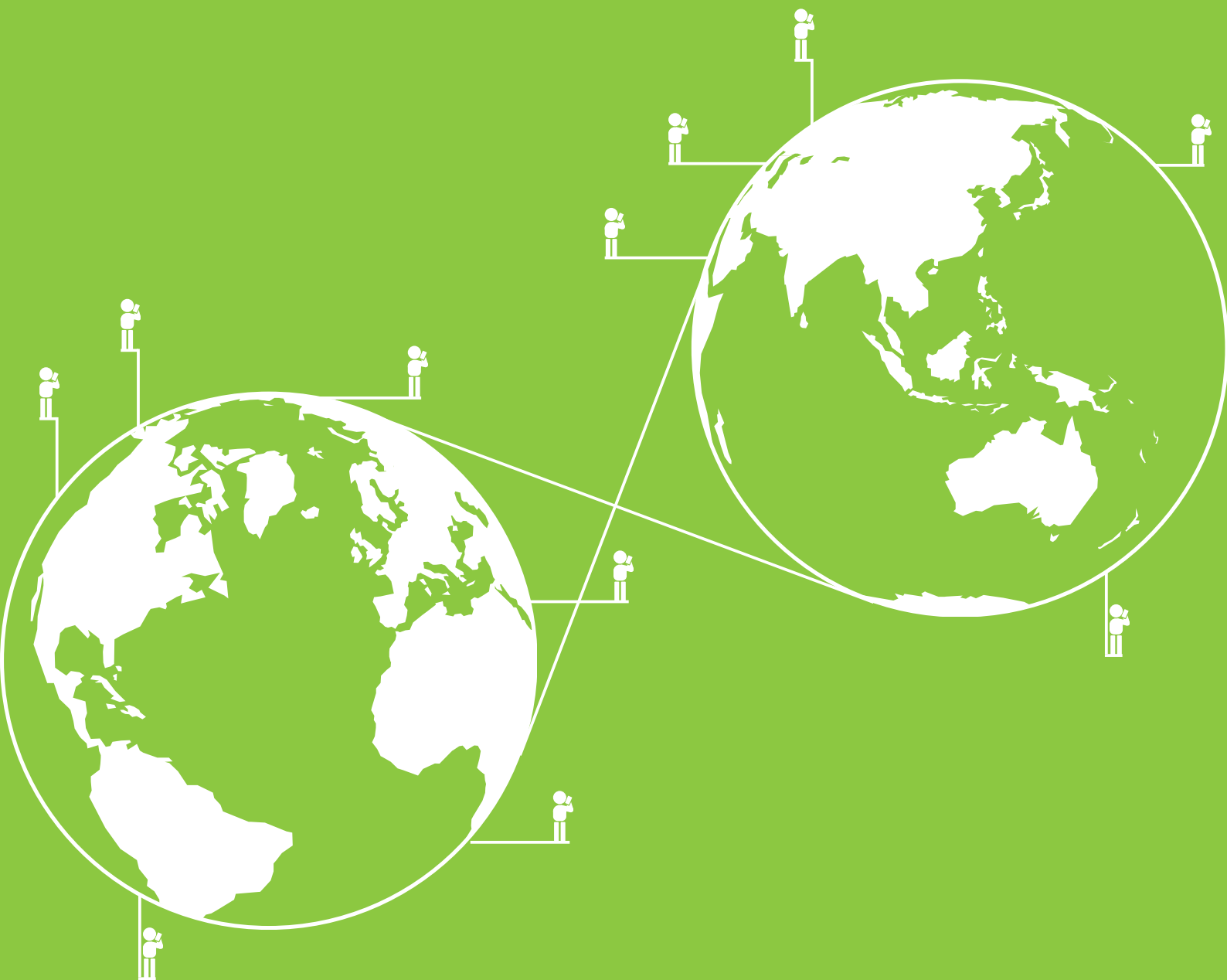


Re-defining Quality-based Routing for Voice Services

2019 Industry White Paper



A new quality paradigm for routing voice traffic	3
Solutions overview	5
Voipfuture Qrystal quality monitoring features	7
NetNumber TITAN CRE routing features	8
Traffic rerouting based on media quality KPIs	9
Summary	12



