        Guidelines for Considering New Performance Metric Development

Abstract

   This document describes a framework and a process for developing
   Performance Metrics of protocols and applications transported over
   IETF-specified protocols. These metrics can be used to characterize
   traffic on live networks and services.

Status of This Memo

   This memo documents an Internet Best Current Practice.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   BCPs is available in Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc6390.

Table of Contents

1. Introduction

   Many networking technologies, applications, or services are
   distributed in nature, and their performance may be impacted by IP
   impairments, server capacity, congestion, and other factors. It is
   important to measure the performance of applications and services to
   ensure that quality objectives are being met and to support problem
   diagnosis. Standardized metrics help ensure that performance
   measurement is implemented consistently, and they facilitate
   interpretation and comparison.

   There are at least three phases in the development of performance
   standards. They are as follows:

   1. Definition of a Performance Metric and its units of measure

   2. Specification of a method of measurement

   3. Specification of the reporting format

   During the development of metrics, it is often useful to define
   performance objectives and expected value ranges. This additional
   information is typically not part of the formal specification of the
   metric but does provide useful background for implementers and users
   of the metric.

   The intended audience for this document includes, but is not limited
   to, IETF participants who write Performance Metrics documents in the
   IETF, reviewers of such documents, and members of the Performance
   Metrics Directorate.

1.1. Background and Motivation

   Previous IETF work related to the reporting of application
   Performance Metrics includes "Real-time Application Quality-of-
   Service Monitoring (RAQMON) Framework" [RFC4710]. This framework
   extends the remote network monitoring (RMON) family of specifications
   to allow real-time quality-of-service (QoS) monitoring of various
   applications that run on devices such as IP phones, pagers, Instant
   Messaging clients, mobile phones, and various other handheld
   computing devices. Furthermore, "RTP Control Protocol Extended
   Reports (RTCP XR)" [RFC3611] and "Session Initiation Protocol Event
   Package for Voice Quality Reporting" [RFC6035] define protocols that
   support real-time Quality of Experience (QoE) reporting for Voice
   over IP (VoIP) and other applications running on devices such as IP
   phones and mobile handsets.

The IETF is also actively involved in the development of reliable
transport protocols, such as TCP [RFC0793] or the Stream Control
Transmission Protocol (SCTP) [RFC4960], which would affect the
relationship between IP performance and application performance.

Thus, there is a gap in the currently chartered coverage of IETF
Working Groups (WGs): development of Performance Metrics for
protocols above and below the IP layer that can be used to
characterize performance on live networks.

Similar to "Guidelines for Considering Operations and Management of
New Protocols and Protocol Extensions" [RFC5706], which is the
reference document for the IETF Operations Directorate, this document
should be consulted as part of the new Performance Metric review by
the members of the Performance Metrics Directorate.

1.2. Organization of This Document

This document is divided into two major sections beyond the "Purpose
and Scope" section. The first is a definition and description of a
Performance Metric and its key aspects. The second defines a process
to develop these metrics that is applicable to the IETF environment.

2. Terminology

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in RFC 2119 [RFC2119].

2.2. Performance Metrics Directorate

The Performance Metrics Directorate is a directorate that provides
guidance for Performance Metrics development in the IETF.

The Performance Metrics Directorate should be composed of experts in
the performance community, potentially selected from the IP
Performance Metrics (IPPM), Benchmarking Methodology (BMWG), and
Performance Metrics for Other Layers (PMOL) WGs.

2.3. Quality of Service

Quality of Service (QoS) is defined in a way similar to the ITU
"Quality of Service (QoS)" section of [E.800], i.e., "Totality of
characteristics of a telecommunications service that bear on its
ability to satisfy stated and implied needs of the user of the
service".

2.4. Quality of Experience

   Quality of Experience (QoE) is defined in a way similar to the ITU
   "QoS experienced/perceived by customer/user (QoSE)" section of
   [E.800], i.e., "a statement expressing the level of quality that
   customers/users believe they have experienced".

      NOTE 1 - The level of QoS experienced and/or perceived by the
      customer/user may be expressed by an opinion rating.

