A Performance Analysis of TMN Systems Using Models of Networks of Queues, Jackson's Theorem, and Simulation

Young Ha Hwang, Sang Wook Chung, Gil-Haeng Lee, and Young Il Kim

We analyze the performance of a telecommunications management network (TMN) system using models of networks of queues, Jackson's theorem, and simulation. TMN systems for managing public asynchronous transfer mode (ATM) networks generally have a four-level hierarchical structure consisting of a network management system, a few element management systems (EMSs), and several pairs of agents and ATM switches. We construct a Jackson's queuing network and present formulae to calculate its performance measures: distributions of queue lengths and waiting times, mean message response time, and maximum throughput. We perform a numerical analysis and a simulation analysis and compare the results.

Manuscript received June 25, 2001; revised June 17, 2002.

Young Ha Hwang (phone: +82 42 860 6698, email: hyh@etri.re.kr) is with R&D and Strategy Planning Division, ETRI, Daejeon, Korea.

Sang Wook Chung (e-mail: swchung@chonnam.ac.kr) is with Chonnam National University, Gwangju, Korea.

Gil-Haeng Lee (e-mail: ghlee@etri.re.kr) is with Network Technology Laboratory, ETRI, Daejeon, Korea.

Young Il Kim (e-mail: yikim@kt.co.kr) is with KT, Daejeon, Korea.

I. INTRODUCTION

The International Telecommunication Union – Telecommunications (ITU-T) has recommended the telecommunications management network (TMN) system as a management network standard [1]. A TMN, which is based on open system Interconnection (OSI) system management concepts, is organized using object-oriented techniques. The managers in managing systems and the agents in managed systems use a standardized information exchange interface to manage communication networks. The manager sends management operations to agents to obtain information on the managed objects and issues management commands using standard communication protocols, such as the common management information service element/common management information protocol (CMISE/CMIP) [2], [3]. The agents analyze the management commands received from the manager and order appropriate actions for the managed objects or managed resources. The agents also send notifications that may be responses to commands from the manager or events from managed resources such as system faults. CMISE/CMIP is a standard communication protocol for the OSI and TMN system to convey management information between the manager and the agents [4].

The TMN system for public asynchronous transfer mode (ATM) network management generally has a hierarchical structure (Fig. 1). There is an agent system for each ATM switch deployed at each region. The element management system (EMS) is a manager that maintains an ATM sub-network; the network management system (NMS) is a high-level manager that manages several EMSs. Usually, several agents in a TMN system are

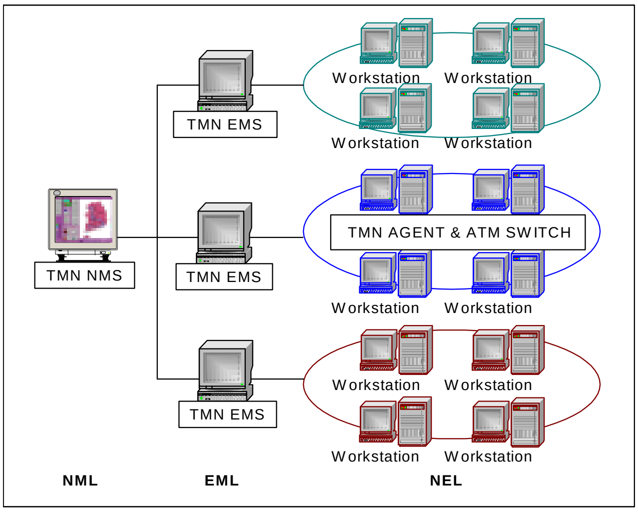

Fig. 1. General structure of a TMN system for public ATM networks.

controlled by a manager [4], [5].

Several authors have studied the problem of analyzing the performance of TMN systems [4], [6]. These previous investigations were concerned only with one EMS, along with several agents and network elements; they did not address the NMS. However, TMN systems consist of many independent sub-systems, and each sub-system plays a key role in the TMN system. Therefore, an analysis of the performance of TMN systems has to contain all sub-systems, such as the NMS, a few EMSs, many agents, and network resources.

Using Jackson's network and simulation models, this paper analyzes the performance of a TMN system that has a four-level hierarchical structure consisting of one NMS, a few EMSs, and several pairs of agents and ATM switches. The feature that makes our investigation different from other studies [4], [6] is that our performance analysis considers all the sub-systems of a TMN system. We construct a queuing network model and present formulae to calculate the performance measures: distributions of queue lengths and waiting times, mean message response time, and maximum throughput. We perform a numerical analysis along with a simulation analysis and compare the results of the numerical analysis with those of the simulation analysis.