A new quality paradigm for routing voice traffic

It is well known that the voice service, both for domestic communications and for the international wholesale service, has been subject to key and deep changes that involve:

- **TECHNICAL PLATFORM:** The IP migration, at least in the developed regions, can be considered almost completed. In addition, the virtualisation of the telecom platform has started generating issues as well as opportunities.
- **STAKEHOLDERS:** The industry players are no more limited to domestic Telcos (FNOs/MNOs) and International Carriers (e.g. IPX Providers), each class mutually playing the role of Customer and Providers. Over The Top (OTT) Providers emerged as alternative service providers and sources of traffic to be terminated and strongly modified the industry scenario in terms of service performance, customer experience and pricing.
- **SERVICES:** In addition to traditional G.711 encoded TDM traffic a variety of traffic types has to be properly managed by FNOs/MNOs and, hence, by IPX Providers: IP traffic, IMS/VoLTE traffic with a variety of narrowband (NB) and wideband codecs (WB) generated by business and consumer customers. In this regard, the provision of high-quality communications, via WB codecs (e.g. AMR-WB) and full-band codecs in the short term (e.g. EVS), has to be considered a primary IMS/VoLTE objective.
- **ECONOMIC CONSTRAINTS:** Domestic Telcos and International Carriers have been heavily investing in LTE/IMS/VoLTE platforms. Additional CAPEX are required for the forthcoming 5G transition. It is mandatory for them to achieve an investment pay-off maximizing the quality performance, drawing more traffic and optimizing the outpayment expenses

- **REQUIREMENTS:** The heterogeneous scenario shortly described above calls for different quality requirements for the various type of traffic to be managed and hence different routing policies to be implemented. It results in a differentiated network performance monitoring of the valuable traffic streams (e.g. WB/SWB/FB coded calls) guaranteeing, for example, a proper routing avoiding transcoding which generates quality degradation.

Traditionally, the concept of quality has been mainly tied to the voice call setup performance since this was relevant in the TDM world. It has been linked to the performance of the very well-known parameters ASR (Answer Seizure Ratio), NER (Network Effectiveness Ratio) and ALOC (Average Length of Call) which have been considered as reference KPIs in the industry as well as in the international standardisation bodies.

This paper describes the solution conceived, designed and implemented by Voipfuture and NetNumber for domestic Telcos and International Carriers (together identified as Communications Service Providers, CSPs) which achieves three basic objectives:

1. To extend the quality monitoring functions to the media plane providing
 - effective, real time and accurate measurement of a wide set of KPIs such as MOS, Good Minute Ratio etc., (see Voipfuture white paper “Why next generation voice services need next generation KPIs”)
2. To provide CSPs with the capability to automatically reroute the traffic in real time (and restore the original configuration) in accordance to the policies set by them for the flow of traffic they consider relevant
3. To guarantee premium services (e.g. for VoLTE traffic with WB codecs, for corporate traffic etc.) an end-to-end path with high performance

The proposed solution does not overlap or substitute the role of OSS/BSS systems or Least Cost Routing (LCR) systems devoted to select the best commercial route for a given traffic destination, but aims to complement that routing providing CSPs with the capability to react in real time to media and signaling degradation.

The business target is to provide domestic Telcos and International Carriers with a real time cost-effective automatic rerouting capability which:

- Improves the customer experience enhancing the overall quality performance of the network
- Attracts additional flows of traffic (e.g. from OTT Providers)
- Selects better downstream operators (domestic and/or Int. Carrier)
- Optimizes termination costs in the Int. wholesale industry
- Allows an improved SLA management and reduces customers’ complaints

Solutions Overview

The proposed service architecture exploits the basic features and capabilities of Voipfuture's Qrystal quality assurance solution and the flexible routing capabilities of NetNumber's TITAN Central Routing Engine (CRE).

