Sensitivity Analysis of Availability of Redundancy in Computer Networks

Rubens de S. Matos Júnior*, Almir P. Guimarães*[†],

Kadna M. A. Camboim^{*}, Paulo R. M. Maciel^{*} *Center of Informatics Federal University of Pernambuco Recife, Brazil Email: {rsmj,apg2,kmac,prmm}@cin.ufpe.br [†]Campus Arapiraca Federal University of Alagoas Arapiraca, Brazil Email: almir@arapiraca.ufal.br Kishor S. Trivedi[‡] [‡]Dept. of Electrical & Computer Eng. Duke University Durham, USA Email: kst@ee.duke.edu

Abstract—In this paper, we investigate the availability modeling of computer networks with redundancy mechanisms. Sensitivity analysis is applied in order to find the bottlenecks of system availability. We use Markov chains for the analytical evaluation of complex scenarios. We apply our proposed modeling approach in a case study to evaluate how sensitive dependability is to failure and recovery times of different components in an enterprise network. The influence of network topologies is also considered in our case study.

Keywords—Availability; Markov Chains; Sensitivity analysis; Computer networks

I. INTRODUCTION

Over the last few years, the use of data networks has significantly increased. This considerable growth is somewhat related to the convergence of many different services on the same transmission technology. These services should be continuously provided even when events like congestion, link failures, routing instabilities, sabotage, natural disasters, hardware or software failures happen. The design, deployment and management of communication network infrastructure ought to meet such requirements. The possibility of identifying points where the unavailability or downtime of these networks may put the business at risk is an interesting track to be followed by organizations.

Recently, much has been done to deal with issues relating to the availability of computer networks. Researchers have used different approaches to deal with these problems, including sensitivity analysis techniques and the development of advanced redundancy mechanisms.

Zou *et al.* [1] discusses algorithmic methods to compute network availability for a given topology and presents two tools for computation of network availability in large and complex networks. Semaan [2] discusses different issues related to network availability. First, the paper presents some of the elements that impact the availability of a solution. Then, it discusses how network designers can calculate the exact availability of their solution and provides means to determine the optimal level of availability. Trivedi *et al.* [3] presents a new classification of dependability and security models for systems and networks. It also presents several individual model types such as availability, confidentiality, integrity, performance, reliability, survivability, safety and maintainability models. Furthermore, it is shown that individual model types can be combined to form composite dependability model types. The dependability/security models can be represented as combinatorial models, state-space models, and hierarchical models.

In this paper, we focus on the availability of data networks, including redundancy mechanisms. Several scenarios are evaluated through analytic-numeric solution of Markov chains [4]. The model parameters used were obtained from manufacturers of network elements, as also from experimental measurements. We evaluate the impact of different component parameters on the overall system availability, by means of differential sensitivity analysis.

The rest of the paper is organized as follows: Section II presents basics of availability of computer networks and sensitivity analysis. Section III describes the proposed availability models. Section IV presents the evaluation of availability and its sensitivity for all the proposed models. Finally, Section V discusses the results of this study and introduces ideas for future research.

II. FUNDAMENTAL CONCEPTS

A. Dependability Requirements for Voice and Data Networks

Standard IP applications traffic is characterized by burstiness. However, such applications are not highly sensitive to delay and jitter. On the other hand, voice applications run continuously and steady, they could thus be strongly affected by long delays and jitter [5]. Providing high quality voice service on IP networks is one of the most pressing issues faced by the VoIP community [6].

Critical services, such as VoIP, have strict QoS requirements for both performance tolerance and service dependability. Dependability of a computer system must be understood as the ability to avoid service failures that are more frequent and more severe than is acceptable [7]. Dependability attributes include the concepts of availability, reliability, safety, integrity and maintainability [7].

Inputs to availability models include component Mean Times to Failure (MTTF) and Mean Times To Repair (MTTR). The hardware component MTTFs are generally supplied by the manufacturer. The MTTRs are tightly related to the maintenance policy adopted by the organization.

