it may be a combination of a device and an application.

We've managed to significantly improve the ability to identify, isolate, and resolve issues, far quicker than traditional solutions can do.

Monica: While you have automated the process of collecting the data, the operator still retains the ability to direct the attention where it's needed.

Paul: Absolutely. We have two main applications.

One operates from QoE almost downwards. You focus in on the areas where poor quality of experience has been detected.

The other one gives the operator complete flexibility as to where he wants to go and look. Is it a particular dimension, location or device? They have that complete flexibility over where to look, and what KPIs to look for.

Monica: There are different concepts of what real time is. Is it one day, is it one millisecond? How do we decide what time resolution we need?

Ronnie: When we're talking about real time, we're talking about of the order of a minute. Traditional solutions would have been 15 minutes up to an hour. We've significantly reduced that. As we move forward, we'd also like to get that reduced even further. Our architecture has the potential to do that.

Monica: Once the operator has the information it needs, how does it move into resolving the performance issues it has identified?

Ronnie: The multidimensional analysis in real time allows them to drill down on those results. If it's a particular device that has an issue, you exclude all the results associated with all of the other devices. You focus only on that device.

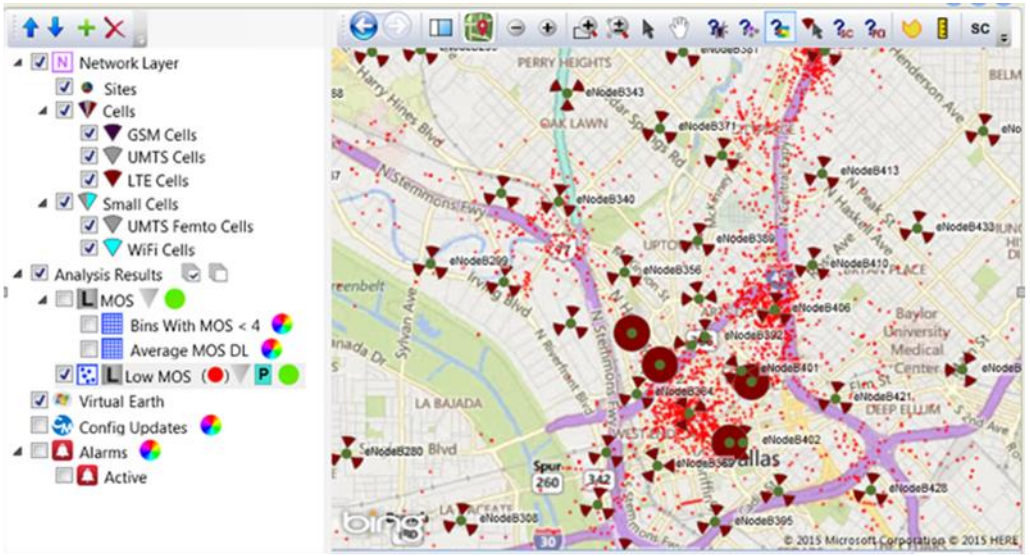
You can drill all the way down to where you can identify network transactions that have caused the problem for that device. It may even be something like a device setting. In particular, we've simplified the troubleshooting process to allow operators to not just identify and prioritize the issues, but to diagnose them and resolve them in a fraction of the time.

Monica: Let's say there is a problem with video performance in the mobile device. How do we know where that one comes from?

Paul: Monica, that's where you need that end-to-end view. We talked about that challenge earlier on. With VoLTE and video, for example, you need to build a segmented view of the network and customer experience to analyze that.

For example, you can have agents on the handset that can look at the RF metrics and can also do active testing to measure VoLTE and video service quality. That's a good measure of what service an end user's got, but what it doesn't tell you is where the problem lies.

You need to go into part of the network. If you get metrics from the RAN, it is important to capture that information with a customer-centric view – not just capturing performance stats, but capturing the information relating to the events that the customer's having with the network. You're getting a customer-centric, location-aware view of this service. Then you can isolate what's going on at the RAN level.



Low VoLTE MOS scores shown by geo-location *Source: Viavi Solutions*

If you go further into the network, the backhaul can often be a blind spot for operators to try and determine what's going on in there. You can embed technology into the backhaul to look at the transport metrics, to see if the transport's working fine, but also then delve into specific packets. As Ronnie mentioned, you might want to capture information on a specific stream.

Let's say you've got RTCP data that says there's a problem with voice quality. You can start to use embedded technology, an agent in your backhaul, to capture that packet stream and do some analysis on it – maybe it's video or voice quality.

Then you move into the core of the network, where you've got, of course, both the signaling and data planes – for example, IMS core control. But then you've also got the user plane, as Ronnie mentioned, in terms of that dimensional analysis, because you might want to look at the network's ability to deliver those packets. Deliver, for example, at the level of a specific R-factor score in VoLTE voice quality.

The end-to-end visibility, from handset through the core, provides a view and a perspective on QoE. It allows you to use that analysis to delve down into where the problem's occurring and hopefully to proactively solve that before the customer even has to phone you up with the problem.

Monica: This allows you to also deal with the non-uniformity of usage, because you can optimize the

network in different ways depending on where in the network you have QoE issues.

