

Availability and QoS Performance Evaluation of Public IP Networks

Tivadar Jakab¹, Gábor Horváth¹, Rozália Konkoly², Éva Csákány²

¹ Department of Telecommunications, Budapest University of Technology and Economics

² Magyar Telekom Plc. – PKI Telecommunications Development Institute

Corresponding Author: Mr Tivadar Jakab

Assistant Professor

Department of Telecommunications,

Budapest University of Technology and Economics,

Postal address: Room 123, Informatics Building,

Magyar tudósok krt. 2, Budapest H-1117, Hungary

E-mail: jakab@hit.bme.hu

Phone: +35 1 463 1010

Fax: +36 1 463 3263

Keywords: availability, quality of service (QoS), analysis, IP over WDM, real size network, application case study.

Abstract

The wide spread of high speed access technologies evolves a mass market of broadband Internet services. The 3play service concept targeting this market covers voice, data and video delivery on the same packet based service platform. To upgrade current IP core and aggregation networks in order to meet enhanced service availability and quality requirements for voice and video services, the evaluation of related currently available network capabilities may provide the starting point. On the other hand, to guide efficient developments, the analysis of designed networking solutions is required, as well. These considerations significantly increase the need for efficient network analysis support in real market oriented processes. This paper introduces an availability and QoS analysis framework developed and applied for real-size networks.

The main scope of the presented work is on the adaptation of theoretical availability and QoS analysis methodology for realistic cases. Both problem size and complexity reduction, as well as efficient worst case estimation approaches are considered to cope with difficulties of large-size real network analysis problems. Adaptation of the analysis methodology for a desktop GRID is overviewed. Numerical examples are presented to illustrate the size, complexity and structure of the problems to be solved. The integration of the analysis support into real network development environment is considered, as well.

1 Motivations

The widespread of high speed access technologies (xDSL, DOCSIS, BWA, GPON, and UMTS) evolves a mass market of broadband Internet services. The 3play service concept targeting this market covers voice, data and video delivery on the same packet based service platform.

The packet based service platform enables advanced service intelligence and flexibility; thus, VoIP and IPTV services can be equipped with attractive advanced service features. However, user experiences of service availability and quality of traditional voice (TDM based PSTN) and TV (e.g.

CATV) define the expectations concerning these new services. To meet these expectations network service requirements for aggregation and IP core networks should be modified significantly.

IP aggregation and core networks were originally designed and developed to carry best effort data services (e-mailing, file transfers, web browsing). The logical topology, the architecture and the operational principles have been tuned to support an extensive traffic growth, i.e. the fast and efficient connection of large number of customers.

The evaluation of related network capabilities may provide the starting point to upgrade this network in order to meet enhanced service availability and quality requirements. On the other hand, to guide the developments efficiently, the analysis of planned and designed networking solutions is required, as well. These considerations significantly increased the need for efficient network analysis support in real market oriented processes.

Research activities in availability and queuing theories provide strong scientific background and valuable achievements, thus, the main scope of the related work is currently focused on the adaptation of these results for realistic problems and on the integration of the analysis support into real network development processes.

This paper introduces an availability and QoS analysis framework developed for and applied at the network development processes of Magyar Telekom, the principal provider of telecommunication services in Hungary.

The structure of the paper is as follows: First the general availability analysis problem is identified in Chapter 2, and then Chapter 3 summarizes the analysis approach, the modeling, and the availability and QoS analysis methodology briefly. Chapter 4 describes the analysis frame; Chapter 5 illustrates the structure of the availability analysis problem with a numerical example. Chapter 6 gives the basic idea of a potential adaptation for a desktop GRID; finally, Chapter 7 concludes the paper.

2 The Availability Analysis Problem

In order to formulate the availability analysis problem, let us assume that the following quantities are given:

- a number of network components (i.e., cables, switches, etc.) $i = 1, \dots, N$

- the failure of component i is a binary random variable denoted by $y_i \in \{0, 1\}$ subject to the distribution

$$P(y_i = 1) = p_i \quad P(y_i = 0) = 1 - p_i ;$$

- the failure state space is composed of binary vectors $\mathbf{y} = (y_1, \dots, y_N) \in \{0, 1\}^N$, and represented by $\mathbf{Y}: \mathbf{y} \in \mathbf{Y}$;

- a product measure over the failure state space

$$p(\mathbf{y}) = \prod_{i=1}^N p_i^{y_i} (1 - p_i)^{1 - y_i} ;$$

- a measure of performance in state \mathbf{y} is $Perf(\mathbf{y})$, and $g(\mathbf{y})$ expresses the loss of system performance due to a failure scenario represented by vector \mathbf{y} is

$$\text{derived as } g(\mathbf{y}) = \frac{Perf(\mathbf{y})}{Perf_{\max}} .$$

$Perf(\mathbf{y})$ may express connectivity (e.g. route exists or can be obtained between given endpoints), degradation of capacity (e.g. a certain transmission capacity is required between the endpoints, and only a portion of it is available).

Since $Perf_{\max} \geq Perf(\mathbf{y})$ in the forthcoming discussion we assume that $0 \leq g(\mathbf{y}) \leq 1$.

The two main availability measures are defined as follows:

1. Average Loss (AL) expressed as

$$E(g(\mathbf{y})) = \sum_{\mathbf{y} \in \mathbf{Y}} g(\mathbf{y}) p(\mathbf{y}) .$$

2. Outage Probability (OP)

$$P(g(\mathbf{y}) > C) = \sum_{\mathbf{y}: g(\mathbf{y}) > C} p(\mathbf{y}) ,$$

where C is a given level of performance degradation.