B. Parametric Sensitivity Analysis

Sensitivity analysis is a method of determining the most influential factors on model results [8], [9]. The effect of changes in data distribution function and the impact of changes in parameter values are examples of study subjects for sensitivity analysis. When dealing with analytic models such as Markov chains, parametric sensitivity analysis is a particularly important technique for assessing the effect of changes in the rate constants on the measures of interest. This approach may be used to find performance or availability bottlenecks in the system, thus guiding an improvement and optimization [10].

There are many ways of conducting sensitivity analyses. The simplest method is to repeatedly vary one parameter at a time, while keeping the others fixed. When applying this method, the sensitivity ranking is obtained by noting the corresponding changes in the model output. Other techniques include factorial experimental design [11], correlation analysis, regression analysis and perturbation analysis (PA). Differential analysis, also referred to as parametric sensitivity analysis or the direct method, is the backbone of nearly all other sensitivity analysis techniques [9]. This method is chosen in this paper, as it can be performed in an efficient computational manner on analytic models commonly used in performance and availability analyses.

Parametric sensitivity analysis is performed by computing the partial derivatives of the measure of interest with respect to each input parameter. For instance, the sensitivity of a given measure Y, which depends on a parameter λ , is computed as in Equation (1), or (2) for a scaled sensitivity.

$$S_{\lambda}(Y) = \frac{\partial Y}{\partial \lambda} \tag{1}$$

$$S_{\lambda}^{*}(Y) = \frac{\partial Y}{\partial \lambda} \left(\frac{\lambda}{Y}\right)$$
(2)

A number of researchers have already demonstrated how to perform parametric sensitivity analysis in a variety of analytic models. In [10], the basics of transient sensitivity analysis in continuous time Markov chains (CTMC) are presented. Sensitivity functions for Markov chains were recently implemented in the SHARPE package [12], making use of the techniques described in the papers we have just cited. Since the reduced reachability graph of a Stochastic Petri Net (SPN) is a Markov chain, this kind of model may also be analyzed, by following the steps indicated in [13]. Their work includes the implementation of sensitivity analysis features in the SPNP package [14]. Queueing systems are another example of analytic models whose sensitivity analysis has been described in [15].

III. PROPOSED AVAILABILITY MODELS

Traditional evaluation techniques for availability use Markov chains and Markov reward models. In this section, we present three CTMC (Continuous Time Markov Chain) availability models (Figures 4, 5 and 6). The first one represents a system without any redundancy. The second one represents a system with aspects of fault-tolerance based on link redundancy (see Figure 2). Then, the last one represents a system with aspects of fault-tolerance based on warmstandby redundancy (see Figure 3). This approach is characterized by fault detection and recovery mechanisms. The dependability models can be evaluated using tools such as SHARPE (Symbolic Hierarchical Automated Reliability and Performance Evaluator) [12].

A. Platform Description

The following three scenarios were used as a basis for the availability analytic models presented in this paper. They also served as experimental testbeds, from which some failure and recovery parameters were obtained, as well as to validate the analytic results obtained from the respective Markov chains.

1) First and Second Scenarios: In the first scenario, the testbed is composed of two machines, a switch and two routers that are connected by a single link (see Figure 1). In the second scenario, the testbed is composed of two machines, a switch and two routers that are connected by redundant links (L0 and L1 - see Figure 2). When the main link (L0) fails, the spare link (L1) assumes the role of the main one. After main link restoration, the system returns to the initial condition.



Figure 2. Test Bed - Scenario 2.

L1 - PPP - 128Kbps

2) Third Scenario: In this scenario, the testbed is composed of two machines, a switch and three routers (see Figure 3). The system uses fault-tolerance based on warm-standby redundancy. When one of the primary components (R0 or L0) fails, the spare components (R1 and L1) assume the role of the primary components. This switchover process takes time for the spare components to start operation, named Mean Time to Activate (MTTA). After restoration of the primary components, the system returns to the initial condition.

B. CTMC Availability Model without redundancy

In Figure 4, the Markov chain represents the first scenario, which is the simplest one, with no redundancy. There is only



System is down





Figure 3. Test Bed - Scenario 3.