Paul: Exactly. Taking a customer-centric view, capturing all events for all services, for all customers all the time allows you to model what's going on in your network, so that you can make some predictions as to what is about to happen.

You can model, over a period of time, the flow of traffic and people in your network, and you can use that data to predict, in a potentially real-time environment, what stresses are in the network and how to deal with them.

Taking into account that non-uniformity, the multidimensional analysis, and the end-to-end perspective is the way that service providers can get the perspective for what sort of service a customer is getting.

Monica: What are the next challenges that we're facing? What are you focusing on right now at Viavi, for the future?

Ronnie: The first thing is data monetization. One of the things that has changed, in the last few years, with our assurance solutions is that the information we gather traditionally was used only by the people in network operations. It was highly technical and mostly dealt with network signaling.

Nowadays, we're gathering information on not only the customers' QoE, but also what's their usage in terms of applications, what devices are they using, what locations are they using.

That, if you think about it, is very useful information that the marketing departments within the operators could use, both internally – perhaps to identify new services to offer to increase their revenues – and externally, to sell information to third-party companies about subscribers: where they're doing things, what applications they're using, what devices they're using, etc. One of the things that we're looking at is the application of the information that we're gathering.

The second one I'd mention is Network Functions Virtualization, or NFV. NFV will have a huge impact on the types of solutions that Viavi sells. Most of the tools that we supply have traditionally been connected to physical network interfaces – either to monitor traffic or indeed to inject traffic, for tools that are active testers.

With NFV, some of these physical interfaces will no longer exist. You'll ask, "Where can I test or monitor a particular interface?" The answer might be, "Well, it's inside that server. It's inside that virtual machine. There is no physical interface." We have to evolve our solutions such that they can be used in virtual environments.

The solutions themselves have to become virtualized. They have to have the ability to connect to virtual network interfaces. These are huge changes for us, in our business, in terms of NFV.

Virtualizing solutions and monetizing data are certainly two of the areas that we're looking at and working on right now.

Paul: The two developments I'd like to touch on are SON and 5G. SON reflects back to that non-uniformity discussion. How do you self-organize? How do you self-optimize? How do you self-heal a network in this environment?

One of the things that we're doing is developing a strong SON strategy around customer-aware, location-aware predictive capabilities within SON, and the capabilities, to develop an end-to-end SON perspective.

SON can't just sit in isolation in the RAN. We need to take an end-to-end approach. We're continuing to develop our SON portfolio, which is already addressing some key challenges in assessing customer QoE today. SON will then evolve into one of the key elements of 5G.

With the end-to-end SON network slicing, you take a slice through the network for a particular service, or particular customer, or particular group of customers. How do you self-optimize, self-configure, self-heal that network?

You can take a predictive approach, so that you can do things in real time. You can take a modeling approach, so that you can model your network and reflect the flow of traffic across the networks. The evolution into network slicing in SON becomes fundamental.

Another aspect of 5G worth bringing up is the ultra-dense network perspective. The evolution of HetNets into ultra-dense networks provides some interesting opportunities, as well as challenges, for operators.

Finally, Mobile Edge Computing brings much more intelligence to the edge of the network, as it moves decisions to the edge, where perhaps operators didn't have visibility before. Again, it's that end-to-end visibility and analytics that will be fundamental to managing QoE.

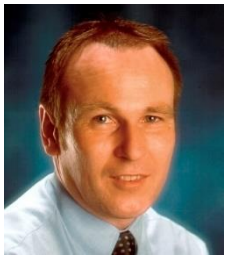
About Viavi Solutions

Viavi Solutions offers fully integrated and interoperable solutions for network testing, performance optimization, and service assurance. Designed to support the most complex IT and communications networks of today and tomorrow, our solutions help you get the best possible performance from your network investments. We deliver the precision intelligence and flexibility you need to cost-efficiently scale operations, transition to next-gen technologies, and diversify revenue opportunities for greater profitability.

About Paul Gowans

Paul is currently the marketing manager for Viavi Solutions' Wireless and RAN solutions. With more than 20 years of experience in the communications and mobile industries, his expertise in IP, Mobile, LTE, VoIP, IMS, and VoLTE comes from developing, supporting, marketing, and consulting on major mobile communications projects. Prior to Viavi, Paul was Global Marketing Manager for Agilent Technologies' Network Systems Division. Paul is a graduate of Edinburgh University in Scotland with a Bachelor of Science in Computer and Management Science and works out of the Viavi Solutions' Edinburgh office in the UK.

About Ronnie Neil



Ronnie is currently marketing manager for the Viavi Solutions' Customer Experience Assurance (CEA) solutions. Prior to Viavi, Ronnie spent over 25 years with HP/Agilent Technologies in various telecom test marketing and R&D management positions. His expertise in customer experience management comes from developing, supporting, marketing, and consulting on large-scale assurance projects worldwide. Ronnie is a graduate of Strathclyde University in Scotland with a Bachelor of Science degree in Physics and works out of the Viavi Solutions' Edinburgh office in the UK.

III. Operator interviews

Three UK

Driving network strategy through customer perception

A conversation with Mony Kochupillai, Head of Network Perception

Monica Paolini: As part of Senza Fili's report on core optimization and traffic management, I am talking today with Mony Kochupillai, Head of Network Perception at Three in the UK.