Since the evaluation of $g(\mathbf{y})$ for a single failure scenario \mathbf{y} can be rather tiresome, the complexity of availability analysis is expressed with the number of how many $g(\mathbf{y})$ must be calculated in the course of evaluating the availability measure. However, from the AL and OP definitions it follows that for an exact evaluation one has to calculate $g(\mathbf{y})$ for each possible binary state vectors, which entails $O(2^N)$ computations. Since, even in the simplest network models the number of components N is in the range of a few hundred, while in case of multilayer models it could go up to a couple of thousands, “taking a full walk” in the state space for evaluating the availability measures is clearly out of reach. Therefore, we want to calculate AL or

OP approximately by using an estimate $\tau = \Psi(g(\mathbf{y}_1), g(\mathbf{y}_2), \dots, g(\mathbf{y}_K))$ which is based only on a few samples $\{(\mathbf{y}_k, g(\mathbf{y}_k)), k = 1, \dots, K\}$ taken from the whole state space. This way we can perform availability analysis by “visiting” only a few number of states in the state space. The general algorithm for availability analysis is given as follows:

Generate K samples $Y_{\text{sample}} = \{\mathbf{y}_1, \dots, \mathbf{y}_K\}$.

Calculate the corresponding loss values for

$$\{g(\mathbf{y}_1), \dots, g(\mathbf{y}_K)\}$$

Use the estimate of the given selection method for evaluating AL as

$$AL \sim \Psi(g(\mathbf{y}_1), g(\mathbf{y}_2), \dots, g(\mathbf{y}_K))$$

This way, $g(\mathbf{y})$ must be calculated fewer times which yields a feasible algorithm. Of course, the underlying question is how to find the most ‘typical’ samples which furnish the most accurate estimation despite the small number of samples.

3 The Availability and QoS Analysis Approaches

To develop an efficient availability and QoS analysis methodology a proper network model is required first. Based on the model processes evaluating the required availability and QoS measures in a single network state can be developed and – since the evaluation of the entire state space is not realistic for real-size networks - a sample selection strategy can be chosen and applied to obtain efficient estimations or bounds

The current Chapter summarizes the network modeling requirements, the main approach to the evaluation of the availability performance of a network state, the considered choice of sample selection approaches, as well as the link-wise QoS analysis approach, the applied traffic and queuing models, and the end-to-end QoS estimations.

3.1 Generation of the Layered Network Model

The networks within the scope of the analysis are multi-layer networks both in technological and logical senses. From technological point of view there are IP and WDM network layers with different functions and capabilities that should be taken into consideration both in failure modeling and performance evaluation. From logical point of view technological network layers consist of different logical layers, like IP-link, Ethernet-link, optical channel, optical multiplex section, optical amplifier section, fiber link, cable layers.

These layers are in client-server relation, since lower layers provide services for the upper layers, i.e. lower layer network resources support the realization of upper layer

logical network elements: optical fibers carry wavelength division multiplex transmission systems, WDM multiplex bundles accommodate optical channels, which transport Ethernet-links carrying IP links. IP-links form a logical network providing different IP services to carry application traffic.

Based on the above network view intra-layer and inter-layer modeling requirements can be identified defining a network model for analysis purposes.

Focusing on the model of a single technological layer the dependencies between the failure of physical network elements and the outages of logical network elements should be represented by the model on the one hand, since failure of a physical network element may result in outage of several logical elements. On the other hand the model should support the evaluation of the network including the impacts of resilience and adaptation mechanisms supported by the technology – for example OSPF adaptation in an IP layer, or optical channel protection switching in a WDM layer –, as well.

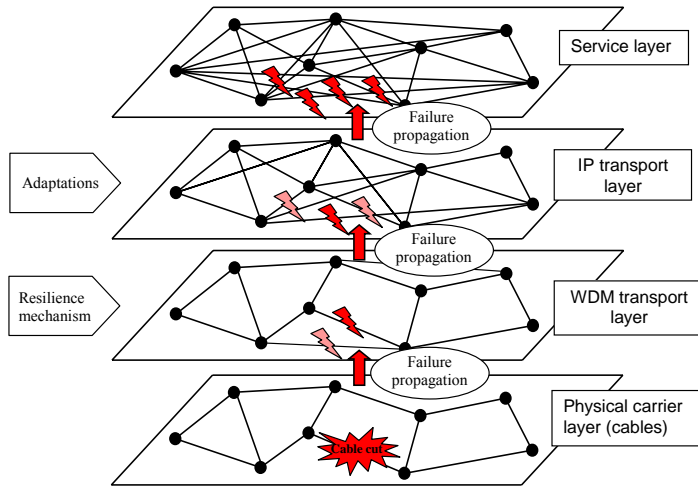


Figure 1 The layered network model and the failure propagation through the layers

Inter-layer modeling requirements define how to combine layer models to form a complete network model. The main issue is to enable the modeling of inter-layer failure propagation: a failure of a logical network element in a given layer propagates towards the client layers, and may result in outage of several logical elements of the client layers (Figure 1). If an adaptation mechanism or a resilience capability eliminates the impact the server layer element failure has on the client layers by reconfiguring the server layer to maintain the services or by replacing the failed resource by a protection switch over action, there is no impact on the client layers. (Note that the analysis is focused on steady state network conditions only.)

The applied technology-independent layered network model is based on [14].

3.2 Availability analysis of real size multi-layer networks

As it has been already emphasized describing the analysis problem the main challenge in the analysis of real-size networks is the sample selection to enable the elaboration of efficient estimations for the targeted availability parameters. The considered choice of solutions for the sample selection in this case includes:

- Monte Carlo method
- Stratified Sampling, and
- Li-Silvester deterministic bounding.

The brief summary of these methods is based on [1].

