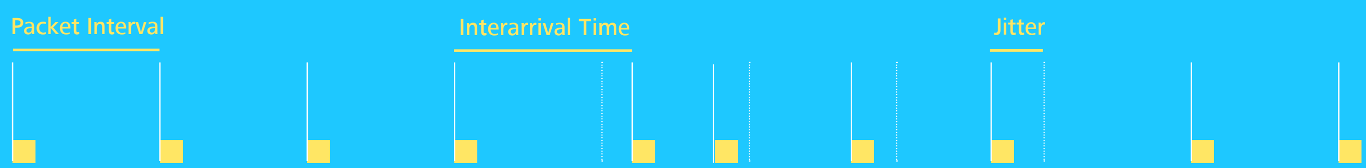
voipfuture

WHITEPAPER

VOIP QUALITY MONITORING BASICS

Jitter Measurement for Voice over IP



The challenge of media monitoring /// Jitter – friend or foe? /// Interarrival Time – key metric to evaluate voice services /// The data acquisition process /// RTCP – a second approach? VoIP quality related data processing /// Presentation of quality – patterns become apparent /// Understanding information content – extracting the root cause /// Late packets /// The distinct quality metric – a new field of voice quality

Wendenstr. 4
20097 Hamburg
Germany

+49 40 688 90 01 0
info@voipfuture.com
www.voipfuture.com

Non-intrusive real-time RTP monitoring addresses two major aspects of the VoIP business: user experience and network performance. This whitepaper will shed light on the thus far mainly untouched topic of RTP monitoring by introducing an efficient, new technology which automatically detects the root causes of VoIP quality degradations. This white paper addresses network engineering and IT experts for carriers, service providers and enterprises.

A short introduction

As far as user experience in the context of VoIP is concerned two things come immediately to mind – waiting for the ring tone (call set-up time) and the sometimes poor speech quality. In IP networks media monitoring is by nature directly tied to the latter.

While call control (SIP has emerged as the ultimate next gen standard here) is an undisputed prerequisite for placing VoIP calls, corresponding tools and monitoring solutions are common practice – the handling of signaling issues is well understood. Surprisingly, the second half of VoIP, media traffic responsible for the connection quality, is not subject to such 24/7 monitoring solutions. RTP monitoring is not yet established. In contrast to traditional switched networks media streams in IP networks are exposed to network originated impairments. Often underestimated, VoIP endpoints may also already contribute to poor quality by generating unstable VoIP traffic, e.g. due to poor performance, improper configuration or implementation issues.

Furthermore, standard conformance might be out of the scope of end users because so far it has to do with latent, not yet audible effects. But it quickly becomes highly interesting if specific thresholds of the media stream are exceeded and problems become noticeable. This is because VoIP endpoints have the ability to compensate for deviations from good quality up to a specific degree by temporarily storing them in a so called jitter buffer (receiver buffer) – eventually showing a somewhat 'digital' behavior. Deceiving, because the audio still sounds good, while the VoIP environment (network, VoIP system) is far from operating in a stable mode. Distortions may add up, eventually becoming critical. However, such violations of conformance indicate unstable operation, or the vulnerability of the voice application, respectively.

Thus, besides monitoring call control in IP networks, permanently observing all media packets in all VoIP calls – RTP monitoring is inevitable to finally ensure and improve end user satisfaction.

The challenge of media monitoring

As RTP monitoring is the ultimate choice to permanently observe media quality, one has to think about processing a massive amount of raw data in real time. This becomes inevitable especially when supporting the 1 GB or 10 GB high speed links of carrier networks in a 24/7 manner. It translates directly into specific tasks to be performed by local passive RTP monitoring probes.