Mony, thanks for talking to us today. You have a cool and insightful title, Head of Network Perception. Can you tell us what you do?

Mony Kochupillai: I'm responsible for ensuring our customers get the best possible network experience, which translates into their perception of our brand.

Monica: From a subscriber point of view, you use your phone, and you're either happy or unhappy. It seems an easy thing to decide as a subscriber. From a network point of view, it's not that easy to qualify or to quantify that level of happiness or unhappiness, because it is subjective. What I might be happy with, you might not be happy with.

Mony: It is an extremely important and a very good question. There is a lack of understanding across the industry, in my opinion, about the linkage between the profitability, perception and experience of customers and the underlying physical performance of the network.

That is an area which the industry should focus on and be working hard to improve. Sometimes engineers try to deliver a level of performance which could be a mismatch from how customers perceive the experience of the network.

Monica: Can you explain how is it that the network may be working well, in the sense that there are no performance issues – all the elements are up and working as promised – but the subscriber's perception might not be as good?

Mony: There are some classic examples. Let me give you one on network coverage.

Today, most of the operators around the world rely on coverage prediction tools to determine the coverage that customers can expect. This means they predict the quality and reliability of what coverage a customer can get inside a particular building. The engineers who do the prediction expect that customers have coverage, but the actual coverage experience for the customers in that building or in that part of the building can be completely different. This is number one.

Number two is that different customer segments have different levels of expectation and tolerance in terms of the quality and reliability of network coverage. For example, I have come across customers who say, "I am happy with your network although I have no coverage anywhere inside my home." I have also come across customers who say, "I'm unhappy with your network, because I haven't got full coverage inside my home." Finally, there are customers who say, "Hey, I'm happy with your network and recommended you to my partner, because I have some coverage inside my home."

So the expectation and tolerance levels among customer segments vary significantly. This is difficult for engineering teams to understand purely from a performance perspective. That's why it is extremely important to bring the customer experience, their perception and profitability into consideration when you invest in and improve the network.

Monica: That's difficult, too, because, as you say, you have to start with the subscribers and end up with engineers, who are more used to optimizing the performance of individual network elements.

Let's start with the subscriber. How do you quantify QoE? As you say, somebody might be happy with just having some coverage. Somebody else might be very unhappy if they cannot have coverage in their basement. How do you quantify that?

At some point, you need to qualify the performance of the network, so you have to average. But if you average the two extremes, you get an average happiness.

Mony: First, we shouldn't go by the average experience. It is extremely important to look at the distribution of customers' experience, their perception and profitability.

Second is to also look at the linkage between the customer perception and the performance of the network.

You will always have customers who are detractors that you want to keep to a minimum. In addition to technology and cost, network strategy has to take into account customers' experience, their perception and profitability. So you keep the proportion of detractors among your customers as low as possible.

What you want is to have most of the customers love your network, your brand; they stay with you longer; and they recommend your network to more people.

Monica: For specific applications, like video or voice – which are real-time and have special requirements – do you set specific targets, in terms of how long it takes to get the video started, or how much stalling you allow?

Mony: Definitely there are performance metrics, which we can have on an individual service basis. But it can be extremely challenging for operators to focus on individual services at a micro level.

The good thing with the internet is that you know what all the key services are. Video is one of the key drivers of data demand. You know what kind of bandwidth and reliability requirements it has; and you try deliver against those requirements.

Just make sure that you provide a level of service for customers, so they can do everything they want to do on the internet, including video, browsing, email, or social networking, without any hindrance. You focus on that, as opposed to just focusing on video. That is a much better and simple approach.

The reason I'm saying this is, again, that different customers have different needs. It is very challenging to optimize a network based on individual customer needs. For example, in my case, say, in Paddington Station in London, I will optimize the network so that the vast majority of the customers get an amazing experience in that location.

Monica: Do you think you have to optimize based on the relative importance of different applications? For instance, if everybody wants to watch a video, do you want to try to optimize more for that? Or, because voice is more important in Paddington, where people may want to tell their friends they are late, you give it priority?

Mony: It has to take into account the demand from the customers, particularly what service the majority of the customers use and how often they use it. That is number one.

Number two, these services also have a certain tolerance limit. For example, voice is less tolerant of any network impairments, so you will have to put more focus on optimizing that service compared to, for instance, browsing or email – even though browsing and email are also used by a lot of customers on a daily basis.

It depends on the customer demand, as well as the quality requirements for that particular service to work.

Monica: In terms of the perception for voice calls, it used to be that, in a switched network, the call either goes through or it doesn't. When it goes through, the latency is relatively stable.

With VoLTE, there is more of a gray zone. The call might go through, but you have higher latency, so the call quality is bad, but there is no dropped call. The dropped call rate is no longer sufficient to represent the perception of voice calls.

Mony: My view is that reliability is the most important thing. When I say reliability, in simple terms I mean that the customers should be able to do, on their phone, what they want to do and when they want to do it, including voice calls. What that means is when I want to be able to make a voice call, I should be able to make a voice call, and to have a quality conversation. Both of them do matter.

If I cannot make a voice call, I will be dissatisfied, but if I cannot have a clear conversation, then also I will be dissatisfied. What we are aiming to provide is, first, that you are able to connect. Then, once you connect, you are able to hold a conversation, and there is a clarity of speech, which allows you to have the conversation you want to have.