One of the oldest methods of availability analysis is the Monte Carlo (MC) method. Its shortcomings as statistical estimation (e.g. slow convergence, low efficiency) are well known. The reason for briefly discussing this method is that it will serve as a reference for the efficiency of more powerful algorithms. In the case of MC algorithm the expected value is estimated by the following form:

$$AL = E(g(\mathbf{y})) \approx \frac{1}{L} \sum_{i=1}^L g(\mathbf{y}_i),$$

where the sequence $\mathbf{y}_i; i=1, \dots, L$ are drawn as random samples subject to the underlying distribution $p(\mathbf{y})$.

The crucial question in Monte Carlo simulations can be posed as how many samples should be taken for a required level of accuracy. This is a very important question indeed, as the ratio $\frac{L}{|Y|}$ determines the percentage of the state space

visited by the method.

To accelerate the Monte Carlo simulations [2] there is a speed-up technique based on grouping the samples into different classes referred to as stratified sampling. Stratified sampling has been extensively investigated in the literature of statistical sampling [3, 4], however some novel adaptive extensions are discussed in this chapter. Again we want to estimate

$$AL = E(g(\mathbf{y})) = \sum_{\mathbf{y} \in Y} g(\mathbf{y})p(\mathbf{y})$$

Stratified (or structural) sampling is defined as follows:

Given

- partition $Y = \{Y_i, i=1, \dots, V\}$ $Y = \bigcup_{i=1}^V Y_i$ and $Y_i \cap Y_j = 0$
- the probability of being in class $i \rightarrow P_i = \sum_{\mathbf{y} \in Y_i} p(\mathbf{y})$
- the average loss expressed in a structured form $E(g(\mathbf{y})) = \sum_{i=1}^V P_i E(g(\mathbf{y}) | \mathbf{y} \in Y_i) = \sum_{i=1}^V P_i m_i$

- a sampling allocation (L_1, L_2, \dots, L_V) $\sum_{i=1}^V L_i = L$
- an estimation of the conditional expected value m_i

$$m_i \approx \frac{1}{L_i} \sum_{k=1}^{L_i} g(\mathbf{y}_k^{(i)})$$
- the overall estimation $\eta := \sum_{i=1}^V P_i \frac{1}{L_i} \sum_{k=1}^{L_i} g(\mathbf{y}_k^{(i)})$

For more details on strata definition, sample size, sample allocation and accuracy under different conditions see [5].

On the other hand, one can approach availability analysis as a deterministic computational problem, namely how to decrease the number of terms in the summation in AL or OP.

The probabilities $p(\mathbf{y}), \mathbf{y} \in \{0,1\}^N$ can be easily calculated and arranged in a monotone decreasing sequence $p(\mathbf{y}^{(1)}) \geq p(\mathbf{y}^{(2)}) \geq \dots \geq p(\mathbf{y}^{(L)})$. This separates Y into two sets, $Y_1 = \{\mathbf{y}^{(k)}, k=1, \dots, N\}$ and $Y_2 = Y \setminus Y_1$, respectively. By using the following two functions:

$$h(\mathbf{y}) = \begin{cases} g(\mathbf{y}) & \text{if } \mathbf{y} \in Y_1 \\ 0 & \text{if } \mathbf{y} \in Y_2 \end{cases}$$

$$q(\mathbf{y}) = \begin{cases} g(\mathbf{y}) & \text{if } \mathbf{y} \in Y_1 \\ 1 & \text{if } \mathbf{y} \in Y_2 \end{cases}$$

one can lower and upper bound AL as follows:

$$\sum_{\mathbf{y} \in Y} h(\mathbf{y}) p(\mathbf{y}) \leq \sum_{\mathbf{y} \in Y} g(\mathbf{y}) p(\mathbf{y}) \leq \sum_{\mathbf{y} \in Y} q(\mathbf{y}) p(\mathbf{y})$$

These inequalities can be rearranged into the following forms, known as Li-Silvester (LS) bounds [6]

$$\sum_{\mathbf{y} \in Y_1} h(\mathbf{y}) p(\mathbf{y}) \leq \sum_{\mathbf{y} \in Y} g(\mathbf{y}) p(\mathbf{y}) \leq \sum_{\mathbf{y} \in Y_1} q(\mathbf{y}) p(\mathbf{y}) + P(Y_2).$$

Here $P(Y_2) = \sum_{\mathbf{y} \in Y_2} p(\mathbf{y})$. One can notice, however, that we

took the first N most significant samples by arranging the probabilities into a decreasing order

$$p(\mathbf{y}^{(1)}) \geq p(\mathbf{y}^{(2)}) \geq \dots \geq p(\mathbf{y}^{(L)})$$

instead of arranging the products

$$g(\mathbf{y}^{(1)}) p(\mathbf{y}^{(1)}) \geq g(\mathbf{y}^{(2)}) p(\mathbf{y}^{(2)}) \geq \dots \geq g(\mathbf{y}^{(L)}) p(\mathbf{y}^{(L)}),$$

which would have resulted in sharper bounds. This cannot be carried out since the calculation of $g(\mathbf{y})$ over the full space (which is necessary to obtain the first N most significant samples) is out of reach. In spite of this deficiency the LS bound can result in rather sharp estimation.

3.3 QoS analysis of end-to-end IP services

To evaluate the performance of the network configurations resulted in different failure cases, these network configurations are assessed in details. The actual available IP logical topology is defined, and the traffic is routed according to the adapted routing tables. Therefore, the detailed traffic load on each link is obtained according to the present service classes.

This detailed network model and description provide proper input for QoS analysis. To elaborate end-to-end QoS parameters for different services link-wise basic analytic models and simple end-to-end worst case estimations are required. This Chapter briefly summarizes the applied methodology.

3.3.1 Link-wise performance analysis

By using the recent results of queuing theory, we are able to compute the most important performance measures like mean and variance of the packet waiting time (delay), packet loss probability and waiting time quantiles under very general conditions.

