The Interval Reliability, its Usage and Calculation for Information and Communication Systems and Networks

Victor Netes
Moscow Technical University of Communications and Informatics
Moscow, Russia
v.a.netes@mtuci.ru

***Abstract*—The quality of information and communication service includes such important aspects as service accessibility and service retainability, which depend on dependability of systems and networks. The paper considers a dependability measure called interval reliability. It is defined as the probability that an item is in a state to perform as required at a given instant and, starting from this moment, continues to be in this state for a required time interval. The interval reliability simultaneously takes into account availability and reliability, describing the impact of dependability on service accessibility and service retainability at once. Special attention is paid to systems with redundancy, where it may be necessary to take into account situations when failover can lead to the premature termination (interruption) of a service session. A method for calculating the interval reliability for such systems is proposed. The presentation is illustrated by numerical examples for a network that can use protection with non-interrupting failover and restoration with interrupting failover.**

I. INTRODUCTION

The Quality of Service (QoS) in the field of Information and Communication Technologies (ICT) includes several aspects, the main of which are: service accessibility, service integrity and service retainability [1], [2]. They relate to different phases of service usage, as shown at Fig. 1 (its source is [1]). These aspects are defined as follows:

- Service accessibility is the ability of a service to be obtained, within specified tolerances and other given conditions, when requested by the user.
- Service integrity is the degree to which a service is provided without excessive impairments, once obtained (an acceptable level of impairments has to be specified).
- Service retainability is the ability of a service, once obtained, to continue to be provided under given conditions for a requested duration. It is important to users of session-oriented services.

A considerable factor affecting QoS is the dependability of systems and networks. It has several attributes (characteristics), among them availability and reliability are most often used in ICT [3]–[6] (their formal definitions will be given later). Availability primarily affects service accessibility (see Fig. 1). Reliability affects service retainability, because failures can lead to an unintentional end (premature termination, cut-off) of the session.

This paper considers a dependability measure called interval reliability. It simultaneously takes into account availability and reliability, describing the impact of dependability on service accessibility and service retainability at once. The interval reliability has been considered in a number of publications ([7]–[12], and others; although in some of them under other names, they will be discussed later). However, in [7], only its definition is given. Some publications ([8]–[10]) provide formulas for calculating the interval reliability of an element and certain types of redundant systems (mostly from identical elements). In [11], its calculation is given for series systems, i.e. the simplest ones. The specific case of discrete time is considered in [12].

The main purposes of this paper are: to draw more attention to the interval reliability and justify its usage in ICT, to clarify its definition and name, and to propose methods for its calculation for various systems. Special attention is paid to systems and networks with redundancy, especially those where it may be necessary to take into account situations when failover can lead to the premature termination of the service session.

The rest of the paper is organized as follows. Section II discussed basic dependability concepts and measures; in particular, it considers the definition, basic properties and standardization of the interval reliability. Section III presents common methods of interval reliability calculation for elements and some systems. In Section IV typical availability enhancement techniques and failover time are discussed; the influence of this time on service retainability is considered. Section V describes proposed method of interval reliability calculation for systems and networks with redundancy in the case of service interrupting failovers. Numerical examples of calculations for a network that can use protection with non-interrupting failover and restoration with interrupting failover are given in Section VI. Concluding Section VII gives main findings of the paper and possible directions for future work.

