Quality of Resilience as a Network Reliability Characterization Tool

Piotr Cholda, Krzysztof Wajda, and Andrzej Jajszczyk,
AGH University of Science and Technology
János Tapolcai and Tibor Cinkler, Budapest University of Technology and Economics

Abstract

With the increased role of resilience in modern networks, the existing quality of service is required to be expanded with service availability and maintainability. Recently, studies have shown the strong limitation of the common availability metrics for measuring the user's quality of experience. In this article a joint specification of QoS definitions with a sophisticated service resilience characterization is proposed, and a concept called quality of resilience is defined. In this unified performance metric, the frequency and length of service interruption are evaluated. It can be used as a tool for characterization of network reliability, as well as comparison and selection of recovery methods. Additionally, by including it in service level agreements, new and more complex requirements of commercial applications can be guaranteed.

Contemporary multilayer networks containing the optical-based transport layer must be protected against failures (faults), as any interruption could cause large loss of data. Thus, the networks are equipped with relevant mechanisms enabling survival under network failures. They are known as recovery methods, providing networks with *resilience*, a general ability to improve network fault tolerance and, as a result, its reliability. An intelligent setup of connections and reactions to faults opens new possibilities and challenges in service differentiation offered by network operators. One of the latter is a proper quantification of resilience.

Recovery mechanisms automatically redirect traffic from working routes affected by failures to bypassing fault-free recovery routes. A sequence of operations is necessary to perform this task: fault detection, fault localization, fault notification, and recovery switching. Although all of them influence quality, the last one can be dominant. A classification of recovery methods can be performed based on the following criteria:

- **Layers in which recovery operates**:
  - –Single-layer-based: in lower layers — fast but usually expensive, or in higher layers — slower but potentially cheaper
  - –Multiple-layer-based: not coordinated — simple but costly, or coordinated—potentially cheaper but complex
- **Recovery resources setup method**:
  - –Computed on demand: flexible but slow, generally known as restoration methods, default for contemporary IP networks
  - –Precomputed: robust and fast, but rather costly, called protection methods, typical for fixed transport networks
- **Recovery resources sharing level**:
  - –Dedicated: very costly but fast
  - –Shared: quite robust and with reasonable cost
  - –With no special resources reservation relevant for restoration: flexible and cost-efficient but slow
- **Scope of recovery**:
  - –Local (single link, node): fast but involving complex optimization and potentially expensive
  - –Global (path): slower, easier to optimize
  - –Segment: intermediate between the above two
- **Domains crossed by recovery**:
  - –Single-domain: fast
  - –Multidomain: slower, hardly enabling quality control

Combinations of the above result in different quality outputs. The operation of those methods is described in several monographs [1, 2]. Although the classifications enable us to elegantly set different approaches, the problem of proper selection of a method arises when an operator is faced with a cost/quality trade-off. To solve that, quantifying metrics must be defined to measure the way a selected method affects the network and applications within. Many measures have been proposed for different environments. The comparison and definitions are given in Table 1.

In this article we deal with the comprehensive characterization of various recovery methods implemented in multilayer networks. The main focus is quantifying the ability of those mechanisms to support continuity in the client applications. The goal is to be able to point out which recovery option is the most appropriate for a specific set of quality requirements by looking at the network as a "black box." Usually, only the steady-state availability or a limit on downtime are used in service level agreements (SLAs) to adequately express the predicted level of resilience when viewed from the perspective of a higher layer. However, we claim that layer-specific steady-state measures fail to provide an appropriate measure for quality of service (QoS) in multilayer networks. This stems from the fact that many interacting mechanisms influence resilience perceived by the users or their applications. They

0890-8044/09/$25.00 © 2009 IEEE

Id	Area	Main metrics	Description, comments, definition of more important metrics
International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Recommendations			
E.800 E.802 E.820 E.850 E.855 E.860 E.862 E.880	Telephone network, ISDN, general (e.g., Internet access)	• Retainability • (Mean) time between interruptions (*MTBI*) • Down time (*MDT*), up time (*MUT*) • Instantaneous (un)availability, steady-state/asymptotic (un)availability (*U*/*A*) • Reliability function ($R(t)$) • Time to failure (*MTTF*) • Time between failures (*MTBF*) • Time to recovery (*MTTR*) • p-fractile repair time • Failure/repair rate ($\lambda(t)/\mu(t)$) • Probability of fault coverage	• The most general recommendations on resilience/QoS terminology • *Retainability*: "probability that a service will continue to be provided" • *Time between interruptions*: "time duration between the end of one interruption and the beginning of the next" • *Down/up time*: "time interval during which an item is in a down/up state" • *Instantaneous availability/unavailability*: "probability that an item is in an up/down state at a given instant of time" • $A = MUT/(MUT + MDT)$, limit of the instantaneous availability • $R(t)$: "probability that an item can perform a required function under stated conditions for a given time interval" • *Time to failure*: " time duration of an item, from the instant of time it goes from a down state to an up state until the next failure" • *(Operating) time between failures*: "time duration between two successive failures of a repaired item" • *Time to recovery (repair)*: "time interval during which an item is in a down state due to a failure"
G.911	Fiber optic systems	• *U*, *A*, *MTBF*, *MTTR*, $\lambda(t)$ • Median life • Standard deviation • Failures in time (*FIT*)	• *Median life*: "point on a lognormal probability plot of time to failure at which 50% of the devices fail earlier and 50% of the devices fail later" • *Standard deviation*: "standard deviation of the natural logarithms of the time to failure" • *FIT*: number of failures per billion device hours
I.350 I.357	ISDN, B-ISDN	• *A* • Dependability • Time between outages (*MTBO*), similar to *MTBI*	• *Dependability*: performance criterion describing the degree of certainty with which the device operation is performed within a given observation interval • Down state is recognized on the basis of thresholds related to: cell loss ratio or severely errored cell block ratio
M.60 M.3342	General	• Retainability, *A*, $R(t)$, *MTBF*, *MTTR*, *MTBO* • Time to restore service (*MTRS*)	• *MTRS* is understood similar to *MTTR* but instead at the level of physical repair, it concerns the logical/service level
M.1301	SDH	• *MTBF*, *MTTR*, *MTRS*	• SDH networks resilience metrics
P.10	Telephone network, general	• Mean opinion score (MOS) • Quality of experience (QoE)	• Mean opinion score is a subjective measurement of the quality. It is used in a survey-based studies when a service is tested by users • QoE: "overall acceptability of an application or service, as perceived subjectively by the end user"
X.641	Data networks	• *A* • Reliability characteristic	• *Reliability characteristic*: "MTBF to maintain a defined QoS requirement"
Y.1540 Y.1541 Y.1542	IP	• IP packet loss ratio (*IPLR*) • Service availability • Percent IP service (un)availability (*PIU*/*PIA*)	• *IPLR*: "ratio of total lost IP packet outcomes to total transmitted IP packets in a population of interest" • *Service availability*: "classifies the total scheduled service time for an IP service into available and unavailable periods," using the threshold on *IPLR* • *PIU*/*PIA*: "percentage of total scheduled IP service time categorized as (un)available using the IP service availability function"
Y.1561	MPLS	• Packet loss ratio • Severe loss block (*SLB*) outcome • Recovery time • Availability service, *PIU*, *PIA*	• *Packet loss ratio* is understood analogously to *IPLR* • *SLB* outcome: "occurs for a block of packets at ingress node when the ratio of lost packets at egress node exceeds some threshold" • *Recovery time*: "count of successive time intervals that form a consecutive *SLB* outcome at ingress node" • *Availability service*, *PIU*, and *PIA* are defined similarly as in Y.1540, but now on the basis of *SLB*
Y.1562	Higher layer protocols	• Service availability	• *Service availability* is defined as in Y.1540, but here it is related to the service transfer delay and service success ratio