During the analysis of the nodes we apply the matrix geometric methods [7]. The application of matrix geometric algorithms makes possible to obtain approximate Markov models even if the system itself is not Markovian, thus, the inter-arrival and service times are not exponentially distributed. With these methods general (even correlated) inter-arrival and service times can be considered, and complex service disciplines like strict priority or weighted fair queuing scheduler policy can be approximated.

3.3.2 Traffic Model

The most popular traffic descriptor in matrix geometric methods is the Markovian Arrival Process (MAP, [7]). A MAP consists of phases with the transitions between the phases described by a Markov chain. The arrival rate can be different in the different phases, and an arrival can cause a phase transition as well. A MAP is able to approximate general inter-arrival time distributions and can capture autocorrelation between the arrivals.

There are several methods published how to construct a MAP to reflect a target behavior. These methods usually obtain the parameters of the MAP by optimization such that the statistics of the traffic it generates matches the measured real traffic statistics.

Using MAPs as a feeding process of a queuing model has the benefit that the model remains Markovian, and can be solved by efficient numerical procedures.

As MAPs can model general arrival processes, phase-type (PH) distributions can be used to approximate general service times. Similar to MAPs, PH distributions represent random variables consisted of a mixture of exponential

distributions. Several PH fitting algorithms have been published to approximate general service time distributions. In our computation method we apply a MAP with only two phases for two reasons. First, with such a small number of phases the state space of the Markov chains modeling the individual queues remains moderate, thus the solution is fast and involves as few numerical issues as possible. Second, the inverse characterization problem of MAPs is solved only for two phases. Thus, given the first three moments and the lag-1 autocorrelation of the inter-arrival times of the packets, explicit formulas are available to compute the MAP that has the same statistical behavior [8]. Similarly, simple and explicit inverse characterization formulas are available for PH distributions with two phases.

To summarize our traffic model, we measure the first three moments and the autocorrelation of the packet arrival stream and construct a 2-state MAP that has the same properties based on the method of [8]. Similarly, the packet size distribution is approximated by a PH distribution constructed by using the three moments of the measured packet sizes using the method of [15].

3.3.3 Best Effort service

The queuing behavior of a best effort traffic class with the packet arrivals described by a MAP and packet sizes described by a PH distribution can be modeled by a MAP/PH/1 queue (or by its finite variant if the buffer is finite), with the following structure:

$$\begin{bmatrix} B_0 & A_0 & & & \\ C_1 & B & A & & \\ & C & B & A & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \quad (1)$$

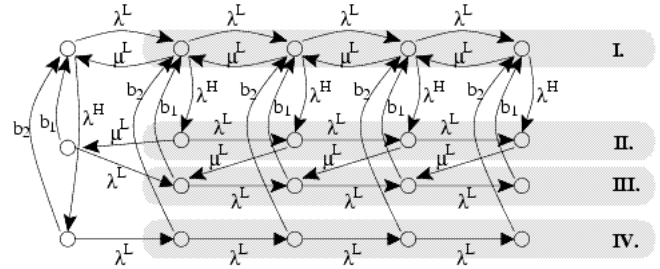
where A_0, B_0, C_1, A, B, C are matrices, computed from the MAP describing the arrivals and from the PH describing the service times (see [1]).

There are many solution techniques with well-studied numerical behavior to obtain the QoS parameters of queuing systems having such a regular structure [7, 9]. We selected the classical matrix-geometric method with the logarithmic reduction algorithm if the buffer length is infinite, and we applied the Folding algorithm when the buffer size is finite.

3.3.4 Strict priority queuing

The difficulty of modeling multi-class queuing systems is that the state space of the resulting model is multi-dimensional (each dimension represents the queue length of a traffic class). In most cases such multi-dimensional Markov chains can not be solved efficiently. If the scheduling policy among the classes is strict priority, the multi-dimensional Markov chain can be solved in an exact

way [10], but numerical problems appear at several points of that algorithm. Therefore we apply an approximate solution based on the separate analysis of the traffic classes. The concept is to approximate the two class system as the classes were separated, and construct a service process for both classes that approximately imitates the behavior of the original server. This special service process is constructed such that it approximates the service the tagged traffic class perceives in the multi class environment.



- I. No high priority customer
- II. There are high pr. customers (blocked by low pr) low pr. customer in the server
- III. There are high pr customers (not blocked any more) high pr. customer in the server
- IV. There are high pr. customers high pr. customer in the server

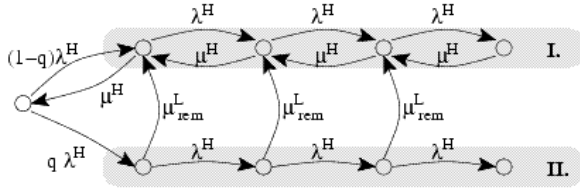
Figure 2 The structure of the approximate Markov chain model of the low priority queue

From the point of view of the low priority customer class, the exact number of high priority customers does not play any role. When there are no high priority customers, the server is available, and when there are high priority customers, the server is not available for low priority customers. Therefore, during the analysis of the low priority queue, the two dimensionally infinity state space is eliminated such a way, that the number of high priority customers is modeled by only 2 states: zero, and more than zero. This approach is reflected by Figure 2, which depicts the structure of the approximate Markov chain model of the low priority queue.

The high priority customers can be affected by the low priority customers only at one point: when the high priority queue is empty at the arrival of a high priority customer, and a low priority customer is in the server. In this case the arrived high priority customer has to wait the remaining service time of the low priority customer, since the service is non-preemptive. The probability of this event (q) will be computed from the queue model of the low priority class. Figure 3 shows the structure of the corresponding Markov chain.

The generator matrices of these Markov chains – with proper state ordering – have the same matrix tri-diagonal structure as show by equation (1). This means that the

solution methods mentioned there can be applied to compute the per-class performance measures. Of course, the matrix blocks of the generators are more complicated to compute compared to the best effort case (it also involves busy period analysis), for the detailed description see [11].