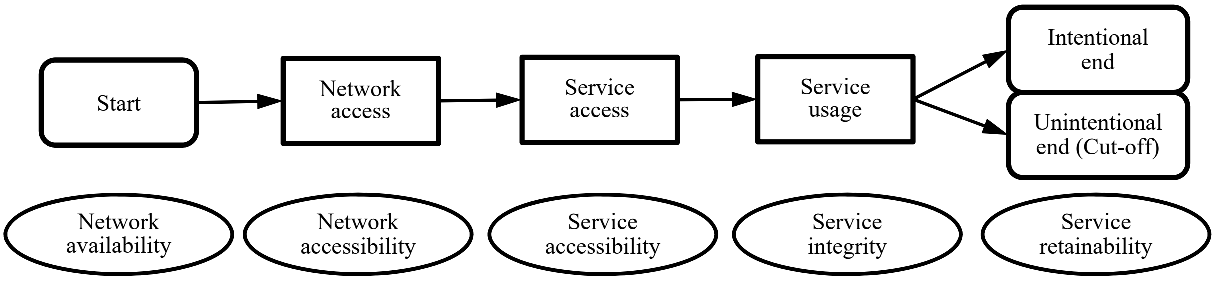

Fig. 1. QoS aspects related to different phases of service usage

II. Basic Dependability Concepts and Measures

A. Dependability and its attributes (characteristics)

Definitions of the basic concepts in the field of dependability are given in the International Standard [12]. It is a part of the International Electrotechnical Vocabulary (IEV) developed by the International Electrotechnical Commission (IEC). IEV has an online version called Electropedia [14], which is freely available. The definitions below are taken from this standard.

All basic concepts are applied to an item that is defined as a subject being considered. It may be an individual part, component, device, functional unit, equipment, subsystem, or system. An item may consist of hardware, software, people or any combination thereof.

Dependability of an item is defined as its ability to perform as and when required. The notes to the definition state that dependability is used as a collective term for the time-related quality characteristics of an item and it includes availability, reliability, recoverability, maintainability, and maintenance support performance and, in some cases, other characteristics. In ICT, the most frequently considered characteristics are availability and reliability [3]–[6].

Reliability of an item is its ability to perform as required, without failure, for a given time interval, under given conditions. Given conditions include aspects that affect reliability, such as: mode of operation, stress levels, environmental conditions, and maintenance.

Availability of an item is its ability to be in a state to perform as required. Availability depends upon the combined characteristics of the reliability, recoverability, and maintainability of the item, and the maintenance support performance.

B. Reliability and availability measures

A dependability measure is a quantitative index of one or more characteristics that make up dependability of an item.

To write mathematical expressions, the following notation will be used:

- The state of an item at time t is denoted by $x(t)$: $x(t) = 1$, if at the moment t the item is in an up state, i.e. it is able to perform as required; $x(t) = 0$, if at the moment t the item is in a down state, i.e. it is unable to perform as required, due to internal fault, or preventive maintenance. Thus, $x(t)$ is a binary random variable.
- $P\{.\}$ denotes the probability of an event enclosed in braces.
- $E[.]$ denotes the mathematical expectation.

1) Reliability measures: Well-known reliability measures are the mean operating time between failures (MTBF) and the failure rate $\lambda(t)$.

Another important measure, called simply reliability, denoted by $R(t_1, t_2)$. It is defined as the probability of performing as required for the time interval (t_1, t_2). It is usually assumed that the item is in an up state at the beginning of the time interval. In other words,

$$R(t_1, t_2) = P\{x(t) = 1, t_1 \le t \le t_2 \mid x(t_1) = 1\}.$$

When $t_1 = 0$ and $t_2 = t$, then $R(0, t)$ is denoted simply as $R(t)$ and termed the reliability function. Through it, the failure rate can be expressed:

$$\lambda(t) = -R'(t)/R(t).$$

If the operating time between failures has an exponential distribution, then the failure rate is constant: $\lambda(t) = \lambda = 1/\text{MTBF}$. In this case, $R(t) = e^{-\lambda t}$.

2) Availability measures: The first among them is the instantaneous (point) availability, denoted by $A(t)$. It is the probability that an item is in a state to perform as required at a given instant:

$$A(t) = P\{x(t) = 1\} = E[x(t)].$$

The steady state (asymptotic) availability denoted by A is the limit, if it exists, of the instantaneous availability when the time tends to infinity:

$$A = \lim_{t\to\infty} A(t).$$

In most cases, it is calculated using the formula

$$A = \text{MTBF}/(\text{MTBF} + \text{MTTR}), \quad (1)$$

where MTTR is mean time to restoration (maintainability measure).