In particular, processing and evaluating all RTP streams in order to deliver application layer voice quality information in real time requires the following:

- Online recognition of each and every RTP packet at wire speed
- High precision time stamping of RTP packets
- De-multiplexing thousands of simultaneous RTP streams
- Recognizing the packet interval of each RTP stream
- Extracting the information elements needed to accurately measure voice quality
- Grouping and aggregating these information elements to keep the size of quality data to a reasonably small amount
- Storing the relevant pre-processed quality data and delivery to a centralized management component in order to get a comprehensive, multi-probe service level view, e.g. of a whole carrier network.

The following figures may demonstrate the severity and challenge of the task to be performed in carrier real life VoIP environments: if monitoring 1 Gbit/s link with 8 hours of peak utilization, equal to around 5 million call-minutes per day, the volume of RTP raw packet data to be processed would be close to 5.49 TB. Knowing how to reduce the measurement data without losing relevant information can reduce the volume to just 7.4 GB of aggregated quality data.

Instead of online processing, another direction, offline analysis, has been given precedence so far. Taking a deep dive, packet by packet, into the protocol layers, the DPI approach, which stands for deep packet inspection, is rather more suited for troubleshooting purposes in small samples of network traces – and is unable to even moderately zoom out to provide a reasonable overview meeting the requirements of online monitoring.

Jitter – friend or foe?

Long ago people all over the world became accustomed to telephony as a semi-natural means of communicating.

Without doubt this can be attributed to the excellent maturity of the switched network architecture – setting the standards of quality in decades past. By means of synchronized clocks in a dedicated infrastructure voice can be transmitted steadily, making people believe that the other party is right next door.

But IP networks do not have such clocking – voice is packetized and carried along with other data traffic competing for the same infrastructure. Thus, higher layer protocols have to take over and provide some kind of synchronization mechanism in order to achieve a similar level of quality. In the case of voice this is the RTP protocol.

Taking the position of an VoIP endpoint it is expected that all packets travel steadily at the same speed arriving in constant intervals. If this is not the case speech on the phone sounds awfully distorted or broken up. And even if no real packet loss has occurred the RTP packets may reach the receiver side too late to contribute to the continuous flow of voice conversation. Though finally arrived, these late packets have to be discarded by the receiving VoIP endpoint.

While the IP network has physically carried the entire VoIP traffic it doesn't really satisfy the tough requirements of the voice application layer. All such deviations from the desired continuous transmission of VoIP traffic which badly affect the user experience are summarized as jitter. Other applications may easily tolerate jitter – VoIP definitely cannot.

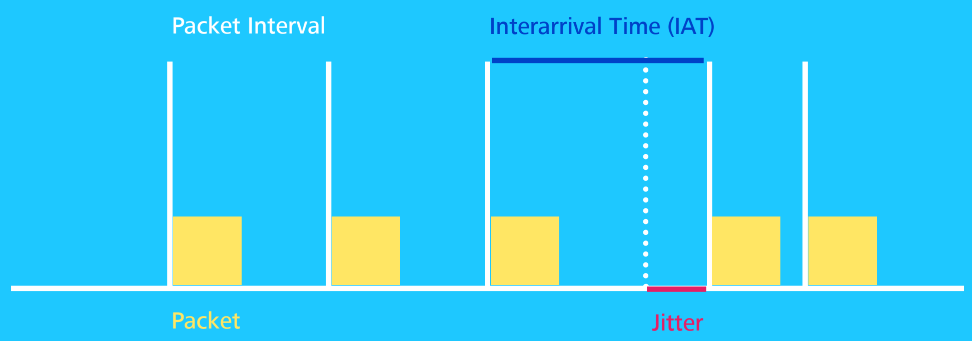
On the other side, viewed from the angle of an IP network, VoIP is just data traffic with lots of packets all of the same size. Powerful mechanisms like MPLS, DiffServ, RSVP, just to name a few, take much care by prioritizing privileged voice traffic. But still there is hardly any certainty about the good quality of the VoIP service.

Thus, permanent non-intrusive RTP monitoring as discussed later is the only way to address VoIP application layer quality in IP networks – by explicitly analyzing jitter or it's 'parent' parameter, the interarrival time, respectively.