In my view, all the three elements are equally important. One is the ability to connect, the second is the ability to hold the call, and the third is the clarity of speech, and that is what we are aiming to provide with VoLTE, as well as any other voice service.

Monica: Obviously, you try to give the best quality to everybody, but let's say that there is a situation where maybe the coverage is not very good: the network is at congestion or near congestion. Do you think that, for the subscriber, a call that might not be perfect is still better than not being able to make a call?

Mony: Definitely, being able to make a call is more important, even if it is at a lower quality. I think that is very, very important.

There are locations where customers may not be able to make a call, or may not be able to send a text message. But if you provide them with some coverage – where they can do something, even at lower quality – that is more important.

Monica: Do you think this is a change that is brought by the wider use of wireless data? We are used to video streams or connections that may get interrupted, and we simply reconnect. Have subscribers developed more tolerance for this type of connectivity for voice as well?

Mony: What the customers are expecting is a reliable service. Reliability is far more important than the coverage or anything else. What that basically means for me is that costumers have a simple expectation: to be able to do what they want to do on their phone, whenever and wherever.

That is in terms of performance. That's a very simple expectation the customers have. If they're unable to achieve that, they should be able to get support they can rely on. When I say support, it is also the information they get when things don't work, which they should be able to rely on.

Monica: You're focusing on what the subscribers tell you. A very difficult part of your job, I would imagine, is to translate that for the network engineers, so that they can take the appropriate action. This is not trivial, not because it's difficult to talk to the engineers, but because it's difficult to translate the input from subscribers to network engineering.

Mony: This is an important point you are making. That is where my job gets more complicated. In terms of what the customers want, we have a very good understanding. They're expecting things to work as reliably as it could. Our job is to translate the complicated technology into something which the costumers don't have to think about too much but they can rely on.

To be able to achieve that, you have to have a model which links customers' profitability to their perception and experience as well as the underlying performance of the physical network. We have a model that we use, mapping the customers profitability to their perception and experience as well as the performance of the physical network, and it's important for all operators to get their model right.

Otherwise, you don't get a grip on how much money to spend on your network, where to spend, and how it impacts the customers' profitability, their perception and experience. Having that understanding is critical for the future of mobile industry, particularly when the revenue is squeezed. You have to be smarter in the way you invest in new technology, as well as in the way you optimize the existing technology. Otherwise, you can get top network performance in benchmark tests, but it doesn't help shift customer profitability and perception.

Monica: How do you take into account subscriber experience to prioritize your investment? It used to be that subscriber perception was separate from network operations. Has that changed the way different teams interact within the company?

Mony: Definitely, this is a change in the industry that I like to see. If you see the operators' role is to be providing a network experience that their customers can rely on, the key is translating that customer experience and perception into better network performance and deliver against that.

There is no standard in the industry to do this translation. But it's not possible to have that standard model, because different operators attract different customer segments, and each of these segments will have different levels of expectation and tolerance for network performance.

Operators will have to build their proprietary model suited to their customer base to translate their customers' profitability into their perception and experience and e into what does it mean in terms of network performance and what does that mean in terms of the priority for investment.

We use such a model to deploy low frequency spectrum to improve indoor coverage and end-end optimization of the network experience.

Monica: It means that you can treat different groups of subscribers differently based on how much they paid.

Mony: Indeed. Again, you come into the micro strategy within your own business to say, "OK, maybe these are my high-value customers and they are giving me majority of my margin, so I have to have a model which addresses the experience and perception of those high-value customers."

That may be a strategy that operators will take. I also personally believe that this is an area which is going to be a differentiation between operators in terms of the way they build and operate their network.

Monica: Traditionally, engineers have relied on KPIs, which are historical measures of network performance, usually averaged. They try to get the best values they can possibly get, and that's how success is measured.

That model doesn't seem to work very well when you start taking into account a subscriber's experience, when you look at different applications, and when you start being able to use real-time data.

How do you relate the data you collect from the subscribers to KPIs? How do they map onto each other, or do they?

Mony: They have to. That's what I'm talking about, the linkage, which is proprietary for each operator. It is extremely important for operators to establish that linkage – what that consumer perception, say in terms of NPS or satisfaction, and profitability equates to in terms of customer experience measures, and what it equates to in operational performance measures.

I'm with you that, historically, the industry has been working on the basis of KPIs as measured by engineers, but the world is changing. It's all about the customer experience and customer perception. That is what is important to drive your profitability, brand perception and equity.

Monica: Is there a way that, as you try to capture the user experience, you learn how to map that into KPIs, or onto something different? Is there a one-to-one mapping, or is there a way that you can map user experience onto the other measures that we're just still learning? Or is it just that we need something else altogether?

Mony: These are the very early days for this concept, and we are still learning. This is an area which needs to develop. Definitely, in terms of outcome, you have the customer perception in terms of NPS, you will also have customer profitability.

Further, there are some lead measures, which maybe the experience measures, and then underneath, there are some performance measures. We, as an operator, have to establish how the business outcome measures, link to the lead measures, and how they, in turn, link to the optimum operational performance measures.

That's how the network needs to be optimized.

Monica: I've talked to some operators that are starting to look into crowdsourcing as a way to understand how networks work, especially in real time, to see what's going on at a given time and a given location. Are you looking at that? Do you think it's useful?