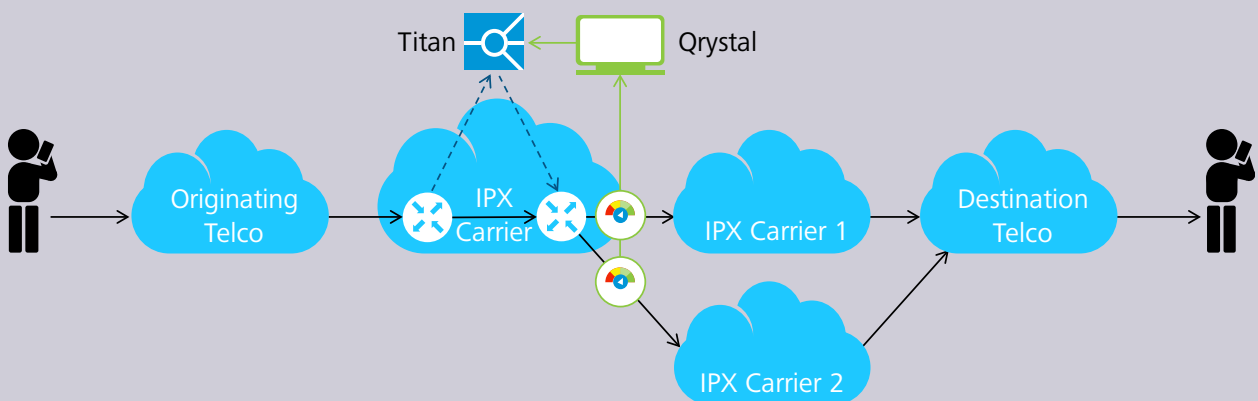
The combined solution allows:

- An integrated real time call monitoring function for user plane (media) and control plane (signaling)
 - It is non-intrusive and uses a unique time slicing mechanism, providing a wide list of information (codec, IP routing information, packet loss, jitter, etc.)
 - Depending where a CSP locates the Qrystal probe the internal network performance and/or the performance of the downstream network can be monitored
- An integrated measurement of signaling KPIs (ASR, NER, ALOC), Mean Opinion

Score (MOS) and a large set of media KPIs (e.g. Good Minute Ratio, Critical Stream Ratio...)

- CSPs to design flexible rerouting/restoration policies without any constraints in terms of:
 - KPIs selection (i.e. which KPI to be used for triggering rerouting)
 - Timing (i.e. after which time interval to activate / deactivate) the policies
- An easy, cost effective and flexible implementation of the policies via TITAN CRE being inherently multiprotocol (e.g. SIP, SS7, Enum/DNS, Diameter)
- To maintain the original high definition codec across downstream networks in compliance with the LCR routing tables
- The setting of an alarm to the CSP's NOC informing that an automatic rerouting took place (or the original configuration has been restored)
- The transmission of the poor network