II. QUEUING NETWORK MODEL

Notation

- *NI* Queue at which management commands to the NMS arrive
- *NO* Queue at which notifications from EMSs or the NMS itself arrive
- EI_i Queue at which management commands to EMS_i arrive (i=1,…,m)
- EO_i Queue at which notifications from agents or EMS_i itself arrive (i=1,…,m)
- AI_{ij} Queue at which management commands to agent j under the control of EMS_i arrive (i=1,…,m, j=1,…,n_i)
- AO_{ij} Queue at which notifications from agent j itself or switch j under the control of EMS_i arrive (i=1,…,m, j=1,…,n_i)
- S_{ij} Queue within switch j under the control of EMS_i (i=1,…,m, j=1,…,n_i)
- λ_k Arrival rate of queue k from internal or external networks
- μ_k Service rate of queue k

This section presents the queuing network model for the performance analysis of a TMN system implemented for ATM networks. Figure 2 illustrates a TMN system that manages an ATM network. The queuing network model presented in this paper is not original. The basic idea for the model is derived from reference paper [4]. To the basic model we added network elements (NMS, switches), arrival rates (λ_{HMIij}), branching probabilities (P_{EXi} , $1-P_{EOi}-P_{EXi}$), and the change of locations of arrival rates (λ_{SRij} , λ_{SOSij}) from agents to switches.

1. Model of Subordinate Systems

The model is organized in four layers of subordinate systems: an NMS, m EMSs, n agents and n switches.

First, let's look into the NMS model. There are two sources of management commands in the NMS. One is the command from the NMS user (λ_{NMS}). The other is what the NMS sends to a queue *NI* according to notifications from the EMS with probability $1-P_{NO}$. The services for some of these commands are completed by the NMS itself with probability P_{EI0}, and other commands are directed to the EMSi system with probability P_{EIi} (i=1,…,m). Of course, $\Sigma_i P_{EIi}=1$ (i=0,…,m) must be satisfied. The other queue (*NO*) in the NMS deals with the notifications from the EMS and the NMS itself. After being processed by the NMS, only the messages with probability P_{NO} exit the network; the messages with probability $1-P_{NO}$ are sent back into the queue *NI* for reprocessing.

Second, in each EMS system there are three sources of management commands. One is the command from the EMS*i* user (λ_{EMSi}); another is from the NMS system. The third is what the EMS_i sends to the queue EI_i according to notifications from the agent with probability $1-P_{EOi}-P_{EXi}$. The services for some of these commands are completed by the EMS*i* itself with probability P_{AIi0} (i=1,…,m), and other commands are sent to the agent j system under the control of the EMS*i* with probability P_{AIij} (j=1,…,n_i). $\Sigma j P_{AIij}=1$ (j=0,…,n_i) must be satisfied. The other queue (EO_i) in the EMS*i* deals with notifications from an agent and the EMS itself. After being processed by the queue EO_i in the EMS_i, the messages with probability P_{EXi} go out of the network and some messages with probability P_{EOi} are directed to the queue *NO* in the NMS. Also some messages with probability $1-P_{EOi}-P_{EXi}$ are sent back into the queue EI_i for reprocessing.