Mony: If you step back years ago, the technology was very, very limited. Now, the technology has developed, and the device capabilities have improved.

Also, customers have become smarter with social media and other applications. This means that there are a lot of informational capabilities out there, which allow us to have things like crowd sourced data, which is one source of data that we use to understand the customer perception and experience. Definitely.

Monica: How important is it to move toward real-time optimization of the network? Instead of just looking at the data from the last month or so, doing it in real time or near-real time?

Mony: It is extremely important, and this is a game changer. You can look at a customer's data once they have churned, and there is nothing that you could do to that customer. It's a reactive approach. The industry as a whole needs to become more proactive, understanding what is happening to the customers in real time, and take real-time actions.

That is an area which needs to develop a lot, and there is a lot of scope for that in the industry.

Monica: If you look at the next five years, what changes do you expect to see, in terms of the way we optimize the network to take into account a quality of experience?

Mony: It is all about measuring customer profitability, customer perception, customer experience, and translating that into the network performance. That, in turn, drives the investment and the focus, based on what drives the best outcome in terms of the revolution to the customer perception, and in turn the profitability, brand perception and equity. That is going to be the key area for the next 5 to 10 years.

About Three UK

Three is a communications company, focused on challenging the industry to make mobile better for everyone. Three wants to give customers a quality mobile experience and address the industry issues that frustrate them. It were the first UK network to introduce all you can eat data and offer 4G at no extra cost. Three lets customers roam abroad in 18 countries with their domestic bundle for no extra charge. The operator continues to look at ways of improving the experience it offers our customers. Three carries 42% of the UK's mobile data. Its network covers 98% of the UK population. Three UK is a member of CK Hutchison Holdings which also has investments in mobile operations in Australia, Austria, Denmark, Hong Kong, Indonesia, Ireland, Italy, Macau, Sri Lanka, Sweden and Vietnam. Three employs over 4,400 people across its offices in Maidenhead, Glasgow and Reading and its 347 retail stores. For more information visit www.three.co.uk.

About Mony Kochupillai

Mony was educated to degree level in Engineering and MBA in marketing and has over twenty years of business, marketing and technical experience. He lived and worked in the UK, Italy and India for lean start-ups as well as large multi-national companies, developing leadership and expertise in business and technology strategy, mobile data economics and profitability, customer experience and perception, as well as go-to-market strategy and execution for major mobile internet and mobile network initiatives, possesses an exceptional understanding of mobile operator business, strategic thinker and passionate with strong communication and influencing skills.

Telefonica Argentina

Leaving the silos behind to embrace end-to-end optimization

A conversation with Adrian Di Meo, CTO

Monica Paolini: As part of our report on optimizing network performance to enhance the QoE, we are talking to Adrian Di Meo, CTO at Telefonica Argentina.

Adrian, can you give us an introduction and tell us what your role at Telefonica Argentina is?

Adrian Di Meo: I have been the CTO of Telefonica Argentina for one year. Previously, I was the CTO of O2 UK, starting in 2011. Previous to that, I was the VP of Access, Transport and Network Coordination at O2 in Germany. That's where I'm coming from. I am Argentinean, but I was lucky to work both in Europe and in Argentina.

Monica: That gives you a lot of experience in many different countries. If we look at the different requirements in different markets, what do you think is different in the way you optimize the core and you manage traffic in Europe and in a market like Argentina?

Adrian: If you will allow me, I would start from the core optimization in general. We are now in the data era. We are managing much more data traffic than we managed before, in the voice era.

We in the industry have been planning networks and managing traffic as if we were in silos. The core guys plan for the core, and then interface with the access. The access guys optimize and do the planning for the access, and then they interface with the core.

Now, in the data era, to be capable of measuring the quality of experience of customers, you need to see the network end to end. You need to plan the network end to end. It's not easy, because each part of a network has its own type of KPIs, measurement systems, and so on.

To be fair, we don't have a perfect end-to-end measurement system. We are very good at measuring the performance of the access. We are very good at measuring the performance of the core. As an industry, we have many years of experience in that. But we don't have something that shows us clearly the quality of experience end to end from the customer point of view.

This is what is changing. It is a must, as an industry, to improve. This is part of the change that we are going through in Telefonica and in the industry.

Now let me go back to the differences among Germany, UK, and Latin America. Germans are very process oriented. The British are normally thinking ahead and trying to have a vision before starting to work. In Latin America, we are a kind of a mix between Spanish and Italian, trying to get the most out of what we have.

In particular, UK is the country that is well ahead in the revolution in the digital, data market. They have really seen this end to end. They are, for example, working in adding resiliency to the core, so that the data core is really on at all times. With this approach that they call DAB, data in a box, they can pool all the resources from the core to be sure that if something isn't working, they have another part of the pool to use. They are very conscious of the measurement of the end-to-end quality with different pools, where there are different vendors.

Based on that, they do not just optimize the radio access, nor the capacity of the different components of the network. Instead, they measure the speed of the DNS resolution of an address, the bandwidth and all the connectivity to the outside on the Internet, and even some services that are inside the network.

I will say that this is the case of O2 in Germany, too, where they still have a concept of centralized core, but that is going to be moving.

