Torwards Performance Modeling and Evaluation of EFT Systems Avaliability

Fábio Chicout, Erica Sousa, Carlos Araújo, Paulo Maciel
Federal Univestity of Pernambuco
Center of Informatics - Performance Evaluation Laboratory Itautec
Recife, PE, Brazil
{fcfmc,etgs,cjma,prmm}@cin.ufpe.br

Abstract

The performance evaluation of Electronic Funds Transfer (EFT) Systems has an enormous importance for companies that implement this service, since the computing resources must be used efficiently in order to ensure high availability, reliability, scalability and security. Many research works have underlined the performance degradation of software application related to cumulative failures and internal conditions that consume system resources. This phenomenon, usually only perceptible after long software running, has been named software aging. This paper presents a first stochastic model for performance evaluation considering aging on EFT software systems. This model aims for planning related to allowing aging monitoring of EFT infrastructure accomplishing specified utilization degrees of storage and processing utilization degree considering load variation range.

1 Introduction

Increasing the capacity integration of payment services and the advances of new technologies have fostered the growth and complexity of electronic transactions. Organizations providing Electronic Funds Transfer (EFT) must not only supply correct services, but also meet customers performance expectations; since failing to achieve expectations jeopardizes profitability, affects the organizations public perception or even sets off uncontrollable economic problems due to the harsh market global competition.

Over the last decade, the EFT market have been massively expanding, thus demanding companies to offer reliable services as well as high availability, scalability, security, and low costs. Furthermore, performance evaluation have been widely adopted as an essential activity for improving services' quality, planning resources' infrastructure, tuning the system components and aging analysis in order to improve performance and reduce costs of services. Additionally, modeling based techniques may represent systems at several abstraction levels, each of them more suitable for solving specific problems [7][12].

Generalized Stochastic Petri Nets (GSPN) [2] have been largely adopted for performance evaluation of systems. Performance metrics such as the probability of a particular condition occurs, the expected number of items in repository and throughput are a few that could be calculated from a stochastic Petri net model.

Different approaches have been proposed in the literature for performance analysis and aging, but, to the best of our present knowledge, none of them focus on performance evaluation of EFT systems considering expolinomial stochastic models [3]. Bobbio et al. propose FSPN based modeling framework for rejuvenation, restoration and checkpointing analysis[4], using failure, load and time in the same framework. Vaidyanathan [13] describes how to include faults attributed to software aging in the framework of Gray's software fault classification (deterministic and transient), and study of faults the treatment and recovery strategies for classes. Garg [6] proposes a methodology for detection and estimation of aging in the UNIX operating system, and they have done that using a tool designed and implemented by them, using the SNMP (Simple Network Management Protocol) protocol. Avritzer et al. [1] present three algorithms for detecting the need for software rejuvenation by monitoring the changing values of a customer-affecting performance metric and their results show that applying the proposed algorithms with suitable choices of control parameters significantly improve system performance as measured by the response time. Kocak [8] proposes an approach to evaluate and characterize Web traffic and server bottlenecks by identifying CPU utilization.

This paper proposes a model based on extended Deterministic Stochastic Petri Nets (eDSPN) [14] for availability analysis of EFT systems at high throughput, focusing on storage (physical disk) and processing resources. The obtained results permit to set the throughput supported by server taking into account a certain amount of operational points of sales and frequency of transactions.

This paper is structured as follows: Section 2 presents presents the methodology conducted. The EFT System eDSPN Model is presented in Section 3. Case studies have been conducted and are presented in Section 4. Concluding remarks are given in Section 5.

2 Evaluation Methodology

This section briefly introduces the adopted process for modeling and evaluating EFT systems through the proposed EFT performance model. The method consist of understanding the environment, workload characterization, measurements, treatment of measured data, eDSPN model development and failure monitoring.

The initial phase consists of understanding the system's interface with the environment; the hardware and software significant components (under performance point of view) and their internal and external interactions. The system is composed by Itautec ETF System [1] running in a Intel Xeon 3GHz Dual Core machine with 4GB of RAM.