This availability measure is most commonly used in practice and is usually called merely availability. In most cases, in ICT, dependability requirements are set and target values are specified just for availability [15]. It is commonly included in Service Level Agreements [5], [15]–[19].

The mean (average) availability is the average value of the instantaneous availability over a given time interval (t_1, t_2):

$$\bar{A}(t_1,t_2) = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} A(t)dt.$$

The instantaneous unavailability is defined as the probability that an item is not in a state to perform as required at a given instant, and it is a complement of instantaneous availability to one: $U(t) = 1 - A(t)$. Accordingly, the steady state (asymptotic) unavailability and the mean (average) unavailability are defined. For them, there are equalities: $U = 1 - A,\ \bar{U}(t_1,t_2) = 1 - \bar{A}(t_1,t_2)$.

C. The interval reliability: the definition, basic properties and standartization

The interval reliability, which is considered in this paper, has been known for a long time. In particular, it was mentioned in the classic monograph [7]. It is defined as the probability that an item is in a state to perform as required at a given instant and, starting from this moment, continues to be in this state for a required time interval. In ICT, the length of the required time interval is usually the duration of the service session.

Denote the interval reliability, as in [10]–[12], by $IR(t,t_0)$, where t is the given instant, and t_0 is the length of the required time interval. Then its definition can be written as:

$$IR(t,t_0) = P\{x(s) = 1,\ t \le s \le t + t_0\}.$$

From this definition, it can be seen that the interval reliability is actually at the junction of availability and reliability. It is clear, that $IR(t,0) = A(t)$, $IR(0,t_0) = R(t_0)$.

Reasoning similarly to how it was done in [12], it is easy to get that the interval reliability is bounded from above by the availability and bounded from below by the reliability:

$$R(t + t_0) \le IR(t,t_0) \le A(t + t_0).$$

Just as for availability, the steady state (asymptotic) interval reliability is mostly considered. It is defined as

$$IR(t_0) = \lim_{t\to\infty} IR(t,t_0).$$

It is apparent that $IR(0) = A$.

The interval reliability is not included in the standard [12] and is only briefly mentioned in [20]. However, it is included in the regional standards [21], [22] adopted by the Interstate (Euro-Asian) Council for Standardization, Metrology and Certification of the Commonwealth of Independent States. In [21] there is the definition of this measure, [22] recommends its usage for intermittently operating items.

Unfortunately, the names of this measure differ in different publications. The term "interval reliability" is used in [7], [10]–[12]. The standard [21] uses a term for it that can literally be translated as "operational availability", this term was used also in [8], [23]. However, operational availability is defined in [12] as availability experienced under actual conditions of operation and maintenance. The opposite concept to it is inherent (intrinsic) availability, which is availability provided by the design under ideal conditions of operation and maintenance. In this sense, the term "operational availability" is actively used in some industries, for example, in aviation [24].

Therefore, [21] specifies the term "interval availability" as the English name for this measure, this term is also used in [9]. However, it is also used in a different sense. In [17], [18] interval availability is defined by the fraction of time during which an item is in up state over a finite observation period. This fraction is a random variable; its mathematical expectation is equal to the mean availability for the same period [20], [25].

III. CONNON METHODS OF INTERVAL RELIABILITY CALCULATION

A. Calculation for an element

First, consider the calculation of interval reliability for an element, i.e., for an item considered as a whole. For the steady state interval reliability there is the formula [8]–[10]:

$$IR(t_0) = \frac{1}{\mathrm{MTBF} + \mathrm{MTTR}}\int_{t_0}^{\infty} R(t)dt.$$

If (and only if) the distribution of operating time between failures is exponential, then there is a simple formula:

$$IR(t_0) = A \cdot R(t_0) = A \cdot \exp(-t_0/\mathrm{MTBF}).$$

It means that in this case interval reliability is equal to the product of availability and reliability. In some works, this formula is incorrectly used as a general one.