      NOTE 2 - QoE has two main components: quantitative and
      qualitative. The quantitative component can be influenced by the
      complete end-to-end system effects (including user devices and
      network infrastructure).

      NOTE 3 - The qualitative component can be influenced by user
      expectations, ambient conditions, psychological factors,
      application context, etc.

      NOTE 4 - QoE may also be considered as QoS delivered, received,
      and interpreted by a user with the pertinent qualitative factors
      influencing his/her perception of the service.

2.5. Performance Metric

   A Performance Metric is a quantitative measure of performance,
   specific to an IETF-specified protocol or specific to an application
   transported over an IETF-specified protocol. Examples of Performance
   Metrics are the FTP response time for a complete file download, the
   DNS response time to resolve the IP address, a database logging time,
   etc.

3. Purpose and Scope

   The purpose of this document is to define a framework and a process
   for developing Performance Metrics for protocols above and below the
   IP layer (such as IP-based applications that operate over reliable or
   datagram transport protocols). These metrics can be used to
   characterize traffic on live networks and services. As such, this
   document does not define any Performance Metrics.

   The scope of this document covers guidelines for the Performance
   Metrics Directorate members for considering new Performance Metrics
   and suggests how the Performance Metrics Directorate will interact
   with the rest of the IETF. However, this document is not intended to
   supersede existing working methods within WGs that have existing
   chartered work in this area.

This process is not intended to govern Performance Metric development
in existing IETF WGs that are focused on metrics development, such as
the IPPM and BMWG WGs. However, this guidelines document may be
useful in these activities and MAY be applied where appropriate. A
typical example is the development of Performance Metrics to be
exported with the IP Flow Information eXport (IPFIX) protocol
[RFC5101], with specific IPFIX information elements [RFC5102], which
would benefit from the framework in this document.

The framework in this document applies to Performance Metrics derived
from both active and passive measurements.

4. Relationship between QoS, QoE, and Application-Specific Performance
    Metrics

   Network QoS deals with network and network protocol performance,
   while QoE deals with the assessment of a user's experience in the
   context of a task or a service. The topic of application-specific
   Performance Metrics includes the measurement of performance at layers
   between IP and the user. For example, network QoS metrics (packet
   loss, delay, and delay variation [RFC5481]) can be used to estimate
   application-specific Performance Metrics (de-jitter buffer size and
   RTP-layer packet loss), and then combined with other known aspects of
   a VoIP application (such as codec type) using an algorithm compliant
   with ITU-T P.564 [P.564] to estimate a Mean Opinion Score (MOS)
   [P.800]. However, the QoE for a particular VoIP user depends on the
   specific context, such as a casual conversation, a business
   conference call, or an emergency call. Finally, QoS and application-
   specific Performance Metrics are quantitative, while QoE is
   qualitative. Also, network QoS and application-specific Performance
   Metrics can be directly or indirectly evident to the user, while the
   QoE is directly evident.

5. Performance Metrics Development

   This section provides key definitions and qualifications of
   Performance Metrics.

5.1. Identifying and Categorizing the Audience

   Many of the aspects of metric definition and reporting, even the
   selection or determination of the essential metrics, depend on who
   will use the results, and for what purpose. For example, the metric
   description SHOULD include use cases and example reports that
   illustrate service quality monitoring and maintenance or
   identification and quantification of problems.

All documents defining Performance Metrics SHOULD identify the
primary audience and its associated requirements. The audience can
influence both the definition of metrics and the methods of
measurement.

The key areas of variation between different metric users include:

o Suitability of passive measurements of live traffic or active
   measurements using dedicated traffic

o Measurement in laboratory environment or on a network of deployed
   devices

o Accuracy of the results

o Access to measurement points and configuration information

o Measurement topology (point-to-point, point-to-multipoint)

o Scale of the measurement system

o Measurements conducted on-demand or continuously

o Required reporting formats and periods

o Sampling criteria [RFC5474], such as systematic or probabilistic

o Period (and duration) of measurement, as the live traffic can have
   patterns

5.2. Definitions of a Performance Metric

A Performance Metric is a measure of an observable behavior of a
networking technology, an application, or a service. Most of the
time, the Performance Metric can be directly measured; however,
sometimes, the Performance Metric value is computed. The process for
determining the value of a metric may assume an implicit or explicit
underlying statistical process; in this case, the Performance Metric
is an estimate of a parameter of this process, assuming that the
statistical process closely models the behavior of the system.

