A Holistic Approach to VoIP Network Call Quality Management & Troubleshooting

Key Tools for Service Provider Control of Quality of Service SLAs in their VoIP Network

February, 2011

Table of Contents

List of Figures

References

[RFC3550] IETF RFC3550, RTP: A Transport Protocol for Real-Time Applications

[P.800] ITU-T P.800, Methods for objective and subjective assessment of quality

[G.107] ITU-T G.107, A computational model for use in transmission planning

[G.109] ITU-T G.109, Definition of categories of speech transmission quality

[RFC3611] IETF RFC3611, RTP Control Protocol Extended Reports (RTCP XR)

[RFC5424] IETF RFC5424, The Syslog Protocol

[P.862] ITU-T P.862, Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs

[P.862.1] ITU-T P.862.1, Mapping function for transforming P.862 raw result scores to MOS-LQO

[P.800.1] ITU-T P.800.1, Mean Opinion Score (MOS) terminology

1 INTRODUCTION

In traditional circuit-switched networks, connections between telephony endpoints are associated with dedicated links for the duration of the call. It is, therefore, relatively simple to pinpoint any QoS (Quality of Service) issues through an analysis of the links traversed by the call. However, in Voice-over-IP (VoIP) networks, matters are little more complicated. The endpoints may be connected through a plethora of network components (such as routers, switches, access nodes, broadband modems etc.) and these linkages may change from one call to the next, or even within a call as routing pathways are updated because of network failures or other issues. IP networks, being packet-based, are also associated with degradations such as jitter and packet loss that would not be experienced by circuit-switched connections. Therefore, despite the undeniable advantages of VoIP over circuit-switched networks, the dynamic nature of IP links and their particular impairments make the task of the service provider in terms of performance monitoring and QoS management that much more challenging.

As far as the end customer is concerned, the quality of the calls that they make is obviously a key determinant of their satisfaction level with their telephony provider. Customers are also not concerned with whether the link is IP-based or circuit-switched, so it is the job of the network operator to provide a user experience for their VoIP customers (in terms of voice quality) that is on a par with, or exceeds, that of a circuit-switched network. In order to be able to do this, the operator cannot just rely on anecdotal evidence of call quality, nor can they merely focus on troubleshooting and correcting problems as they arise – as by then the customer is already likely to be dissatisfied. Instead, the service provider needs to take a proactive role in managing the quality offered by their network. They must be able to constantly monitor the health of the network and take remedial action when early warning signs arise. If and when customer problems are reported, they also need to be able to home in on the problem area rapidly and then resolve the issue. To accomplish all of this effectively, it is imperative that the operator has the right tools in place to do all of the following:

- Provide quantitative measures of call quality both during the call and at the end of each call. This will allow the operator to determine whether key Service Level Agreement (SLA) criteria are being met for their customers.
- Aggregated versions of these quality measures can also offer an insight into the general health of the network and alert the operator to any impending issues.

- When a specific customer complains, it is important to be able to accurately and quickly probe that connection and gather test call statistics that can shed light on where the problem may lie.

Therefore, these tools need to meet the twin, but divergent, objectives of (a) gathering data from thousands of calls and storing the information for future retrieval so individual problem cases can be addressed, while also (b) presenting that data in a coherent and succinct form so that network operations staff can maintain a birds-eye view of QoS within the whole network.

This is the challenge facing service providers today as the sizes of their VoIP networks grow and ad-hoc call quality management solutions are no longer viable. Instead, a holistic approach is required that builds in QoS gathering and reporting capabilities to each and every VoIP endpoint in their network, and then ties that into central management servers specifically dedicated to the accumulation and presentation of that data in a lucid fashion. This approach can be likened to a nervous system where the VoIP endpoints act as the 'eyes and ears' of the network and constantly gather information that is sent back to the nerve center, or brain, that processes that information into an intelligible form. Luckily, these tools exist today and this whitepaper will introduce the reader to what is possible with the latest in call quality management technology.

2 Call Quality Parameters

As discussed above, the telephony endpoints will need to collect call quality information. However, a key question remains: 'What parameters must be gathered by the endpoints that can then be used as a proxy for call quality?' This is not as simple as it sounds, since traditional IP network degradations such as packet loss, jitter and delay can certainly be measured, but these do not necessarily impinge upon voice quality in a direct manner – and even if they do, the relationship is not a linear or even a deterministic one. It is, therefore, vital that more sophisticated measures are employed that can unequivocally be related to speech quality. This section outlines some of the parameters that are available in the toolkit of the VoIP network management designer and discusses their advantages and disadvantages.

2.1 RTP and RTCP

Most VoIP traffic today consists of packets of voice data encapsulated with RTP headers. RTP [RFC3550] is a transport protocol that is designed for the transmission of voice and video data across IP networks. RTP is supplemented by a control protocol, RTCP. One of the roles of RTCP is to provide feedback on the quality of the RTP packet stream. The way in which it accomplishes this task is as follows. RTP packets are sent from each side of the communication link towards the other side. In addition, RTCP packets are sent by both sides at a regular rate, but much less frequently than RTP packets[1]. The RTP packets going across the network path between each RTP endpoint will be affected by network impairments such as network jitter, delay and packet loss. The RTCP packet includes parameters within it which allow both endpoints, as well as anyone able to intercept the packets mid-stream, to assess the seriousness of these network degradations. Some of the relevant parameters included in RTCP reports are:

RTCP Parameter(s)	Associated Function
Fraction Lost, Cumulative Packets Lost	Allows receiver of RTCP packets to identify packet loss experienced by far-end endpoint
Interarrival Jitter	Allows receiver of RTCP packets to obtain a smoothed estimate of the network jitter experienced by far-end endpoint
Delay since last RTCP report	Allows receiver of RTCP packets to calculate round-trip network delay between the two endpoints

[1] Typically, RTP packets are sent by an endpoint every 10-30 milliseconds. In contrast, RTCP packets are only sent every 5-10 seconds.

As can be seen, this information can be very useful to both the participants in the RTP session as well as the network operator providing the transmission path. Not only are these parameters important as end-of-call statistics to give a concise view of the overall call quality but, since the RTCP reports are continually being sent during the call, they can also highlight any transient effects that may have occurred and impacted speech quality over a shorter period of time.

Some of these parameters are determined by the endpoints based on only the RTP packets they receive, while others (such as round-trip delay) require the cooperation of both endpoints and, therefore, need both endpoints to be sending RTCP reports regularly to each other during a call.

Although the usefulness of these parameters is self-evident, they do need to be treated with a degree of caution. After all, the ultimate determinant of customer satisfaction is the subjective quality of a conversation between two parties – and degradations such as network packet loss, jitter and delay can only be regarded as proxies for call quality in this context. In fact, these network parameters do NOT, in themselves, provide a one-to-one mapping to voice quality. That is to say, they are not necessarily directly correlated to more traditional subjective quality measures such as MOS (Mean Opinion Score)[2]. This is because of two separate factors:

1. Firstly, modern speech coders can very effectively mask many of these distortions. For instance, coders often include Packet Loss Concealment (PLC) mechanisms that alleviate packet loss on the network, as well as jitter buffer control to overcome network jitter. Therefore, poor packet statistics may mislead the observer into believing that the listener speech quality is worse than it actually is.
2. Conversely, other degradations such as incorrectly set signal levels, noise and speech coder-related distortions will not be captured by network-related statistics such as packet loss and jitter and so, analyzing these proxy measures may give a false sense of security that all is well with the quality of the speech that the subscriber hears when it actually isn't.

It is clear that another approach is necessary – one that doesn't merely rely on how well RTP packets are being transmitted across the network.

[2] The methodology for deriving MOS scores has been standardized by ITU in specification [P.800]. Listeners are asked to rate degraded speech samples on a scale of 1 (Bad) to 5 (Excellent). These scores are then aggregated across all the listeners to arrive at an average MOS score between 1.0 and 5.0 for that test.

2.2 The E-model

The ITU-T has come up with a computational model (known as the E-model) that attempts to replicate the combined effects of the large variety of transmission-related factors that can have an impact upon conversational speech quality. This E-model is standardized in ITU-T specification [G.107].