 TABLE I

 States of CTMC Model with link redundancy

State	Description
UUU	The System is UP
DUU	Down state, Router R0 failed
UDU	Down state, Router R1 failed
UUD	Down state, Link L0 failed

one link, named L0, connecting router R0 and router R1. In this model, the normal operation of a component is denoted by the label U (up), and a failed component is represented by label D (down). A state in the Markov chain is defined by a sequence of labels, representing router R0, router R1 and link L0, respectively. We assume the failure and repair time of each component are exponentially distributed. λ_{R0} , λ_{R1} and λ_{L0} are the respective failure rates of R0, R1 and L0. In a similar notation, μ_{L0} , μ_{R0} and μ_{R1} are the respective repair rates of each system component. Once any component (R0, R1, or L0) has failed, the overall system is in a down state and subsequently no additional failures occur until the component is repaired, so that the expansion of state space stops at the first down state. In Table I, a description of each state is given. For this model, the system is up and running only in the state UUU. All the other states are shaded gray in Figure 4, representing the system down states.

C. CTMC Availability Model with link redundancy

In Figure 5, we consider a system that has redundancy only at the link level, as illustrated in Figure 2. The normal operation of a component is denoted by the label U (up), and a failed component is represented by label D (down). A state in the Markov chain is also defined by a sequence of labels, representing router R0, router R1, link L0, and

 TABLE II

 States of CTMC Model with link redundancy

State	Description
UUUU	The System is UP
DUUU	Down state, Router R0 failed
UDUU	Down state, Router R1 failed
UUDU	The System is UP, Link L0 failed
UDDU	Down state, R1 and L0 failed
DUDU	Down state, R0 and L0 failed
UUDD	Down state, L0 and L1 failed
UUUD	The System is UP, L1 failed
DUUD	Down state, R0 and L1 failed
UDUD	Down state, R1 and L1 failed

link L1, respectively. The ideal condition for this system is denoted by state UUUU, in which all components are in nonfailed condition. Failure transitions have rates λ_X , and repair transitions have rates μ_X , where $X \in \{R0, R1, L0, L1\}$, representing each system component. In states shaded gray, the system has failed, due to a failure in one of the routers, or a failure in both links. We assume that in those states no additional failures occur in the remaining components, since they are in an idle condition. Another assumption for this model is that there is a repair policy, that prioritizes the repair of link L0 over link L1 when both are failed. We do not consider any priority in the repair of routers because it is not possible in this model to have both routers down, since a failure in any of them brings the overall system to a down state. In Table II, a description for each state is given.

D. CTMC Availability Model with router redundancy

We present in Figure 6 a Markov chain that represents the system illustrated in Figure 3. The failure and repair rates of each component are represented by λ_X and μ_X , where $X \in \{R0, R1, R2, L0, L1\}$. Rates α_R and α_L are the inverse of mean time to activate the spare router and the spare link, respectively.

For simplicity, we have made some simplifications that do not significantly affect the results we obtain from the analysis. One of the assumptions for this model is that there is a priority in the repair of components. Router R2 has the higher priority, followed by router R0, link L0, router R1, and link L1, in descending order. We also consider that no failure is possible



Figure 5. Markov chain for the availability of link-redundant network

when a component is in waiting condition.

The nomenclature used for the states in this model is based on the current condition of each system component, in the following order: router R0, link L0, router R2, router R1, link L1. A letter U indicates the up condition, when component is active. Letter D denotes a down condition for that component, meaning that it has failed, and that a repair is needed. Letter W represents a waiting condition, in which the component is not being used, but is ready to enter in active mode, as soon as it is needed. Therefore, state UUUWW denotes router R0, link L0, and router R2 are active, router R1 and link L1 are on waiting condition. For an active state, the system must present one of the following combinations: UUUWW, DWUUU, WDUUU, UUUDW or UUUWD (see Figure 6). Particularly, in the DWUUU state, router R2 with spare router and link (R1 and L1) are active, while R0 is in down state and L0 is waiting, since it only works together with R0. A similar situation happens in WDUUU state, where R2, R1 and L1 are active, but L0 is down, leaving R0 in a waiting condition.

Figure 6 shows the initial state of the system, UUUWW. With rate λ_{R0} a failure happens and brings the system to down state (DUUWW). Similarly, with rate λ_{L0} a failure happens and brings the system to down state (UDUWW). Likewise, with rate λ_{R2} a failure happens and brings the system to down state (UUDWW). In this case, the repair with rate μ_{R2} brings the system back to state (UUUWW).