Continued on next page...

■ Table 1. *Measurable metrics for resilience quantification. (Table 1 continued on next page.)*

Id	Area	Main metrics	Description, comments, definition of more important metrics
Internet Engineering Task Force (IETF) Requests for Comments (RFCs)			
2330	IP	• Packet loss rate	• Determination of IP-related performance metrics definitions/measurement • Advises resignation from probabilistic metric definitions in favor of the deterministic ones • *Packet loss rate*, analogous to *IPLR*, is used as an example
3386	Multilayer networks	• Protection switch time • Restoration time	• *Protection switch time*: "time interval from the occurrence of a network fault until the completion of the protection-switching operations" • *Restoration time*: "time interval from the occurrence of a network fault to the instant when the affected traffic is either completely restored, or until spare resources are exhausted, or no more extra traffic exists" • Definitions show the difference in the approaches of ITU-T and IETF, where the former is more general, and the latter more focused on particular methods
3469 4378	MPLS	• Loss • Recovery time • Full restoration time • Setup vulnerability • Number of concurrent faults • Percentage of coverage • Availability	• *Recovery time*: "time required for a recovery path to be activated (and traffic flowing) after a fault" • *Full restoration time*: "time required for traffic to be routed onto links, which are capable of or have been engineered sufficiently to handle traffic in recovery scenarios" • *Setup vulnerability*: "amount of time that a working path is left unprotected during such tasks as recovery path computation and recovery path setup" • *Number of concurrent faults/percentage of coverage* determines a number of failures/ratio of faults that can be covered by a selected recovery scheme • *Availability* is defined as "measure of the percentage of time that a service is operating within a specification," giving a practical consequence of the definition of A as in E.800
3945 4427 4428	GMPLS	• Recovery time (down time) • Recovery ratio	*Recovery ratio*: "quotient of the actual recovery bandwidth divided by the traffic bandwidth that is intended to be protected"
Except for $R(t)$, $\lambda(t)$, and $\mu(t)$, most common denotations for *mean* values are given as abbreviations; see, for instance, *MTBI*.			

■ Table 1 (continued). *Measurable metrics for resilience quantification.*

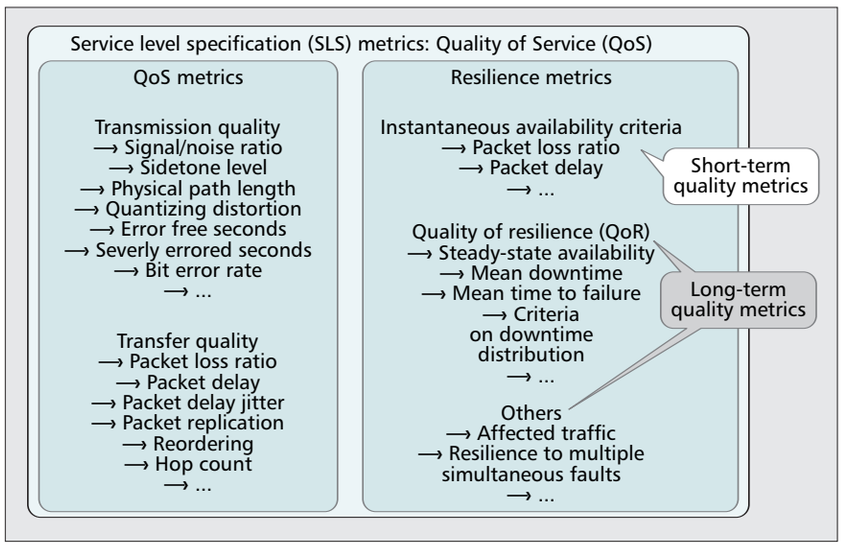
cannot be captured with a too simple description based on availability, but rather on some probability density function [3]. Here, we present a novel method to quantify resilience in today's multilayered multiservice networks as better matching the needs of both operators and clients. Therefore, we lump the behavior of multiple layers in a single metric describing a downtime histogram from the perspective of a higher layer. To do that, we propose advanced QoS definitions, which include service resilience measures. This novel concept is called *quality of resilience* (QoR). The next section overviews the relation between the QoR approach and QoS metrics. Subsequently, we present the idea of recovery characterization with the QoR downtime histogram. Finally, a detailed numerical example dealing with optical layer recovery methods is given.