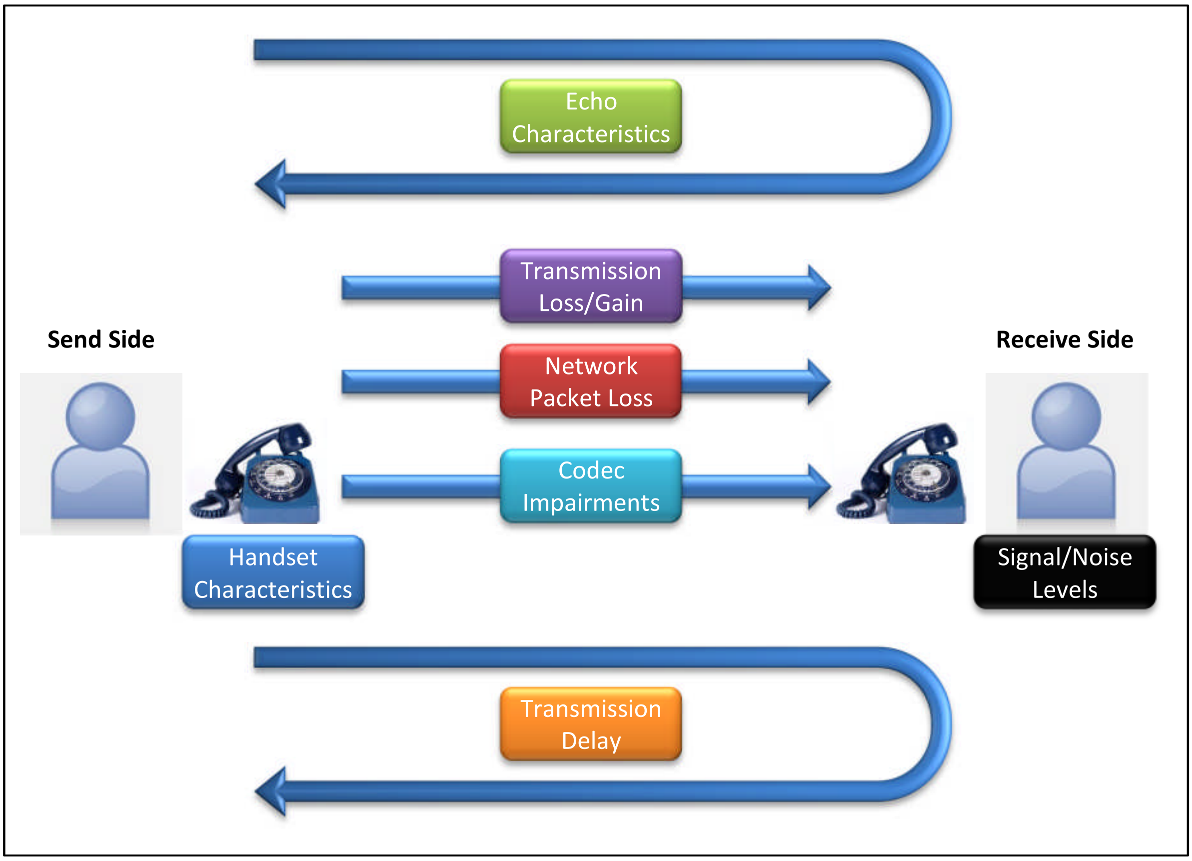

Figure 1. Parameters of the E-model

As shown in Figure 1, the inputs to the E-model bring together a number of different parameters including:

• **Telephone Handset Characteristics**	• **Signal Levels**
• **Transmission Loss/Gain**	• **Circuit Noise, Background Noise**
• **Transmission Delay**	• **Codec-related Impairments**
• **Echo Characteristics**	• **Network Packet Loss**

Therefore, some of the types of parameters used by the E-model are similar to those reported by RTCP, but the E-model list goes much further. Altogether, around 20 different inputs are fed into the E-model. The model itself consists of a complex sequence of mathematical operations which attempt to characterize the effects of these inputs on the subjective quality of the speech heard by the listener. To do this, the E-model incorporates an acoustic model of speech production, a psychoacoustic model of how speech is detected by the ear and then perceived by the brain as well as how speech signals are affected by network impairments.

This intricate analysis within the E-model manages to boil down the wide array of inputs into a small set of criteria that then go to form a single output known as the rating factor (R), or simply R-factor:

$$R = R_o - I_s - I_d - I_{e,eff} + A$$

where:

R_o	Signal-to-Noise Ratio incorporating effects of Signal Levels, Transmission Loss/Gain, Circuit Noise and Background Noise
I_s	Effect of all impairments that occur simultaneously with speech, eg incorrectly set Transmission Loss/Gain, Sidetone Levels, Quantization effects
I_d	Delay-related impairments, including the effects of Echo
$I_{e,eff}$	Impairments caused by the use of low-bit-rate codecs such as G.729, G.723.1 etc. Also includes the effect of Network Packet Loss
A	An Advantage Factor that is intended to encapsulate the benefits accrued to the listener from being mobile (cell phone or satellite phone) compared to a fixed-line connection[3].

The R-factor can lie anywhere between values of 0 and 100. [G.107] also specifies a one-to-one mapping between the computed R-factor and a MOS score[4]. The relationship between the two is expressed as a fairly complicated mathematical transformation, but can be shown graphically as in Figure 2.

[3] This Advantage Factor tries to account for the fact that a listener may be prepared to tolerate a poorer quality connection for the sake of being mobile, by biasing the R-factor to a higher value if it is a mobile connection.

[4] Strictly speaking, the MOS score that can be calculated from the R-factor is known as a MOS-CQE score (Mean Opinion Score, Conversational Quality Estimated). The type of MOS result obtained from a listening test where listeners rate short, degraded speech samples is referred to as a MOS-LQS (Mean Opinion Score, Listening Quality Subjective) score.

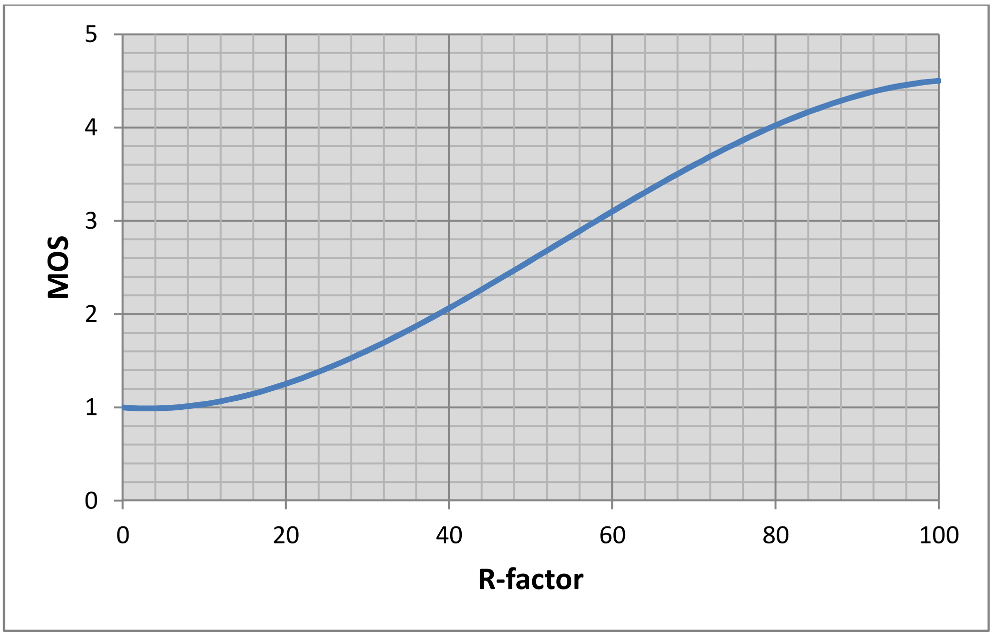

Figure 2. Relationship between R-factor and MOS

Therefore, a perfect communication link would rate an R-factor of 100 which equates to a MOS score of 4.5, while the worst possible R-factor would be 0 which is equivalent to a MOS score of 1.0. To provide further detail, the ITU-T has also come up with a rough guide as to what quality can be expected if the R-factor is within certain bands [G.109]:

R-factor	Below 60	60-70	70-80	80-90	90-100
MOS	< 3.10	3.10-3.60	3.6-4.02	4.02-4.34	4.34-4.50
Quality Category	Poor	Low	Medium	High	Best
User Satisfaction	Nearly all users dissatisfied	Many users dissatisfied	Some users dissatisfied	Satisfied	Very satisfied

In practice, if all the input parameters to the E-model are left at their default values, a G.711 PCM codec is used and zero delay is assumed, then an R-factor of 93.2 is achieved.

The E-model is, therefore, a very useful tool as it can assist the network designer in determining if the end-to-end quality of a call placed across a particular VoIP network will be acceptable or not. Its utility is further enhanced by the fact that the rather esoteric R-factor can be turned into a MOS score – which is perhaps more easily understood and widely recognized. The above methodology can be used quite effectively in determining call quality in that, if a customer raises a complaint, the network operator can set up a test call for a short duration of time towards this customer and then the two RTP endpoints can gather statistics that would allow them to compute R-factor/MOS for the duration of the call and thereby verify any potential problems.

However, for some of the purposes that were being discussed in section 1, there are still gaps to be filled. For instance, it would be very useful to the operator if the R-factor could be computed for every call placed across his or her network. Furthermore, it would be even better if R-factor could be sent at regular intervals during the call, so any transitory deterioration in voice quality could be immediately detected. These features would allow service providers to maintain both an overall assessment of network conditions at different parts of their network, as well as pinpoint any problem areas or specific customers that might be experiencing quality issues. The mechanisms discussed in the next section aim to overcome these concerns.