A Performance Metric should serve some defined purposes. This may
include the measurement of capacity, quantifying how bad some
problems are, measurement of service level, problem diagnosis or
location, and other such uses. A Performance Metric may also be an
input to some other processes, for example, the computation of a
composite Performance Metric or a model or simulation of a system.
Tests of the "usefulness" of a Performance Metric include:

   (i) the degree to which its absence would cause significant loss
   of information on the behavior or performance of the application
   or system being measured

   (ii) the correlation between the Performance Metric, the QoS, and
   the QoE delivered to the user (person or other application)

   (iii) the degree to which the Performance Metric is able to
   support the identification and location of problems affecting
   service quality

   (iv) the requirement to develop policies (Service Level Agreement,
   and potentially Service Level Contract) based on the Performance
   Metric

For example, consider a distributed application operating over a
network connection that is subject to packet loss. A Packet Loss
Rate (PLR) Performance Metric is defined as the mean packet loss
ratio over some time period. If the application performs poorly over
network connections with a high packet loss ratio and always performs
well when the packet loss ratio is zero, then the PLR Performance
Metric is useful to some degree. Some applications are sensitive to
short periods of high loss (bursty loss) and are relatively
insensitive to isolated packet loss events; for this type of
application, there would be very weak correlation between PLR and
application performance. A "better" Performance Metric would
consider both the packet loss ratio and the distribution of loss
events. If application performance is degraded when the PLR exceeds
some rate, then a useful Performance Metric may be a measure of the
duration and frequency of periods during which the PLR exceeds that
rate (as, for example, in RFC 3611).

5.3. Computed Performance Metrics

5.3.1. Composed Performance Metrics

Some Performance Metrics may not be measured directly, but can be
composed from base metrics that have been measured. A composed
Performance Metric is derived from other metrics by applying a
deterministic process or function (e.g., a composition function).
The process may use metrics that are identical to the metric being
composed, or metrics that are dissimilar, or some combination of both
types. Usually, the base metrics have a limited scope in time or
space, and they can be combined to estimate the performance of some
larger entities.

Some examples of composed Performance Metrics and composed
Performance Metric definitions are as follows:

Spatial composition is defined as the composition of metrics of
the same type with differing spatial domains [RFC5835] [RFC6049].
Ideally, for spatially composed metrics to be meaningful, the
spatial domains should be non-overlapping and contiguous, and the
composition operation should be mathematically appropriate for the
type of metric.

Temporal composition is defined as the composition of sets of
metrics of the same type with differing time spans [RFC5835]. For
temporally composed metrics to be meaningful, the time spans
should be non-overlapping and contiguous, and the composition
operation should be mathematically appropriate for the type of
metric.

Temporal aggregation is a summarization of metrics into a smaller
number of metrics that relate to the total time span covered by
the original metrics. An example would be to compute the minimum,
maximum, and average values of a series of time-sampled values of
a metric.

In the context of flow records in IP Flow Information eXport (IPFIX),
the IPFIX Mediation Framework [RFC6183], based on "IP Flow
Information Export (IPFIX) Mediation: Problem Statement" [RFC5982],
also discusses some aspects of the temporal and spatial composition.

5.3.2. Index

An index is a metric for which the output value range has been
selected for convenience or clarity, and the behavior of which is
selected to support ease of understanding, for example, the R Factor
[G.107]. The deterministic function for an index is often developed
after the index range and behavior have been determined.

5.4. Performance Metric Specification

5.4.1. Outline

A Performance Metric definition MUST have a normative part that
defines what the metric is and how it is measured or computed, and it
SHOULD have an informative part that describes the Performance Metric
and its application.