Interarrival Time – key metric to evaluate voice services

First things first. Before we start discussing RTP monitoring, the appropriate terminology must be defined first.

Relation of Interarrival time and jitter in an isochronous RTP media stream with a given packet interval



RTP packets are supposed to travel at the same speed and at the same distance (isochronous). So the definition of jitter, interarrival time and their mutual dependencies must be made crystal clear.

While the packet interval represents the previously mentioned isochronous regular state of a VoIP call sending RTP packets e.g. every 20 ms, the jitter – or more precisely the interarrival jitter – is usually taken as the absolute value of the deviation from this regular state, even though it may be negative or positive. Finally, the real distance between two subsequent packets is the interarrival time.

In order to keep all parameters separate some simple framing conditions apply to the calculation of IAT. Only if the RTP sequence numbers of two subsequent packets show an increment of 1, they will be taken into account. So sequence errors, aka packet order, or packet loss providing a decrement or an increment greater than 1, respectively, are discarded. These effects are covered by supplemental metrics (avoiding redundant information) without adulterating the IAT. Referring to the above definition of the terminology it is quite evident that interarrival time (IAT) can be measured directly – without the need to know anything about the RTP stream such as the packet interval.

Actually, it's the only way to cope with the tremendous flood of information from a high number of simultaneous calls. On the other hand, jitter is the deviation from the regular state of the IAT, i.e. from the packet interval. It can be calculated only if the packet interval is known and the IAT has been ascertained by observing a sequence of RTP packets.

Thus, we can assign attributes to these parameters. IAT is a primary parameter (directly measureable) while jitter appears to be a secondary (derived) parameter. When introducing appropriate parameter handling we will see that secondary parameters are characterized by loss of information content. Consequently, we will be following the IAT path here.

Experts who have to judge the quality of a VoIP environment are faced with several challenges that will be discussed briefly in the following sections: selection of relevant parameters, accurate measurement, correct handling of parameters, extracting and understanding the information content.

The data acquisition process

Two questions typically arise when talking about measurement or, more precisely, the data acquisition process: what (in a given VoIP environment) and how (with which granularity) to capture.

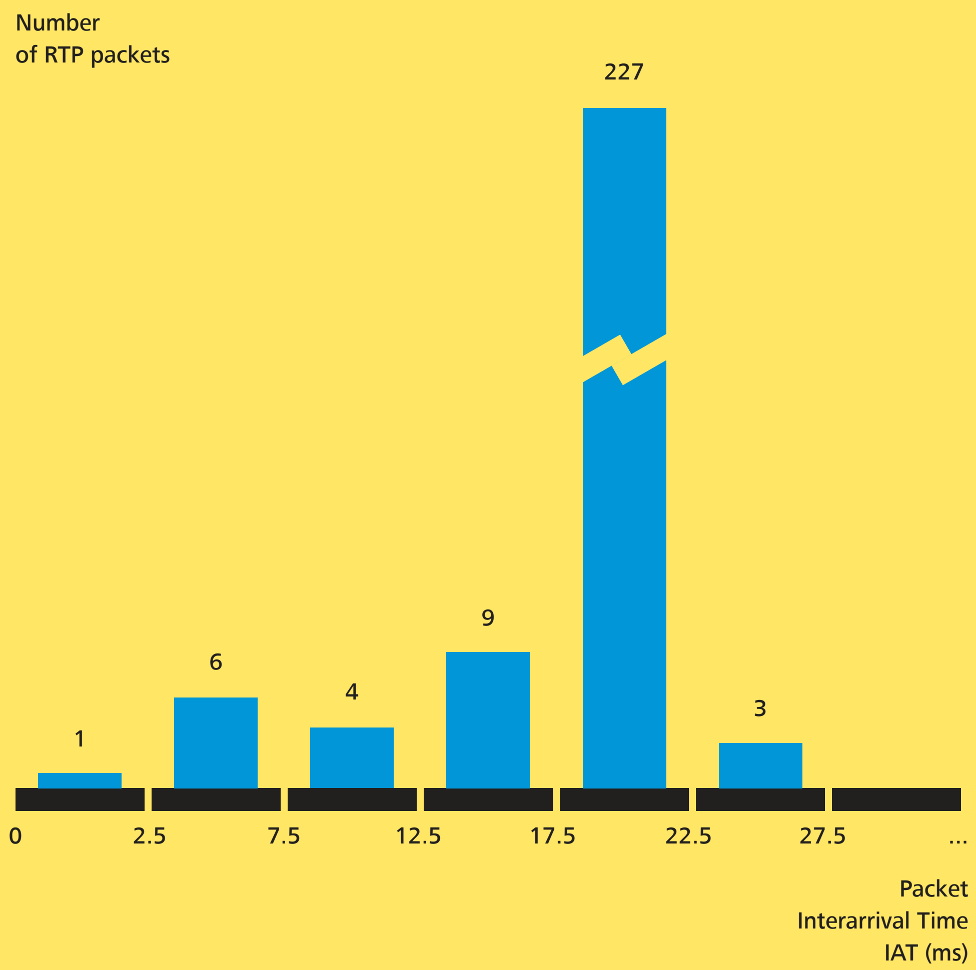
Jitter, the most valuable information in order to address user experience, conformance and root cause diagnosis can be obtained by jitter measurement, actually the IAT analysis. Hence, a reliable capturing process, optimized for tough voice application level requirements is absolutely indispensable. IAT can be measured at any single point on an IP network with RTP traffic.