In Argentina now, with the deployment of 4G, customers are starting to use data. They don't complain any more about "I have no data." Now, they complain about "It's buffering." Or they even measure the ping for some application and they say, "I have more than x ms latency in the application." The conversation of the customers has changed. For them it is not just about whether the access is up and running.

You need to have a look deep inside of your proxy platform and see if you are caching or not. In an optimized network, you should have no buffering for the service that doesn't need to have buffer. And you should have a quick DNS resolution of the

address, as well as the lowest latency that you can achieve.

Monica: Is QoE changing because the subscriber experience is changing? Even if the network performance is the same, the customer may perceive it differently.

Adrian: Yes. We see quality of service as a way of monetizing data. Data growth is exponential, and revenues growth in the industry is not exponential. The decrease in costs is not exponential, and its pace is too slow to counter the rate of increase in traffic.

We always try to see quality of service as a way of monetizing, by using things like customer segmentation. Quality of service has become a must to offer a QoE that can keep your customer on the network. You need to have end-to-end quality of service to be sure that the customer who is streaming video has enough capacity for a good experience. If you have VoLTE or you have a voice service in an LTE network, you have to be able to keep a good voice experience and avoid deteriorating the quality of the call.

Monica: There is a constant trend toward paying more attention to the QoE, what the subscriber sees. But as you said before, that's difficult. It's an end-to-end process. You can measure how different elements perform in your network, but to measure QoE is more difficult, because it is a more subjective measure. How do you deal with that?

Adrian: We're using different types of probes and DPIs to get to a proper measure of QoE. We are still not there. Now, we're making a trial with two big companies, trying to use probes and trying to differentiate what the customers are doing and what their experience is with different applications – for instance, with streaming versus social media.

There is another challenge that is coming. Most of the traffic is becoming encrypted, so you do not know what type of traffic the network is transmitting. If you are on Facebook and you are just reading posts or you are watching a video, from the network point of view, both are nearly the same. You cannot differentiate it, so to measure the different quality of experience for each applications is difficult.

Of course, you can guess from the volume of traffic, and you can use some kind of algorithm to see if you are watching a video, and to guarantee a certain throughput. But it's really complex. To be fair, I don't think that we're there yet.

If we're going to seriously measure quality of service or the quality of experience, you need to find a way to take all the raw data from the network. Based on that, you need to work it out, which is the quality of experience.

Monica: Do you think it is crucial to have information collected from the handset itself, or can you just collect it just from the network? Do you think you need to have emulated data, or can you just use passive testing or passive monitoring of a network?

Adrian: Putting applications on the device is, from the technical point of view, the easiest way to get the proper measurement of QoE. But who wants to have this application on their device?

In the last five years or seven years, many companies have come up with the idea of putting some application on the device. Truly, I don't see any that has been successful. It's the right place, but no one is willing to accept that. Google and Facebook track everything you do. But subscribers are not willing to let the telco do the same for optimization purposes. That's a matter of fact. We might like it or not, but this is how it is.

Monica: Is it valuable, then, to simulate data, as if it was coming from specific handsets? This approach allows operators to try specific use cases.

Adrian: There are two options. One is field testing – the option that most people in the industry use to test for a particular place or a particular event. You put your own device there and test things like file transfers or even voice calls, and you can measure from there the QoE for this particular device in this particular position.

The other is testing with probes. But I think that we need to move beyond the probes that you put on the network, from the different probes that you put across the network to try to get information from there.

There is a big game to play from our big-data experts. In addition to the probe measurements, we need to have KPIs for the QoE that are different from those that we normally have in the network access, such as dropped calls.

Monica: That's a complex issue. As a mobile operator, you have access to large data sets on network performance. In fact, you have too much data, so you need to have a big-data approach to sort through the massive amount of data. But isn't it difficult to find what's relevant in that huge amount of data?

Adrian: We're working on that. For instance, we have some good use cases to try to find out if a customer with a 4G device is on 3G, and why he is on 3G instead of 4G, and how he can have a better experience. We have some use cases on that. But I think there is a long way to go before we find a good solution to this challenge.

Monica: Traditionally, when you monitor the network or you try to figure out what's going on, you look at historical data. Knowing the QoE of a subscriber a week ago does not really help you in optimizing the network, because you want to do it in real time. How important you think it is to move to real time? And what does real time mean? Is it an hour, one day, one minute?

Adrian: We need real-time optimization for big events. There, real time is 15 or 10 minutes for big events – i.e., when you have a football match in Argentina, a concert, or during the Olympics in London in 2012.

I believe that for the overall network, a 24-hour window is good enough. When we get to the point that we are optimizing the network every 24 hours with the measurement of the day before, then we can start to think about a lower time resolution.

I love the example of the highway. In the morning people come to downtown, and in the evening they go out of downtown. We still have a long way to go to optimize a network every 24 hours and to make some decision on the basis of that, as we would do for traffic moving on the highway. But we will get there.

Monica: Right now, the first target is the 24 hours, making sure you can capture those data variations. Is that good enough for special events?

Adrian: In that case, you need to have a 10- to 15-minute window, no more than that.

Monica: What do you do in that case? You have a stadium, and you're never going to have enough capacity to meet whatever everybody wants to do. What do you do there?