If the element is "aging", that is it has an increasing failure rate (IFR) or an increasing failure rate in average (IFRA) [25], the following lower and upper bounds can be written [8]:

$$A \cdot (1 - t_0/\mathrm{MTBF}) \le IR(t_0) \le A \cdot \exp(-t_0/\mathrm{MTBF}). \qquad (2)$$

For the difference δ between the upper and lower bounds, the inequality holds

$$\delta \le (t_0/\mathrm{MTBF})^2/2, \qquad (3)$$

so it is quite small when $t_0 \ll$ MTBF.

B. Calculation for systems

The simplest type of systems is series ones. They do not have any redundancy, and failure of any element causes the system to fail. For a series system with independent elements its interval reliability is calculated in the same simple way as availability and reliability, i.e. as the product of the corresponding values for all elements [11]:

$$IR(t,t_0) = \prod_{i=1}^{n} IR_i(t,t_0),$$

where $IR_i(t,t_0)$ is interval reliability of the ith element, n is the number of elements in the system.

Unfortunately, for more complex systems, such a simple calculation based only on interval reliability of elements is not possible. This does not work even for parallel systems. Therefore, the following approach to assessing the interval reliability can be proposed. A fairly general class of monotone systems (coherent structures) [7], [25] will be considered. It includes serial and parallel structures, their various combinations, and more complex network structures. It is known [25] that if all elements have failure rates from the IFRA class, then the same is true for a monotone system as a whole consisting of such elements. Therefore, inequalities (2) can be applied to the system. To calculate the availability and the MTBF of the system included in (2), well-known techniques can be applied [8]–[11], [26]. In particular, when distributions of elements' operating time between failures and time to restoration are exponential, Markov techniques can be used [27]. An example of such a calculation will be given below.

Essentially the same method was used in [8]–[10] for some redundant systems (although sometimes without proper justification). However, this did not take into account the failover time that is the time of switching from failed primary facilities to redundant ones. Meanwhile, this can significantly affect the interval reliability calculation. This issue will be discussed in the following sections.

IV. Availability Enhancement Techniques and Failover Time

Redundancy is a widely used way to ensure dependability. However, if the failover time is not negligible, redundancy increases availability, but does not increase reliability. Therefore, the recovery time is an important characteristic. Typical situations are discussed in this section.

There are two main strategies which may be used to enhance the availability of a transport network [28], [29]:

- Protection that uses pre-assigned capacity between nodes. The simplest architecture has one dedicated protection entity for each working entity (1 + 1); the most complex architecture has m protection entities shared amongst n working ones (m:n).
- Restoration that uses any capacity available between nodes. In general the algorithms used for restoration will involve traffic rerouting.

A significant advantage of protection in comparison with restoration is typically is its shorter recovery time. On the other hand, restoration is usually more flexible with regard to failure scenarios and lower requirements for backup capacities. So, each of these two mechanisms has its own scope, depending on the situation.

A typical recovery time is within 50 ms for Automatic Protection Switching in SDH/SONET and OTN, Fast Reroute in MPLS, Linear/Ring Protection Switching in Ethernet. In general, this threshold (50 ms) is usually considered as a requirement for carrier-grade services. Many network and customer devices build this level of buffer into their operation so that these short interrupts are entirely unnoticed. In more modern software defined networks using OpenFlow protocol the average recovery time of the fast failover mechanism based on the pre-established paths can be less than 40 ms, compared to hundreds of milliseconds in the fast restoration mechanism [30]. The failover time for such availability enhancement mechanisms used in LAN as Link Aggregation and Rapid Spanning Tree Protocol is also in hundreds of milliseconds [31]. Layer 3 routing protocols (such as RIP, OSPF, ISIS and BGP) include the ability to reroute IP traffic in case of link or node failures. However, these techniques can take seconds to complete depending on the size and complexity of the network.

In [15] it was shown that achieving high availability of cloud services requires redundancy for network connections that provides customer interaction with data centers and data centers themselves. In particular, a combination of the two connections from different providers can be used. In this case, it is not possible assuring a recovery time in milliseconds. When switching to a redundant data center, the recovery time can be minutes, hours, or even days [32] (the shorter the recovery time, the higher the expenses of building and managing data centers).