2.3 RTCP-XR

The IETF specification [RFC3611] describes a method for transmitting reports on the quality of an RTP stream. These extended reports (known as RTCP-XR) operate in a similar manner to conventional RTCP reports in that they supplement the RTP traffic sent from each endpoint to the other and are transmitted every few seconds. However, they contain a much larger set of parameters than standard RTCP reports. There are many types of parameters that can be sent in RTCP-XR reports, but the ones that are of most interest for this discussion are those contained in the 'VoIP Metrics Report Block'. This block of data is itself divided into several sub-blocks as shown below:

VoIP Metrics Sub-block	Included Parameters
Packet Loss and Discard Metrics	Overall packet loss rate and the rate at which packets have been discarded by the receiver
Burst Metrics	Parameters which given an indication of not only how many packets are lost during transmission, but also how bursty the losses are
Delay Metrics	Round-trip delay and how much delay is incurred in the RTP endpoint's internal processing
Signal Related Metrics	Signal level, Noise level, Echo Canceller characteristics
Call Quality Metrics	R-factor, MOS
Configuration Parameters	Type of Packet Loss Concealment and Jitter Buffer used
Jitter Buffer Parameters	Delay incurred due to the jitter buffer

As can be seen, some of these parameters are similar to those discussed in relation to RTCP reports in section 2.1, but the detail provided by RTCP-XR is much greater. However, some of the metrics in the above table are totally new. In particular, the call quality metrics such as R-factor/MOS discussed in section 2.2 with reference to the E-model can be incorporated into RTCP-XR reports. This is very powerful as a mechanism now exists for an endpoint to report both to its peer RTP endpoint as well as to the network as a whole what the call quality is not only at the end of each call, but also at regular intervals within the call, if need be.

At this point, it may seem as if we have reached the holy grail and the quest to find a set of parameters that can accurately and quickly detect fluctuations in call quality is complete. A method of transmitting these parameters within the network has also been described. However, the picture is still not quite complete. Although VoIP endpoints can use RTCP-XR and other mechanisms to collect data, calculate a suitable set of parameters and then transmit them, the amount of data that will be gathered in this way will be immense. If each and every call generates RTCP-XR reports every few seconds and there may be thousands or even tens of thousands of calls every hour in a network, the operator is likely to be overwhelmed. How is the network operations staff to make any sense of this mountain of information and get an appreciation of network conditions as a whole, or indeed find the needle in the haystack they need to debug a particular customer's complaint?

As was discussed in section 1, this challenge requires a very special set of tools. These tools need to be part of the network control center and must meet the dual objectives of gathering and collating all this data and then displaying it in a visually appealing form that can allow the operator to detect wholesale network degradations quickly, whilst also zeroing in on specific customers who may be experiencing call-quality related issues. Fortunately, such a set of tools does exist – as will be described in the next section.

3 THE INNOMEDIA DEVICE MANAGEMENT SYSTEM (DMS)

InnoMedia is a leading provider of VoIP devices and management systems for both residential and business voice applications. These devices can be used with either cable HFC access networks or those that are Ethernet-based. InnoMedia's products include VoIP endpoints such as Multimedia Terminal Adapters (MTA's), IP phones as well as business voice solutions such as IP-PBX's and Enterprise Session Border Controllers (ESBC's). Section 4 provides an overview of the InnoMedia product portfolio with particular emphasis on the vital role that these devices play in generating and transmitting call quality-related metrics. At this point, it is sufficient to say that InnoMedia devices are purpose-built with the need to monitor call quality in mind and support protocols such as RTCP & RTCP-XR fully in order to achieve these aims.

The entire suite of InnoMedia products can be managed from a central management platform known as the Device Management System (DMS). This system will typically be located in the service provider's network operations center and consists of a highly available server configuration, with each DMS system being capable of managing tens of thousands of devices. The DMS is responsible for a variety of features such as fault-finding individual devices and alarm management. However, in this document, we will focus the discussion on those aspects of the DMS that pertain specifically to monitoring and managing call quality in the network.

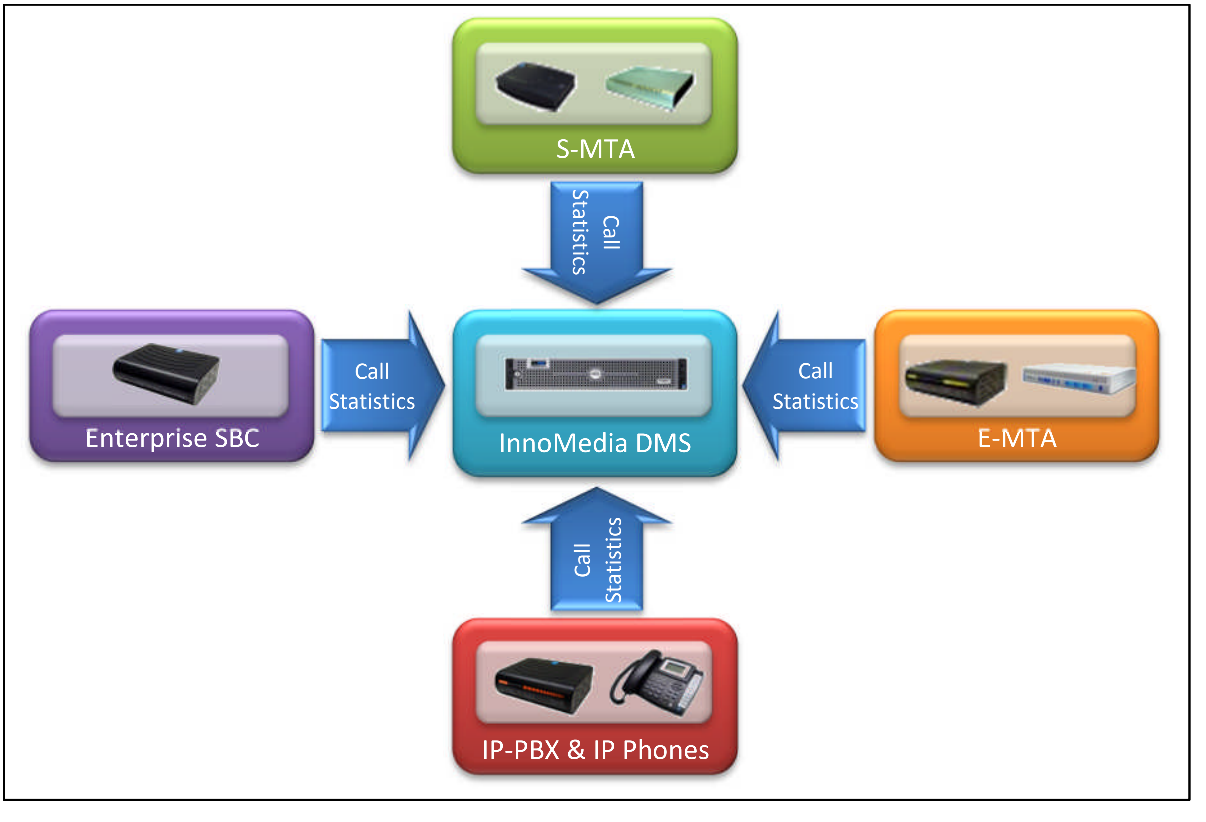

Figure 3. The InnoMedia Device Management Eco-System

As shown in Figure 3, the DMS plays a central role in gathering and collating input from the VoIP devices deployed in the network. To fulfil the requirements outlined in previous sections, this data is then used for two different purposes:

1. As an aid in assessing the overall health of the network by providing aggregated statistics which enable the operator to quickly gain an assessment of problem areas and thereby to focus in on these potential issues.
2. When problematic devices have been identified, to assess the network quality being provided to that device in particular and use this to determine further action.

The following sections discuss the implications of these two requirements and how the DMS is architected to meet these challenges.

3.1 Aggregated Network Quality Management

3.1.1 The Dashboard View

The first view that the operator sees upon logging onto the DMS system GUI is the 'Dashboard'. This is a user-specific, customizable view that brings all the network-related information that a particular user may be interested in onto a single view. It allows the operator to keep their favorite set of key assessment criteria on hand at all times – even when they may be engaged in other tasks.

To fully understand the ways in which network quality can be managed, it is important to first introduce the concept of regions. When a new VoIP device is first registered onto the network, it is assigned to a specific region (geographic or otherwise). This then allows a large and complex network to be compartmentalized, and so permits the service provider to (i) assign responsibility of specific regions to certain operations staff (in which case they may only be able to view the specific devices under their own control) as well as (ii) realize that problems may be only occurring in a particular part of their network and thereby trace the underlying cause. Each region can also be further divided up into sub-regions (eg cities or states within a region) to gain greater granularity.

Figure 4. The DMS Dashboard

As explained earlier, the dashboard is fully customizable, both in terms of content and layout. In the particular dashboard screen shown in Figure 4, several views are shown on the same display. These include:

- **Network Map:** Overall view of the number of devices in each region
- **Device Type:** Shows how many device of each type are present in the network
- **Device Version Panel:** How the devices are divided up by software version number
- **Alarms:** Illustrates the alarms detected by the DMS by region
- **Device Status:** How many devices are on-line/off-line within each region
- **Voice Quality – MOS:** Perhaps most importantly in respect of call quality management, the average MOS scores for all devices in a specific region, or across all regions, can be viewed graphically over a period of time. Similar graphs can also be produced for R-factor as well.

It is important to note that this is not a static display. For instance, a number of the items above can be presented either graphically, or as a list view. Also, by clicking on one of the segments of the pie chart in the network map, the specific region that has been chosen will explode to show a similar network map for all the sub-regions within this region. A further click will allow the user to punch through to a list of devices within that sub-region, or across a number of regions as shown in Figure 5.

