Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 27 April 2022

G. Mirsky
J. Halpern
Ericsson
X. Min
ZTE Corp.
L. Han
China Mobile

24 October 2021

Error Performance Measurement in Packet-switched Networks
draft-mirsky-ippm-epm-04

Abstract

This document describes the use of the error performance metric to characterize a packet-switched network's conformance to the pre-defined set of performance objectives. In this document, metrics that characterize error performance in a packet-switched network (PSN) are defined, as well as methods to measure and calculate them. Also, the requirements for an active Operation, Administration, and Maintenance protocol to support the error performance measurement in PSN are discussed, and potential candidate protocols are analyzed. All metrics and measurement methods are equally applicable to underlay and overlay networks.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Operations, Administration, and Maintenance (OAM) is a collection of methods to detect, characterize, localize failures in a network, and monitor the network's performance using various measurement methods. Traditionally, the former set of OAM tools identified as Fault Management (FM) OAM. The latter - Performance Monitoring (PM) OAM. Some OAM protocols can be used for both groups of tasks, while some serve one particular group. But regardless of how many OAM protocols are in use, network operators and network users are faced with multiple metrics that characterize the network conditions. This document describes a new component of packet-switched network (PSN) OAM.

Error performance measurement (EPM) is a part of an OAM toolset that provides an operator with information related to network measurements for a uni-directional or a bidirectional connection between two systems. In current technology, EPM has been defined only for data communication methods that have a constant bit-rate transmission

[ITU.G.826] and not for PSN, where transmissions are statistically random. As a statistically multiplexed network in a PSN, a receiver node does not expect a packet to arrive from a sender node at a specific moment, less from a particular sender. That is what differentiates PSN from networks built on a constant bit-rate transmission, where a stream of bits between two nodes is always present, whether it represents data or not. That provides the receiver with a predictable number of measurements in a series of measurement intervals. In PSN, on-path OAM methods, i.e., measurement methods that use data flow, cannot provide such predictability and thus be used for EPM. In PSN, EPM needs to use active OAM methods, per definition in [RFC7799]. This document identifies metrics that characterize PSN error performance and methods to measure and calculate them. Also, the requirements for an active OAM protocol to support EPM in PSN are discussed, and potential candidate protocols are analyzed.

2. Conventions used in this document

2.1. Terminology and Acronyms

OAM Operations, Administration, and Maintenance

EP Error Performance

EPM Error Performance Measurement

ES Errored Second

ESR Errored Second Ratio

SES Severely Errored Second

SESR Severely Errored Second Ratio

EFS Error-Free Second

PSN Packet-switched Network

FM Fault Management

PM Performance Monitoring

2.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Error Performance Metrics

When analyzing the error performance of a path between two nodes, we need to select a time interval as the unit of EPM. In [ITU.G.826], a time interval of one second is used. It is reasonable to use the same time interval for EPM for PSNs. Further, for the purpose of EPM, each time interval, i.e., second, is classified either as Errored Second (ES), Severely Errored Second (SES), or Error-Free Second (EFS). These are defined as follows:

* An ES is a time interval during which at least one of the performance parameters degraded below its optimal level threshold or a defect was detected.

* An SES is a time interval during which at least one the performance parameters degraded below its critical threshold or a defect was detected.

* Consequently, an EFS is a time interval during which all performance objectives are at or above their respective optimal levels, and no defect has been detected.

The definition of a state of a defect in the network is also necessary for understanding the EPM. In this document, the defect is interpreted as the state of inability to communicate between a particular set of nodes. It is important to note that it is being defined as a state, and thus, it has conditions that define entry into it and exit out of it. Also, the state of defect exists only in connection to the particular group of nodes in the network, not the network as a domain.

3.1. Measure Error Performance Metrics

The definitions of ES, SES, and EFS allow for characterization of the communication between two nodes relative to the level of required and acceptable performance and when performance degrades below the acceptable level. The former condition in this document referred to as network availability. The latter - network unavailability. Based on the definitions, SES is the one-second of network unavailability while ES and EFS present an interval of network availability. But since the conditions of network are everchanging periods of network availability and unavailability need to be defined with duration larger than one-second interval to reduce the number of state changes while correctly reflecting the network condition. The method to determine the state of the network in terms of EPM OAM is described below:

* If ten consecutive SES intervals been detected, then the EPM state of the network determined as unavailability and the beginning of that period of unavailability state is at the start of the first SES in the sequence of the consecutive SES intervals.

* Similarly, ten consecutive non-SES intervals, i.e., either ES or EFS, indicate that the network is in the availability period, i.e., available. The start of that period is at the beginning of the first non-SES interval.

* Resulting from these two definitions, a sequence of less than ten consecutive SES or non-SES intervals does not change the EPM state of the network. For example, if the EPM state is determined as unavailability, a sequence of seven EFS intervals is not viewed as an availability period.