Voipfuture / Netnumber architecture for quality-based rerouting

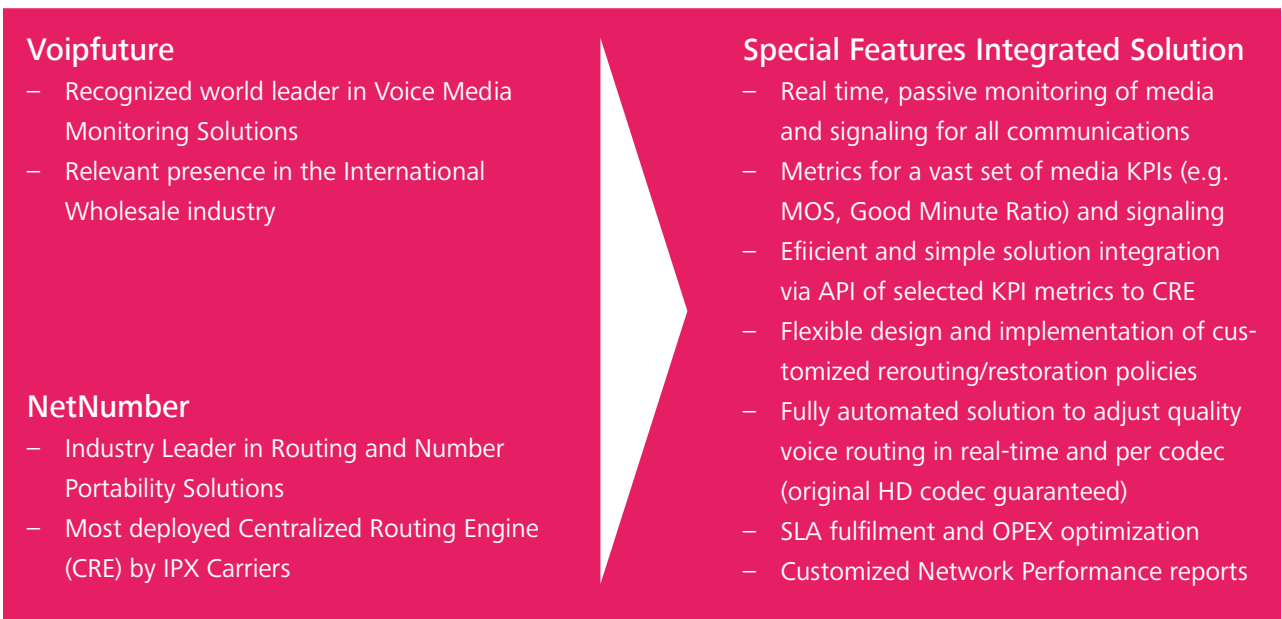


- performance (internally or downstream) to the OSS/BSS or LCR BB systems, depending on CSP's requirement
- To complement the real time information with off-line reports of overall network performance; the large set of collected data allows Qrystal to accurately calculate a very large set of KPIs whose measure can be properly "assembled" in reports according to the Customer's requirements.

The final aim is to guarantee the best voice traffic routing guaranteeing Customers' satisfaction meeting SLA targets and optimising the OPEX for traffic delivery.

As pointed out, the proposed solution can simply be added in networks where the Net-Number TITAN platform is already deployed as the Centralized Routing Engine (CRE). In other situations, TITAN provides easy integration in the signaling control layer and offers additional value as a carrier-grade platform as explained in the box devoted to introduce the TITAN CRE features.

Key characteristics of the joint solution



Voipfuture Qrystal Quality Monitoring Features

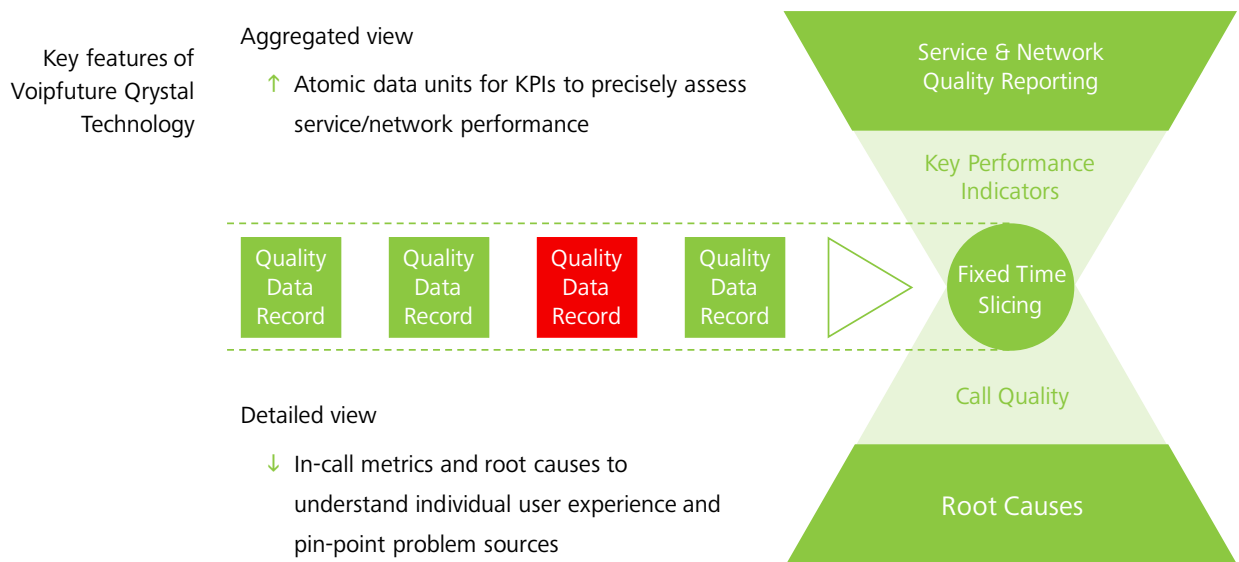
Voipfuture's mission is to unfold the full potential of VoIP analytics enabling CSPs to take control over every aspect of voice service quality. Qrystal is a complete vertical solution delivering business critical information to all organizational units of a CSP.