Select All | Add Selected to Prov | Delete Selected | Re-Prov Selected | Reset All

	ST	MAC	IP	Region	Device Type	Version	Prov	
1		00:10:99:09:b0:46	209.133.49.172:1139	US/San Diego, CA	MTA 6328-8 - MGCP	1:0:6		
2		00:10:99:00:fd:03	61.218.105.10:5200	China	MTA 6328-4 SIP	1:0:5		
3		00:10:99:22:86:43	209.133.49.172:6880	US/San Jose, CA	MTA 6328-4 SIP	1:0:6		
4		00:10:99:01:ac:3e	121.6.65.253:5200	Singapore	MTA 3328 SIP	4:11:1	-	
5		00:10:99:01:c8:6d	209.133.49.18:27123	US/San Jose, CA	MTA 3328 SIP	4:2:67		
6		00:10:99:09:92:92	209.133.49.18:14944	US/Alviso	MTA6328-8e	4:2:69		
7		00:10:99:05:a4:af	209.133.49.18:17496	US/Alviso	MTA6328-8e	4:2:69		
8		00:10:99:09:92:29	209.133.49.18:24523	US/Alviso	MTA 3328-2Re - MGCP	4:2:69		
9		00:10:99:09:91:5d	61.218.105.11:5200	886	MTA 6328-4 SIP	1:0:9	-	
10		00:10:99:30:95:7e	12.22.51.80:5200	US/San Jose, CA	MTA 3328 SIP	4:2:69		
11		00:10:99:02:eb:55	209.133.49.18:32845	US/San Jose, CA	MTA 3328 SIP	4:2:69		
12		00:10:99:01:c8:1b	122.152.136.93:5200	Singapore	MTA 3328 SIP	4:2:67	-	
13		00:10:99:06:92:8e	122.152.136.93:5200	Singapore	MTA6328-8e	4:2:69	-	
14		00:10:99:05:a7:d5	209.133.49.18:7037	US/Alviso	EMTA-6528-2 SIP	4:2:69		

Select All | Add Selected to Prov | Delete Selected | Re-Prov Selected | Reset All

Figure 5. Device List View

This view is important as it forms the bridge between looking at regions as a whole, and then focusing on the investigation of issues with specific devices. So, for example, the operator may detect unexpected fluctuations within the R-factor or MOS scores within a specific region or sub-region. They are then able to drill down from region to sub-region to a set of devices (or even a single device).

3.1.2 Parametric Voice Quality Analysis

Although the dashboard provides a high-level depiction of network call quality, a more detailed analysis is often warranted. As was discussed in section 2.1, network degradations such as packet loss, jitter and delay can often influence the quality of speech reaching the end-user. Therefore, it is useful to have an overall sense of how these parameters are behaving at any point in time. The DMS achieves this through a spatial representation of these parameters as shown in Figure 6.

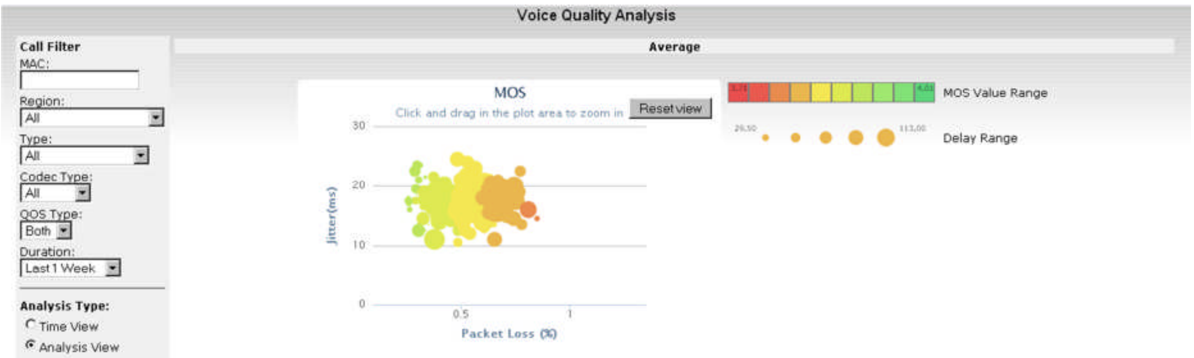

Figure 6. System Level Voice Quality Analysis

This analysis is an extremely effective way of showing the impact of four separate quality-related variables in a single, visually informative graph:

- Packet Loss is measured along the x-axis
- Jitter is shown on the y-axis
- The size of each circle represents the delay associated with that call
- The color of each measurement indicates the quality of that call either in terms of MOS score or R-factor[5]. Green indicates good quality, while red indicates poor quality

This analysis can be viewed for the entire network, a specific region or a particular sub-region. Again, the graph is interactive in that the operator can click-and-drag to zoom into any part of the graph that may be of interest for closer analysis. By focusing on areas where any one of the four parameters shown in the graph are outside acceptable limits, the operator can get a better picture of what particular network degradations might be causing quality issues.

With these capabilities and with those associated with the dashboard, it is relatively simple for the operator to keep an eye on call quality within the network as a whole and then, when he or she needs to really understand what sort of problems might be occurring (and how often) with a particular device, more detailed information is available as discussed in the next section.

3.2 Device-Specific Quality Management

Once the decision has been made to investigate a particular VoIP device in the network, the arsenal of tools available within the DMS opens up to provide a number of possibilities. Although the overall view of R-factor or MOS within a region/sub-region is useful, the analysis required to isolate problems with a specific device needs to be much more precise. It is this set of features that are discussed in this section.

3.2.1 Temporal Analysis of Voice Quality

Identifying that a specific VoIP endpoint may be suffering from network degradations is important. However, the next set of questions that immediately need to be answered are: (a) is network quality to the device fluctuating or steady, (b) if it is fluctuating, how severe are these variations, when did they occur and which aspects of network quality are being affected? To address these concerns, an essential prerequisite is a time-based analysis of call quality experienced by the device. The DMS is

[5] In this case, the R-factor and MOS are calculated according to the E-model as discussed in section 2.2.

well-equipped to perform this function and can be used to bring up a graphical display of the variation of various network metrics with time as shown in Figure 7.

Figure 7. Device-Specific Temporal Analysis of Voice Quality

This particular graph shows how MOS, R-factor, jitter, packet loss and delay have varied over a configurable period of time. As before, it is possible to select any time interval of interest and zoom into it simply by clicking and dragging. With a variety of network degradations included in this analysis, it is a relatively straightforward task to isolate which one might be impacting voice quality.

3.2.2 Analysis of Call Statistics

Another common scenario when debugging call quality problems is that the operator will be alerted that the end-customer has experienced poor quality for a specific call or set of calls within a certain time window. This is where yet another component of the network quality feature set available on the DMS comes into play. For any device, call statistics can be easily displayed as shown in Figure 8.

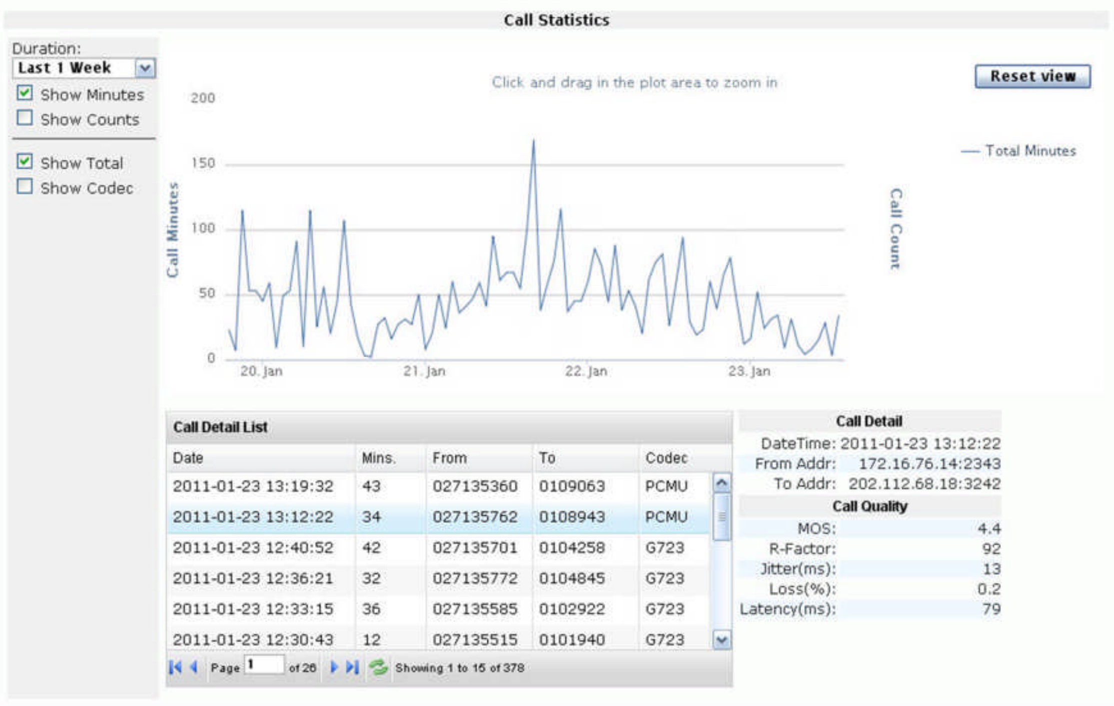

Figure 8. Call Statistics View

This view provides graphical information on calling trends through this particular device. However, even more importantly, it also shows individual call details. This list gives the operator data on not just important attributes associated with the call such as call time, call length, caller/callee information etc., but also, by highlighting a particular call, quality-related metrics can be seen in the lower right-hand sub-window. For this specific call, MOS, R-factor, jitter, packet loss and delay are provided and can be used to pinpoint what particular issue may have caused this particular call to experience quality problems.