After detecting a failure, a switchover occur making the spare components (R1 and L1) active. In the states DUUWW and UDUWW, the system activates the spare components with rates α_R and α_L corresponding to router R0 and link L0 failures, respectively. After switch-over, the standby components are able to take over the failed components, bringing the system to and active state (DWUUU or WDUUU). The repair with rates μ_{R0} and μ_{L0} bring the system back to

initial state. Before a repair happens, another failure with rate λ_{R1} , λ_{L1} or λ_{R2} may bring the system to a down state (DWUDU, DWUUD, DWDUU, WDDUU, WDUDU, WDUUD). From states DWDUU and WDDUU, with rate μ_{R2} , the system comes back to active states (DWUUU and WDUUU). Similarly, from states DWUDU and DWUUD, with rate μ_{R0} , the system comes back to active states (UUUDW and UUUWD). Likewise, from states WDUDU and WDUUD, with rate μ_{L0} , the system return to active states (UUUDW and UUUWD). From the active states, the repair with rate μ_{R1} or μ_{L1} brings the system to initial state. Finally, from UUUDW and UUUWD, the system can return to down state with rate λ_{R0} , λ_{L0} or λ_{R2} . The repair with rate μ_{R0} , μ_{L0} or μ_{R2} brings the system back to active states (UUUDW and UUUWD).

In Table III, we see the system availability condition for each state of this Markov chain. Note that only on 5 states the system is operational: UUUWW, DWUUU, WDUUU, UUUDW, UUUWD.

IV. CASE STUDY

We concentrate our attention on parametric sensitivity analysis, as a technique to compute the effect of changes in the rate constants of a Markov model on the measures of interest. Parametric sensitivity analysis helps: (1) to guide system optimization, (2) to find availability, performance, and performability bottlenecks in the system, and (3) to identify the model parameters that may produce significant modeling errors. In this paper, (1) and (2) are the main purposes, although the identification of errors in the early versions of the proposed models was also possible through such analysis.

We consider the testbed shown in Section III-A to perform a parametric sensitivity analysis using the proposed dependability CTMC models. The MTTFs of components used in this work are respectively: 131,000 hours for routers and 11,988



Figure 6. Markov chain for the availability of router-redundant network

TABLE III STATES OF CTMC MODEL WITH ROUTER REDUNDANCY

-	
State	Description
UUUWW	The System is UP
DUUWW	Down state, Switchover Started
UDUWW	Down state, Switchover Started
UUDWW	The System is Down
DWUUU	The System is UP
WDUUU	The System is UP
DWUDU	The System is Down
DWUUD	The System is Down
DWDUU	The System is Down
WDDUU	The System is Down
WDUDU	The System is Down
WDUUD	The System is Down
UUUDW	The System is UP
UUUWD	The System is UP
DUUDW	The System is Down
UDUDW	The System is Down
UUDDW	The System is Down
DUUWD	The System is Down
UDUWD	The System is Down
UUDWD	The System is Down

hours for links. We use a mean time to repair (MTTR) equal to 12 hours, for all components. Those values shall be considered as the base case throughout this section, unless another value is specified in each specific analysis. Notice that all λ_i in the models of previous section are equal to $1/MTTF_i$, and all μ_i are equal to $1/MTTR_i$. Sensitivity analysis of steady-state availability is carried out by computing $S^*_{MTTF_i}(A)$ as the scaled sensitivity of availability with respect to $MTTF_i$, and $S^*_{MTTR_i}(A)$ as the corresponding measure with respect to $MTTR_i$.

In the base case, using the values we have just mentioned, the steady-state availability for the first scenario is 0.998817194. The second scenario, in which link redundancy is added, presents an availability of 0.999815827. In the third scenario, which has redundant router, the availability increases to 0.999906968.