Thus, monitoring all RTP packets in a given VoIP environment is the only way to obtain reliable quality data. Audio, being the focus of this white paper, is sent in equal intervals (isochronously). This allows us to focus on the receipt (arrival) time. Since IAT is defined to be the distance of two subsequent packets in an RTP stream, measuring the IAT is quite simple. All measured packets (arrived at the point of measurement) have to be time stamped (arrival time).

IAT is simply the difference in the arrival times of an RTP packet and it's predecessor (reference point). Needless to say, RTP monitoring should be performed in real time.

Finally, the important topic of accuracy shall be addressed. RTP packet intervals

Well defined characteristics for a frequence distribution model (histogram) for voice analysis



have values at multiples of 10 milliseconds with typical values of 20 or 30 ms. The measurement process should thus be able to support the correct recognition of a packet's interval and potential deviations. Precise RTP monitoring relies on an accurate, packet capturing process. Packet capturing has to provide measurement timestamps with sufficient accuracy and granularity of one microsecond under real-time conditions. After the real-time processing of RTP packets the microsecond values have to be grouped into clusters of 5 milliseconds in size in order to avoid massive amounts of stored data and, even more importantly, to deliver reasonable, relevant values. More details will be discussed in the 'data processing' section.

As always in practical life, finding the right balance is the ultimate secret. Sticking with extreme granularity (some tools even proudly present nanosecond timestamps) quickly becomes a disadvantage. It does not reflect the needs of practical measurement.

RTCP – a second approach?

Theoretically, VoIP offers another, indirect way of gathering information about the temporal progression of an RTP stream: RTCP reports. They are sent by the receivers of an RTP stream in the opposite direction – to inform the sending endpoint about the quality received.

Besides the fact that RFC 3550 already gives a clear warning not to take jitter values from RTCP reports (see chapter 6.4.4 – quote: "The interarrival jitter field is only a snapshot of the jitter at the time of a report and is not intended to be taken quantitatively."), countless practical projects have proven that RTCP jitter does not reflect the real quality of an RTP stream. As none of the important aspects discussed in this white paper are met, (real) RTP quality can be good while RTCP reports poor quality and vice versa. Unfortunately, any combination of RTP/RTCP quality information can be seen. Furthermore, it's only jitter – a derived parameter anyway.