- I. No low pr. customer at first arrival
- II. There was a low pr. customer at first arrival (high pr. customer is blocked)

Figure 3 The Markov chain of the queue model of the low priority class

3.3.5 Weighted fair queuing

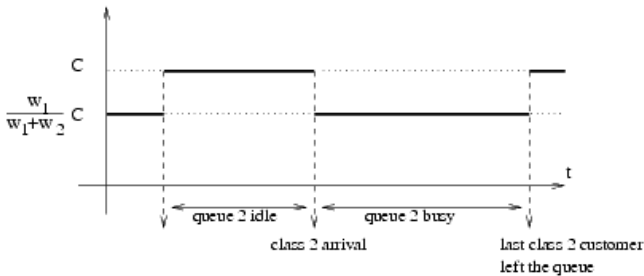


Figure 4 The varying capacity of the server

The Markov model of the system is multi-dimensional besides weighted fair queuing (wfq) as well. No numerically applicable solutions are known for this service discipline. An approximate solution can be found for the two class case in [12].

The idea of the approximation is to separate the classes like in case of the analysis of the priority system. From the point of view of a customer class, the capacity of the server is varying depending on the presence of customers belonging to the other class (Figure 4).

The busy period analysis of the traffic classes provides the durations the service process spends in the full and in the reduced capacity states.

This capacity-switching behavior is reflected on Figure 5, which shows the structure of our approximating Markov chain. The states can be partitioned into two groups: one belongs to the full capacity, the other to the reduced capacity state.

During the analysis the inter arrival times are characterized by MAPs and the service times are phase type distributions (the phases of these processes belong to the state space as well, the figure above show only the macro structure of the

Markov chains). The generator of the resulting Markov chain has a block tri-diagonal structure again (the exact definition of these matrices is available in [12]), whose performance analysis is studied extensively in the literature.

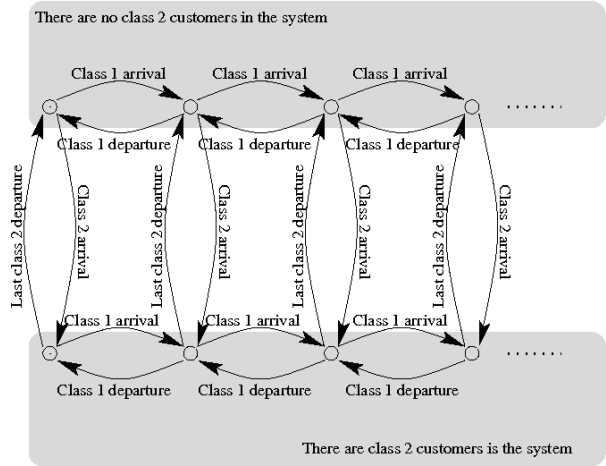


Figure 5 The capacity-switching behavior

3.3.6 Extension to more than two traffic classes

The case of more than two traffic classes is reduced to the two class analysis.

In case of strict priority policy it is performed as follows. During the analysis of the highest priority queue all the traffic of the lower priority queues are aggregated and considered as the only low priority class in the system. During the analysis of the lower priority classes the effect of even lower priority classes is neglected, and all the higher priority classes are aggregated and considered as the only high priority class of the system.

In case of weighted fair queuing (wfq) the idea is similar. We analyze the traffic classes one by one. During the analysis of a tagged traffic class all the traffic of the other classes are aggregated and considered as one traffic class. The weight of this aggregated class in the wfq scheme is computed by the average of the weights of the components weighted by their traffic intensity.

3.4 End-to-end QoS parameters

The end-to-end QoS parameters are approximated by the aggregation of the per-node performance measures along a route. Since the correlation between the QoS measures at the different nodes can not be taken into account by such a heterogeneous network model, we assume that the waiting time and the loss probability perceived by the packets at the consecutive network nodes are independent. Thus, the mean and the variance of the waiting time can be computed by summing these measures along a route:

$$E(W_{E2E}) = \sum_{n \in route} E(W_n)$$

$$\sigma_{W_{E2E}}^2 = \sum_{n \in \text{route}} \sigma_{W_n}^2.$$

The loss probability is approximated similarly, by assuming independence:

$$P_{\text{loss}_{E2E}} = 1 - \prod_{n \in \text{route}} (1 - P_{\text{loss}_n}).$$

4 The analysis framework and environment

Having an overview of the considered methodology in the previous Chapters, let us focus on the implemented analysis framework, and the business environment.

The network to analyze and evaluate is a service provider's network consisting of access, aggregation and core network domains.

The majority of the access networks is realized by ADSL technology, and provides an order of Mbyte per sec downlink and a few hundreds of kbyte per sec uplink. The aggregation is a switched carrier class Ethernet network with typical tree-like topology.

The core part of the IP network is directly routed, there is no switched Layer 2, Ethernet-links carried by WDM optical channels interconnect the IP routers directly. (From analysis point of view this architecture eases the QoS analysis problem, since simplifies the assessment of traffic load over a given physical resource.)

On the current extend of the work the main focus is on core part, however, the extension of the analysis for the aggregation is foreseen in the near future, as well.

The detailed network data for modeling the WDM layer is provided by the network management databases (equipment, systems, cable infrastructure). The description, configuration and neighborhood information of the IP routers are available from SNMP downloads. Based on these detailed information the required network model can be obtained.

Deciding on the sample selection methodology the considerations are as follows: Besides the general description in Chapter 3.3 the illustration on Figure 6 based on a rather small, but realistic network example [5] represents the high performance of Stratified Sampling based estimation of availability parameters. Despite of this fact the significantly less efficient Li-Silvester deterministic bounding was applied in the implemented analysis process. The reasons are twofold: the practical one is that strict bounds are more preferred to describe the availability performance of a network service for business purposes than statistical estimations. "The service is not worse than ..." is a clear worst case characteristic to be compared with the requirements or expectations. On the other hand, there are some difficulties to apply Stratified Sampling based estimations to describe availability performance of a single end-to-end service.