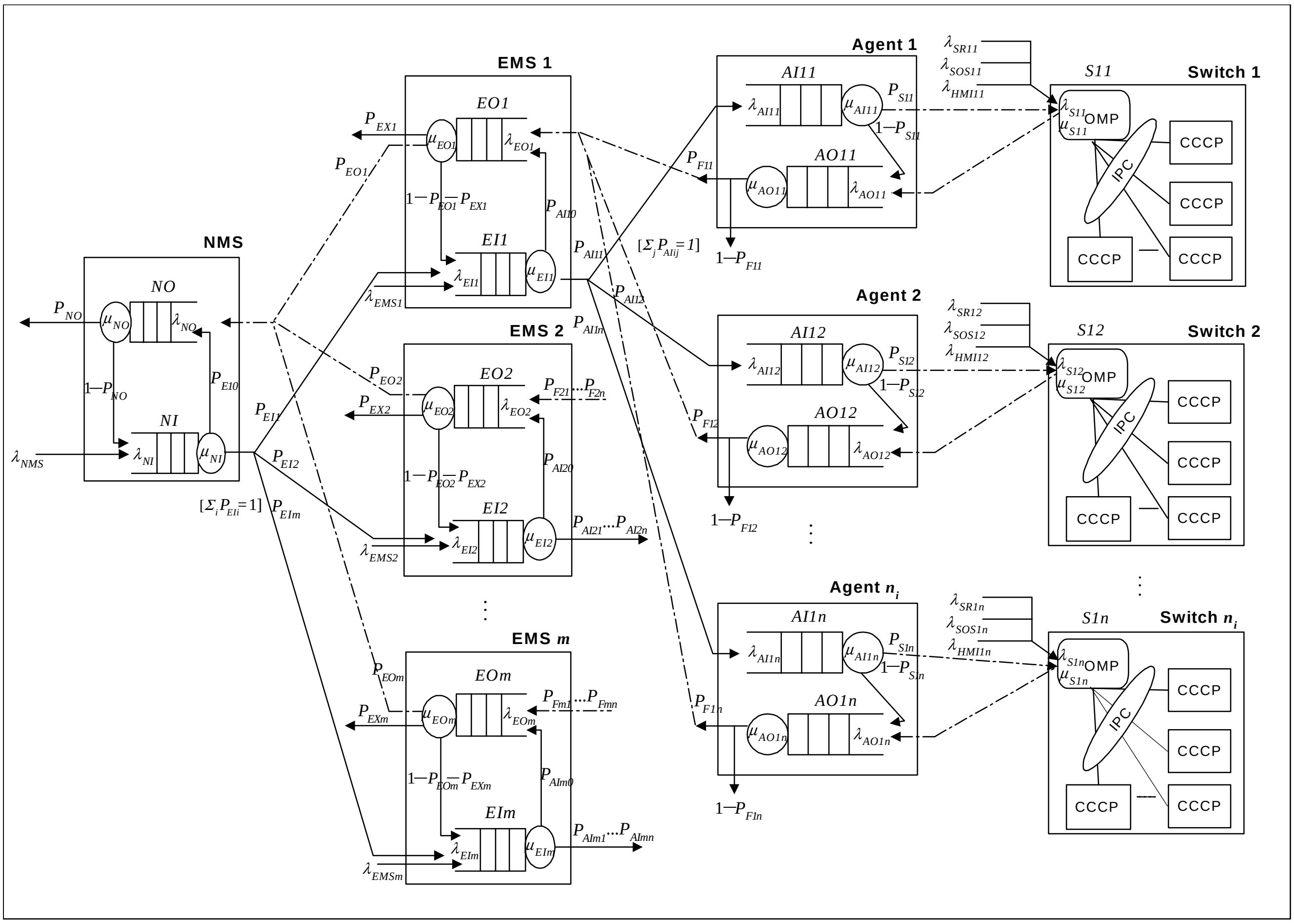

Fig. 2. Queuing network model.

Third, in each agent system the source of management commands is from the EMS system. After being processed by the queue AI_{ij} in the agent j system under the control of the EMS$_i$, these messages are sent to the switch j with probability P_{Sij} and to the queue AO_{ij} in the agent j system itself with probability $1–P_{Sij}$. There are two kinds of responses the agents (AO_{ij}) may receive: The first is from the agent system itself, and the second is the results of the management commands processed by the ATM switch. After being treated by the queue AO_{ij} in the agent j system under the control of the EMS$_i$, some of these responses are not delivered to the EMS system because of the filtering and scoping action of the agent (with probability $1–P_{Fij}$), and others are sent to the EMS system with probability P_{Fij}.

Fourth, in each switch system there are four sources of messages that have to be handled by the Operation and Maintenance Processor (OMP, S_{ij}) within switch j under the control of the EMS$_i$: The first is from agent j; the second from the internal processors within switch j under the control of the EMS$_i$ (λ_{SRij}) by, for example, fault notifications; the third from the operation system that monitors and administers the ATM switch (λ_{SOSij}); the last from the human-machine interface (HMI) of the ATM switch system (λ_{HMIij}). After being handled by the OMP, these notifications are sent to the queue AO_{ij} in the agent j system under the control of the EMS$_i$.