5.4.2. Normative Parts of Performance Metric Definition

   The normative part of a Performance Metric definition MUST define at
   least the following:

   (i) Metric Name

      Performance Metric names are RECOMMENDED to be unique within the
      set of metrics being defined for the protocol layer and context.
      While strict uniqueness may not be attainable (see the IPPM
      registry [RFC6248] for an example of an IANA metric registry
      failing to provide sufficient specificity), broad review must be
      sought to avoid naming overlap. Note that the Performance Metrics
      Directorate can help with suggestions for IANA metric registration
      for unique naming. The Performance Metric name MAY be
      descriptive.

   (ii) Metric Description

      The Performance Metric description MUST explain what the metric
      is, what is being measured, and how this relates to the
      performance of the system being measured.

   (iii) Method of Measurement or Calculation

      The method of measurement or calculation MUST define what is being
      measured or computed and the specific algorithm to be used. Does
      the measurement involve active or only passive measurements?
      Terms such as "average" should be qualified (e.g., running average
      or average over some interval). Exception cases SHOULD also be
      defined with the appropriate handling method. For example, there
      are a number of commonly used metrics related to packet loss;
      these often don't define the criteria by which a packet is
      determined to be lost (versus very delayed) or how duplicate
      packets are handled. For example, if the average PLR during a
      time interval is reported, and a packet's arrival is delayed from
      one interval to the next, then was it "lost" during the interval
      during which it should have arrived or should it be counted as
      received?

      Some methods of calculation might require discarding some data
      collected (due to outliers) so as to make the measurement
      parameters meaningful. One example is burstable billing that
      sorts the 5-min samples and discards the top 5 percentile.

Some parameters linked to the method MAY also be reported, in
order to fully interpret the Performance Metric, for example, the
time interval, the load, the minimum packet loss, the potential
measurement errors and their sources, the attainable accuracy of
the metric (e.g., +/- 0.1), the method of calculation, etc.

(iv) Units of Measurement

The units of measurement MUST be clearly stated.

(v) Measurement Point(s) with Potential Measurement Domain

If the measurement is specific to a measurement point, this SHOULD
be defined. The measurement domain MAY also be defined.
Specifically, if measurement points are spread across domains, the
measurement domain (intra-, inter-) is another factor to consider.

The Performance Metric definition should discuss how the
Performance Metric value might vary, depending on which
measurement point is chosen. For example, the time between a SIP
request [RFC3261] and the final response can be significantly
different at the User Agent Client (UAC) or User Agent Server
(UAS).

In some cases, the measurement requires multiple measurement
points: all measurement points SHOULD be defined, including the
measurement domain(s).

(vi) Measurement Timing

The acceptable range of timing intervals or sampling intervals for
a measurement, and the timing accuracy required for such
intervals, MUST be specified. Short sampling intervals or
frequent samples provide a rich source of information that can
help assess application performance but may lead to excessive
measurement data. Long measurement or sampling intervals reduce
the amount of reported and collected data such that it may be
insufficient to understand application performance or service
quality, insofar as the measured quantity may vary significantly
with time.

In the case of multiple measurement points, the potential
requirement for synchronized clocks must be clearly specified. In
the specific example of the IP delay variation application metric,
the different aspects of synchronized clocks are discussed in
[RFC5481].

5.4.3. Informative Parts of Performance Metric Definition

   The informative part of a Performance Metric specification is
   intended to support the implementation and use of the metric. This
   part SHOULD provide the following data:

   (i) Implementation

      The implementation description MAY be in the form of text, an
      algorithm, or example software. The objective of this part of the
      metric definition is to help implementers achieve consistent
      results.

   (ii) Verification

      The Performance Metric definition SHOULD provide guidance on
      verification testing. This may be in the form of test vectors, a
      formal verification test method, or informal advice.

   (iii) Use and Applications

      The use and applications description is intended to help the
      "user" understand how, when, and where the metric can be applied,
      and what significance the value range for the metric may have.
      This MAY include a definition of the "typical" and "abnormal"
      range of the Performance Metric, if this was not apparent from the
      nature of the metric. The description MAY include information
      about the influence of extreme measurement values, i.e., if the
      Performance Metric is sensitive to outliers. The Use and
      Application section SHOULD also include the security implications
      in the description.