Qrystal is a passive, mid-point monitoring solution, i.e. it measures live voice over IP (VoIP) traffic in real-time. It consists of distributed Qrystal Probes and the central Qrystal Manager. Non-intrusive test access ports (TAPs) connect the probes to the network.

Qrystal Probes determine the signaling and media stream characteristics at multiple points in the network. The probes analyze all SIP, RTP, and RTCP packets on a link in real time and generate unique metric data for further processing by the Qrystal Manager. Voipfuture's patented fixed time slicing technology creates quality data records (QDR), containing highly condensed information for every 5-second segment of an RTP stream. QDRs hold several hundred values, ratios, KPIs, and automatic root cause indicators. The QDRs complement xDRs, which summarize a call's signaling characteristics at a monitoring point. Qrystal correlates all QDR and xDR related to a call to create quality-enriched call detail records (CDR).

The Qrystal Manager is at the heart of the solution, featuring a Telco grade data warehouse. It collects all metric data from the Qrystal Probes for post-processing and aggregation. KPIs, quality-enriched CDRs, and other data are stored in the data warehouse. The solution conforms to ITU-T P.564, meaning its MOS estimates are accurate and comparable to results from PESQ and other conformant systems. There are no dependencies on the network configuration or on data from 3rd party sources.

A variety of interfaces to network management, customer experience and other OSS and BSS systems, facilitates its integration with existing IT infrastructure. These system characteristics open up a range of use cases for fixed and mobile network operators, wholesale carriers and enterprises. Qrystal covers the entire spectrum of functionality needed for effective voice quality assurance. It can be used for monitoring, troubleshooting, reporting and alarming for VoLTE, VoWifi, IMS, NGN and VoIP Wholesale services.



NetNumber TITAN CRE Routing Features

In the heterogenous voice service scenario described above, the inter-network routing complexity increases, and operating costs continue to rise as new functionality is added into the network. Further routing and operational complexities are introduced with the coexistence of existing networks with NFV-based platforms.