The identification of aging on EFT systems is important for preventing data corruption, software misfunction caused by long-term running environments, or resource exhaustion on the use of various workload levels that are credit and debit commercial transactions while the EFT system is characterized by workload intensity and service demand parameters at each resource. These parameters are obtained directly from measurements and derived from other parameters that are measured directly.

The measures obtained from the system include: processor percentage of idle time, physical disk percentage of idle time, physical disk transfers per second and physical disk average time of transfer. Some metrics as utilization and service time for physical disk and processor are indirectly estimated by basic equations from the Queueing Theory Concepts [7, 11].

After data exploration, statistics are generated and analyzed. The mean and the respective standard deviation of processor and disk service time are considered for deciding which Expolinomial distribution best fits the measured data distribution. Afterwards, the suitable Expolinomial model (Refined Model) is generated taking into account the Abstract Model and the analysis result of the measured data. Later on, the Refined Model should be validated by t-paired test [9]. This process is conducted by observing a set of pairs of values obtained from the system (by measurements) and through the Refined Model. These pairs of values are related to a carefully chosen set of scenarios.

[1]from this point until the end of this paper, we treat EFT system as system. However, EFT is only one element of the system components (hw and sw)

After validating the model, scenarios' evaluations may be performed. Such scenarios could be quite complex, and consider, for instance, several classes of clients' workloads. System's bottlenecks might also be analyzed, as well as system capacity planning studies could be conducted.

3 EFT Systems and Aging

The EFT system is composed by a transaction authentication server, that manages receipts and forwards trade transactions, an authorizer entity which reviews if the buyer really have funds to pay and authorizes (or not) the transaction, and the clients (here called points of sale) which sends to the server credit and debit transactions. In our case, the points of sale are implemented in a workload generator that is capable of sending trade transaction to various servers. These servers are responsible for managing contracts that are links to each authorizer, receiving or forwarding authentication trade transaction. The EFT system aging phenomenon shows up when the system exhibits increased failure rate and performance degradation over time. The consequences for this degradation are data corruption and accumulation of errors.

Transaction throughput on EFT Systems are an accelerator to the aging fenomenon. The system was executed on a range of transaction throughputs, what brought the effects of aging in a earlier time. Fixed the transaction throughput, The Table 1 shows the aging on an execution of a ETF system at 2200 tpms. It is verified that the utilization level degrades along the time, due the disconnections of the points of sale. With a measure time of 500 seconds the processor utilization observed is 65%. When the same system is in execution with the same transaction throughput on a time of 6500 seconds, the processor utilization goes down to 50%.

Table 1. Aged EFT System

Running Time(sec)	Processor Utilization(%)
500	65%
4500	60%
6500	50%

4 EFT System Performance Model

This section presents the eDSPN model conceived for EFT performance evaluation and aginning monitoring. First the proposed EFT Abstract Model is described via its 'submodels' (subnets that describes system´s components). Then, a refined version of the model is generated for validation purposes. The validation is carried out considering seven significant scenarios and a set of metrics whose values were obtained by measurements and through the refined model. Client subnets represent different company traffics such as those related to drugstores, supermarkets, gas stations, shopping and stores or even workload related to particular period or season. These subnets could be refined to represent a vast range of traffic, from simple Poison processes to complex modulated Poisson processes related to a vast sort of burst traffic. These models might represent a large number of points of sale forwarding credit and debit trade transactions with distinct occurrence frequencies trends and variations. Hence, service demands can be represented by varying the transfer transaction frequencies and other parameters. The N_i markings assigned to places on subnets representing clients (drugstore, supermarket, gas station, shopping and store) specify the number of points of sale, and the generic stochastic transitions represent the transmission delay between total time of commercial transitions. The place buffer represents the temporarily hold transactions waiting to be served. Its dual place ($P12$) represent the buffer storage capacity.