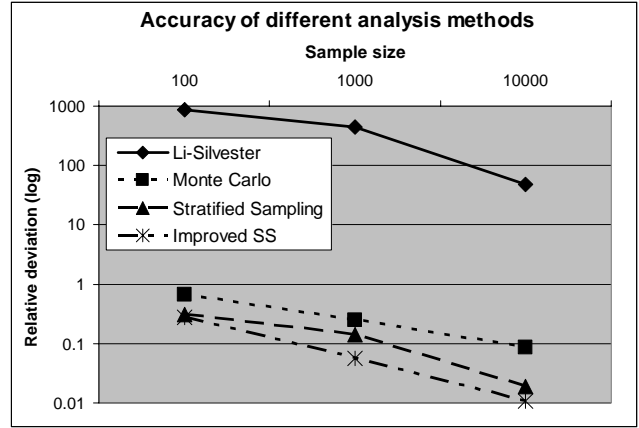


Figure 6 Efficiency of different sample selection methods

Stratified Sampling based sample selection strategy is proved to be efficient for estimation of overall network level parameters [5]. These characteristics are for strategic purposes to evaluate and compare architectural solutions and strategic networking alternatives. Since each network element contributes to the overall network level performance, thus, an obvious definition of strata is based on the number of failed components (transmission links and nodes) in the given network state. To extend the approach for a layered network structure these components can be categorized by technology.

Targeting to estimate the availability of a single point to point network service it is quite difficult to distinguish which network element contributes to the accommodation of the given service traffic even if dynamic mechanism is present in the network like OSPF adaptation. Lack of separation of these network element types may decrease the efficiency of stratified sampling significantly.

An improvement of the basic Li-Sylvester bounds for estimating the average loss or the outage probability theoretically can be obtained quite obviously. In order to do that let us introduce the following definitions and notations:

binary vector \mathbf{y} is said to be the "successor" of binary vector $\tilde{\mathbf{y}}$ (denoted by $\mathbf{y} \prec \tilde{\mathbf{y}}$), if the following condition holds:

$$\text{if } y_i = 1 \text{ then } \tilde{y}_i = 1 \text{ and } w(\mathbf{y}) \geq w(\tilde{\mathbf{y}}),$$

where $w(\mathbf{y})$ represents the weight of the binary vector \mathbf{y} (i.e. $w(\mathbf{y})$ yields the number of ones in binary vector \mathbf{y}).

We assume that if $\mathbf{y} \prec \tilde{\mathbf{y}}$ then $g(\mathbf{y}) \leq g(\tilde{\mathbf{y}})$ (this assumption holds for the large majority of systems, since it asserts that having additional failures in a certain failure scenario, then the performance is the same or it can only degrade). Based on this consideration the bounds can be improved. Unfortunately, to obtain the groups of states for such an improved bound calculation implies combinatorial

difficulties in implementation, and real gains in efficiency are rather difficult to realize.

Some reduction of $g(\mathbf{y})$ calculation needs can be achieved taking into account the specialties of the availability model structure of layered networks.

If there are two failure states \mathbf{y}_i and \mathbf{y}_j with identical performances $g(\mathbf{y}_i) = g(\mathbf{y}_j)$ it is well enough to perform the calculation of the related $g(\mathbf{y})$ only once. If the failure states with definitely identical performances can be identified and group properly, and the need for $g(\mathbf{y})$ calculations can be reduced. Is it possible to identify and group failure states with identical performance by reasonable efforts? In our case, when the analysis targets to evaluate IP network services in a multi-layer network, the answer is definitely yes. The performance of IP services depends on the configuration of IP logical topology (subset of failure free links and routers) and the underlying Ethernet-link physical topology. The IP logical topology defines the conditions for traffic routing, the Ethernet-link physical topology accommodates the routed traffic. How can different failure cases result in identical IP logical topology and Ethernet-link physical topology network configuration?

For a simple illustration assume that there is a GbE port card in a router with one active port only. A transponder to perform electrical/optical signal conversion and a WDM terminal multiplexer is installed, and a fiber carries this single 1GbE connection. It is a simple serial availability structure to realize the given IP link, thus, the failure of any element interrupts the link. If it is a single failure in the network the resulted IP logical topology is the same for each failed single element from this structure. Therefore these failure configurations can be grouped, and the performance can be calculated only once. Without any morphological analysis of availability models, just comparing the resulted IP-link logical and the corresponding Ethernet-link physical topologies these types of coincidences can be identified efficiently.

5 A numerical example

To illustrate the size and structure of a real size network analysis problem a certain part of the Magyar Telekom IP network is discussed here. The network model consists of about 1000 active model elements, summarized in Table 1.

A typical requirement for service availability is 99,99% (i.e. of 1 hour per year outage approximately). It is an obvious expectation that the accuracy of the analysis should be better by an order of magnitude: 0.99999 of the state space should be covered, i.e. according to the Li-Silvester bounding $P(Y_2) \leq 0.000001$.

Based on the analysis of the first 1000000 most probable failure states Table 2 summarizes the problem statistics.

As it can be seen from Table 2 the first 1000000 most probable failure cases cover 0.99997 of the state space.

Applying the above described grouping of failure states according to their performance, the 1000000 failure cases imply 445874 different network configurations regarding IP logical topology and Ethernet-link physical topology.