Relationship between QoS and QoR

The goal of a carrier is to provide a service to its customer. It is supported by settlement of an agreement satisfactory to both parties. The aim is to meet the complex requirements of profit earning applications. The impact of resilience on service quality and liability issues has gained more attention [2]. The necessity to treat resilience as one of the factors for selected services or customers has appeared. Consequently, the need to directly compare and characterize recovery mechanisms has become more urgent. The idea that a customer should be aware of the risk related to failures and actively participate in the process of defining recovery conditions is also highly relevant. We call it "resilience risk/responsibility sharing." We believe that in future networks a larger amount of information on network status should be transferred from an operator to its customer.

QoS is the umbrella representing the measurable (objective) requirements of the users regarding the service. According to ITU-T Recommendation E.800, it can be partitioned to transmission/transfer quality, dependability (resilience) performance, and, additionally, security performance, which is not considered here. Figure 1 shows this model with our extensions. It forms a basis for the service level specification (SLS), the technical part of SLAs. There are substantial differences between those groups. The first one has been defined on the basis of metrics such as bit error rate (BER), delay, packet loss probability, available bandwidth, traffic load, and throughput. They can be directly perceived by users. In this article we refer to them collectively as *QoS metrics*. If they encompass a transport layer, they are called "transmission quality," whereas for a service layer they are known as "transfer quality." They are quantified according to a relevant performance model (e.g., E-model defined in ITU-T Recommendation G.107 for voice over IP [VoIP]).

In SLAs the long- and short-term quality characteristics can



■ Figure 1. *Service class metrics in SLA.*

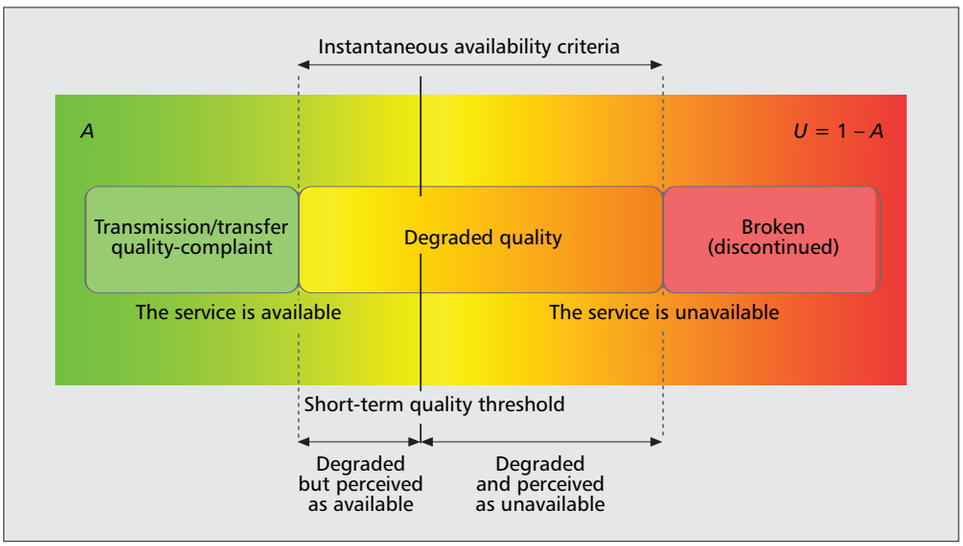
be defined. The short-term quality metrics are related to the instant perceived quality of the service provided to the user, while the long-term measure is the overall service quality during the whole length of service operation. While the QoS metrics group captures short-term characteristics, we note that most of the resilience measures are long-term in nature. Additionally, the QoS metrics can be precisely measured, while resilience, in many cases, is not perceived by a user directly. Moreover, it cannot be distinguished just from quality degradation. For instance, an end user recognizes increased delay or packet loss in the service layer leading to a certain level of TCP throughput degradation, resulting in longer response times. The reason for those impairments might be either network congestion due to uncontrolled traffic variability, or even a hardware failure. Failures usually have a larger impact than pure congestion and are often more severe, as they may cause total packet loss over a considerable period. Only in some cases, such as in a virtual private network (VPN) service, will the user be influenced by fault directly (i.e., a complete, perceptible breakdown). Additionally, as was shown in [4] for voice and video streaming, the frequency of short time service interruption significantly influences the quality of experience (QoE), which is the perceived quality of the multimedia stream. As networks can recover failures in a few tens of milliseconds, simply measuring service availability by the number of downtime minutes per year is far from sufficient to estimate the possible quality degradation for the mentioned case. To unambiguously quantify both QoE and the effects of resilience, some kind of separation in the short-and long-term quality metrics must be carried out. Clearly, resilience can be measured only with long-term quality measures based on end-to-end measurements. They can be performed for certain network layers separately, where each layer contains the characteristics of layers beneath itself. There is no commonly adopted approach that can be tailored to contemporary multilayer networks. On the basis of existing ITU-T Recommendations related to telephony — synchronous digital hierarchy (SDH), integrated services digital network (ISDN), or IP networks (Table 1) — practically only steady-state availability or mean recovery/repair time are used as resilience metrics. However, they are neither general nor flexible enough.

With the separation of short- and long-term quality metrics, we can unambiguously quantify the effects of resilience and cover a whole range of different *resilience metrics* as a large dimension of SLAs. Although they have long-term significance, in practice they cannot be measured or interpreted without the short-term quality metrics [5]. The reason is that it is hard to assess any resilience metrics at the service layer, since perceiving a network fault therein is a rather ambiguous task, as in the above TCP-related example. Additionally, we try to measure resilience from the logical perspective. Thus, we are interested in how the service with a required quality level is perceived, not in a more traditional "physical" perspective associated with hardware faults. Thus, short-term metrics are also present in the resilience part of Fig. 1. Figure 2 shows the general relation between transmission/transfer quality and the availability of a whole service. If a service meets all transmission/transfer requirements (the left part of Fig. 1), it is considered QoS-compliant. If any of those requirements are violated, the service is considered degraded. If they are significantly violated, the service is assessed as unavailable.