Adrian: It depends. I and my team covered the Olympic stadium for the opening ceremony. You could watch BBC live from the stadium.

You can have enough capacity. We put around 40 sectors in the Olympic stadium. You can have capacity. But sometimes, you have the capacity and it's not optimized, and you have problems with an unbalanced network. This is why you need a 10- to 15-minute window.

What you can do there is adjust some parameters. For instance, how many simultaneous customers can you keep live? How much do you have in each carrier? How much do you put on 3G? Do you prioritize voice over data? There are a few things you can do.

This does not mean that in every special event you're going to have fantastic service. But if you plan and you put in enough capacity, you can. If you don't have enough capacity, at least you can keep the network at a level at which most of the capacity is available.

Another thing that may happen, depending on how you configured your network, is that you get to a certain point at which you cannot keep the maximum capacity. Capacity goes down because of collision, reconnection attempts, and so on. In this scenario, you may have to use parameter settings that are not optimal for a regular situation.

You need to know when you trigger these parameters to guarantee the maximum experience that you can offer during a period with low traffic, a period with peak traffic, and when traffic goes back to normal levels.

Monica: You have to protect the network from degradation.

Adrian: And to keep the maximum throughput that you can offer from a BTS.

Monica: You're trying to optimize basically the RAN performance, but you do it also from the core. The core and the RAN seem to be much closer to each other than they were in the past, when core and RAN were isolated from each other. How important is it, in terms of optimizing the network end to end, to have the core and RAN talk to each other?

Adrian: As I mentioned at the beginning, the end-to-end view is a must. To be fair, when you have this type of peak of traffic, the core is not something that you are touching in real time. What you really affect in real time is the RAN; it's not the core. But you need to guarantee that the end-to-end view is there.

Caching is a good example of this. How much caching do you need to do? In a country such as Argentina, where most of the OTT services are based in the USA and you are 10,000 km from there, the caching policy is very important. Caching is part of the core. If you don't have a proper core, even if your RAN is great, you will not see proper service. This is one example why end-to-end is

critical. But the truth is that you cannot define the caching policy in real time. This must be planned in advance, and must be fully aligned with your end-to-end view of the capacity. For data services, it's critical.

Monica: I guess that you also need to set the policies and parameters in the core in such a way that they are sensitive to what happens both in the RAN and on the subscriber's end. What that means is that the marketing and the operations are coming together with the optimization of the core. It's different groups within the operator's company – RAN, core, marketing – all different groups that traditionally have been more in silos and that now need to work together.

Adrian: That's true. For instance, the caching policy and the decision of which type of service do we want to cache or not is an area of big discussion with our marketing colleagues.

Monica: How you do deal with that?

Adrian: We sit down. We have conversations. For me, there is a clear discussion between the OTT model and the telco model that I cannot go into now.

Monica: Let me ask you a final question. What do you see changing over the next five years?

Adrian: First of all, I don't want to see us running four or five different access networks – 2G, 3G, 4G and 5G. This does not mean that it doesn't matter what access customers have to the network. We are moving to a seamless type of connectivity to the customer.

Within three to five years, we need to simplify the access network, because the current network architecture is not sustainable. It's not sustainable from a landlord point of view, from the power point of view, from the capacity point of view. That's one big change.

The second change is we still don't know what it means to have traffic that grows 100% year over year. We still don't know. This is going to change a lot of things: the way we watch TV, the way we consume, and the way we buy.

Monica: Mobile is becoming the primary way for people to access the Internet to stay connected. In the past it was not like that.

Adrian: I see my kids. They are not watching broadcasting TV anymore. They are watching all the time on the device, using streaming from different providers. They even buy from mobile devices. They see something, they go to the mobile, they choose it, and they send it to me via any social network. In five years, this is going to be huge.

But I'm sure from our point of view, we need to put bigger and bigger pipes. We need to guarantee ubiquitous, continuous connectivity, because people want to be connected all the time.

Monica: It's a big opportunity, but also it's a lot work that you have ahead.

Adrian: No doubt. Look at it. It took nearly 20 years to move beyond 2G, and for 3G to explode. In just 3 to 4 years, 4G will cover most of the country. With the quality of networks, and with the ecosystem of devices needed to forge ahead to 3G, it took 3G nearly 10 years to get to a similar stage. The speed, the pace is increasing and is truly exponential.

About Telefonica

Telefónica is one of the largest telecommunications companies in the world in terms of market capitalization and number of customers. With its best in class mobile, fixed and broadband networks, and innovative portfolio of digital solutions, Telefónica is transforming itself into a 'Digital Telco', a company that will be even better placed to meet the needs of its customers and capture new revenue growth. The company has a significant presence in 21 countries and a customer base that amounts more than 341 million accesses around the world. Telefónica has a strong presence in Spain, Europe and Latin America, where the company focuses an important part of its growth strategy. Telefónica is a 100% listed company, with 1.5 million direct shareholders. Its capital traded on the continuous market on the Spanish Stock Exchanges (Madrid, Barcelona, Bilbao and Valencia) and on those of London, New York, Lima and Buenos Aires.