Table 1 Network model elements

Element category	Number of elements	Typical order of DTR
Router	119	$10^{-4} \dots 10^{-5}$
Router port card	286	10^{-5}
Optical channel	32	10^{-5}
Optical multiplex section	36	10^{-4}
Optical amplifier section	56	10^{-5}
Cable link	437	$10^{-4} \dots 10^{-6}$
Network node (e.g. common functions like power supply)	33	10^{-7}

Table 2 Failure statistics based on the analysis of the first 1000000 most probable failure cases

Failure depth	Number of states	Accumulated probability
failure free	1	0.94284
single	999	0.055498212
double	489745	0.001630879
triple	509256	1.47E-06
not covered (triple, quadruple and higher)	n. a.	2.94E-05

Figure 7 depicts the productivity of the two different performance calculation approaches. The productivity is expressed as the average number of failure states covered by the evaluation of a single network configuration.

If the grouping of failure states according to identical IP logical topology and Ethernet-link physical topology would not be applied, the evaluation of $g(\mathbf{y})$ should be performed for all failure configurations, and thus the performance of each failure configuration should be evaluated individually, i.e. the productivity would be 1. If the grouping of failure states is applied, the first 10 evaluation covers more than 500 failure states, the first 100 covers more than 3000, i.e. their productivity is 50 and 30, respectively (failure states are taken according to their decreasing probability). For the 1000000 analyzed failure case, we take into account 445874 different network configurations concerning IP logical topology and Ethernet-link physical topology, and so, the productivity is above 2, i.e. applying the grouping of failure states the processing load of needed $g(\mathbf{y})$ calculations became halved.

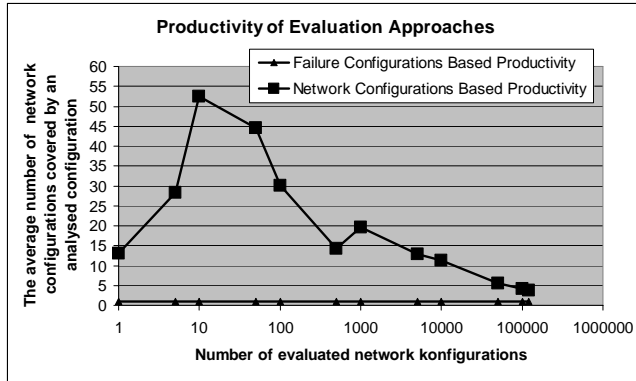


Figure 7 Productivity of different performance calculations

The gain involved in the necessary $g(y)$ calculations can be achieved by the proper grouping of failure cases, it is significant. However, the processing load is very high from practical application point of view, since tens of days are required to perform the analysis of a network of illustrated complexity on a good desktop PC.

Li-Silvester bounds provide a scalable approach for availability performance estimation; however the efficiency of the method is weak. The presented illustration shows that the accuracy specified by the realistic application requires the evaluation of a significant part of the failure state space, and the processing load is high.

Since there are no promising further improvements in the efficiency of sample selection and the reduction of failure state performance calculations, the processing capacity should be increased to fulfill the requirements of practical applications.

6 Adaptation for a desktop GRID Frame

A proper and economical option for a scalable increase of the processing capacity is a desktop GRID. A desktop GRID environment organizes tens or even hundreds of desktop PCs into a commonly controlled processing frame [13]. Since in the network analysis there are some tasks that can be processed independently, and in parallel without strict scheduling requirements, the adaptation of the analysis process for a desktop GRID environment seems to be promising.

In the analysis problems and approaches discussed in the current paper there are two groups of tasks of that type. On the one hand the elaboration of network configurations according to the different failure cases and their evaluation can be done independently from one another. On the other hand the evaluation of link-wise QoS models to provide input for end-to-end estimation in a given failure case can be done independently from one another, as well (Figure 8). The gridification of these analysis processes is efficient, and due to the low communication and control overhead the expected gain in processing time is 0.8 – 0.9 by each involved processing unit, thus a few tens of PCs involved

into a desktop GRID may decrease the required processing time below one hour, what seems to be acceptable to solve a problem of that complexity.

7 Summary and conclusions

Based on the identification of analysis requirements and the overview of available methodology the current paper describes a network availability and QoS analyzing frame. Some considerations and illustrations are discussed how the efficiency of the analysis can be tuned to the practical applications.

Based on the numerical illustration of the problem size and structure the need for high processing power is identified, and the adaptability for a desktop GRID is shown.

The integration of the developed analysis approach into the network development processes of Magyar Telecom is overviewed. Currently, the developed analysis frame is successfully applied in the Magyar Telecom's network development processes. There are two directions for further work identified:

The current availability modeling of IP routers is based on the hardware models only, failures caused by software and configuration errors has not been covered by the model, yet. Thus, to refine the availability modeling of IP routers trouble ticketing based modeling of software and configuration related failures is a research and development issue for the near future.

New services and the permanently increasing capacity demands require the implementation of further technological changes. The introduction of IP-MPLS technology with Traffic Engineering, and All Optical Network technology will imply the extension of the modeling capabilities in the near future.

8 Acknowledgement

The research work presented in this paper has been initiated and partially granted by OTKA 048985 Project titled "Dimensioning and reliability analysis of fault-tolerant networks in Differentiated Reliability (DiR) environment". The adaptation of the developed methodology for real case applications was performed within the frame of the *XPLANET Network Planning and Analysis Tool development* in co-operation with the PKI Telecommunications Development Institute of Magyar Telekom.

9 References

- [1] J. Levendovszky: Novel statistical methods in network reliability analysis, Proc. of 6th International Workshop on Rare Event Simulation, RESIM'06, October 2006, Bamberg Germany;
- [2] W. E. Deming, Some Theory of Sampling, Dover Publ., Inc. New York, 1966

[3] J. Carlier, Y. Li, J. Lutton, Reliability Evaluation of Large Telecommunication Networks by Stratified Sampling Method. Proceedings of the 6th International Network Planning Symposium, Networks'94, Budapest, Hungary, Sept 1994, pp. 113-118.