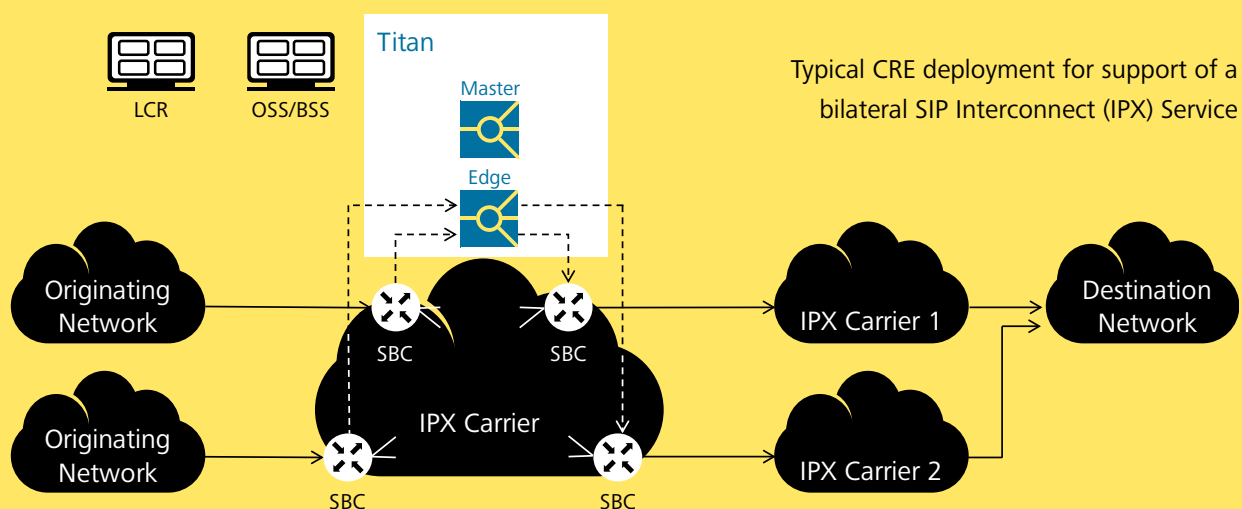
Some service providers attempt to improve the quality of service of their traffic with complicated multi-step process orchestration via the OSS/BSS Layer. Based on CDR data and/or monitoring of the data streams at the data network layer, CSPs can use this feedback to change their routing policies; however, via OSS/BSS orchestration processes the solution is not real-time (and thus not very effective), labor intensive, and error prone.

The NetNumber Centralized Routing Engine (CRE) application on the TITAN platform removes this entire burden with its centralized provisioning of routing data and centralized routing control. This dramatically simplifies routing across networks and technologies. TITAN dynamically chains customer-defined services and delivers exactly what the CSP wants. In a major architectural shift, it can now realize a substantially lower cost for their networks due to the removal of all internal signaling burdens.

The unique TITAN Master/Edge architecture provides a carrier grade distributed platform. The TITAN Master is the central point of provisioning and OSS/BSS integration for a TITAN server deployment and is responsible for the replication of its database to a constellation of TITAN Edge servers/instances. The TITAN Edge provides multi-protocol query support and support of multiple application services. The TITAN platform is highly scalable and can be deployed either on standard hardware or in various virtualized deployment modes.

In the figure given below a basic TITAN CRE routing function is shown in a multiprotocol environment. TITAN CRE monitors the state of the outgoing trunk groups when TITAN detects that a link is down, all calls to affected destinations will be rerouted via alternative trunk groups. As a result, it will massively enhance the Answer Seizure Rate (ASR) of the calls in the CSP's network while also avoiding any revenue loss caused by down links.

With the real time quality data provided by Voipfuture's Qrystal, the described routing capabilities are enhanced to cover additional use cases. For example, voice degradation may have ceased before the routing control changes initiated by the OSS/BSS layer take effect. A real-time acting routing solution is enabled by invoking dynamic policy routing adjustments in the signaling control layer based on the actual voice quality route assessments in the media layer. This eliminates OSS/BBS complexity, achieves shorter turnaround delays for repairs of voice quality imperfections, and creates savings in IT/process and operational costs.



Traffic rerouting based on media quality KPIs

It is worth providing a detailed example of how a CSP can exploit the operational advantages of the combined Voipfuture / NetNumber proposal.