The Server subnet is composed of subnets Processor and Disk where the subnet Server represents the computational resources of the EFT system server. The subnet Processor represents the transaction´s processing and the Disk subnet represents the physical disk reading and writing operations related to transactions. The Processor place marking (Np) denotes the processing capacity, that is the number of concurrent transactions supported by the processing resource (in other words, the concurrency degree). The Disk place marking (Nd) denotes the number of concurrent disk transactions supported by the storage resource (its concurrency degree). The transitions T_ip and T_id firing represent the transaction processing time and the storage (reading and writing related to a transaction) operation time (See Figure 1). The transitions *Clock* and $T6$ are part of a subnet representing the aging time of the system. Furthermore, the time on transition $T6$ must be extremely higher than the time on *Clock*. The subnet is a erlang-based clock, tending to a deterministic one. Additionally, the number of tokens on place $P17$ represents a aging time counting. The immediate transitions in EFT system model have priority and weight 1. The qualitative analysis of eDSPN model shows that the EFT model has some desired properties, such as no deadlock, conservativeness and boundedness [10].

4.1 Model Validation

This section presents the evaluation results related to seven (7) scenarios. This experiment was conducted in environment set in the CIn-Itautec Performance Evaluation Laboratory.

These scenarios were chosen, since real measures were already available. Hence, it is expected the model providing equivalent results for each of these particular scenarios. Each scenario describes a supermarket with various points of sale where the demand register occurred at 7 different rates, that is 100, 200, 300, 400, 500, 600 e 700 tpms (transactions per minute). In these scenarios, the concurrency degrees related to processing and storage elements are equal to one, that is NP=1 and ND=1.

The measured data were analyzed for deciding which Expolinomial distribution best fit the processing and storage operations (represented by transitions Tp and Td). The respective processing and storage time means (μ_D) and standard deviations (σ_D) were analyzed, according the process described in [5], and these operations should be refined according the results presented in Table 2.

Table 2. Distribution

Transition	μ_D (s)	σ_D (s)	Distribution
T_ip	0,001311	0,000508	Hypoexponential
T_id	0,002756	0,000353	Hypoexponential

After defining which distribution is suitable for representing the measured data, the related distribution parameters have to be calculated. Since the hypoexponential model was chosen for refining both Tp and Td, μ_1, μ_2 and γ should be computed. These values are calculated using Equations 1, 2 and 3. Table 3 shows the respective values of μ_1, μ_2 and γ for the models that refine Tp and Td.

$$(\frac{\mu}{\sigma})^2 - 1 \leq \gamma < (\frac{\mu}{\sigma})^2 \qquad (1)$$

$$\mu_1 = \mu \pm \frac{\sqrt{\gamma(\gamma+1)\sigma^2 - \gamma\mu^2}}{\gamma+1} \qquad (2)$$

$$\mu_2 = \mu \pm \frac{\sqrt{\gamma(\gamma+1)\sigma^2 - \gamma\mu^2}}{\gamma+1} \qquad (3)$$

Table 3. μ_1, μ_2 and γ

Transition	μ_1 (s)	μ_2 (s)	γ
T_ip	0,000080	0,00054	6
T_id	0,000001	0,00005	61

The respective values of transitions Clock, T15 and weight of arcs are presented in Tables 4 and 5.

With the refined eDSPN model generated, the processor and disk utilization are obtained using the following expressions: $UProc = (P\{\#Processor = 0\}/NP)$ and $UDisk = (P\{\#Disk = 0\}/ND)$, respectively. Figure 2 shows a visual comparison between measured processor utilization and values obtained through the model by the