[4] A. Kiss, J. Levendovszky, L. Jereb, Stratified Sampling Based Network Reliability Analysis. 8th International Conference on Telecommunications Systems, Nashville, Tennessee, USA, 2000

[5] L. Jereb, F. Unghváry, T. Jakab, "A Methodology for Reliability Analysis of Multi-Layer Communication Networks", Optical Networks Magazine, 2 (2001), pp. 42-51

[6] V. O. K. Li, J. A. Silvester, Performance Analysis of Networks with Unreliable Components. IEEE Transactions on Communications, Vol.COM-32, 1984, pp. 1105-1110.

[7] G. Latouche, V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, American Statistical Association and the Society for Industrial and Applied Mathematics, 1999.

[8] L. Bodrog, A. Heindl, G. Horváth, M. Telek, A Markovian Canonical Form of Second-Order Matrix-Exponential Processes, Journal of Appl. Math. and Stoch. Analysis, 2007, submitted paper.

[9] J. Ye, S. Li, A computational method for finite QBD processes with level-dependent transitions, IEEE Trans. Comm., 42(2):625-639, 1994.

[10] Tetsuya Takine, The Non-preemptive Priority MAP/G/1 Queue, Operation Research, 47(6):917-927, 1999.

[11] G. Horváth, A Fast Matrix-Analytic Approximation for the Two Class GI/G/1 Non-Preemptive Priority Queue, Analytical and Stochastic Modeling Techniques and Applications, 2005.

[12] G. Horváth, Approximate Analysis of Two Class WFQ Systems, Workshop on Performability Modeling of Computer and Communication Systems - PMCCS, 2003.

[13] A. Cs. Marosi, G. Gombás, Z. Balaton, Secure application deployment in the Hierarchical Local Desktop Grid, Proceedings of the 6th Austrian-Hungarian Workshop on Distributed and Parallel Systems (DAPSYS) pp. 145-154. 2006, Innsbruck, Austria.

[14] L. Jereb, P. Bajor, A. J. Kiss, "Network Reliability Analysis Based on Multilayer Models", in Proceedings of the 7th International Conference on Telecommunication Systems/Modeling and Analysis, Nashville, USA, 1999, pp. 490-500

[15] M. Telek and A. Heindl, Matching moments for acyclic discrete and continuous phase-type distributions of second order, *International Journal of Simulation Systems, Science & Technology*, vol. 3:3-4, Dec. 2002. Special Issue on: Analytical & Stochastic Modeling Techniques.

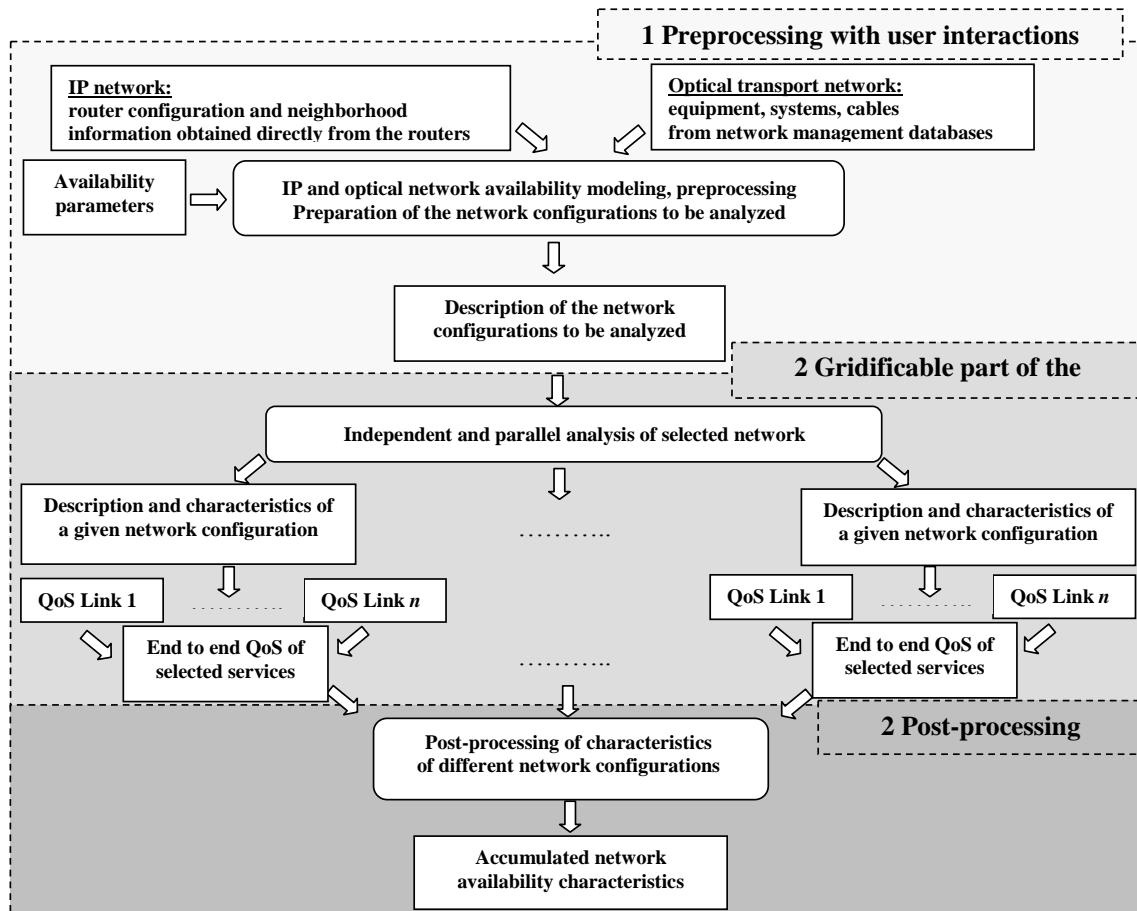


Figure 8 The analysis process and the gridifiable part of the process