      For example:

      (a) it is fairly intuitive that a lower packet loss ratio would
           equate to better performance. However, the user may not know
           the significance of some given packet loss ratio.

     (b) the speech level of a telephone signal is commonly expressed
          in dBm0. If the user is presented with:

          Speech level = -7 dBm0

          this is not intuitively understandable, unless the user
          is a telephony expert. If the metric definition explains
          that the typical range is -18 to -28 dBm0, a value higher
          than -18 means the signal may be too high (loud), and
          less than -28 means that the signal may be too low
          (quiet), it is much easier to interpret the metric.

   (iv) Reporting Model

      The reporting model definition is intended to make any
      relationship between the metric and the reporting model clear.
      There are often implied relationships between the method of
      reporting metrics and the metric itself; however, these are often
      not made apparent to the implementor. For example, if the metric
      is a short-term running average packet delay variation (e.g., the
      inter-arrival jitter in [RFC3550]) and this value is reported at
      intervals of 6-10 seconds, the resulting measurement may have
      limited accuracy when packet delay variation is non-stationary.

5.4.4. Performance Metric Definition Template

   Normative

      o Metric Name

      o Metric Description

      o Method of Measurement or Calculation

      o Units of Measurement

      o Measurement Point(s) with Potential Measurement Domain

      o Measurement Timing

   Informative

      o   Implementation

      o   Verification

      o   Use and Applications

      o   Reporting Model

5.4.5. Example: Loss Rate

   The example used is the loss rate metric as specified in RFC 3611
   [RFC3611].

   Metric Name: LossRate

   Metric Description: The fraction of RTP data packets from the source
      lost since the beginning of reception.

   Method of Measurement or Calculation: This value is calculated by
      dividing the total number of packets lost (after the effects of
      applying any error protection, such as Forward Error Correction
      (FEC)) by the total number of packets expected, multiplying the
      result of the division by 256, limiting the maximum value to 255
      (to avoid overflow), and taking the integer part.

   Units of Measurement: This metric is expressed as a fixed-point
      number with the binary point at the left edge of the field. For
      example, a metric value of 12 means a loss rate of
      approximately 5%.

   Measurement Point(s) with Potential Measurement Domain: This metric
      is made at the receiving end of the RTP stream sent during a Voice
      over IP call.

   Measurement Timing: This metric can be used over a wide range of
      time intervals. Using time intervals of longer than one hour may
      prevent the detection of variations in the value of this metric
      due to time-of-day changes in network load. Timing intervals
      should not vary in duration by more than +/- 2%.

   Implementation: The numbers of duplicated packets and discarded
      packets do not enter into this calculation. Since receivers
      cannot be required to maintain unlimited buffers, a receiver MAY
      categorize late-arriving packets as lost. The degree of lateness
      that triggers a loss SHOULD be significantly greater than that
      which triggers a discard.

   Verification: The metric value ranges between 0 and 255.

   Use and Applications: This metric is useful for monitoring VoIP
      calls, more precisely, to detect the VoIP loss rate in the
      network. This loss rate, along with the rate of packets discarded
      due to jitter, has some effect on the quality of the voice stream.

   Reporting Model: This metric needs to be associated with a defined
      time interval, which could be defined by fixed intervals or by a
      sliding window. In the context of RFC 3611, the metric is
      measured continuously from the start of the RTP stream, and the
      value of the metric is sampled and reported in RTCP XR VoIP
      Metrics reports.

5.5. Dependencies

   This section introduces several Performance Metrics dependencies,
   which the Performance Metric designer should keep in mind during
   Performance Metric development. These dependencies, and any others
   not listed here, SHOULD be documented in the Performance Metric
   specifications.

5.5.1. Timing Accuracy

   The accuracy of the timing of a measurement may affect the accuracy
   of the Performance Metric. This may not materially affect a sampled-
   value metric; however, it would affect an interval-based metric.
   Some metrics -- for example, the number of events per time interval
   -- would be directly affected; for example, a 10% variation in time
   interval would lead directly to a 10% variation in the measured
   value. Other metrics, such as the average packet loss ratio during
   some time interval, would be affected to a lesser extent.