In short, RTCP reports made by endpoints are intended to be used by endpoints, and are outside the control of a carrier's OSS/BSS. All one can expect here is a fragmented, invalid view, never suited for carrier grade voice service assurance.

VoIP quality related data processing

At this point another aspect must be introduced – the information content, aka entropy. Technical parameters show very distinct levels of entropy, depending on their nature. The IAT is one of those parameters holding extremely valuable information requiring proper handling.

Min./max./average/standard deviation have so far failed to satisfactorily serve the presentation of IAT. Only histograms can deliver what is needed. Histograms are not completely new, but apart from statistical purposes, using them to obtain and derive application layer information about voice quality is definitely new. This starts with the grouping of values.

Often seen as a rather practical demand to cope with mass data, it results in a formal grouping of values which do not necessarily belong together, just by statistical relevance. Viewed from a different (application) angle, the grouping of values has to follow the nature of the parameter.

So the center point (!) of an IAT histogram bar must correspond to the grid of the RTP packet interval as well as multiples and fractional parts of this (see histogram). Suddenly, the same raw data appears to look quite different at the presentation level, revealing essential details which, as we will see later, directly lead to the root cause of poor quality.

The second detail of IAT histograms to be taken account of is the selection of the relevant subset of possible IAT values. Keeping in mind that regular packet intervals are around 20 or 30 ms, a maximum of 100 ms perfectly meets the needs of RTP monitoring to show all distortion effects. Values above 100 ms are summarized in the 100 ms histogram bar (≥ 100 ms). As will later be illustrated in the pattern recognition section, IAT histograms exhibit specific patterns which are directly related to the root cause of VoIP quality degradation.

Presentation of quality – patterns become apparent

Another rule applies when creating IAT histograms – choosing an adequate time interval to be covered by such a histogram. By doing so, patterns which are directly linked to the VoIP user experience become readily apparent.

The reason is simple. Histograms may hold all IAT values of entire VoIP streams of different length and still have the same size. But it is quite evident that short sequences (e.g. 20 seconds) of VoIP quality distortions are hardly discernible in the histogram of a 5-minute VoIP stream. Measurable effects are compensated if the time intervals covered by an IAT histogram that have been selected are too large. On the other hand, very small intervals covering just a few RTP packets are unable to show any reasonable, recognizable pattern at all. The truth again lies somewhere in the middle. Intervals of 5 seconds in size are the absolute optimum to reflect the needs of VoIP quality presentation. Despite the packet interval values in practical use, typically selected from a range between 10 ms and 40 ms, 5-second intervals with the corresponding number of RTP packets they represent (see adjacent table) are robust enough to allow for automatic pattern recognition while remaining as small as possible to prove a sufficient granularity of the temporal progression of a real-time VoIP stream.

Packet Intervall	10 ms	20 ms	30 ms	40 ms
Packets / 5 s	500	250	167	125
Frequency of use	~5 %	~80 %	~14 %	~1 %