2. Performance Measures

The performance measures to evaluate the performance of the TMN system using the above model are as follows:

- Distribution of queue length (magnitude of waiting messages)
- Distribution of waiting time (waiting time for receiving service)
- Mean message response time (processing time of messages)
- Maximum throughput (maximum magnitude of messages processed per unit time)

The performance measures we use are means to analyze the performance of TMN systems, not ultimate objects. These measures are very important because they are basic measures for analyzing the performance of the systems. Our numerical and simulation analyses are possible by the calculation of these basic measures, and they also make it possible for us to compare the performance of the system for specific cases.

III. PERFORMANCE ANALYSIS

Using the above queuing network model and Jackson's Theorem, we developed formulae to calculate the following performance measures: distribution of queue length and waiting time, mean message response time, and maximum throughput. We made the following assumptions: that all inter-arrival times of each queue are independently and identically distributed according to an exponential distribution (i.e., the input process is Poisson); that all service times of each queue are independently and identically distributed according to another exponential distribution; that the number of all servers of each queue is one; and that all queues are infinite queues (consequently, the network of *M/M/1* queues).

Most mathematical evaluations of the performance of telecommunications systems, such as TMN systems, assume that the traffic characteristics follow Poisson arrival and the service time of an element of the system is exponentially distributed [4], [8]-[10]. It is hard to consider the operations of systems that violate the assumption of the Poisson process. For service system level analysis, the assumption of the Poisson process shows few significant deficiencies. Whenever the Poisson arrival assumption cannot be preserved [11]-[13], other evaluation techniques, such as simulation and measurement, are more appropriate for analyzing the performance of a system.

1. Jackson's Theorem

A Jackson's network is a system of m service queues where queue u (u=1,2,…,m) has

– an infinite queue,

– customers arriving from outside the system according to a Poisson input process with parameter a_u, and

– s_u servers with an exponential service-time distribution with parameter μ

The customers visit the queues in different orders or might not visit them all. A customer leaving queue u is routed next to queue v (v=1,2,…,m) with probability p_{uv} or departs the system with probability

$$q_u = 1 - \sum_{v=1}^{m} p_{uv} \ .$$

Under steady-state conditions, each queue v (v=1,2,…,m) in a Jackson's network behaves as if it were an independent *M/M/s* queuing system with arrival rate

$$\lambda_v = a_v + \sum_{u=1}^{m} \lambda_v \, p_{uv} \ \text{, where } s_v\mu_v > \lambda_v.$$

In such a Jackson's network, a simple form for the solution, called the product form solution, can be used to obtain measures of performance for the network [14]-[16].

In this TMN system, there are $2+2m+m(2n_i+n_i)$ infinite service queues. The parameter a_u of a Jackson's network corresponds to arrival rates, λ_{NMS}, λ_{EMSi}, λ_{SRij}, λ_{HMIij}, λ_{SOSij}. The server s_u of each queue is one. Probabilities p_{uv} and q_u correspond to branching probabilities, P_{NO}, P_{EIi}, P_{EOi}, P_{EXi}, P_{AIij}, P_{Sij}, P_{Fij}. Parameter μ_u corresponds to the service rate μ_k of queue k. Finally, the arrival rate λ_v of queue v in a Jackson's network corresponds to the arrival rate λ_k of queue k. Thus, the queuing network of the TMN system that we illustrated in Fig. 2 could be a Jackson's network.

2. Arrival Rate and Utilization Factors

Table 1 shows the arrival rates and utilization factors for each queue. The utilization factor ρ_k of queue k is an important parameter called the traffic intensity of the system [17].

Table 1. Arrival rates and utilization factors for each queue.

Queue	λ_k	ρ_k
NI	$\lambda_{NMS}+(1-P_{NO})\lambda_{NO}$	λ_{NI}/μ_{NI}
NO	$P_{EI0}\lambda_{NI}+\sum_{i=1}^{m} p_{EOi}\lambda_{EOi}$	λ_{NO}/μ_{NO}
EI_i	$\lambda_{EMSi}+P_{EIi}\lambda_{NI}+(1-P_{EOi}-P_{EXi})\lambda_{EOi}$	λ_{EIi}/μ_{EIi}
EO_i	$P_{AIij}\lambda_{EIi}+\sum_{j=1}^{n} p_{Fij}\lambda_{AOij}$	λ_{EOi}/μ_{EOi}
AI_{ij}	$P_{AIij}\lambda_{EIi}$	$\lambda_{AIij}/\mu_{AIij}$
AO_{ij}	$\lambda_{Sij}+(1-P_{Sij})\lambda_{AIij}$	$\lambda_{AOij}/\mu_{AOij}$
S_{ij}	$\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij}+P_{Sij}\lambda_{AIij}$	λ_{Sij}/μ_{Sij}