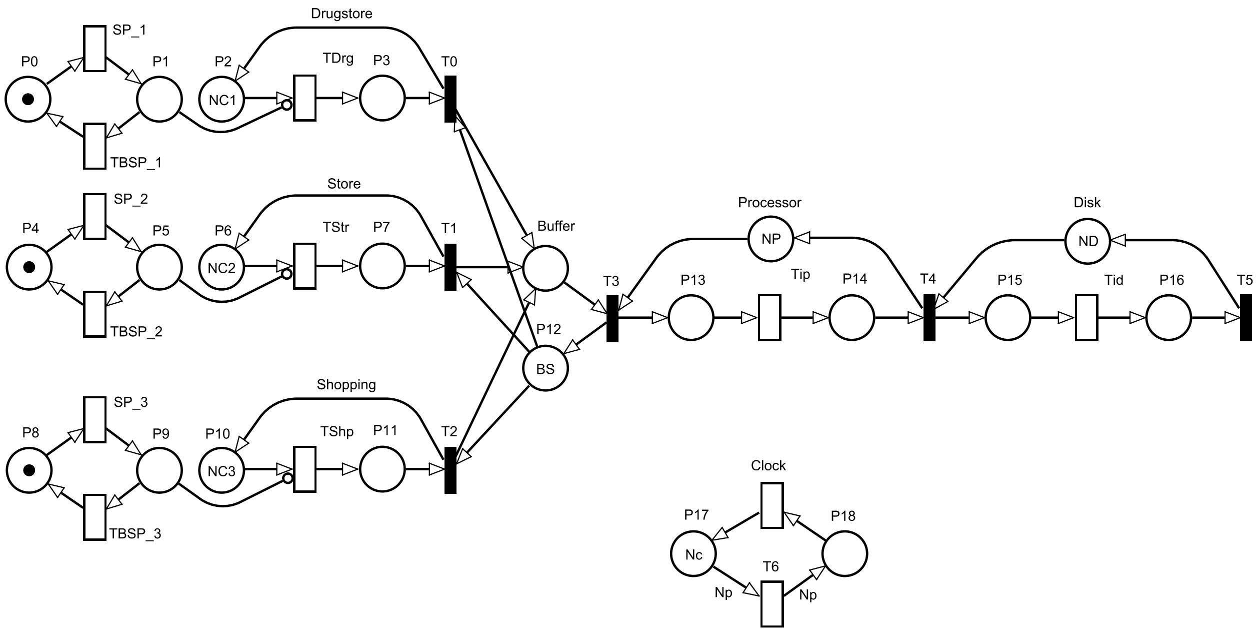

Figure 1. Abstract eDSPN Model

Table 4. Clock

Transition	Delay (s)
Clock	1
$T6$	1000000

Table 5. Clock Places

Places	# Tokens
$P17$	20
$P18$	0

metrics $UProc = (P\{\#Processor = 0\}/NP)$. The result presents a maximum error of $9,88\%$.

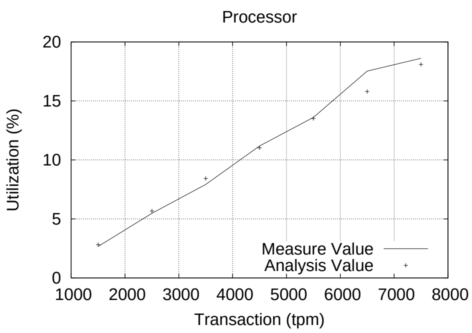

Figure 2. Processor Utilization.

Figure 3 depicts similar comparison related to disk utilization. The values presented are those obtained by measurements and by model's evaluation through the metric $UDisk = (P\{\#Disk = 0\}/ND)$. The result presents a maximum error of $12,30\%$.

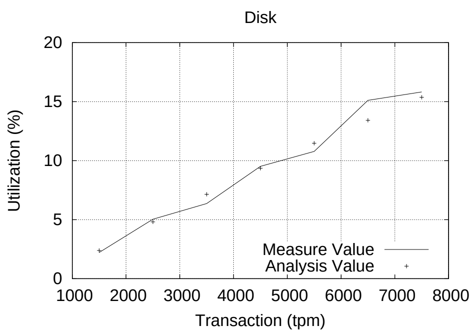

Figure 3. Disk Utilization.

The results obtained by stationary analysis measurements were compared by means of paired T-test [9], taking into account 95% confidence degree, and as result no evidence highlights difference between measured data and the respective values obtained through the model.

5 Case Study

This section presents a case study by customizing the abstract model to particular configurations and a scenario.