3.3 External Call Quality Monitoring Mechanisms

The methods outlined thus far in this section for monitoring call quality have relied upon the DMS system to perform the required analysis and display the results. In this subsection, another group of techniques will be introduced that can be used for gathering, collating and analyzing quality metrics from the VoIP endpoints either using servers external to the DMS by themselves, or in conjunction with the DMS.

3.3.1 Syslog Reporting

Although the methods described in section 2 such as the E-model and RTCP-XR can be extremely useful in ascertaining whether there are any call quality issues in real-time, they are transient in nature in that they are only available around the time they are sent and are not generally stored by each endpoint. However, it can be important to store in bulk and later retrieve details of how calls have performed, perhaps long after the fact. For this purpose, the InnoMedia call quality reporting eco-system relies upon the syslog mechanism.

Syslog (documented in [RFC5424]) provides a wrapper within which many types of information can be transmitted from a syslog-capable device to a syslog server. However, in this context, the focus is on using syslog to convey call-related information. Several members of the InnoMedia device family are able to send the information mentioned below within a syslog message at the end of each and every call:

• **Device MAC Address**	• **Call R-factor**
• **Callee/Caller Identification**	• **Call MOS Score**
• **Call Start-Time/End-Time**	• **Average Packet Loss**
• **Call Direction (inbound/outbound)**	• **Average Jitter**
• **Codec Used**	• **Average Delay**

Therefore, the information sent to the syslog server is not just call quality-related, but also contains vital data pertaining to the characteristics of the call itself. The syslog server collects this information for all calls within the network, if required. This data is then available for later retrieval and offline analysis. Armed with this capability, the operator can go back at any time and view the details of either a specific call, or compare calls from the same user at different times, or across different users.

3.3.2 Call Quality Traps via SNMP

The procedures discussed up till now have allowed the operator to work in two different ways:

- to go back and check what has happened in the past to affect call quality in any part of the network,
- to see adverse trends in network performance that may be impacting quality at the current moment as they develop in real-time.

However, even in the second case, it requires the operator to be actively monitoring quality metrics on the DMS system and then react to them. This subsection takes this one step further by describing a mechanism whereby the operator can be alerted to impending problems automatically.

A number of InnoMedia devices support the provision of quality thresholds that, once breached, will trigger the system to generate SNMP traps to either an external trap collector/alarm manager, or towards the DMS itself. For instance, the partial screenshot below is taken from the web GUI of the ESBC product and shows how thresholds can be set independently for MOS score, delay and packet loss on the device.

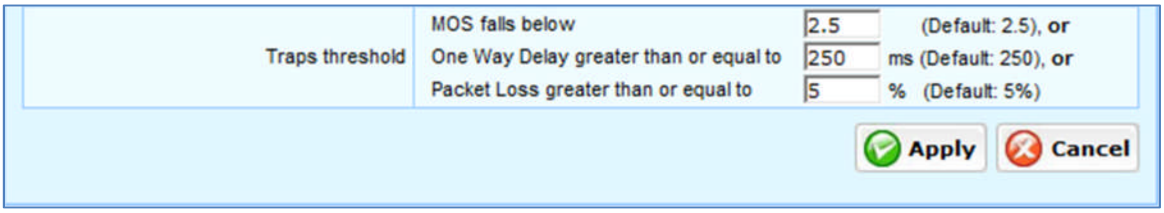

Figure 9. Setting of Call Quality Thresholds on ESBC

If these thresholds are ever exceeded, the device will not only send an appropriate SNMP trap, but the DMS system can also be configured to send the network operator an e-mail alerting them to the situation, if necessary. This process allows the operator to be proactive in addressing quality concerns and perhaps react before the end customer is even aware that a problem affecting call quality has occurred.

3.4 Active Call Quality Reporting

The previous discussion has given details of a number of modes of performing what can be termed as ‘passive’ monitoring, since they focus upon determining the quality of live calls in the network, either at the current time or in the past. The two techniques outlined in this subsection add yet another arrow into the quiver of the service provider by delivering ‘active’ call quality reporting procedures. These aim to test out the specific network link to a particular subscriber by placing test calls to this end customer and measuring the quality of these calls as they proceed.

3.4.1 Voice Quality Monitoring (VQM)

Most of the techniques discussed so far have relied upon voice quality metrics such as network packet loss, jitter, delay, R-factor and MOS in judging whether a particular link or region of the network may be degrading speech to an extent that may raise quality concerns. As discussed in section 2.1, while network packet loss, jitter and delay can be useful in providing indicative markers to call quality, users of these metrics may be mislead by them just as easily. Measures such as R-factor and MOS are better in that the underlying models incorporate many different factors, including details of the human speech production and auditory perception mechanisms.

However, at the end of the day, all of these parameters are simply proxies for what is the fundamental basis for call quality determination – whether the speech sounds degraded to the network subscriber or not. None of the metrics mentioned above rely upon analyzing actual speech samples themselves, and this can be an elemental limitation of their use.

This is where the VQM module comes in. It is a server-based software feature that is part of the DMS platform and operates as follows:

1. The VQM module contacts the device and places it into RTP loopback mode
2. The VQM sends a reference speech file (in the form of RTP packets) downstream to the device
3. Speech is looped back at the device and returns to the VQM module
4. The VQM compares the reference and degraded speech samples using PESQ analysis
5. VQM generates an estimated MOS score.

This procedure is shown graphically in Figure 10.

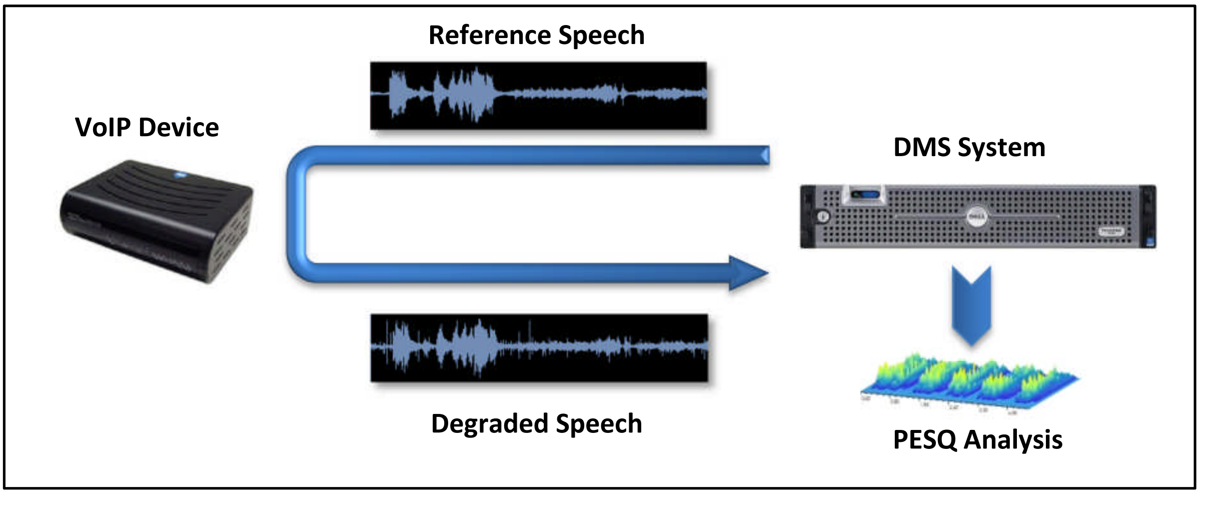

Figure 10. Operation of the VQM Module

The PESQ (Perceptual Evaluation of Speech Quality) algorithm is an objective speech quality measure that is defined by ITU-T in [P.862]. It incorporates a perceptual model of the human auditory mechanism as well as a cognitive model of how the brain perceives speech quality. Most importantly, as shown in Figure 11, it takes both reference and degraded speech inputs and compares them using signal processing techniques to arrive at a PESQ score on a scale of -0.5 to 4.5.