Initially, we consider the sensitivity analysis regarding the third scenario, in Section III-A2. The values for $S_k^*(A)$, where A is the system steady-state availability and k is each of the components' MTTF and MTTR, were computed using sensitivity analysis features developed for the SHARPE package. In Table IV, we see that parameters $MTTF_{R2}$ and $MTTR_{R2}$ assume the greatest importance in system steadystate availability, since they have the highest sensitivity values. Any change in these parameters will have a major impact on system availability, but in opposite directions. Sensitivity with respect to $MTTF_{R2}$ is positive, since the availability increases when this parameter increases. In contrast, $S^*_{MTTR_{B2}}(A)$ is negative, because a smaller repair time of R2 implies an increased availability. In Table IV, we can also notice that time to repair spare components $(MTTR_{R1} \text{ and } MTTR_{L1})$ are the ones with smallest impact on the system availability. This result matches the results from the established repair policy, since failed spare components are repaired only after main components have returned to normal operation.

Figure 7 depicts a plot for the system availability, in which the MTTF parameters for the system links (L0 and L1) were changed one at a time, and the analytic model of Figure 6 was solved. This plot confirms that efforts in expanding the time to failure of link L0 have more impact on system availability than increases in $MTTF_{L1}$ do. Figure 8 also validates the

TABLE IV Sensitivity of Availability for scenario with router redundancy

Parameter k	$S_k^*(A)$
$MTTF_{R2}$	9.15946628e-05
$MTTR_{R2}$	-9.15946627e-05
$MTTR_{L0}$	-2.18237573e-06
$MTTF_{L0}$	1.31711540e-06
$MTTF_{L1}$	1.09121018e-06
$MTTR_{R0}$	-1.99712368e-07
$MTTF_{R0}$	1.20531140e-07
$MTTF_{R1}$	9.98582261e-08
$MTTR_{L1}$	-1.19080014e-09
$MTTR_{R1}$	-1.08971849e-10



Figure 7. Effect of each link MTTF on system availability (scenario 3)

results from sensitivity ranking. If we make $MTTF_{R2}$ bigger the benefits will be much higher than that resulting from enhancements on either R0 or R1 MTTFs.

MTTR is an important factor for system availability since it will affect the downtime of network elements. For a redundant topology (third scenario), Figure 9 shows the system availability as a function of each component MTTR. Router R2 is the component whose time to repair causes the biggest effect on the steady-state availability, followed by link L0. This information can also be obtained comparing the corresponding values for each MTTR in Table IV.

Then, we will analyze the impact of R0 and L0 MTTFs on the system availability in each proposed scenarios. The idea is to observe the real impact of these parameters with respect



Figure 8. Effect of each router MTTF on system availability (scenario 3)



Figure 9. Effect of each component MTTR on system availability (scenario 3)

to system availability in different redundancy schemes (i.e. different scenarios).

Comparing Tables V and VI - also obtained through SHARPE's sensitivity analysis features - we see that the redundancy mechanism makes the system availability less sensitive to failures of primary link (L0), while the sensitivity with respect to $MTTF_{R0}$ is almost the same. The results from differential sensitivity analysis can also be confirmed by comparing, in a scatter plot, the effect of changes in $MTTF_{L0}$ and $MTTF_{R0}$ for each one of our three scenarios. In Figures 10 and 11, we see that in second and third scenarios availability is less affected by increases in $MTTF_{L0}$ and in $MTTF_{R0}$. As in the previous plots, those values were obtained by repeatedly varying the value of each parameter at a time, and solving the Markov chain for that configuration. It is important to state that this comparison using scatter plots becomes more difficult as the number of parameters in the model increases. The ranking obtained from differential sensitivity analysis allows a direct view of the importance order of all parameters.

The reduction of main router impact on availability is confirmed in Table IV, that shows a higher sensitivity with respect to the failure of spare link $(MTTF_{L1})$ than with respect to main router failure $(MTTF_{R0})$. So, actions to increase the MTTF of link L1 should be considered more important than additional efforts to enhance the MTTF of router R0. This kind of decision could not be easily made without an accurate sensitivity analysis, as we have performed in this case study.

V. CONCLUSION

In this paper, we proposed Markov chain models to evaluate several dependability aspects of computer networks in different scenarios. The models support the analysis of system availability along with its services, based on different topologies, redundancy mechanisms and network elements. Sensitivity analysis was applied in order to guide system optimization in terms of steady-state availability.