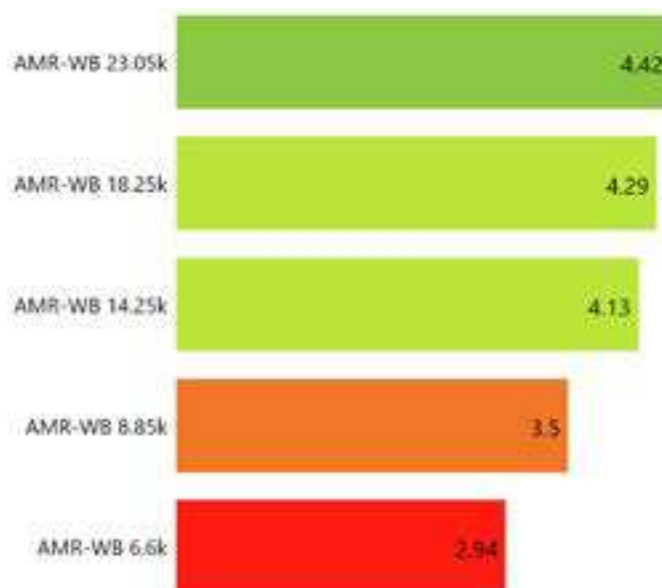
The analysis starts with the KPI selection i.e. which quality parameter has to be monitored in order to effectively trigger a traffic rerouting outlining the impact of WB codecs on the quality level perceived by the talkers. Then it discusses how the related policy can be implemented in terms of thresholds which trigger the traffic rerouting and thresholds which activate the restoration of the original route. The next step gives some technical outlines on the communication link between Crystal Manager and TITAN Master, the alarms generated towards NOC, and optionally the information send to the CSP's routing system. The possibility to receive a wide set of off-line reports with an exhaustive description of the overall network performance is dealt with in the last step.

Step1- KPI selection

As described in the previous sections, the IP migration is now a reality and a fully IP world creates new challenges in terms of voice quality of service having to face packet loss and jitter. In addition, the IMS/VoLTE services have enabled a variety of high definition codecs (e.g. from AMR-WB to EVS) which make it much more difficult to estimate the real user experience. Indeed, modern codecs offer several modes and bit rates, which allow to adapt voice transmission to the available

channel conditions and bandwidths. Such codec mode changes can occur on a packet-by-packet basis without any indication in the SIP signaling. VoIP monitoring thus needs to be aware of such mode changes, because the user experience strongly depends on the bit rate.

A graphical proof of the impact of codecs on the perceived quality is given by the graphic below which shows the MOS depending on different AMR-Wideband bitrates.



AMR-WB codec modes and their maximum quality (MOS)

The 8.85kb bitrate of the AMR-WB codec can achieve a MOS of 3.51 on the wideband scale. In the context of traditional telephony this would be nearly perfect quality since, for

comparison, the G.711 MOS can reach 3.69 on the wideband scale. In a WB environment the listeners' expectation is higher and the same AMR-WB bitrate will only be considered "acceptable". In addition, the 23.05kb bitrate of AMR-WB has a MOS of 4.42, which most listeners will consider to be "excellent". This huge variability leads to the need for Service Providers to monitor the offered media quality.

All the above considerations lead to extending the traditional signaling-based KPIs with KPIs taking into account the media quality, which capture the in-call user experience. For the sake of this example MOS is considered since it is widely known but it is stressed that the Int. Carrier can select any quality parameter (or combination of parameters) it likes depending on signaling messages as well as on media performance without any constraints.

Step 2 - Policy definition

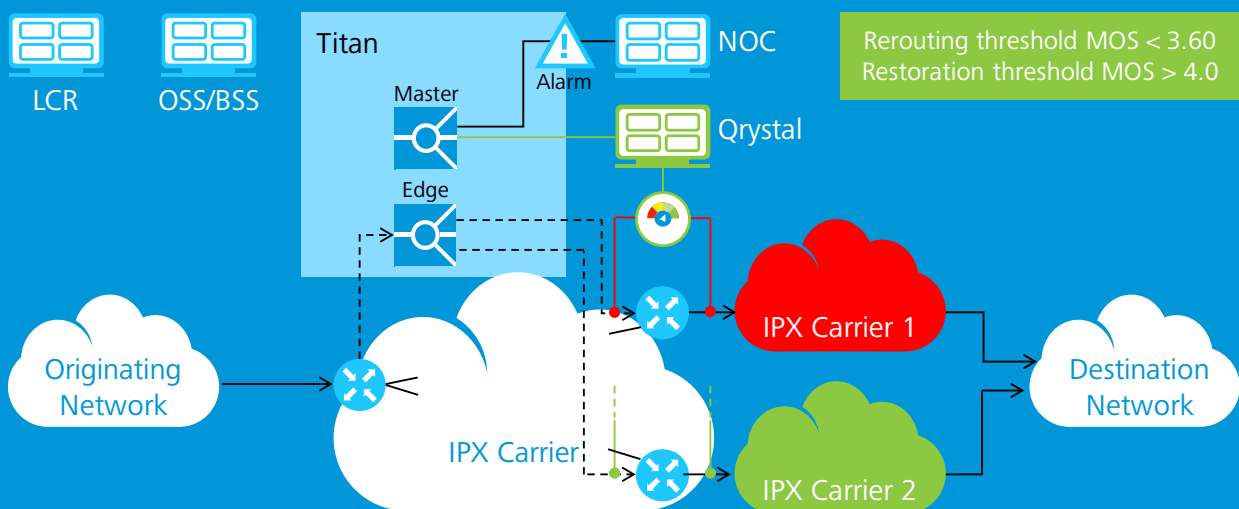
Following the ITU MOS value classification in 5 different bands, as an example, it can be set that the rerouting function is triggered when the MOS values drop below a specified quali-