   If it is necessary to correlate sampled values or intervals, then it
   is essential that the accuracy of sampling time and interval start/
   stop times is sufficient for the application (for example, +/- 2%).

5.5.2. Dependencies of Performance Metric Definitions on Related Events or Metrics

   Performance Metric definitions may explicitly or implicitly rely on
   factors that may not be obvious. For example, the recognition of a
   packet as being "lost" relies on having some method of knowing the
   packet was actually lost (e.g., RTP sequence number), and some time
   threshold after which a non-received packet is declared lost. It is
   important that any such dependencies are recognized and incorporated
   into the metric definition.

5.5.3. Relationship between Performance Metric and Lower-Layer
        Performance Metrics

   Lower-layer Performance Metrics may be used to compute or infer the
   performance of higher-layer applications, potentially using an
   application performance model. The accuracy of this will depend on
   many factors, including:

      (i) The completeness of the set of metrics (i.e., are there
      metrics for all the input values to the application performance
      model?)

      (ii) Correlation between input variables (being measured) and
      application performance

      (iii) Variability in the measured metrics and how this variability
      affects application performance

5.5.4. Middlebox Presence

   Presence of a middlebox [RFC3303], e.g., proxy, network address
   translation (NAT), redirect server, session border controller (SBC)
   [RFC5853], and application layer gateway (ALG), may add variability
   to or restrict the scope of measurements of a metric. For example,
   an SBC that does not process RTP loopback packets may block or
   locally terminate this traffic rather than pass it through to its
   target.

5.6. Organization of Results

   The IPPM Framework [RFC2330] organizes the results of metrics into
   three related notions:

   o singleton: an elementary instance, or "atomic" value.

   o sample: a set of singletons with some common properties and some
      varying properties.

   o statistic: a value derived from a sample through deterministic
      calculation, such as the mean.

   Performance Metrics MAY use this organization for the results, with
   or without the term names used by the IPPM WG. Section 11 of
   RFC 2330 [RFC2330] should be consulted for further details.

5.7. Parameters: the Variables of a Performance Metric

   Metrics are completely defined when all options and input variables
   have been identified and considered. These variables are sometimes
   left unspecified in a metric definition, and their general name
   indicates that the user must set and report them with the results.
   Such variables are called "parameters" in the IPPM metric template.
   The scope of the metric, the time at which it was conducted, the
   length interval of the sliding-window measurement, the settings for
   timers, and the thresholds for counters are all examples of
   parameters.

   All documents defining Performance Metrics SHOULD identify all key
   parameters for each Performance Metric.

6. Performance Metric Development Process

6.1. New Proposals for Performance Metrics

   This process is intended to add more considerations to the processes
   for adopting new work as described in RFC 2026 [RFC2026] and RFC 2418
   [RFC2418]. Note that new Performance Metrics work item proposals
   SHALL be approved using the existing IETF process. The following
   entry criteria will be considered for each proposal.

   Proposals SHOULD be prepared as Internet-Drafts, describing the
   Performance Metric and conforming to the qualifications above as much
   as possible. Proposals SHOULD be deliverables of the corresponding
   protocol development WG charters. As such, the proposals SHOULD be
   vetted by that WG prior to discussion by the Performance Metrics
   Directorate. This aspect of the process includes an assessment of
   the need for the Performance Metric proposed and assessment of the
   support for its development in the IETF.

   Proposals SHOULD include an assessment of interaction and/or overlap
   with work in other Standards Development Organizations (SDOs).
   Proposals SHOULD identify additional expertise that might be
   consulted.

   Proposals SHOULD specify the intended audience and users of the
   Performance Metrics. The development process encourages
   participation by members of the intended audience.

   Proposals SHOULD identify any security and IANA requirements.
   Security issues could potentially involve revealing data identifying
   a user, or the potential misuse of active test tools. IANA
   considerations may involve the need for a Performance Metrics
   registry.