The service interruption time depends on a traffic type. For example, a half-second interruption will be unnoticed in a web page download or peer-to-peer transfer, annoying in a video download, and unacceptable in a voice call. Thus, when calculating the interval reliability, we should be able to take into account both possibilities: when failover does not interrupt service, i.e. it is not considered as a system failure, and when failover interrupts service, i.e. it is a system failure. Appropriate examples will be discussed below.

V. Interval Reliability Calculation in the Case of Interrupting Failover

Consider a system in which there are k paths numbered in order of preference for their use. This means that if the first path is available, then it is used; if the first path is not available, but the second path is available, then the second path is used; etc. Denote by I_j the set of element numbers and by H_j the probability of using for the j-th path ($j = 1,\dots, k$). The sum $H_1 +\dots+ H_k$ is equal to the probability that at least one path is available, i.e. the system availability.

In this context, the word "path" for a network can be taken literally. For any system, a path is the minimum set of elements that satisfies the condition: if all its elements are in up state, then the entire system is also in up state.

Then the interval reliability, using the idea from [23], can be written as:

$$IR(t_0) = \sum_{j=1}^{k} H_j \prod_{i\in I_j} R_i(t_0), \qquad (4)$$

where $R_i(t_0)$ is the reliability of the i-th element (the product is equal to the reliability of the j-th path).

The following expressions can be used to calculate the probabilities H_j:

$$H_1 = E\left[\prod_{i\in I_1} x_i\right] = \prod_{i\in I_1} A_i, \qquad (5)$$

and for $j = 2,\dots,k$

$$H_j = E\left[\left(1-\prod_{i\in I_1} x_i\right)\cdots\left(1-\prod_{i\in I_{j-1}} x_i\right)\prod_{i\in I_j} x_i\right], \qquad (6)$$

where x_i is the state indicator of the i-th element ($x_i = 1$, if the i-th element is in up state, $x_i = 0$, if the i-th element is in down state), and A_i is the availability of the i-th element.

To calculate the probabilities H_j in accordance with (6), the expression in the brackets in the right member of (6) should be transformed so that there are no repeated variables x_i in it, i.e. that they are all different. This can be done by using the following equalities:

$$(1 - xy)\cdot x = (1 - y)\cdot x, \quad (1 - xy)\cdot(1 - x) = (1 - x), \quad (7)$$

$$(1 - xy)\cdot(1 - xz) = 1 - x\cdot(y + z - yz).$$

They are valid for any variables $x, y, z \in \{0,1\}$ since they are idempotent.

After that, the final result is obtained by substituting A_i instead of x_i in the resulting expression. This follows from the properties of the mathematical expectation and the equality $A_i = E[x_i]$. An example of such calculation will be given below.

In the case where the j-th path ($j = 2,\ldots,k$) is disjoint with each of the paths with smaller numbers, the calculation of the probability H_j becomes much easier since it is expressed in an explicit form:

$$H_j = \left(1 - \prod_{i \in I_1} A_i\right) \ldots \left(1 - \prod_{i \in I_{j-1}} A_i\right) \prod_{i \in I_j} A_i \quad (8)$$

(the expressions in parentheses are the unavailabilities of the first $j - 1$ paths, the last product is equal to the availability of the j-th path).

Note also that the reliability according to (4) depends on the order of path selection. It can be maximized by ordering the paths in descending order of their reliability.

VI. EXAMPLES

For examples consider the network shown in the Fig. 2, where the ingress and egress nodes are highlighted with a fill. This is the so-called bridge system, which is the simplest irreducible to combinations of serial and parallel configurations.