The measurement or interpretation can be performed by dividing the whole observation interval (e.g., duration of the connection or session) into Δt-long intervals (Fig. 3). Then resilience is evaluated within each using short-term quality metrics called here *instantaneous availability criteria*. Note that they may involve different thresholds than in the QoS metrics part, because having a failure and the subsequent recovery of a connection may lead to serious service degradation for a limited period of time. However, since it is very rare, the user might be satisfied in the long run. Sometimes those requirements are defined so loosely that only a broken service would be treated as unavailable. Defining these criteria is a difficult task; they strongly depend on application. For instance, according to ITU-T Recommendation Y.1540 (Table 1), the basis for this assessment for an IP service is a threshold on the *IPLR* performance. An example would be an SLA with packet loss probability of 5 percent for the QoS metric, while packet loss temporarily exceeding 20 percent causes the service to be considered unavailable.

Those short-term metrics measure the ability of a network to perform a required function at a given instance of time of Δt length. When instantaneous availability criteria and Δt are properly chosen, the QoR metric is a powerful high-level observation of how the application "experiences the performance." The ability to provide the required quality is assessed according to the binary function shown in Fig. 3: either present in a network in a given interval (1 = up/available service) or not (0 = down/unavailable service). In other words, a user is either fully satisfied with a service for this Δt interval or is not at all. Considering a very fine time granularity, partially satisfied user cases are not relevant, and service users restrict their judgment to only these two options.

The choice of time granularity is based on the type of application and the measurement method. Δt should be very short to minimize the harmful influence of erroneous decisions related to the quality assessment in a single interval. On the other hand, it must be long enough to decide in a "binary manner" if the service is satisfactory. For example, Δt might be a few tens of milliseconds for telesurgery applications, real-time control or emergency services, and hundreds of milliseconds for VoIP [4] or audio/video transmissions (e.g., video on demand), and seconds for traditional Internet applications like FTP transfer or VPN connectivity. Note that although a



■ Figure 2. *The service state model.* A*: availability of a service,* U*: unavailability of a service.*

real SLS agreed with customers is not likely to be as detailed as the one presented in Fig. 1, at least the operator must have its own technical specification of the provided services of this type.

Quality of Resilience

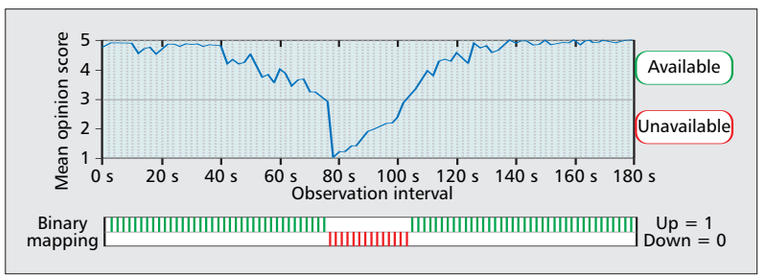
We propose a QoR metric to compare all types of recovery strategies that can be offered to a client. It is a fractile (quantile) representation of downtimes over the long run. It serves different goals. First, it can be used for characterization purposes to report failure intensities and recovery times for management purposes; second, as another provider-oriented approach (i.e., in active engineering) for decision making to select recovery strategies on the basis of simulated QoR histograms; and third, as a marketing approach becoming a part of more detailed SLAs than those currently in use. The last goal is quite challenging, as exhaustive resilience metrics are not agreed on. The adoption of such a usage of QoR involves risk sharing between an operator and a client. The operator reveals more information than steady-state availability or packet loss ratio. In return, a client equipped with more comprehensive knowledge cannot claim to not be thoroughly informed of the risks. This is different from the common practice of making resilience-related information confidential with the carrier taking all the risk. The two former goals are quite realistic to be adopted over the short term. They do not involve many practical changes, only a more detailed way of fault reporting and more thorough simulation of network behavior. On the other hand, operators or regulatory bodies are likely to be reluctant about the third type of QoR usage. Thus, we focus on the first and second goals.

Recall that provisioning time of a selected service is partitioned into Δt-long intervals. The time duration for which the service instantaneous availability criterion is assessed as 0 is called *downtime*. Measurements of this type can be used to derive the frequency distribution of different lengths of consecutive intervals assessed as 0, representing the *QoR downtime density histogram*. For easy interpretation, the range of downtime is split into periods (bins) of different sizes and ends (t_i). The histograms are evaluated by measuring the length of all consecutive unavailability intervals and counting the frequency of each downtime period. Finally, the chart is normalized with respect to the whole observation interval. Note that the bins are not identical to the Δt intervals. Similarly, the uninterrupted time (steady-state availability of the physical connection) can be estimated when lengths of all intervals assessed as 1 are summed up and divided by the overall number of intervals. An example is given in Fig. 4.

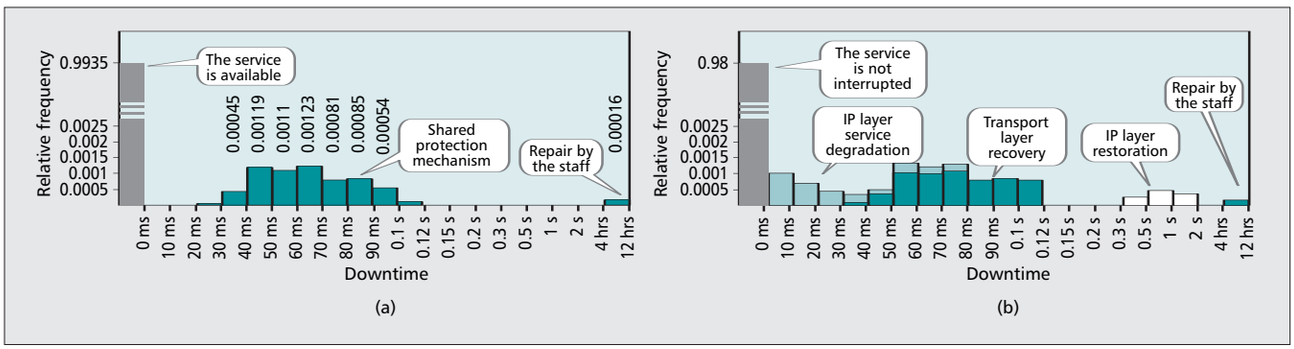
Using this method, an operator can obtain some overall profiles for its network to compare downtime durations and their relative frequencies adhering to various recovery methods applied to different service classes. For instance, it is then possible to assess if there are either many unrelated intermittent faults/errors or some long lasting tendency. QoR will show that a long lasting tendency generates different outputs (histograms with "longer tails"), while short interruptions generate some mass around zero in the histogram. In addition to this characterization usage, decision-enabling use can be envisaged: various alternative recovery methods can be compared to find the desirable profiles for the lowest cost. They can also be used to study interdependencies between different connections and their resilience metrics (e.g., when the sharing level is changed).