For future work, we plan to extend these models to include network availability with redundant topologies, different re-

 TABLE V

 Sensitivity of Availability for scenario with link redundancy

Parameter k	$S_k^*(A)$
$MTTF_{R0}$	9.15861826e-05
$MTTF_{R1}$	9.15861826e-05
$MTTR_{R0}$	-9.15861826e-05
$MTTR_{R1}$	-9.15861826e-05
$MTTF_{L0}$	1.00081664e-06
$MTTR_{L0}$	-1.99963265e-06
$MTTF_{L1}$	9.99815827e-07
$MTTR_{L1}$	-9.99815827e-10

 TABLE VI

 Sensitivity of Availability for scenario with no redundancy

Parameter k	$S_{k}^{*}(A)$
$MTTF_{L0}$	9.99817011e-04
$MTTR_{L0}$	-9.99817011e-04
$MTTF_{R0}$	9.14947048e-05
$MTTR_{R0}$	-9.14947048e-05
$MTTF_{R1}$	9.14947048e-05
$MTTR_{R1}$	-9.14947048e-05



Figure 10. Availability Vs. $MTTF_{L0}$



Figure 11. Availability Vs. $MTTF_{R0}$

covery strategies as well as taking into account different failure modes.

REFERENCES

- W. Zou, M. Janic, R. Kooij, and F. Kuipers, On the availability of networks, in Proc. of BroadBand Europe 2007, Antwerp, Dec. 2007.
- [2] G. Semaan, *Designing networks with the optimal availability*, in National Fiber Optic Engineers Conference, Optical Society of America, pp. 1–6, 2008.
- [3] K. Trivedi, D. Kim, A. Roy, and D. Medhi, *Dependability and security models*, Proc. DRCN, pp. 11–20, 2009.
- [4] Kolmogorov, A. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung (in German). Mathematische Annalen. Springer-Verlag, 1931.
- [5] C. Hoene, B. Rathke, and A. Wolisz, On the importance of a VoIP packet, In Proc. Of ISCA Tutorial and Research Workshop on the Auditory Quality of Systems, 2003.
- [6] H. Sze, S. Liew, J. Lee and D. Yip, A multiplexing scheme for h.323 voice-over-ip applications, IEEE J. Select. Areas Commun., vol. 20, no. 7, pp. 1360–1368, 2002.
- [7] A. Avizienis, J. Laprie, B. Randell, and C. Landwehr, *Basic concepts and taxonomy of dependable and secure computing*, IEEE Transactions on Dependable and Secure Computing, vol. 1, no. 1, pp. 11–33, 2004.
- [8] P. Frank, Introduction to System Sensitivity Theory. Academic Press Inc, Sept. 1978.
- D. Hamby, A review of techniques for parameter sensitivity analysis of environmental models, Environmental Monitoring and Assessment, pp. 135–154, 1994.
- [10] J. Blake, A. Reibman, and K. Trivedi, Sensitivity analysis of reliability and performability measures for multiprocessor systems, in SIGMET-RICS '88: Proceedings of the 1988 ACM SIGMETRICS conference on Measurement and modeling of computer systems. New York, NY, USA: ACM, 1988, pp. 177–186.
- [11] R. Jain, The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling, John Wiley, 1991.
- [12] K. Trivedi and R. Sahner, Sharpe at the age of twenty two, SIGMET-RICS Perform. Eval. Rev., vol. 36, no. 4, pp. 52–57, 2009.
- [13] J. Muppala and K. Trivedi, GSPN models: sensitivity analysis and applications, in ACM-SE 28: Proceedings of the 28th annual Southeast regional conference.New York, NY, USA: ACM, 1990, pp. 25–33.
- [14] C. Hirel, B. Tu, and K. Trivedi, SPNP: Stochastic Petri Nets. version 6.0, Mar. 2010. [Online]. Available: http://www.ee.duke.edu/~chirel/PAPER/ paperSpnp.pdf
- [15] B. Yin, G. Dai, Y. Li, and H. Xi, Sensitivity analysis and estimates of the performance for M/G/1 queueing systems, Performance Evaluation, vol. 64, no. 4, pp. 347–356, 2007.