About Adrian Di Meo

Adrian Di Meo is the Chief Technology Officer at Telefónica Argentina, where he leads working teams managing the mobile (GSM/UMTS/4G) and fixed networks, with a focus on network strategic definition and implementation, network roll out plan and operation of the network. He was responsible for the 4G service launch in Argentina, with more than 1.000 enodeB live in 6 months. Previously, he was the CTO at Telefónica UK, where he was responsible for the Olympic program for o2 in London, and for the network-sharing agreement with Vodafone in the UK. He also worked at o2 Germany and Movistar. He holds a degree in electronic engineering from the Universidad Tecnológica Nacional, and an MBA from the Universidad Argentina de la Empresa UADE.

Glossary

3G	Third generation
3GPP	Third Generation Partnership Project
4G	Fourth generation
4K	4,000 pixels [UHD resolution]
5G	Fifth generation
8K	8,000 pixels [UHD resolution]
AAA	Authentication, authorization and accounting
ABM	Adaptive bit rate management
ABR	Adaptive bit rate
ARPU	Average revenue per user
BSS	Business support system
BTS	Base transceiver station
CAGR	Compound annual growth rate
CDN	Content delivery network
CDR	Charging data record
C-RAN	Cloud RAN
CRM	Customer relationship management
CSCF	Call Session Control Function
DAB	Data in a box
DAS	Distributed antenna system
DL	Downlink
DNS	Domain name system
DPI	Deep packet inspection
DSL	Digital subscriber line
E2E	End to end
eNB	Evolved NodeB
eNodeB	Evolved NodeB
EPC	Evolved Packet Core
eSIM	Embedded Subscriber Identity Module
ETSI	European Telecommunications Standards Institute
FM	Fault management
GGSN	Gateway GPRS support node
GPRS	General packet radio service
GSM	Global System for Mobile Communications
GW	Gateway
HD	High-definition
HetNet	Heterogeneous network
HSS	Home subscriber server
HTTP	Hypertext Transfer Protocol
HTTPS	HTTP Service
I-CSCF	Interrogating CSCF
IMS	IP Multimedia Subsystem
IoT	Internet of things
IP	Internet Protocol
ISP	Internet service provider
KPI	Key performance indicator
LTE	Long Term Evolution
M2M	Machine to machine
MME	Mobility management entity
MOS	Mean opinion score
MRF	Media Resource Function
MTTR	Mean time to repair
MVNO	Mobile virtual network operator
NAT	Network address translation
NFV	Network Functions Virtualization
NPS	Net promoter score
O&M	Operations and management
OCS	Online charging system
OSS	Operations support systems
OTT	Over the top
PCRF	Policy and charging rules function
P-CSCF	Proxy CSCF
PE	Provider Edge
PESQ	Perceptual Evaluation of Speech Quality
PGW	Packet gateway
PIM	Passive intermodulation
PRB	Physical resource block
QCI	QoS class indicator
QoE	Quality of experience
QoS	Quality of service
RAN	Radio access network
RF	Radio frequency
R-factor	Rating factor
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
SaaS	Software as a service
S-CSCF	Serving CSCF
SD	Secure Digital
SDN	Software-defined networking
SEG	Security gateway
SGSN	Serving GPRS support node
SGW	Serving gateway
SIP	Session initiation protocol
SLA	Service level agreement
SON	Self-organizing network
TAS	Telephony Application Server
TCP	Transmission Control Protocol
UE	User equipment
UHD	Ultra-high definition
UI	User interface
UMTS	Universal Mobile Telecommunications System
vEPC	Virtual EPC
VoIP	Voice over Internet Protocol
VoLTE	Voice over LTE
vRAN	Virtual RAN
WAP	Wireless Application Protocol

About RCR Wireless News

Since 1982, RCR Wireless News has been providing wireless and mobile industry news, insights, and analysis to industry and enterprise professionals, decision makers, policy makers, analysts and investors. Our mission is to connect, globally and locally, mobile technology professionals and companies online, in person, in print and now on video. Our dedication to editorial excellence coupled with one of the industry's most comprehensive industry databases and digital networks leads readers and advertisers to consistently choose RCR Wireless News over other industry publications.

About Senza Fili

Senza Fili provides advisory support on wireless data technologies and services. At Senza Fili we have in-depth expertise in financial modelling, market forecasts and research, white paper preparation, business plan support, RFP preparation and management, due diligence, and training. Our client base is international and spans the entire value chain: clients include wireline, fixed wireless and mobile operators, enterprises and other vertical players, vendors, system integrators, investors, regulators, and industry associations.

We provide a bridge between technologies and services, helping our clients assess established and emerging technologies, leverage these technologies to support new or existing services, and build solid, profitable business models. Independent advice, a strong quantitative orientation, and an international perspective are the hallmarks of our work. For additional information, visit www.senzafiliconsulting.com or contact us at info@senzafiliconsulting.com or +1 425 657 4991.

About the author

Monica Paolini, PhD, is the founder and president of Senza Fili. She is an expert in wireless technologies and has helped clients worldwide to understand technology and customer requirements, evaluate business plan opportunities, market their services and products, and estimate the market size and revenue opportunity of new and established wireless technologies. She has frequently been invited to give presentations at conferences and has written several reports and articles on wireless broadband technologies. She has a PhD in cognitive science from the University of California, San Diego (US), an MBA from the University of Oxford (UK), and a BA/MA in philosophy from the University of Bologna (Italy). You can search her at monica.paolini@senzafiliconsulting.com.