3.2. Calculate Error Performance Metrics

Determining the period in which the path is currently EP-wise is helpful. But because switching between periods requires ten consecutive one-second intervals, conditions that last shorter intervals may not be adequately reflected. Two additional EP OAM metrics can be used, and they are defined as follows:

* errored second ratio (ESR) is the ratio of ES to the total number of seconds in a time of the availability periods during a fixed measurement interval.

* severely errored second ratio (SESR) - is the ratio of SES to the total number of seconds in a time of the availability periods during a fixed measurement interval.

4. Requirements to EPM

TBA

5. Active OAM Protocol for EPM

Digital communication methods characterized as the constant-bit rate digital paths and connections allow measurement of the error performance without using an active OAM. That is possible because a predictable flow of digital signals is expected at an egress system. That is not the case for packet-switched networks that are based on the principle of statistical multiplexing flows. The latter usually improves the utilization of the communication network's resources, but it also makes the flow unpredictable for the egress system. For that reason, an active OAM has to be used in measuring the error performance in a network. A combination of OAM protocols can provide the necessary for EPM functionality. For example, Bidirectional Forwarding Detection (BFD) [RFC5880] can be used to monitor the continuity of a path between the ingress and egress systems. And STAMP [RFC8762] can be used to measure and calculate performance metrics that are used as Service Level Objectives. But using two protocols and correlating the state of the network from them adds to the complexity in network operation.

6. Availability of Anything-as-a-Service

Anything as a service (XaaS) describes a general category of services related to cloud computing and remote access. These services include the vast number of products, tools, and technologies that are delivered to users as a service over the Internet. In this document, the availability of XaaS is viewed as the ability to access the service over a period of time with pre-defined performance objectives. Among the advantages of the XaaS model are:

* Improving the expense model by purchasing services from providers on a subscription basis rather than buying individual products, e.g., software, hardware, servers, security, infrastructure, and install them on-site, and then link everything together to create networks.

* Speeding new apps and business processes by quickly adapting to changing market conditions with new applications or solutions.

* Shifting IT resources to specialized higher-value projects that use the core expertise of the company.

But XaaS model also has potential challenges:

* Possible downtime resulting from issues of internet reliability, resilience, provisioning, and managing the infrastructure resources.

* Performance issues caused by depleted resources like bandwidth, computing power, inefficiencies of virtualized environments, ongoing management and security of multi-cloud services.

* Complexity impacts enterprise IT team that must remain in the process of the continued learning of the provided services.

The framework and metrics of the EPM defined in Section 3 allow a provider of XaaS and their customers to quantify, measure, monitor for conformance what is often referred to as an ephemeral - availability of the service to be delivered. There are other definitions and methods of expressing availability. For example, [HighAvailability-WP] uses the following equation:

Availability Average = MTBF/(MTBF + MTRR),
where:
MTBF (Mean Time Between Failures) - mean time between individual component failures. For example, a hard drive malfunction or hypervisor reboot.
MTTR (Mean Time To Repair) - refers to how long it takes to fix the broken component or the application to come back online,

While this approach estimates the expected availability of a XaaS, the EPM reflects near-real-time availability of a service as experienced by a user. It also provides valuable data for more accurate and realistic MTBF and MTTR in the particular environment, and simplifies comparison of different solutions that may use redundant servers (web and database), load balancers.

In another field of communication, mobile voice and data services, the definition of service availability is understood as "the probability of successful service reception: a given area is declared "in-coverage" if the service in that area is available with a pre-specified minimum rate of success. Service availability has the advantage of being more easily understandable for consumers and is expressed as a percentage of the number of attempts to access a given service." [BEREC-CP]. The definition of the availability used in the EPM throughout this document is close to the quoted above. It might be considered as the extension that allows regulators, operators, and consumers to compare not only the rate of successfully establishing a connection but the quality of the connection during its lifetime.

7. IANA Considerations

TBA

8. Security Considerations

TBA

9. Acknowledgments

TBA

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

10.2. Informative References

[BEREC-CP] Body of European Regulators for Electronic Communications, "BEREC Common Position on information to consumers on mobile coverage", Common Approaches/Positions BoR (18) 237, June 2018, <https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/common_approaches_positions/8315-berec-common-position-on-information-to-consumers-on-mobile-coverage>.

[HighAvailability-WP] Avi Freedman, Server Central, "High Availability in Cloud and Dedicated Infrastructure", <https://www.deft.com/wp-content/uploads/pdf/SCTG-High-Availability-White-Paper-Part-2.pdf>.

[ITU.G.826] ITU-T, "End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections", ITU-T G.826, December 2002.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

Authors' Addresses

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com

Joel Halpern
Ericsson

Email: joel.halpern@ericsson.com

Xiao Min
ZTE Corp.

Email: xiao.min2@zte.com.cn

Liuyan Han
China Mobile
32 XuanWuMenXi Street
Beijing
100053
China

Email: hanliuyan@chinamobile.com