It is interesting to see that this method is backward compatible with the commonly used resilience metrics. The following measures are most typical for today's resilience description (Table 1):

- *Steady-state availability* and *unavailability*
- *Mean time to recovery*, usually identified with *mean downtime*
- *Mean time to failure*, usually identified with *mean up time*



■ Figure 3. *Binary mapping of user satisfaction with a service according to instantaneous availability criteria for each time interval. MOS measuring the satisfaction (see Table I) has its short-term quality counterpart covered in SLS.*



■ Figure 4. *Example of QoR downtime density histograms: a) transport layer; b) service layer.*

QoR encompasses all of them. When the length of the downtime is denoted T, it is a discrete random variable (since we base it on finite time intervals), with the distribution function F where

$$F(x) = \Pr\{T < x\} = \sum_{i|t_i \le x} \Pr\{t_{i-1} < T \le t_i\}$$

is the probability of having the service interrupted for at most x. $\Pr\{t_{i-1} < T \le t_i\}$ is the relative frequency of downtime period t_i. Then the probability of having uninterrupted service (the availability of the physical connection) is determined as the probability that a downtime is at most 0 long [4]:

$$P_0 = \Pr\{T = 0\} = F(0).$$

The availability of the service is related to the threshold value of time α, exceeding which makes this service down (unavailable):

$$A = \sum_{i|t_i \le \alpha} \Pr\{t_{i-1} < T \le t_i\} = F(\alpha).$$

The availability of the physical connection is lower than the availability of the service, $P_0 \le A$, as interruptions shorter than α do not harm the service due to successful recovery mechanism operation. Mean time to recovery is the average value of the pdf by definition, normalized for non-zero downtimes:

$$MTTR = E[T] = \frac{\sum_i T \times \Pr\{t_{i-1} < T \le t_i\}}{U} \approx MDT.$$

Taking into account the mutual relation between A, $U = 1 - A$, MDT, and MUT (Table 1), the mean time to failure of the recovered connection is expressed as

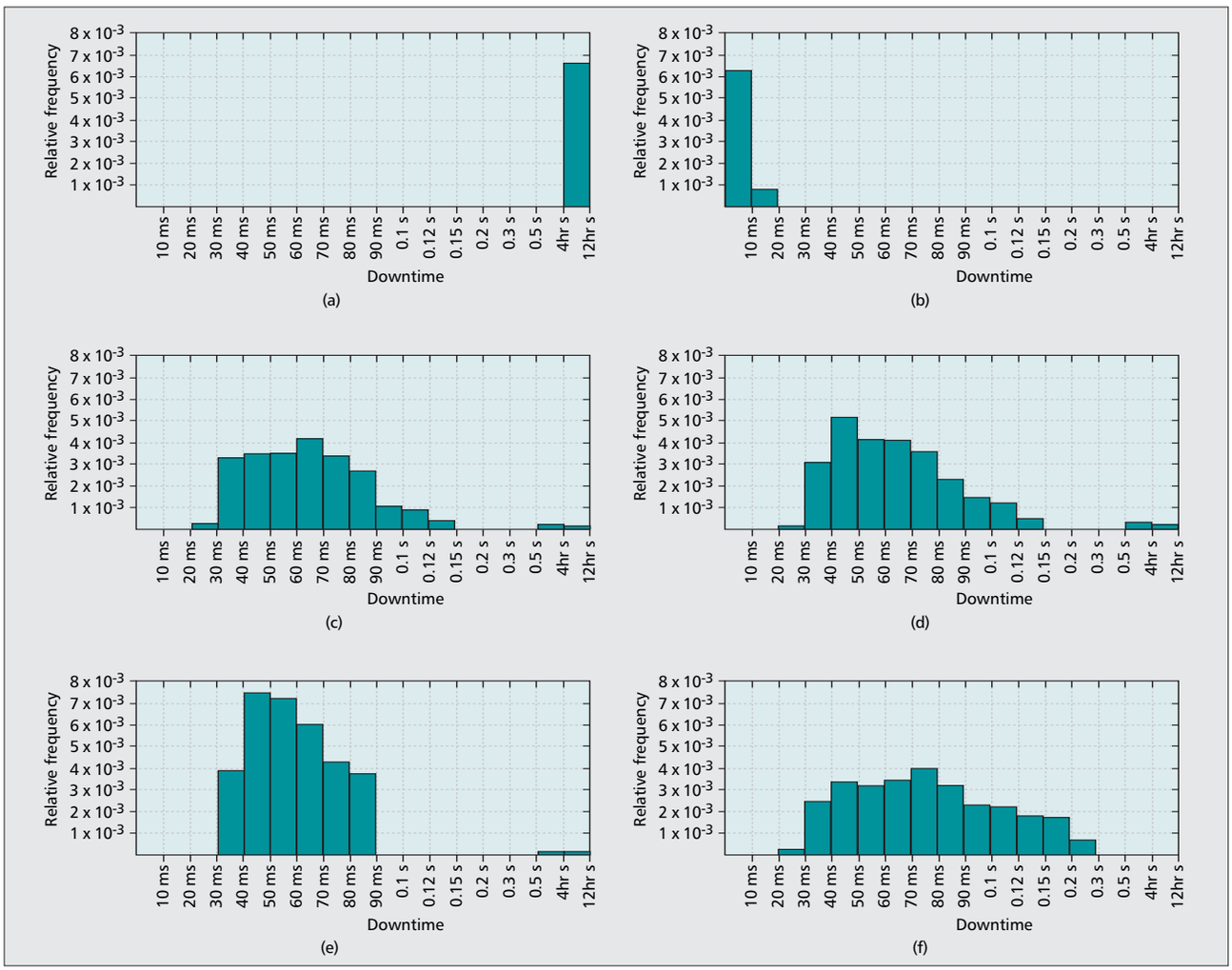
$$MTTF \approx MUT = \frac{A}{U} \times MDT.$$

It is especially useful to point out the predicted number of service instances disconnected due to ineffective recovery operation by using the conditional p-fractile notion, $t_{p\%|>0}$. It is defined as the p-fractile of the QoR histogram truncated for time values larger than 0. The value of $t_{p\%|>0} = \alpha$ will be of utmost interest to an operator.