ty, e.g. below MOS 3.6, for more than consecutive 15 minutes. MOS 3.6 is chosen in this example, because according to ITU-T G.107 most users will be dissatisfied below this threshold.

The new network destination of the rerouted traffic can be selected as the second choice identified by the routing system and regularly fed into the TITAN CRE or according to the CSP requirement (see graphic above).

It is desirable that the original first choice configuration is properly restored, while avoiding fluctuations of the routing around a threshold value. A hysteresis cycle has to be considered, e.g. the CSP can specify that if the MOS value is higher than 4.0 for 15 consecutive minutes the traffic goes back to the original network. According to ITU-T G.107 this threshold will ensure that most users are satisfied.

The rerouting policy can be designed and activated for a whole trunk (i.e. a stream of traffic containing a mix of NB and WB codecs) or, depending on CSP's requirement, only on a portion, for example only on calls using high definition voice codecs in order to guarantee the best performance for the most valuable traffic. Of course, the rerouting action is performed on the new calls without any cut-off of the already established communications.



Step 3 - System integration and alarm generation

The various systems mentioned above have to be properly integrated to guarantee the requested level of response time, availability and flexibility in operation. Considering that both the Qrystal Manager and TITAN CRE are available as Virtual Machine, they can be easily installed as colocated VMs into the CSP's Virtual Infrastructure. The integration uses a unidirectional data feed from the Qrystal Manager towards the TITAN CRE with tested and certified standard APIs to guarantee flexibility in passing the measurements of the selected KPIs.

The generation of specific rerouting alarms from TITAN CRE towards NOC take place via the standard SNMP communication line.

The same approach has been applied to the communication from the domestic Telco's routing system / International Carrier's LCR system towards TITAN CRE. In this way, the routing system provides TITAN CRE with the routing rules and related routing policies. As additional option it is possible to include extra information in the CDR like the selected codec and rerouting details or implement a direct data feed that allows TITAN CRE to inform the routing system in real-time that a rerouting and/or restoration action took place.

Step 4 - Quality reports

Separate from the actual VoIP routing process, which takes place in near real time, Qrystal can also provide off-line data to document and report on the service quality and network performance over longer periods. Such reports can be used for more long-term assessment of network performance and/or to help CSPs to evaluate the quality of the upstream/downstream networks.

Exploiting the correlation function performed by Qrystal Manager for all QDRs and XDRs related to a call, reports provide:

- An end-to-end view
- Signaling and media quality by origination, destination, codecs (NB, WB, SWB, FB) and other dimension
- Trunk statistics, which capture the quality at a specific point in the network, i.e. typically at an interconnection point

The reports can make use of all standard signaling KPIs (ASR, ALOC, etc.) as well as the unique Voipfuture media plane KPIs, such as Good Minute Ratio and Critical Minute Ratio. Raw report data can be exported to the CSP's own reporting engine or it can be visualized using a Voipfuture-supplied business intelligence tool. Numerous available templates provide guidance for creating custom reports meeting the CSP's requirements. The large set of collected data allows Voipfuture Qrystal to provide a comprehensive and effective view on all aspects of VoIP service quality.