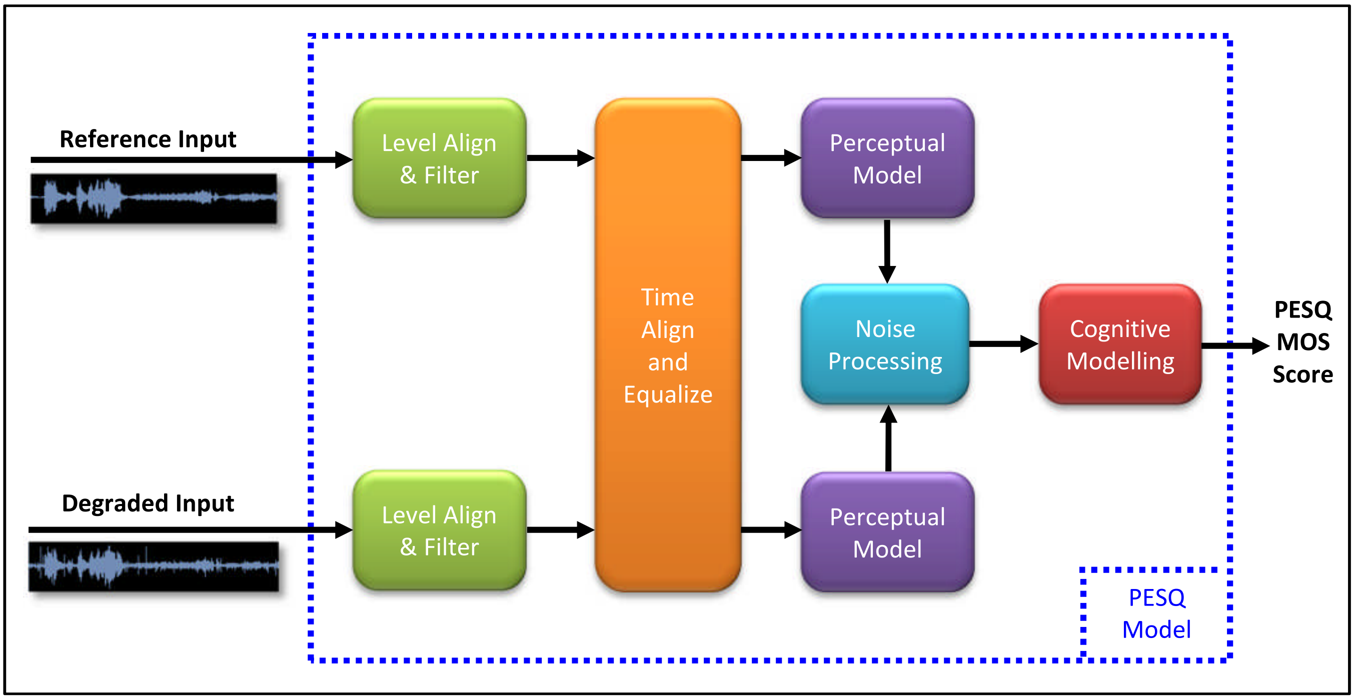

Figure 11. The PESQ Analysis Model

These PESQ scores have been found to be highly correlated to MOS scores derived from actual subjective tests with real listeners. Furthermore, the PESQ scores can be mapped as described in [P.862.1] to a MOS score that can be reported through the DMS[6].

The significance of this unique VQM analysis procedure should not be underestimated. If the end customer approaches the service provider complaining of poor call quality, literally at the touch of a button, the operations team can set up a quick test call and almost instantaneously arrive at a measure for the quality of the network link from the operator's core network to the subscriber, directly from the network operations center without any customer intervention being required. Moreover, the quality measure produced is based on comparing actual speech waveforms and so corresponds closely to the subjective call quality heard by the subscriber.

3.4.2 Embedded Test Agent

The final call quality reporting mechanism to be highlighted in this document is similar in certain respects to the VQM module described in the previous subsection in that it is based on a quality analysis of test calls, but is also different in some ways. It relies upon a special test agent function

[6] Strictly speaking, the mapping from PESQ score to MOS score produces a MOS-LQO (Mean Opinion Score – Listening Quality Objective) result as described in [P.800.1].

embedded in the VoIP device that is being monitored. This test agent can be used by the network operator in one of two distinct ways:

1. **Test Agent Originating Calls.** The test agent can initiate test calls to a remote party telephone number. Once the call is set up, a speech file is sent to the remote number for a period of time and the call quality is monitored to accumulate values of R-factor and MOS using RTCP-XR both during the call and for the call as a whole.
2. **Test Agent Terminating Calls.** The test agent can also be placed into loopback mode so that an external server can utilize it in performing call quality measurements. In fact, the functionality discussed in section 3.4.1 for the VQM module on the DMS server can be performed using the embedded test agent on the device as the test agent is used in loopback mode and the VQM drives the delivery of a speech file to the device and back in performing its PESQ analysis.

For case 1. above where the test agent originates the test calls, the calls can either be initiated manually, or be pre-configured to occur on a regular basis (daily, weekly, monthly) and then build up a picture of the variation of test call quality over time. Figure 12 shows a typical test call in progress (manually initiated by the test agent) as well as the results of the last such test call shown on the device's web GUI.

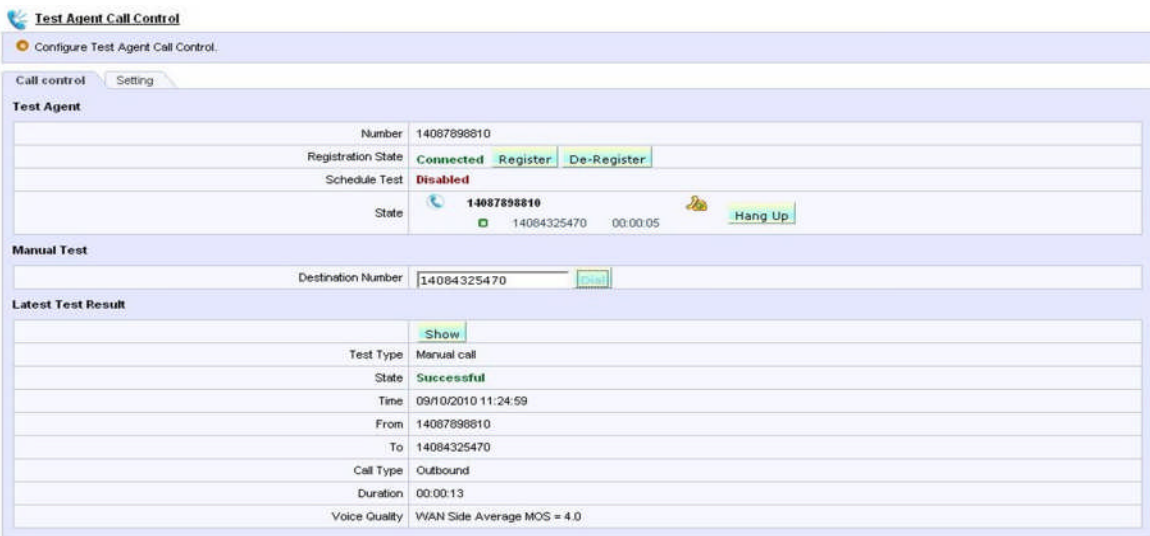

Figure 12. Embedded Test Agent Operation

4 The InnoMedia Device Portfolio

This section offers some high-level information on the members of the InnoMedia family of VoIP products. The aim is not to go into depth on product details and features, but rather to provide a high-level overview and also discuss how these devices support the call quality monitoring and reporting techniques presented in previous sections. These products cover both residential and business voice solutions and are engineered from the ground up with call quality measurement in mind. Further information on these devices and their specific applications is available at: www.innomedia.com/products.shtml.

4.1 Multimedia Terminal Adapters (MTA's)

InnoMedia MTA's are compatible with either cable HFC access networks (requiring Embedded-MTA's or E-MTA's including an embedded cable modem) or those that are Ethernet-based (requiring Standalone-MTA's or S-MTA's). They cover the range from 1 analog voice port through to 24 voice ports to deliver solutions for all types of subscribers from the smallest residential customer to medium-sized business customers. All devices include integrated LAN ports to allow Internet as well as voice traffic to pass through the device. Triple-play residential deployments with voice, data and video through IPTV are supported using VLAN tagging, allowing prioritized QoS-based service delivery. Business Voice Service MTA's incorporate business-friendly features such as ground start/loop start signaling, low-speed modem support for credit-card readers/Point-of-Sale terminals, GR-1089 lightning protection support etc.

In terms of call quality management, all of these devices support the RTCP protocol and many also support the E-model and RTCP-XR and, therefore, are capable of producing R-factor and MOS scores. Several devices also allow the termination of test calls that may be initiated by the DMS system VQM module in carrying out its PESQ analysis as in section 3.4.1. Syslog reporting for end-of-call quality metrics is also supported.

4.2 IP-PBX and IP Phones

The InnoMedia IP-PBX solution consists of a SIP Trunking-enabled IP-PBX together with a suite of IP phone devices. Together these form an integrated package that is simple to install and easy to manage through its unique hierarchical management system that allows service providers to provision both the IP-PBX and IP phones remotely, as well as auto-discovery and configuration of the IP phones as soon as they are plugged into the enterprise network. A powerful self-care GUI provides the end-user a rich set of business telephony features and a totally configurable set of multiple auto-attendants offer the enterprise ultimate control in managing their customer-facing persona.