Figure 4a shows a sample histogram of the downtime perceived at the transport layer. A multilayer recovery is simulated for a single failure in the European reference network. Simulation results are averaged for all connections (for details see [5]). The downtime is mainly caused by a physical fault and is recovered within a short period of time. In case of a single fault, a shared protection mechanism is triggered first. In this example 30 ms is the shortest time to recovery. However, after some (not automatically recoverable) network failures, the service is down for several hours until the network is repaired. Figure 4b shows the histogram of the downtime perceived by the user in the same network. IP restoration is applied after a 0.5–2 s outage. Thus, a difference between the downtime (outage) and the perceived downtime can be seen. For instance, for 1 + 1 protection we can have outage of one channel, but it will not be perceived as we have an operating backup channel. On the other hand, short outages will not be perceived as, for instance, higher-layer burstiness makes them not visible (e.g., when there is an off period during the downtime related to protection switching). Poor performance perceived by a user is usually caused by network congestion and consequently some packet layer service degradation like delays due to rerouting; see the lighter green part of Fig. 4b. The connection service unavailability inherited from the transport layer is drawn in Fig. 4b as the darker green part.

The impact of faults in IP networks has been extensively studied in [6] where the authors show that the length of the unavailability period can be much longer than that of the service outage at the transport layer. TCP segments in the transport layer lost during the outage are resent after the service is restored, leading to possible congestion and degraded transfer quality (the additive latency in Table 1). This phenomenon is also illustrated in the figures by repeating the dark part from Fig. 4a in Fig. 4b after shifting mass to longer downtimes. Some of the faults that cannot be recovered at the transport layer might be restored at the service layer (by IP rerouting). It reduces the probability of extremely long unavailability periods where the user is I senforced to wait until manual repairs take place. Such recovery typically requires a few seconds of downtime (due to the holdoff waiting time, link state advertising, and calculating new routes). It is shown by white bars in Fig. 4b.

This example shows that we are interested in multilayer recovery not in the sense of the cooperation between different layers in resilience operations, but in the sense of recovery influence on the quality perceived by higher (logical, service) layers. Thus, QoR is useful even if the recovery methods have no multilayer character. This rationale also explains why some parts of the frequency histogram are more (or less) important in relation to the particular network or service (application) for which different recovery methods are analyzed. The "interesting part" of the histogram/distribution is related to the application requirements related to connections investigated with the usage of QoR. It is the part related to the interval of downtimes that are lower than the critical threshold value for service breaking. It differs among applications. For example, thresholds related to acceptable downtimes (α) that do not harm transmission for different applications are 50 ms for no TCP fallback, 200 ms for voiceband connections, or 2 s for switched connections, and so on. [1, 7]. Then, on the basis of histograms as in Fig. 4, an operator can see the probability



■ Figure 5. *Example recovery time histograms: a) no recovery, MTTR = 11.76 h,* $t_{25\%|>0} = t_{50\%|>0} = t_{75\%|>0} = 12$ *h; b) dedicated protection of single link/node faults, MTTR = 11.01 ms,* $t_{25\%|>0} = t_{50\%|>0} = t^{75\%|>0} = 10$ *ms; c) shared path protection of single link faults with the usage of the Dijkstra algorithm, MTTR = 6.6 min,* $t_{25\%|>0} = 50$ *ms,* $t_{50\%|>0} = 70$ *ms,* $t_{75\%|>0} = 80$ *ms; d) shared path protection of single link faults with the usage of ILP, MTTR = 6.2 min,* $t_{25\%|>0} = 50$ *ms,* $t_{50\%|>0} = 70$ *ms,* $t_{75\%|>0} = 80$ *ms; e) shared segment protection of single link faults with recovery time constraints, MTTR = 3 min,* $t_{25\%|>0} = 50$ *ms,* $t_{50\%|>0} = 60$ *ms,* $t_{75\%|>0} = 70$ *ms; f) dual fault with the Dijkstra algorithm, MTTR = 90 ms,* $t_{25\%|>0} = 60$ *ms,* $t_{50\%|>0} = 80$ *ms,* $t_{75\%|>0} = 100$ *ms.*

(also using more sophisticated approaches, e.g., *p*-fractile-based) that by using a selected recovery method some transport services are lost. Such histograms evidently give more insight into the downtime character, conveying more information than plain availability or mean downtime measures.

Numerical Example: The Impact of Protection Mechanisms on QoR

This section shows simulated recovery time histograms for various recovery schemes for a single optical layer in a generalized multiprotocol label switching (GMPLS)-controlled network. It presents the potential for differentiation of the quantitative behavior of various recovery schemes to be used as a characterization and decision tool for operators. Some subset of methods representative for the optical network is selected. Here, we do not focus on multilayer recovery methods, as IP restoration may be better from the cost and management plane viewpoint, but it is always poorly competitive as it involves long downtimes due to at least fault detection and new path finding times [2] (white bars of the histogram in Fig. 4b). The details of the design and simulation can be found in [5]. The network components fault/repair model, the recovery time model, timing sequence, and traffic matrices adopted to assess the average recovery times are described in [5]. Simulations were conducted on the model of the Pan-European GMPLS fiber optic network defined by COST 266. It contains 28 nodes and 57 bidirectional links; 253 connections are established. Generally, the design of recovery methods has a long-term character where we can assume that traffic matrices do not change over a very long time and are known to operators (i.e., they use some good prediction methods). This is a standard assumption used in planning of long-haul wired networks. A dynamic traffic pattern is generated according to the traffic matrix such that an *interrupted Poisson process* and *Pareto interarrival times* are used with the *exponential holding time*. No connection is blocked. The bandwidth of the connections was reduced to obtain a lightly loaded network (in our case it is about 12 percent of the average load per link), where the effects on routing caused by shortage of spare resources are not significant. The *inverse capacity proportion rule* is applied as a traffic engineering method to calculate link costs for each connection.