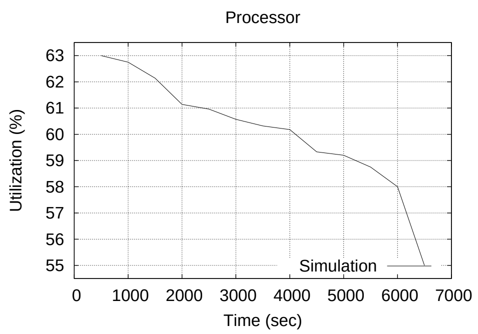

Figure 4. Aging Processor

This model can be used to forecast a aging event, but here it was just validated. The evaluated scenario describes a set stores with containing 3345 points of sale registering the credit and debit transactions. This scenario was evaluated aiming to find the frequency where the system does not work correctly, and better observe the aging. We adopted 880, 1100, 1320, 1540, 1760, 1980, 2178, 2222 tpms in order to accelerate the effect of aging on the system. This fenomenon was evaluated considering the server utilization degree at each rates. At 2222, the processor utilization goes down, taking all the points of sale to disconnection, which is the sign that the system got aged.

Then the execution time held on the frequency 2200 on the model was changed, to identify where it gets aged. The times simulated were 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500 seconds. The Figure 4 depicts the decrease of processor utilization resources on the time due to disconection of points of sale.

The scenario study the system (see Figure 5). Then, the token in place $P2$ represents 3345 points of sale forwarding credit and debit trade transactions with 880, 1100, 1320, 1540, 1760, 1980, 2178, 2222 tpms. The $P4$ place represents the buffer storage capacity. A large number was assigned as the buffer size. The reader should bear in mind that a large number is meant to be a buffer size that allows storing a much larger number of tokens (representing trade transactions) than the real system actually may store on the real buffer. The number of tokens in place Processor represents the concurrency degree related EFT transaction executed by the processing component (server), that is, the number of transactions that could be simultaneously carried out by the application server. The marking assigned to place Disk, likewise, also represents storage operations related to an EFT transaction. In these scenarios, the concurrency degrees related to processing and storage elements are equal to one, that is NP=1 and ND=1. Also, the transitions of subnets Processor and Disk firing represent the transaction processing time and the storage (reading and writing related to a transaction) operation time. Furthermore, the transitions *Clock* and $T6$ are part of a subnet representing the aging time of the system. Furthermore, the time on transition $T15$ must be extremely higher than the time on *Clock*. The subnet is a erlang-based clock, tending to a deterministic one. Additionally, the number of tokens on place $P18$ represents a aging time counting.

6 Conclusions

This paper presents a model based on extended Discrete Stochastic Petri Nets (eDSPN) for aging and performance evaluation of EFT systems, focusing on processing resources. The obtained results permit to forecast the time that the system is alive without the consequences of aging.

As future works, we intend to model the aging of disk resources and some rejuvenation politics to be executed on the system, to get better performance. Another task is to implement aging with a fault model.

7 Acknowledgement

Authors would like to thank the CIn-Itautec Performance Evaluation Laboratory for supporting the development of this work and Projeto Universal 476545/2007-2. Paulo Maciel acknowledges partial support of this research by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant 305527/2007-0.

References

[1] A. Avritzer, A. Bondi, M. Grottke, K.S. Trivedi, and E.J. Weyuker. Performance assurance via software rejuvenation: Monitoring, statistics and algorithms. *Dependable Systems and Networks, 2006. DSN 2006. International Conference on*, pages 435–444, 2006.

[2] G. Balbo. Introduction to Stochastic Petri Nets. *Lectures on Formal Methods and Performance Analysis: First EEF/Euro Summer School on Trends in Computer Science, Berg en Dal, The Netherlands, July 3-7, 2000: Revised Lectures*, 2001.

[3] G. Bloch, S. Greiner, H. de Meer, and S. Trivedi. Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications, 1998.

