Computer Communications 36 (2013) 708–717

Contents lists available at SciVerse ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

SLA success probability assessment in networks with correlated failures ☆

Andres J. Gonzalez *, Bjarne E. Helvik

Centre for Quantifiable Quality of Service in Communication Systems, Norwegian University of Science and Technology, O.S. Bragstads Plass 2E, N-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Available online 30 August 2012

Keywords:
Network dependability
Failure dependencies
Failure correlation
SLA risk
Trace driven simulation

ABSTRACT

Service Level Agreements (SLAs) are used to define obligations between network/service providers and customers in business relationships. The terms that define the guaranteed availability for a given period are fundamental to these contracts. The appropriate selection of the availability to be promised is still an open challenge for network operators due to: (i) SLAs are defined for finite periods, and hence the stochastic properties of the availability have to be considered. (ii) Real operational networks have not the Markovian properties. (iii) The way that correlation affects the interval availability in operational networks is unknown. In this work, we show the impact of dependent failures on SLAs, based on operational failure data obtained from the UNINETT network. Using these data, we simulate the behavior of network connections that use shared backup protection. We evaluate the SLA success probability using two different methods. First, we apply trace driven simulation combined with random circular shifting. Second, we develop a model that uses Monte Carlo techniques. This approach includes the characterization of up and down times of each network component and the use of a model that generates correlated samples based on fitted marginal distributions. Finally, we analyze the probability density function of the interval availability for different observation periods under independent and correlated failures.

1. Introduction

Network operators and customers use Service Level Agreements (SLAs) to define a contracted QoS where availability is a significant element. The violation of the agreed value may have economic and reputation consequences for both parts. Under this scenario, a natural question for network operators is: how to assess the availability level that my network is able to provide? In order to answer this question, many factors have to be considered. First of all, it is important to notice that the availability offered to a connection during a contract (SLA) is a stochastic variable. In this case, the assessment of expected values, assuming steady state conditions, will provide misleading estimates that may dramatically increase the risk of failing the SLA. Therefore, the study of the entire distribution of the interval availability is necessary. Under Markovian assumptions, the interval availability can be obtained by numerical methods using uniformization techniques, as is presented in [1].

The probability that a network operator meets the contracted interval availability α can be calculated directly from the interval availability Probability Density Function PDF as $P(\alpha \leqslant$ Interval Availability $\leqslant 1)$, and it will be referred as *SLAs success probability*. This concept was first raised for Markovian systems in [2] by Goyal and Tantawi. They observed that if α is larger than the *steady state availability* (A), the success probability decreases continuously. However, they also showed that there is a considerable risk even when $\alpha < A$. Real operational networks do not follow Markovian assumptions. The inter-failure times are not exponentially distributed, neither any restoration time. In addition, the failure processes of two or more network components may be correlated. In such scenario, the assessment of the interval availability and the SLA success probability represents an open challenge that needs to be addressed. The way of addressing that challenge is the main point of this paper.

The work presented in [3] highlights the importance of assessing the SLA risk in WDM mesh networks, and the dangers implied by dealing only with steady state probabilities. They developed a method to assess the SLA risk when the stochastic properties of the networks are partially unknown. However, they assume independence between failure processes, Poison failure arrivals, and they do not consider the existence of overlapping failures.

Analysis of real network failure processes is mandatory in order to get the appropriate information for availability dimensioning and to deal with the risks associated with SLAs. In spite of this, for a number of reasons, among them that failures of their networks are not what operators like to have exposed in a competitive commercial marketplace, the access to such failure log information is very limited. A study of the failure behavior in an operational

☆ Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence appointed by The Research Council of Norway, funded by the Research Council, NTNU, UNINETT and Telenor. http://www.q2s.ntnu.no.

* Corresponding author. Tel.: +47 735 92783.
E-mail addresses: andresgm@q2s.ntnu.no (A.J. Gonzalez), bjarne@q2s.ntnu.no (B.E. Helvik).

0140-3664/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comcom.2012.08.007

backbone network is reported by Iannaccone et al. [4]. They examine the frequency and duration of failure events and discuss various statistics, like the distribution of inter-failure times and distribution of link failure durations. This work was continued by Markopoulou et al. in [5], where failures and repairs in the Sprint IP backbone Network are classified and analyzed.

The objective of this paper is to evaluate the SLA success probability of network connections that use shared backup protection, based on operational data obtained from the UNINETT's network management system [6]. For the assessment, we use trace driven simulation (*TDS*) and Monte Carlo methods, taking into account dependencies between events. To the authors' knowledge, this is the first time that a study of the SLA success probability is made based on real correlated failure processes.

Our first approach (TDS) uses filtered ON/OFF processes measured during two years. We extend the number of observations using randomly generated circular shifts. With this method, we capture directly the properties of failure and repair processes. Additionally, when correlation is detected, the circular shift is made in blocks. This approach is a bootstrapping technique to gain more information out of the logged network behavior. Nevertheless, in order to obtain a more accurate estimation of the success probability, we propose the use of a Monte Carlo method that generates failure and repair times based on fitted models. In addition, the correlation is handled by using the Marshall and Olkin copula proposed in [7]. Finally, we analyze the behavior of the probability density function of the interval availability for different observation periods.