3. Distribution of Queue Length

Let $P_k(n)$ be the probability of exactly n messages in queue k. The probability of exactly n messages in queue k is

$$P_k(n)=(1-\rho_k)\rho_k^{\,n}. \tag{1}$$

The expected number of messages (mean queue length) of queue k is

$$L_k=\rho_k/(1-\rho_k)=\lambda_k/(\mu_k-\lambda_k). \tag{2}$$

The expected number of messages (mean queue length, excluding messages being served) of queue k is

$$L_{qk}=\rho_k^2/(1-\rho_k)=\lambda_k^2/[\mu_k(\mu_k-\lambda_k)]. \tag{3}$$

The expected total number of messages in the entire system then is

$$L_{total}=\sum_k L_k = \sum_k \frac{\rho_k}{1-\rho_k}. \tag{4}$$

Using (1) and Jackson's Theorem [14], [15], the joint distribution of the expected number of messages (mean queue length) in a system can be obtained by multiplying the probability of exactly n_k messages in queue k. Thus,

$$P(n) = P_{NI}(n_{NI})P_{NO}(n_{NO})P_{EII}(n_{EII})P_{EOI}(n_{EOI})\ldots P_{Smn}(n_{Smn}) = \sum_k (1-\rho_k)\rho_k^{n_k}, \tag{5}$$

where the state of system n is the vector (n_{NI}, n_{NO}, n_{EI1}, n_{EO1}, …, n_{Smn}) that denotes the number of messages at each queue.

4. Distribution of Waiting Time

The expected waiting time (including service time) of messages in queue k, W_k, can be calculated from (2) and Little's formula [16] as

$$W_k=1/(\mu_k-\lambda_k). \tag{6}$$

Also, the expected waiting time (excluding service time) of messages in queue k, W_{qk}, is

$$W_{qk}=\lambda_k/[\mu_k(\mu_k-\lambda_k)]=\rho_k/[\mu_k(1-\rho_k)]. \tag{7}$$

Obtaining W_{total} (the expected total waiting time including service time in the entire system for a message) is more complicated. The expected waiting times at the respective queues cannot be simply added, because a message does not necessarily visit each queue exactly once. However, Little's formula can still be used, where the system arrival rate λ_{total} is the sum of the arrival rate from outside to the queues [16],

$$\lambda_{total}=\lambda_{NMS}+\sum_{i=1}^{m}\lambda_{EMSi}+\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})$$

Thus, $$W_{total}=L_{total}/\lambda_{total}=\frac{\sum_k \frac{\rho_k}{1-\rho_k}}{\lambda_{NMS}+\sum_{i=1}^{m}\lambda_{EMSi}+\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})}. \tag{8}$$

5. Mean Message Response Time

When the message response time is defined as the time that the response for a management command invoked by an NMS user takes to arrive in the NMS user after being processed by the EMS, Agent, and Switch, its expected value W_{NMS} can be obtained. The expected message response time of a management command by an NMS user, W_{NMS}, is

$$W_{NMS}=W_{NI}+P_{EI0}W_{NO}+(1-P_{EI0}) \times\left\{\sum_{i=1}^{m}\frac{P_{EIi}}{1-P_{EI0}}[W_{EIi}+P_{AIi0}(W_{EOi}+W_{NO})+(1-P_{AIi0})a]\right\},$$

where

$$a=\sum_{j=1}^{n_i}\frac{P_{AIij}}{1-P_{AIi0}}(W_{AIij}+W_{AOij}+P_{Sij}W_{Sij}+W_{EOi}+W_{NO}). \tag{9}$$

6. Maximum Throughput

In the above queuing network model, as the arrival rate increases, a queue k with a larger value of ρ_k will introduce instability. Hence, the queue with the largest value of ρ is called the bottleneck of a TMN system [17].

Eq. (7) shows that as the traffic intensity ρ_k approaches 1, the waiting time for messages approaches infinity. Therefore, the maximum throughput of the system can be predicted by evaluating the traffic intensity of the bottleneck, $\rho_{bottleneck}=1$ ($\mu_{bottleneck}=\lambda_{bottleneck}$) [17].