6.2. Reviewing Metrics

   Each Performance Metric SHOULD be assessed according to the following
   list of qualifications:

   o Are the performance metrics unambiguously defined?

   o Are the units of measure specified?

   o Does the metric clearly define the measurement interval where
      applicable?

   o Are significant sources of measurement errors identified and
      discussed?

   o Does the method of measurement ensure that results are repeatable?

   o Does the metric or method of measurement appear to be
      implementable (or offer evidence of a working implementation)?

   o Are there any undocumented assumptions concerning the underlying
      process that would affect an implementation or interpretation of
      the metric?

   o Can the metric results be related to application performance or
      user experience, when such a relationship is of value?

   o Is there an existing relationship to metrics defined elsewhere
      within the IETF or within other SDOs?

   o Do the security considerations adequately address denial-of-
      service attacks, unwanted interference with the metric/
      measurement, and user data confidentiality (when measuring live
      traffic)?

6.3. Performance Metrics Directorate Interaction with Other WGs

   The Performance Metrics Directorate SHALL provide guidance to the
   related protocol development WG when considering an Internet-Draft
   that specifies Performance Metrics for a protocol. A sufficient
   number of individuals with expertise must be willing to consult on
   the draft. If the related WG has concluded, comments on the proposal
   should still be sought from key RFC authors and former chairs.

   As with expert reviews performed by other directorates, a formal
   review is recommended by the time the document is reviewed by the
   Area Directors or an IETF Last Call is being conducted.

   Existing mailing lists SHOULD be used; however, a dedicated mailing
   list MAY be initiated if necessary to facilitate work on a draft.

   In some cases, it will be appropriate to have the IETF session
   discussion during the related protocol WG session, to maximize
   visibility of the effort to that WG and expand the review.

6.4. Standards Track Performance Metrics

   The Performance Metrics Directorate will assist with the progression
   of RFCs along the Standards Track. See [IPPM-STANDARD-ADV-TESTING].
   This may include the preparation of test plans to examine different
   implementations of the metrics to ensure that the metric definitions
   are clear and unambiguous (depending on the final form of the draft
   mentioned above).

7. Security Considerations

   In general, the existence of a framework for Performance Metric
   development does not constitute a security issue for the Internet.
   Performance Metric definitions, however, may introduce security
   issues, and this framework recommends that persons defining
   Performance Metrics should identify any such risk factors.

   The security considerations that apply to any active measurement of
   live networks are relevant here. See [RFC4656].

   The security considerations that apply to any passive measurement of
   specific packets in live networks are relevant here as well. See the
   security considerations in [RFC5475].

8. Acknowledgements

   The authors would like to thank Al Morton, Dan Romascanu, Daryl
   Malas, and Loki Jorgenson for their comments and contributions, and
   Aamer Akhter, Yaakov Stein, Carsten Schmoll, and Jan Novak for their
   reviews.

9. References

9.1. Normative References

   [RFC2026] Bradner, S., "The Internet Standards Process --
              Revision 3", BCP 9, RFC 2026, October 1996.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC2418] Bradner, S., "IETF Working Group Guidelines and
              Procedures", BCP 25, RFC 2418, September 1998.

   [RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M.
              Zekauskas, "A One-way Active Measurement Protocol
              (OWAMP)", RFC 4656, September 2006.

9.2. Informative References

   [E.800]   "ITU-T Recommendation E.800. E SERIES: OVERALL NETWORK
              OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN
              FACTORS", September 2008.

   [G.107]   "ITU-T Recommendation G.107. The E-model: a computational
              model for use in transmission planning", April 2009.

   [IPPM-STANDARD-ADV-TESTING]
              Geib, R., Ed., Morton, A., Fardid, R., and A. Steinmitz,
              "IPPM standard advancement testing", Work in Progress,
              June 2011.

   [P.564]   "ITU-T Recommendation P.564. Conformance Testing for
              Voice over IP Transmission Quality Assessment Models",
              November 2007.

   [P.800]   "ITU-T Recommendation P.800. Methods for subjective
              determination of transmission quality", August 1996.

   [RFC0793] Postel, J., "Transmission Control Protocol", STD 7,
              RFC 793, September 1981.