[4] A. Bobbio, S. Garg, M. Gribaudo, M. Sereno, A. Horvath, and M. Telek. Modeling software systems with rejuvenation, restoration and checkpointing through fluid stochastic petri nets. *pnpm*, 00:82, 1999.

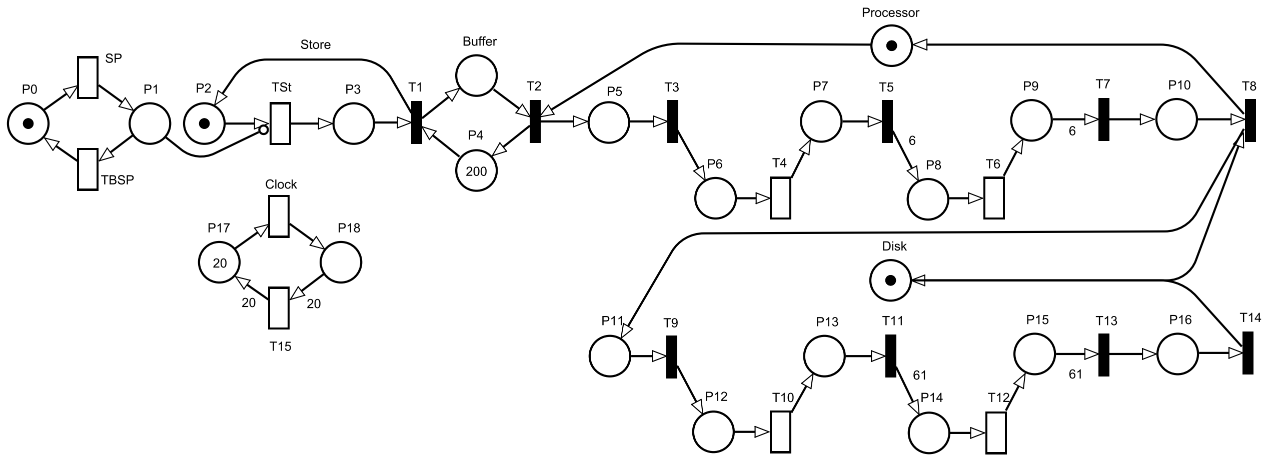

Figure 5. Refined Model

[5] A.A. Desrochers and R.Y. Al-Jaar. *Applications of Petri Nets in Manufacturing Systems: Modeling, Control, and Performance Analysis*. IEEE Press, 1995.

[6] S. Garg, A. van Moorsel, K. Vaidyanathan, and K.S. Trivedi. A methodology for detection and estimation of software aging. *Software Reliability Engineering, 1998. Proceedings. The Ninth International Symposium on*, pages 283–292, Nov 1998.

[7] R. Jain. *The Art of Computer Systems Performance Analysis:: Techniques For Experimental Design, Measurement, Simulation, and Modeling.* John Wiley and Sons, New York, Chichester, 1991.

[8] Engel J. Kocak, T. and Buboltz J. Characterization of Server Performance Bottlenecks in Distributed Interactive Simulation Environments. *Simulation Modeling Practices and Theory*, 2008.

[9] D.J. Lilja. *Measuring Computer Performance: A Practitioner's Guide*. Cambridge University Press, 2000.

[10] T. Murata. Petri nets: Properties, Analysis and Applications. *Proceedings of the IEEE*, 77(4):541–580, 1989.

[11] K.S. Trivedi. *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*. John Wiley and Sons, 2nd Edition, 2006.

[12] K. Vaidyanathan and K.S. Trivedi. A measurement-based model for estimation of resource exhaustion in operational software systems. *Software Reliability Engineering, 1999. Proceedings. 10th International Symposium on*, pages 84–93, 1999.

[13] Kalyanaraman Vaidyanathan and Kishor S. Trivedi. A comprehensive model for software rejuvenation. *IEEE Transactions on Dependable and Secure Computing*, 02(2):124–137, 2005.

[14] Armin Zimmerman. *Stochastic Discrete Event Systems*. Springer Berlin Heidelberg, 2008.