The recovery time histograms for the no recovery case and dedicated protection scenarios are presented in Fig. 5a and 5b, respectively. Without recovery, the probability of losing the service is 0.00334. This means that the availability is 1 – 0.00334 = 0.99666. It cannot be seen in the pictures; we skipped the values of the availability due to the scaling problems (extremely high peaks related to the uninterrupted work in 0). The dedicated 1 + 1 protection gives very fast restoration in less than 20 ms and reduces the probability of losing the service more than 80 times (4.067×10^{-5}). The dedicated 1 + 1 protection has a desirable recovery time histogram: it has good abilities to support services with high QoR requirements. However, it is characterized by considerable cost. Other schemes try to balance quality and cost effectiveness.

The recovery time histograms for two shared path protection approaches are provided in Fig. 5c and 5d. The working path is routed and established with the shortest path first principle. Subsequently, the disjoint protecting path is computed and selected to enable sharing of the protection bandwidth for connections that use disjoint working paths. This method requires two shortest path searches and is referred to as "shared path protection of single link faults with the usage of the Dijkstra algorithm" (SPPD). This scheme is currently perceived by the IETF as desirable for both MPLS-layer protection and MPLS-controlled optical path protection. The second shared path protection method is a single step approach, which jointly calculates the minimum cost working and protection paths. It uses integer linear programming (ILP) to ensure the optimality of the solution. It is referred to as "shared path protection of single link faults with the usage of ILP." The optimization can save, on average, approximately 15 percent of network resources and can decrease the blocking probability due to lack of resources by ~ 5 percent. However, it leads to lower availability. This stems from the fact that working paths are routed on slightly longer routes to save the shared spare capacity for protection routes. This phenomenon related to cost optimization, not resilience, is frequently observed (see, e.g., [8]. Consequently, a longer working route has a higher chance of being affected by faults. When nodes are also protected against faults, this difference in availability value is 0.000216, which shows some risk of using highly capacity-efficient methods as the increase in sharing (related to savings) could involve a decrease in resilience performance, especially when multiple faults are present in a network. As simultaneous failures are not uncommon in contemporary networks [9], this phenomenon can be important for some sensitive applications.

Figure 5e shows the recovery time histograms for the shared segment protection. The approach is similar to the two-step approach introduced for shared path protection, where first the working path is routed on the shortest path, and in the second step a disjoint shared protection route is selected with a heuristic similar to the one presented in [10]. The chart illustrates the benefits of segment protection, which results in a significant decrease in recovery time. Since both SPPD and shared segment protection of single link faults with recovery time constraints use the shortest path as the working path, their service availability is similar. The shared segment protection with recovery time constraints shows its excellent ability to guarantee short recovery traded off with longer working paths leading to slightly lower service availability.

Figure 5f presents the recovery time histogram of shared protection scenarios providing resilience against dual faults. The method referred to as "dual fault with the Dijkstra algorithm" is a generalization of SPPD. The probability of having a fault that cannot be restored is ~ 10^{-6} when dual link faults are protected and ~ 10^{-7} when all combinations of two network elements (a link/node) are protected. The resilience against dual faults requires 40–50 percent more network resources leading to higher network utilization and longer working routes. As a consequence, lower service availability than in the single link protection case is obtained.

Mean time to recovery and some conditional fractile values are given in the description of the recovery scheme in Fig. 5. We can see that this gives us better insight on the quality of resilience offered by a selected method in a particular network, as we can better track the behavior of the method in relation to service discontinuity thresholds. In this sense QoR is a superior way to define QoS than methods commonly used, as the mean values can be misleading. For instance, compare the diagrams in Fig. 5e and 5f. Although the mean value of time to recovery is better for the latter case, we can see that fractiles can describe resilience more precisely, and in their light the former method is better. It is especially relevant if the difference is around some threshold value like 50 ms, the standard required recovery time for SDH networks. We can see that in the latter case we have a longer "tail," and it can be important in some situations (i.e., sensitive applications). Such a comparison assumes that an operator using some conditions in their networks tries to simulate different recovery scenarios as shown in Fig. 5. Each scenario is related to a selected recovery method and generates a corresponding QoR histogram. Then, on the basis of the application that is dominant, one of the alternative recovery methods can be selected (decision making usage of QoR). Then on the basis of long run measurement it can be checked if the predicted QoR conforms with the one obtained from practice (reporting and characterization usage of QoR).

Conclusions

The article presents extensions of the QoS quantification concepts focused on resilience issues. We propose a more thorough reliability and resilience description, which is called QoR, and allows for qualitatively and quantitatively comparing network recovery schemes deployed in a given network architecture. This task can be performed with the proposed QoR downtime histograms, enabling a provider to get a clear view of the characteristics of each recovery scheme. This method can be used by operators to avoid three types of pitfalls:

- Choosing an improper recovery method not adequate for user or application quality requirements
- Dealing with resilience in too narrow a scope based on only averaged metrics to describe a multidimensional quality problem
- Adopting too simple resilience agreements with customers that might lead to liability problems

The method of preparing QoR histograms and how to measure the well-known metrics with them was given. The method of measurement of the perceived instantaneous availability performance in the service layer is presented. It enables the determination of QoR measures for the logical layer, which is more important for a client and has not been satisfactorily covered in the studies on the quality of multilayer networks. Those extensions let operators study different possibilities related to recovery methods, prepare a more comprehensive portfolio, and make more reasonable decisions about whether to upgrade or redesign their networks. The given numerical examples illustrate the use of the QoR concept in quantitative terms.