   [RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis,
              "Framework for IP Performance Metrics", RFC 2330,
              May 1998.

   [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
              A., Peterson, J., Sparks, R., Handley, M., and E.
              Schooler, "SIP: Session Initiation Protocol", RFC 3261,
              June 2002.

   [RFC3303] Srisuresh, P., Kuthan, J., Rosenberg, J., Molitor, A., and
              A. Rayhan, "Middlebox communication architecture and
              framework", RFC 3303, August 2002.

   [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.
              Jacobson, "RTP: A Transport Protocol for Real-Time
              Applications", STD 64, RFC 3550, July 2003.

   [RFC3611] Friedman, T., Ed., Caceres, R., Ed., and A. Clark, Ed.,
              "RTP Control Protocol Extended Reports (RTCP XR)",
              RFC 3611, November 2003.

   [RFC4710] Siddiqui, A., Romascanu, D., and E. Golovinsky, "Real-time
              Application Quality-of-Service Monitoring (RAQMON)
              Framework", RFC 4710, October 2006.

   [RFC4960] Stewart, R., Ed., "Stream Control Transmission Protocol",
              RFC 4960, September 2007.

   [RFC5101] Claise, B., Ed., "Specification of the IP Flow Information
              Export (IPFIX) Protocol for the Exchange of IP Traffic
              Flow Information", RFC 5101, January 2008.

   [RFC5102] Quittek, J., Bryant, S., Claise, B., Aitken, P., and J.
              Meyer, "Information Model for IP Flow Information Export",
              RFC 5102, January 2008.

   [RFC5474] Duffield, N., Ed., Chiou, D., Claise, B., Greenberg, A.,
              Grossglauser, M., and J. Rexford, "A Framework for Packet
              Selection and Reporting", RFC 5474, March 2009.

   [RFC5475] Zseby, T., Molina, M., Duffield, N., Niccolini, S., and F.
              Raspall, "Sampling and Filtering Techniques for IP Packet
              Selection", RFC 5475, March 2009.

   [RFC5481] Morton, A. and B. Claise, "Packet Delay Variation
              Applicability Statement", RFC 5481, March 2009.

   [RFC5706] Harrington, D., "Guidelines for Considering Operations and
              Management of New Protocols and Protocol Extensions",
              RFC 5706, November 2009.

   [RFC5835] Morton, A., Ed., and S. Van den Berghe, Ed., "Framework
              for Metric Composition", RFC 5835, April 2010.

   [RFC5853] Hautakorpi, J., Ed., Camarillo, G., Penfield, R.,
              Hawrylyshen, A., and M. Bhatia, "Requirements from Session
              Initiation Protocol (SIP) Session Border Control (SBC)
              Deployments", RFC 5853, April 2010.

   [RFC5982] Kobayashi, A., Ed., and B. Claise, Ed., "IP Flow
              Information Export (IPFIX) Mediation: Problem Statement",
              RFC 5982, August 2010.

   [RFC6035] Pendleton, A., Clark, A., Johnston, A., and H. Sinnreich,
              "Session Initiation Protocol Event Package for Voice
              Quality Reporting", RFC 6035, November 2010.

   [RFC6049] Morton, A. and E. Stephan, "Spatial Composition of
              Metrics", RFC 6049, January 2011.

   [RFC6183] Kobayashi, A., Claise, B., Muenz, G., and K. Ishibashi,
              "IP Flow Information Export (IPFIX) Mediation: Framework",
              RFC 6183, April 2011.

   [RFC6248] Morton, A., "RFC 4148 and the IP Performance Metrics
              (IPPM) Registry of Metrics Are Obsolete", RFC 6248,
              April 2011.

Authors' Addresses

   Alan Clark
   Telchemy Incorporated
   2905 Premiere Parkway, Suite 280
   Duluth, Georgia 30097
   USA

   EMail: alan.d.clark@telchemy.com

   Benoit Claise
   Cisco Systems, Inc.
   De Kleetlaan 6a b1
   Diegem 1831
   Belgium

   Phone: +32 2 704 5622
   EMail: bclaise@cisco.com