This paper is organized as follows. In Section 2, issues related to the distribution of the interval availability and their relation to SLAs are introduced and discussed. In Section 3, the UNINETT's IP backbone network and the information collection method are presented. Section 4 shows the simulation setup and the criteria used in order to allocate connections in the UNINETT's network. Section 5 describes the methods used to estimate dependencies in failure and repair processes. Section 6 studies the success probability in connections that use shared backup path protection, using trace driven simulations. Section 7 describes the Monte Carlo approach used to evaluate the success probability and show the evolution of the probability density of the interval availability with the interval duration. Finally, Section 8 concludes the paper.

2. SLA success probability

A *network connection* is a group of interconnected routers and links that provide end to end service. Maintaining it operational (up/working) is salient to offer a good quality of service. Its performance as a function of time can be modeled with the random process $O(t)$ defined as follows:

$$O(t) = \begin{cases} 1 & \text{If the connection is working} \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

The Interval Availability $\widehat{A}(\tau)$ is an important element of the SLA. $\widehat{A}(\tau)$ is a stochastic variable that measures the time that the connection has been working during a defined period τ, i.e.:

$$\widehat{A}(\tau) = \frac{1}{\tau}\int_0^{\tau} O(t)dt \quad (2)$$

When the transient behavior of a connection is studied, the expected interval availability $A(\tau) = E[\widehat{A}(\tau)]$ is usually evaluated, given that the first moment of a random variable is simpler to obtain than the entire probability distribution. $A(\tau)$ converges always after a long period to the steady state availability A defined as:

$$A = \lim_{\tau\to\infty} A(\tau) \quad (3)$$

When an SLA is defined, the provider promises an availability α for a given period τ (the duration of the contract). Under this scenario, to know the probability that the availability after some observation period τ will be larger than or equal to the defined guarantee is crucial. In this case, the evaluation of the expected interval availability is not enough for the estimation of such probability, and hence the entire probability distribution has to be considered. In this paper, the SLA Success Probability is defined as follows:

$$S(\tau,\alpha) = Pr[\widehat{A}(\tau) \geqslant \alpha] \quad (4)$$

Additionally, the *risk* will be defined as the probability that the specified availability α will not be met, which can be expressed as $1 - S(\tau,\alpha)$. Fig. 1(a) shows the general shape of the probability density function (PDF) of the interval availability and how the risk and the success of the SLA can be estimated from this information.

The success probability as a function of the observation period was first discussed by Goyal and Tantawi [2]. It has basically two different behaviors depending on whether the guaranteed value can be met in the asymptotic case or not. In the first case, $\alpha \leqslant A,\ S(\tau,\alpha)$ drops below one for a period, but converges back to one. In the second case, $\alpha > A$, it decreases continuously to zero. In order to illustrate this effect, we simulate the behavior of a network connection, evaluating the success probability for different values of τ. Fig. 1(b) shows the behavior of $S(\tau,\alpha)$ for a network connection with five independent links having Poisson failure arrivals with expectation of one year, and i.i.d. negatively exponentially distributed down times with expected repair time of 12 h, obtaining a path steady state availability of 0.99315. From Fig. 1 one can observe that the shape of the interval availability PDF changes considerably with the duration of the SLA. In Section 7.3 the evolution of the probability density of $\widehat{A}(\tau)$ with the increase of τ will be described.

$S(\tau,\alpha)$ converges to one or zero after certain value of τ. The convergence speed depends on the ratio between the guarantee α and the asymptotic availability A, as well as on the burstiness of the failure and repair processes that affect the network devices. For instance, in [8], a study of $S(\tau,\alpha)$ under Weibull distributed time to failure and time to repair processes shows that the risk increases considerably faster in both cases ($\alpha \leqslant A,\ \alpha > A$) when the shape parameter is shorter than one.

The accumulated down time over τ $t(\tau)$ is associated with $\widehat{A}(\tau)$ as: $t(\tau) = \tau[1 - \widehat{A}(\tau)]$, where $\Omega(\tau,t)$ will be defined as the CDF of $t(\tau)$. A general expression for $\Omega(\tau,t)$ was derived by Takács in [9] as follows

$$\Omega(\tau,t) = \sum_{n=0}^{\infty} H_n(t)[G_n(\tau - t) - G_{n+1}(\tau - t)] \quad (5)$$

where the failure and repair processes are described by i.i.d. up and down times with CDF $G(t)$ and $H(t)$ respectively, and the subindex n represents the n-fold Stieltjes convolution of a given function.

Eq. (5) characterizes a problem with general distributions. For the case of failure and repair processes negatively exponentially distributed, a complete result was obtained by Takács as

$$\Omega(\tau,t) = e^{-\lambda(\tau-t)}\left[1 + \sqrt{\lambda\mu(\tau-t)}\int_0^{\infty} e^{-\mu y}\sqrt{y}I_1(2\sqrt{\lambda\mu(\tau-t)y})dy\right] \quad (6)$$

where λ and μ are the respective failure and repair rates and I_1 is the Bessel function of order 1. However, $\Omega(\tau,t)$ is difficult to compute for other kind of distributions due to: (i) $G(t)$ and $H(t)$ represent the CDF of up and down times of the entire connection which depends on the corresponding failure and repair distributions of all the network elements involved in the connection. (ii) To obtain the n-fold Stieltjes convolution of a generally distributed function