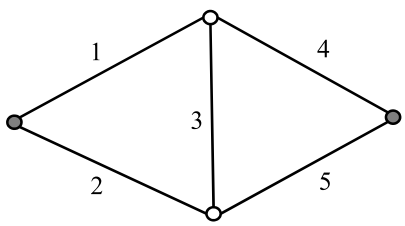

Fig. 2. The example of a network (bridge system)

Let all lines have an exponential distribution of operating time between failures, $\mathrm{MTBF}_i = 1000$ h, and $\mathrm{MTTR}_i = 5$ h ($i = 1,\ldots,5$). Then according to (1), $A_i = 0.995025$.

There are four paths between the ingress and egress nodes: $I_1 = \{1, 4\}$, $I_2 = \{2, 5\}$, $I_3 = \{1, 3, 5\}$, $I_4 = \{2, 3, 4\}$. The availability and the MTBF of the j-th path can be calculated using well-known formulas for series systems:

$$A_{\mathrm{P}j} = \prod_{i \in I_j} A_i, \qquad \mathrm{MTBF}_{\mathrm{P}j} = \left(\sum_{i \in I_j} \frac{1}{\mathrm{MTBF}_i}\right)^{-1}.$$

Two scenarios will be considered:

1) Protection 1 + 1: The protected connection has two disjoint paths (working $\{1, 4\}$ and protection $\{2, 5\}$). In this case, there is a fast recovery, so failover is non-interrupting.

2) Restoration: All four possible paths can be used for traffic transfer depending on the failure locations. However, the recovery is not so fast, so failover is interrupting.

For both cases, take the session duration equal to one hour, i.e. $t_0 = 1$ h.

In the case of using protection 1 + 1, for the 1st and 2nd paths we get:

$$A_{\mathrm{P1}} = A_1 A_4 = 0.990075, \; A_{\mathrm{P2}} = A_2 A_5 = 0.990075,$$

$$\mathrm{MTBF}_{\mathrm{P1}} = (1/\mathrm{MTBF}_1 + 1/\mathrm{MTBF}_2)^{-1} = 500 \text{ h},$$

$$\mathrm{MTBF}_{\mathrm{P2}} = (1/\mathrm{MTBF}_2 + 1/\mathrm{MTBF}_5)^{-1} = 500 \text{ h}.$$

In this case, the entire system is a parallel composition of these two paths. Using the formulas for the parallel system, we get:

$$A = 1 - (1 - A_{\mathrm{P1}})(1 - A_{\mathrm{P2}}) = 0.999902,$$

$$\mathrm{MTBF} = A[A_{\mathrm{P1}}(1 - A_{\mathrm{P2}})/\mathrm{MTBF}_{\mathrm{P1}} + A_{\mathrm{P2}}(1 - A_{\mathrm{P1}})/\mathrm{MTBF}_{\mathrm{P2}}]^{-1} =$$

$$= 25440 \text{ h}.$$

The value of δ in accordance with (3) is negligible, so we can take

$$IR(1 \text{ h}) = A\cdot(1 - 1/\mathrm{MTBF}) = 0.999863.$$

In the case of using restoration, according to (4),

$$IR(t_0) = H_1 R_{\mathrm{P1}}(t_0) + H_2 R_{\mathrm{P2}}(t_0) + H_3 R_{\mathrm{P3}}(t_0) + H_4 R_{\mathrm{P4}}(t_0),$$

where $R_{\mathrm{P}j}(t_0)$ is the reliability of the j-th path.

According to (5) and (8) respectively,

$$H_1 = A_1 A_4 = 0.990075, \; H_2 = (1 - A_1 A_4) A_2 A_5 = 0.009827.$$

On the basis of (6) and using (7), we obtain:

$$H_3 = E[(1 - x_1 x_4)(1 - x_2 x_5) x_1 x_3 x_5] = E[(1 - x_4)(1 - x_2) x_1 x_3 x_5] =$$

$$= (1 - A_4)(1 - A_2) A_1 A_3 A_5 = 0.000024,$$

$$H_4 = E[(1 - x_1 x_4)(1 - x_2 x_5)(1 - x_1 x_3 x_5) x_2 x_3 x_4)] =$$

$$= E[(1 - x_1)(1 - x_5) x_2 x_3 x_4] = (1 - A_1)(1 - A_5) A_2 A_3 A_4 = 0.000024.$$