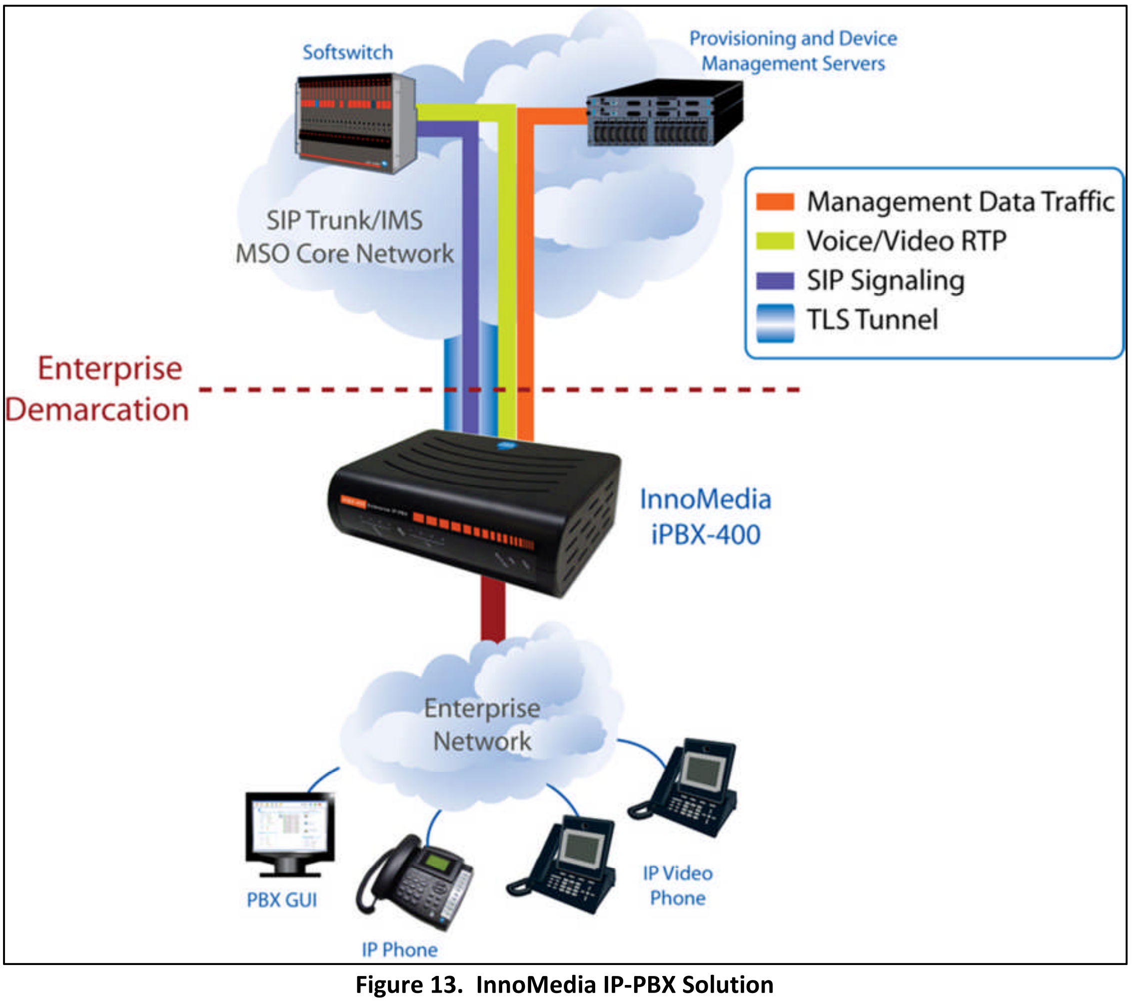

Figure 13. InnoMedia IP-PBX Solution

The IP-PBX itself comes equipped with the capability to handle up to 200 users and 60 concurrent calls as well as integrated NAT router, stateful firewall and TLS-based security. The IP phones range from entry-level devices to full-featured, large screen phones offering HD voice support among other features.

The devices are also integrated into the call quality framework described in this document with all offering RTCP and some also providing RTCP-XR capability as well as syslog reporting for end-of-call statistics.

4.3 Enterprise SBC (ESBC)

The InnoMedia family of ESBC products offers broadband service providers and Cable MSO's the ideal product for delivering scalable, QoS-managed SIP Trunking and Hosted PBX services to their business customers. As shown in Figure 14, the ESBC sits at the demarcation point between the operator's core network and the enterprise that it is serving. As such, it provides the following functions:

- **SIP Normalization** between the service provider's SIP signaling and that employed by the enterprise's IP-PBX. InnoMedia ESBC devices simplify this function through the use of both LAN-side and WAN-side SIP profiles as well as pre-configured templates for most popular PBX's.
- Operation in either **B2BUA** mode for SIP Trunking services or **SIP ALG** for Hosted PBX networks, with full support for NAT Traversal when separated from the PBX by a NAT/firewall. B2BUA operation also brings benefits of **Topology Hiding** to the service provider.

- End-to-end **QoS** support either using ToS/DSCP marking in broadband Ethernet deployments or, for Cable networks, InnoMedia's ***Smart*-DQoS**™ technology allows the ESBC to initiate and manage UGS service flows autonomously without the need for PacketCable Multimedia (PCMM).
- **Security** features such as TLS support for SIP signaling from the ESBC to the operator's core network and Intrusion Detection/Prevention Systems (IDS/IPS).
- **Other Features** such as DTMF translation and special handling of emergency calls.

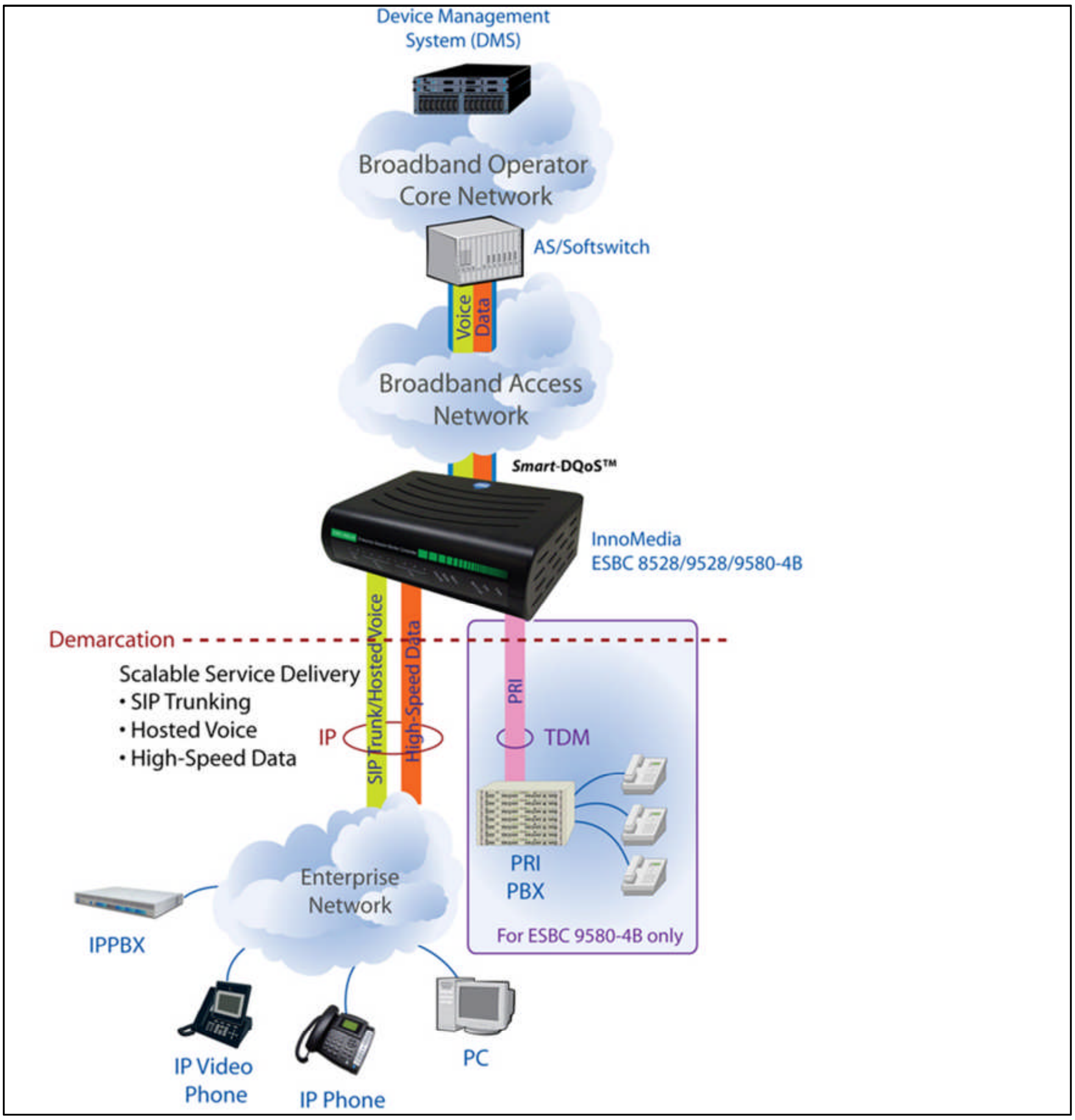

Figure 14. Cable-Based Deployment Configuration for InnoMedia ESBC

The suite of InnoMedia ESBC products is also extremely integrated with each member of the family having an internal battery, 4x FXS ports, 4x LAN ports (one of which may be configured as a high-speed pass-through data port for Internet traffic) as well as an optional PRI interface (supporting up to 2x T1/E1s) to interface with TDM-based PBX's. Those models intended for cable deployments also include an embedded Cable Modem supporting either DOCSIS 2.0 or DOCSIS 3.0 – with the DOCSIS 3.0 devices providing 8x4 channel bonding with 24 SIDs.