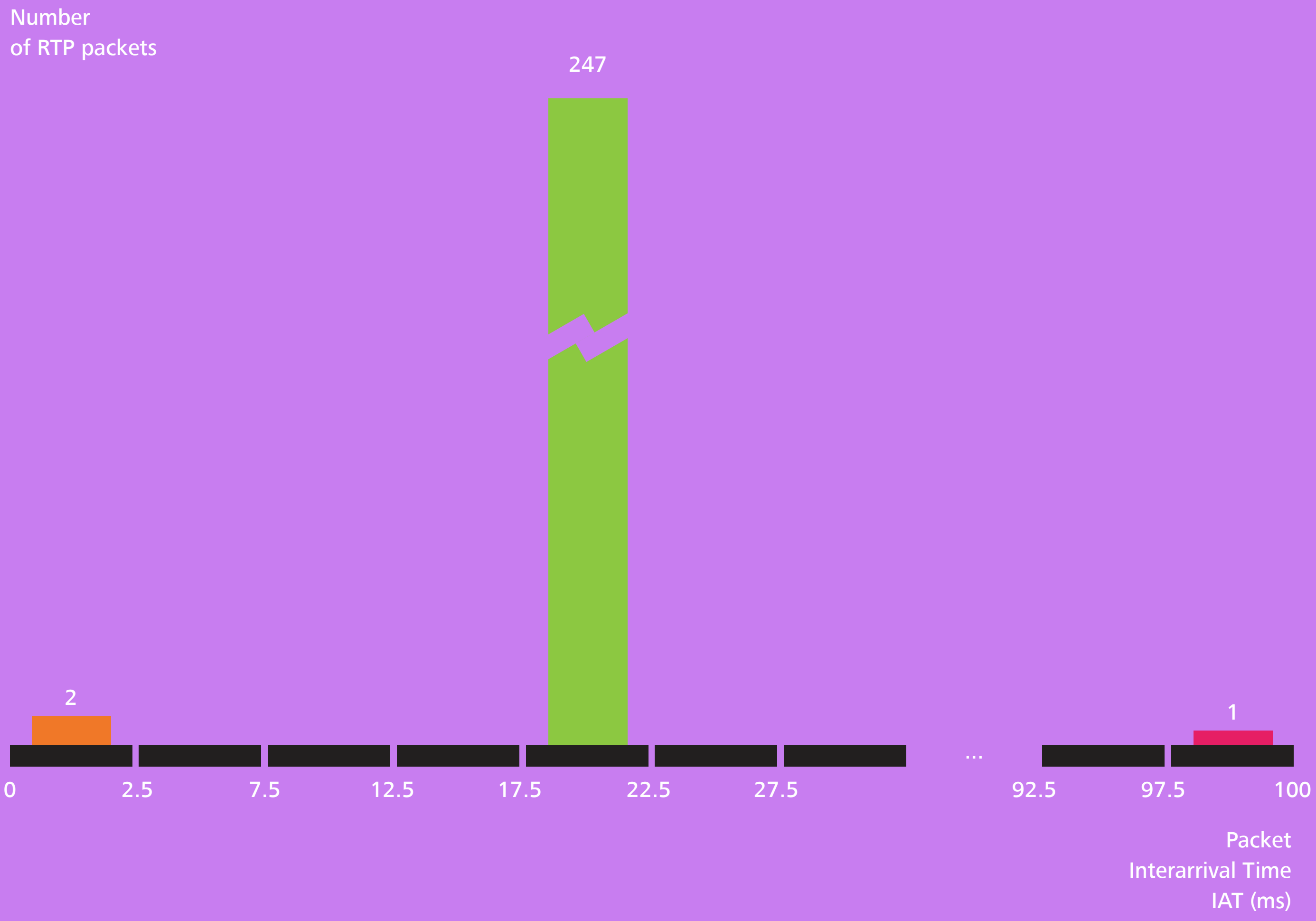
Understanding information content – extracting the root cause

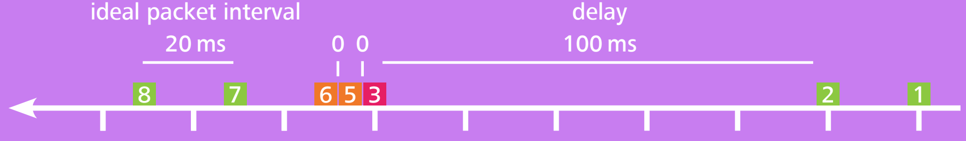
The graph below shows an IAT histogram with the right center points for the histogram bars covering the relevant IAT range and generated for the right interval of time.

It reveals essential details of the quality of the RTP stream. At a glance experts are informed that the RTP stream was temporarily buffered in an active component (switch, router) on the IP network. This pattern is an example that illustrates a root cause diagnosis outcome called ‘network overload’. But what’s behind this pattern and how is it related to the VoIP user experience?