Acknowledgments

This work was done within the EU FP6 IP IST-NOBEL II

(http://www.ist-nobel.org) framework. The reported work was also supported by the Polish Ministry of Science and Higher Education under grant N517 013 32/2131; High Speed Network Laboratory (HSNLab), the Hungarian National Research Fund, and the National Office for Research and Technology (Grant Number OTKA 67651). János Tapolcai was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The authors thank Gyäkon Lønsethagen, Ingeorgy Lajtha, Stefan Bodamer, Achim Autenrieth, Didier Colle, Håkon Løngsethagen, Einar Svinnset, Dominique Verchere, Arpád Szlávik, Attila Mihály, Csaba Antal, András Császár, Rafal Stankiewicz, and Przemyslaw Pawelczak for their helpful comments in the work on this topic.

References

[1] W. D. Grover, *Mesh-Based Survivable Networks: Options and Strategies for Optical, MPLS, SONET, and ATM Networks*, Prentice Hall PTR, 2004.
[2] J.-P. Vasseur *et al.*, *Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*, Morgan Kaufmann, 2004.
[3] Norros *et al.*, "Downtime-Frequency Curves for Availability Characterization," *Proc. DSN 2007*, Edinburgh, U.K., June 2007.
[4] J. Tapolcai *et al.*, "Quantification of Resilience for Voice-over-IP Applications," *Proc. ISBAT 2006*, Niagara Falls, Canada, Oct. 2006.
[5] J. Tapolcai *et al.*, "Joint Quantification of Resilience and Quality of Service," *Proc. IEEE ICC '06*, Istanbul, Turkey, June 2006.
[6] G. Iannaccone *et al.*, "Analysis of Link Failures in an IP Backbone," *Proc. ACM IMW '02*, Marseille, France, Nov. 2002.
[7] J. Tapolcai *et al.*, "Quality of Resilience (QoR): NOBEL Approach to the Multi-Service Resilience Characterization," *Proc. IEEE/CreateNet GOSP '05*, Boston, MA, Oct. 2005.
[8] M. Pioro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*, Morgan Kaufmann, 2004.
[9] J. Doucette *et al.*, "On the Availability and Capacity Requirements of Shared Backup Path-Protected Networks," *Optical Net.*, vol. 4, no. 6, Nov./Dec. 2003.
[10] D. Xu *et al.*, "Protection with Multi-Segments (PROMISE) in Networks with Shared Risk Link Groups (SRLG)," *Proc. 20th Allerton Conf.*, Monticello, IL, Oct. 2002.

Biographies

PIOTR CHOLDA [S'04, M'07] (piotr.cholda@agh.edu.pl) received a Ph.D. degree in telecommunications from AGH University of Science and Technology, Kraków, Poland, in 2006. Then he joined the Department of Telecommunications at the same university. His research interests focus on design and resilience of multilayer optical networks, as well as reliability and quality modeling concepts, including overlay networking. He is the co-author of 17 refereed technical papers and two tutorials on resilient networks. He is the recipient of the Communications QoS, Reliability, and Performance Modeling Symposium Best Paper Award from ICC '06. Now he is involved in the networking reliability and resilience area in two EU projects, Euro-NF and SmoothIT.

JANOS TAPOLCAI [M'05] (tapolcai@tmit.bme.hu) received his M.S. degree in technical informatics in 2000, and his Ph.D. in computer science in 2005 Budapest University of Technology and Economics (BME), Hungary. Currently he is an associate professor at the High-Speed Networks Laboratory at the Department of Telecommunications and Media Informatics at BME. His research interests include applied mathematics, combinatorial optimization, linear programming, linear algebra, routing in circuit-switched survivable networks, availability analysis, and distributed computing. He has been involved in several related European and Canadian projects. He is an author of over 40 scientific publications, half of them as first author, and is the recipient of the Best Paper Award at ICC '06.

TIBOR CINKLER [M'96] (cinkler@tmit.bme.hu) received M.S. (1994) and Ph.D. (1999) degrees from BME, where he is currently an associate professor in the Department of Telecommunications and Media Informatics. His research interests focus on optimization of routing, traffic engineering, design, configuration, dimensioning, and resilience of IP, Ethernet, MPLS, ngSDH, OTN, and particularly heterogeneous GMPLS-controlled WDM-based multilayer networks. He is an author of over 180 refereed scientific publications and four patents. He has been involved in numerous related European and Hungarian projects including ACTS METON and DEMON; COST 266, 291, 293; IP NOBEL I and II and MUSE; NoE e-Photon/ONe, NoE e-Photon/ONe+ and BONE; CELTIC PROMISE and Tiger II; NKFP, GVOP, ETIK; and he is member of ONDM, DRCN, BroadNets, AccessNets, IEEE ICC and GLOBECOM, EUNICE, CHINACOM, Networks, WynSys, ICTON, other Scientific and Program Committees. He has been guest editor of a Feature Topic in *IEEE Communications Magazine*, and a reviewer for many journals and conferences.

KRZYSZTOF WAJDA [M'99] (wajda@kt.agh.edu.pl) received his M.S. and Ph.D. in telecommunications from AGH University of Science and Technology in 1982 and 1990, respectively, and currently is an assistant professor in the Department of Telecommunications of the same university. Since 1991 he has been involved in a few international projects: COST 242, Copernicus ISMAN, ACTS BBL, IST LION, NOBEL and NOBEL2, e-Photon/ONe, e-Photon/ONe+, and BONE, and many national projects. He has been a consultant to private telecommunication companies and Polish Telecom. His research interests are traffic engineering, architecture and implementation of broadband packet networks, multimedia services, and network reliability issues.

ANDRZEJ JAJSZCZYK [M'91, SM'95, F'99] (jajszczyk@kt.agh.edu.pl) is a professor at AGH University of Science and Technology. He received M.S., Ph.D., and Dr.Hab. degrees from Poznań University of Technology in 1974, 1979, and 1986, respectively. He is the author or co-author of seven books and more than 240 papers, as well as 19 patents in the areas of telecommunications switching, high-speed networking, and network management. His current research interests focus on control plane architectures for transport networks and quality of service, as well as network resilience and reliability.