Coming up next,

$$\mathrm{MTBF}_{\mathrm{P3}} = (1/\mathrm{MTBF}_1 + 1/\mathrm{MTBF}_3 + 1/\mathrm{MTBF}_5)^{-1} = 333.333 \text{ h},$$

$$\mathrm{MTBF}_{\mathrm{P4}} = (1/\mathrm{MTBF}_2 + 1/\mathrm{MTBF}_3 + 1/\mathrm{MTBF}_4)^{-1} = 333.333 \text{ h};$$

$$R_{\mathrm{P1}}(1 \text{ h}) = \exp(-1/\mathrm{MTBF}_1) = \exp(-1/500) = 0.998002,$$

$$R_{\mathrm{P2}}(1 \text{ h}) = \exp(-1/\mathrm{MTBF}_2) = \exp(-1/500) = 0.998002,$$

$$R_{\mathrm{P3}}(1 \text{ h}) = \exp(-1/\mathrm{MTBF}_3) = \exp(-1/333.333) = 0.997005,$$

$$R_{\mathrm{P4}}(1 \text{ h}) = \exp(-1/\mathrm{MTBF}_4) = \exp(-1/333.333) = 0.997005.$$

Finally we get:

$$IR(1\text{ h}) = 0.990075 \cdot 0.998002 + 0.009827 \cdot 0.998002 +$$

$$+ 0.000024 \cdot 0.997005 + 0.000024 \cdot 0.997005 = 0.997952.$$

The results of the calculations are presented in the Table I. They show that in the second case, despite the use of a greater number of paths, the probability of an adverse event, which is a complement of the interval reliability to one, is almost 15 times greater.

TABLE I. INTERVAL RELIABILITY OF THE NETWORK

Scenario	***IR*(1 h)**	**1 – *IR*(1 h)**
1) Protection: two disjoint paths, non-interrupting failover	0.999863	0.000137
2) Restoration: all four possible paths, interrupting failover	0.997952	0.002048

VII. CONCLUSION

The main findings of this paper are the following. The interval reliability is a useful dependability measure. It simultaneously takes into account availability and reliability, describing the impact of dependability on service accessibility and service retainability at once. In some publications, it appeared under other names: operational availability and interval availability. However, these terms also have other meanings.

When calculating the interval reliability for redundant systems, it is necessary to understand what the situation is: whether failover will interrupt the system's ability to perform as required (for example, to provide a service) or not. This usually depends on the failover time. In the first case, failover should be considered as a system failure, in the second case it should not. This circumstance has a significant impact on the interval reliability. In this paper, a method of calculation the interval reliability in the situation of interrupting failover is proposed.

Further work could be devoted to more detailed analysis. In particular, scenarios where both non-interrupting and interrupting failovers occur can be considered, the failover time can be taken into account also in availability calculation, etc.

REFERENCES

[1] Recommendation ITU-T E.804 (02/2014). *QoS Aspects for Popular Services in Mobile Networks.*

[2] E. Bauer, R. Adams, *Service Quality of Cloud-Based Applications.* New York: John Wiley & Sons, 2013.

[3] M.L. Shooman, *Reliability of Computer Systems and Networks: Fault Tolerance, Analysis, and Design.* New York: John Wiley & Sons, 2002.

[4] E. Bauer, *Design for Reliability: Information and Computer-Based Systems.* Hoboken: John Wiley & Sons, 2010.

[5] M.L. Ayers, *Telecommunications System Reliability Engineering, Theory, and Practice.* Hoboken: John Wiley & Sons, 2012.

[6] V. Netes, "Modern network technologies and dependability", *in Proc. of the 3d Intern. Science and Technology Conf. "Modern Network Technologies – 2020".* Moscow, October, 2020, pp. 104-113.

[7] R.E. Barlow, F. Proshan, *Mathematical Theory of Reliability.* New York: John Wiley & Sons, 1965.