Once again it’s the result of a simple process of permanent RTP monitoring. The temporary buffering of an RTP stream will appear to a receiving VoIP endpoint as a sequence of one large, audible gap (IAT) followed by at least two subsequent very small

Sample of "Network Overload" pattern for root cause analysis





Schematic packet flow of impaired voice transmission by network (near zero) IATs caused by the fast-as-fast-can emptying of the component's ingress/ egress buffer. Such network components are transparent in terms of higher layer protocols like RTP, thus incapable of reproducing the original packet interval.

This sequence (large IAT and subsequent very small IATs), as easily observed from the flow of RTP packets, is reflected in the pattern of the IAT histogram. Of course this is just one example of many obvious root cause diagnosis patterns which are reliably detectable by automatic passive RTP monitoring probes only. Sometimes root causes can be obtained by simply checking thresholds in the IAT histogram but it often gets far more complex. As discussed with the 'network overload' effect in addition to the histogram itself also the original order of the appearance of it's values in the RTP stream matters. In other cases ratios of values in the IAT histogram are of interest instead of just thresholds.

However, it pretty well explains why the distinction between negative and positive deviations from the regular packet interval is essential. Absolute values of such deviation as provided by the jitter help no one.

Late packets

A valuable side effect of using IAT histograms is the ability to efficiently detect late packets – RTP packets are not in

reality being lost, but have to be treated as virtually lost by the VoIP application of the receiver – they are arriving too

late to contribute to the decoding of the audio signal. In the IAT histogram late packets are those who are situated

beyond a specific point on the x-axis. E.g. a jitter buffer of 40 ms would mean that all packets above 60 ms (exceeding the regular packet interval of 20 ms by 40 ms) can be deemed as being too late.

Discarding those late packets is directly linked to the configuration of the jitter buffer size which may compensate for slight delays in RTP packets. Such buffers work by simply adding delay to the entire stream of RTP packets in order to allocate a specific amount of time for the re-arrangement or even re-ordering of RTP packets, if necessary. Since delay disrupts two-way conversation and in extreme cases makes it sound like a satellite link it should be kept as small as possible. So the buffer size is subject to optimization, big enough to compensate for jitter and small enough to avoid audible delay. Some VoIP endpoints even deploy dynamic jitter buffers. Examining the IAT histogram helps to identify late packets and to optimize jitter buffers as well.

The distinct quality metric – a new field of voice quality

As outlined in the previous sections VoIP service assurance requires media (RTP) monitoring. But the benefits of permanent media monitoring can only be capitalized by introducing an appropriate quality metric that permits expression of the apparent root cause information, making sustainable improvements in the quality of the VoIP service a reality.

This is where years of theoretical research and practical experience gained in particular from numerous customer projects around the world, are paying off. Details discussed above, such as:

- Taking account of each and every VoIP packet

- Supporting carrier grade high speed links in real time
- Defining the appropriate relevant VoIP quality parameters
- Keeping the parameters separate
- Treating each parameter individually according to it's information content
- Understanding the information content and extracting the root cause

provide a rock solid foundation for the new metric and are of course not limited to the interarrival time/jitter.

They also apply to other parameters as well, e.g. packet loss. Just gathering packet loss data, even if expressed as consecutive packet loss, does not provide enough valuable information. The histograms mentioned in this white paper help in particular to describe the characteristic of packet loss including the distance between two packet losses to better understand it's eventual root cause.

To sum up, this new VoIP metric makes reliable, 24/7 monitoring of the user experience in carrier environments possible.

Voice connection quality can be assured and network utilization optimized.

About

Founded in 2007, Voipfuture is the leading provider of RTP Monitoring as a standard network solution.

Live traffic RTP Monitoring provides insight into customer experience and network performance of voice services in IP networks.

Voipfuture delivers leading edge technology to monitor and evaluate voice quality in Next Generation Networks. The products are characterized by performance, precision and vendor independence.

Our solutions offered to carriers, enterprises and system integrators open up a new dimension in the efficient operation of VoIP networks. Secure your VoIP future with Voipfuture.