Now turning to the call quality management aspects of the ESBC, all of the products described above support both RTCP and RTCP-XR – thereby supplying R-factor and MOS scores. As these devices are installed at the demarcation point between the operator's network and that of the enterprise and so connect two separate IP networks, they are able to generate call quality metrics both for their LAN interface towards the enterprise as well as their WAN interface towards the service provider's core network. Therefore, both LAN-side and WAN-side R-factor and MOS scores will be created. Each ESBC will also generate syslog reports and SNMP traps when call quality thresholds are exceeded. An embedded test agent is also integrated into the device and can be used for generating test calls and also in conjunction with the VQM module on the DMS platform for generating PESQ-based MOS scores. In short, the ESBC product family supports the entire range of call quality solutions discussed in section 3 and thereby allows the service provider to deliver premium business voice solutions to enterprise customers, secure in the knowledge that they have the full gamut of call quality management solutions at their disposal in order to cope with a wide variety of network problems.

5 SUMMARY

This document has hopefully made it clear that call quality management in a large VoIP network is not an easy task. After all, the aim is to try and manage something that is inherently ephemeral and insubstantial in nature – the subjective quality of speech perceived by the customer on their handset. To make matters worse, poor call quality cannot usually be blamed upon a single causal factor. There are usually a multitude of network components that may affect the traversal of VoIP packets from their source to their destination and also perhaps multiple paths that are taken by these packets within a single call. Finally, the sheer scale of the problem can be overwhelming. Of the thousands or even millions of calls that are placed on a substantial VoIP network, the charge of finding a few that are impacted by network impairments may seem daunting. Fortunately, it is possible for the network operator to overcome these hurdles – with the right tools in their armoury.

First of all, the user needs to decide which particular network degradations are going to be considered as key markers – those which will be monitored and used as determinants of network quality. Unfortunately, again there is a choice of many different metrics. Some, such as those used in RTCP, are easy to compute and readily available. However, they do not necessarily map directly into how the customer would perceive call quality. RTCP-XR uses a more sophisticated set of parameters such as R-factor and MOS (as prescribed by the E-model) that are perceptually-based and, therefore, offer a more accurate proxy for subjective speech quality. R-factor and MOS scores are widely used in the DMS system for this reason but, for those occasions where users are complaining specifically of voice quality, perhaps the most accurate method is the VQM module which relies on a direct comparison of reference and degraded speech samples using PESQ analysis.

Through the previous sections, we have described what may seem like a bewildering array of techniques aimed at solving the problems facing the service provider in network performance monitoring and call quality management. However, this complexity is there for a purpose, as the operator is not always presented with the same types of problems over and over again. Take the analogy of a doctor's diagnosis of a patient for a moment. Initially, the physician will employ simple procedures such as measuring the heart rate or blood pressure of the patient. However, as the diagnosis becomes more precise, the doctor's tools will also change and become more specific to the disease at hand. In the same way, the operator sometimes only wishes to monitor the 'heart-beat' of the network and so tools such as the dashboard are more than adequate. However, as the operator feels their way through a customer-specific or region-specific problem, the tools need to provide more detail on the exact nature of the problem.

It is instructive to view the DMS methodologies developed in section 3 as a progression: starting from being able to gain a network-wide (or perhaps region-wide) understanding of how metrics such as MOS or R-factor are behaving, then to see how these metrics are being affected by network degradations such as packet loss, jitter and delay. From the insight gained here, the operator is able to isolate which devices might be experiencing problems, and then dive into a device-specific investigation offered by either temporal analysis, call statistics, syslog or SNMP traps. The ultimate arbiters of call quality are tools such as VQM or the test agent which allow actual test calls to be placed and then analyzed for a particular device. This gradual transition from the general to the specific is shown in Figure 15, from network-wide analysis to device-specific investigations, from a high-level overview of a few network parameters to detailed scrutiny of a variety of metrics. This palette of tools is crucial to the operator in keeping their finger on the pulse of network quality and dealing with impairments as they occur – whether they might be affecting thousands of customers, or just a single user.

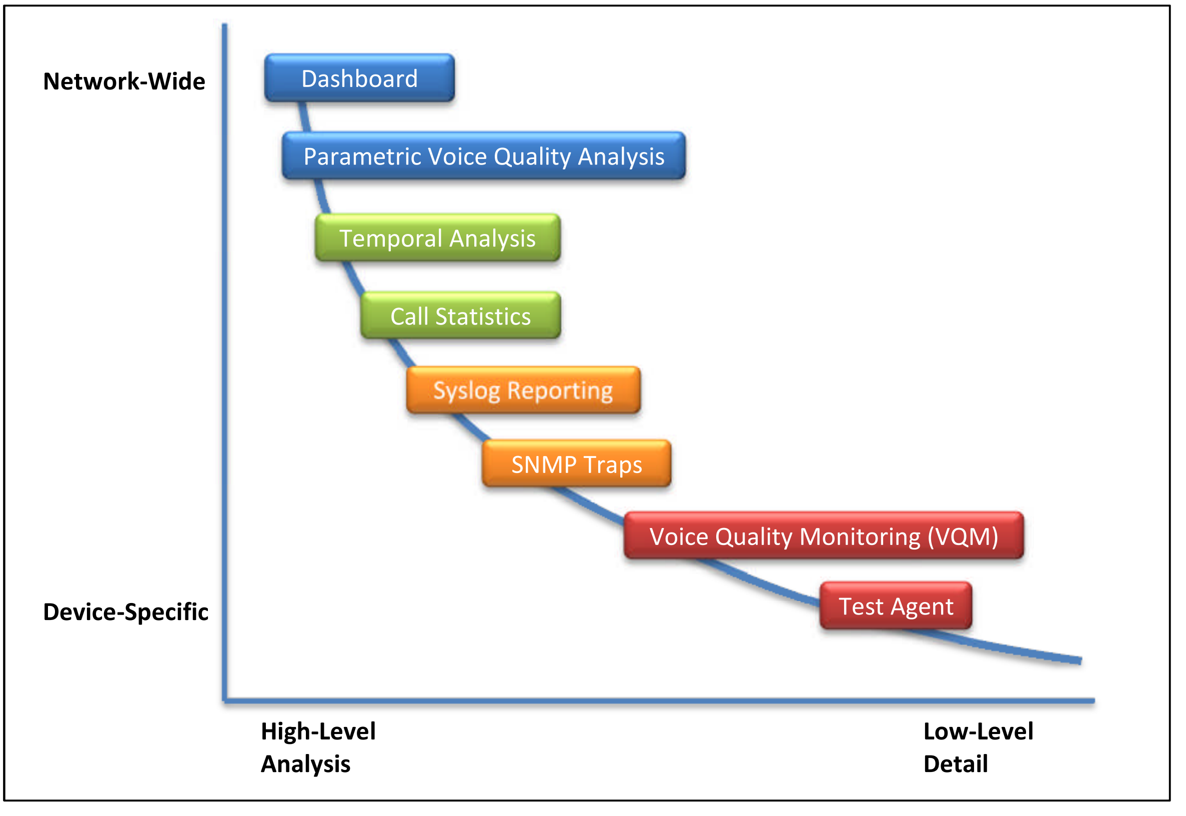

Figure 15. Hierarchy of Call Quality Management Tools

However, the role of the DMS does not just end at isolating those devices which might be experiencing network quality problems. The DMS also incorporates a powerful feature which allows the operator to go directly from a DMS screen to either the Web GUI of the device or even through to its console screen via telnet access. This functionality is possible even when the device is separated from the DMS by a firewall or NAT. This unique capability has two specific uses. Firstly, the operator can view call quality details on the device that may not be available on the DMS – such as the results of test agent calls as described in section 3.4.2. Secondly, it allows the operator direct access to the product as if they were sitting beside it in order to check or change the configuration of the device or to investigate any issues associated with it at first-hand.

Ultimately, the success or failure of a network management operation is highly dependent on the care and foresight adopted by the network's designers in terms of the suitability and capability of network equipment, the redundancy of links and nodes in the network as well as bandwidth provision for the most demanding load scenarios. However, any experienced network operations team will still be aware that problems inevitably do occur, affecting specific customers, certain regions, or perhaps even network-wide. It is for these occasions that the operator needs to be armed with a powerful set of tools that will give them the assurance that they can, at any time, monitor network health, identify problem regions or devices as well as probe deeply into the nature and causes of network impairments. The InnoMedia call quality management solution provides this level of surety through an integrated and holistic framework that encompasses devices that support the generation and transmission of call quality metrics as well as dedicated server platforms that can gather, collate and present this data in a coherent and lucid form to allow the service provider to seamlessly manage their VoIP networks.

6 BIOGRAPHY OF THE AUTHOR

Harprit S. Chhatwal is Vice President, Technology for Innomedia, Inc. – a leading supplier of residential and enterprise IP telephony solutions for Cable MSO's and Broadband telephony providers. He is involved in setting strategic technology directions for the company and communicating with key partners and customers. Prior to joining Innomedia, Dr. Chhatwal was Chief Technical Officer for Nuera Communications, a provider of VoIP telephony media gateways for the cable MSO market. He has also been active in numerous industry standards organizations and has most recently co-authored the PacketCable 1.5 and 2.0 Speech Codec specifications as well as the PacketCable Enterprise SIP Gateway (ESG) specification. Dr. Chhatwal holds a B.Sc. and Ph.D. from Imperial College, London for his work on speech coding, an MBA from Warwick University, UK as well as 5 patents in digital signal processing.