[8] *Handbook of Reliability Engineering,* Ed. I.A. Ushakov. New York: John Wiley & Sons, 1994.

[9] B. Gnedenko, I. Ushakov, *Probabilistic Reliability Engineering.* New York: John Wiley & Sons, 1995.

[10] A. Birolini, *Reliability Engineering. Theory and Practice. 5th ed.* Berlin, Heidelberg: Springer-Verlag, 2007.

[11] K.S. Trivedi, A. Bobbio, *Reliability and Availability Engineering: Modeling, Analysis, and Applications,* Cambridge: Cambridge University Press, 2017.

[12] S. Georgiadis, N. Limnios, "Interval reliability for semi-Markov systems in discrete time", *Journ. de la Societe Francaise de Statistique & Revue de Statistique Appliquee,* vol. 153, 2014, No. 3, pp. 152-166.

[13] IEC 60050-192:2015. *International Electrotechnical Vocabulary – Part 192: Dependability.*

[14] Electropedia: The World's Online Electrotechnical Vocabulary, Web: http://www.electropedia.org/.

[15] V. Netes, "End-to-end availability of cloud services", *in Proc. of the 22d Conf. of Open Innovations Association FRUCT,* Jyvaskyla, Finland, May 2018, pp. 198-203.

[16] ITU-T Recommendation E.860 (06/2002). *Framework of a Service Level Agreement.*

[17] J. Mi, "On measure of system interval availability", *Probability in the Engineering and Informational Sciences*, vol. 13, 1999, No 3, pp. 359-375.

[18] A.J. Gonzalez, B.E. Helvik, "A study of the interval availability and its impact on SLAs risk", *in Advances in Computer Science, Engineering & Applications. Advances in Intelligent and Soft Computing*, vol. 166. Berlin, Heidelberg: Springer, 2012, pp. 879-890.

[19] V.A. Netes, "Service level agreements and dependability", *Dependability*, vol. 17, 2017, No 4, pp. 27-30.

[20] IEC 61703:2016. *Mathematical Expressions for Reliability, Availability, Maintainability and Maintenance Support Terms.*

[21] GOST 27.002–2015. *Dependability in Technics. Terms and Definitions* (in Russian).

[22] GOST 27.003–2016. *Industrial Product Dependability. Contents and General Rules for Specifying Dependability Requirements* (in Russian).

[23] V.A. Netes, B.P. Filin, "Calculating the operational availability factor of network systems", *Automation and Remote Control,* vol. 53, 1992, No 9, pp. 1465-1471.

[24] *Aircraft Operational Availability.* Montreal-Geneva: International Air Transport Association, 2018.

[25] R.E. Barlow, F. Proshan, *Statistical Theory of Reliability and Life Testing: Probability Models.* New York: Holt, Rinehart and Winston, 1975.

[26] IEC 60300-3-1:2003. *Dependability management – Part 3-1: Application guide – Analysis techniques for dependability – Guide on methodology.*

[27] IEC 61165:2006. *Application of Markov Techniques.*

[28] Recommendation ITU-T G.805 (03/2000). *Generic Functional Architecture of Transport Networks.*

[29] J.-P. Vasseur, M. Pickavet, P. Demeester, *Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS.* Amsterdam etc.: Elsevier, 2004.

[30] Y. Lin, H. Teng, C. Hsu, C. Liao, Y. Lai, "Fast failover and switchover for link failures and congestion in software defined networks," *in Proc. of the 2016 IEEE Intern. Conf. on Communications,* Kuala Lumpur, Malaysia, 2016, pp. 1-6.

[31] J.T. Yu, "Measuring failover time for high availability network", *in Proc. of the 2018 Intern. Conf. on Scientific Computing.* 2018, Las Vegas, USA, pp. 34-39.

[32] H.M. Brotherton, J.E. Dietz, "Data center site redundancy", *in Proc. of the 10th Intern. Conf. of the Intern. Inst. for Infrastructure Resilience and Reconstruction,* May 2014, pp. 2-6.