In this system it can be predicted that the bottleneck is the queue *NO* in the NMS, because all notifications to management commands from the NMS or EMS user and all notifications from several agents or switch systems are concentrated in the queue *NO* through some EMSs, agents, and switches. Therefore, the maximum throughput can be obtained by using [18]

$$\mu_{NO}=\lambda_{NO}=P_{EI0}\lambda_{NI}+\sum_{i=1}^{m}p_{EOi}\lambda_{EOi}. \tag{10}$$

IV. NUMERICAL ANALYSIS

This section presents our numerical analysis for the performance measures of a TMN system composed of one NMS, m EMSs, n agents, and n switches. We assumed that each value of the parameters in all EMSs, agents, and switches is the same. Thus, for example,

$n_1=n_2=\ldots=n_m=n$, $P_{EO1}=P_{EO2}=\ldots=P_{EOm}$,
$P_{EI1}=P_{EI2}=\ldots=P_{EIm}$, $P_{Si1}=P_{Si2}=\ldots=P_{Sin}$,
$\lambda_{EMS1}=\lambda_{EMS2}=\ldots=\lambda_{EMSm}$, $\lambda_{SR11}=\lambda_{SR12}=\ldots=\lambda_{SR1n}$,
$\mu_{EO1}=\mu_{EO2}=\ldots=\mu_{EOi}$.

In fact, the above assumption is not practical in real TMN systems. However, if we don't make this assumption it is not easy to solve mathematically.

The formulae of the numerical analysis for a TMN system under the above conditions are as follows:

• Arrival rates of each queue

$$\begin{aligned}\lambda_{NO}=\Big(&P_{EI0}\lambda_{NMS}+mnP_{EOi}P_{Fij}(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})\\&+\{mP_{EOi}(nP_{Fij}+1)\cdot P_{AIij}\cdot[\lambda_{EMSi}+P_{EIi}\lambda_{NMS}\\&\quad+nP_{Fij}(1-P_{EOi}-P_{EXi})(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})]\}\\&\quad/\{1-(1-P_{EOi}-P_{EXi})P_{AIij}[1+nP_{Fij}P_{Sij}+nP_{Fij}(1-P_{Sij})]\}\Big)\\/\Big(&1-P_{EI0}(1-P_{NO})-mP_{EOi}(nP_{Fij}+1)P_{AIij}P_{EIi}(1-P_{NO})\\&\quad/\{1-(1-P_{EOi}-P_{EXi})P_{AIij}\\&\quad\times[1+nP_{Fij}P_{Sij}+nP_{Fij}(1-P_{Sij})]\}\Big)\end{aligned}$$

$\lambda_{NI}=\lambda_{NMS}+(1–P_{NO})\lambda_{NO}$,

$$\lambda_{EIi} = \{\lambda_{EMSi} + P_{EIi}[\lambda_{NMS} + (1-P_{NO})\lambda_{NO}] + (1-P_{EOi}-P_{EXi})nP_{Fij}(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})\} / \{1-(1-P_{EOi}-P_{EXi})P_{AIij}[1+nP_{Fij}P_{Sij}+nP_{Fij}(1-P_{Sij})]\}$$

$\lambda_{EOi} = P_{AIij}\lambda_{EIi} + nP_{Fij}\lambda_{AOij}$,
$\lambda_{AIij} = P_{AIij}\lambda_{EIi}$,
$\lambda_{AOij} = \lambda_{Sij}+(1–P_{Sij})P_{AIij}\lambda_{EIi}$,
$\lambda_{Sij} = \lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij}+P_{Sij}\lambda_{AIij}$,
$\lambda_{total} = \lambda_{NMS}+m\lambda_{EMSi}+mn(\lambda_{SRij}+\lambda_{SOSij}+\lambda_{HMIij})$.

• Traffic intensity of queue k

$\rho_k = \lambda_k/\mu_k$.

• The expected number of messages (mean queue length) of queue k

$L_k = \rho_k/(1–\rho_k)$.

• The expected number of messages (mean queue length, excluding messages being served) of queue k

$L_{qk}=\rho_k^2/(1–\rho_k)=\lambda_k^2/\{\mu_k(\mu_k–\lambda_k)\}$.

• The expected total number of messages in the entire system

$L_{total}=L_{NI}+L_{NO}+m(L_{EOi}+L_{EIi})+mn(L_{AOij}+L_{AIij}+L_{Sij})$.

• The expected waiting time (including service time) of messages in queue k

$W_k = 1/(\mu_k–\lambda_k)$.

• The expected waiting time (excluding service time) of messages in queue k

$W_{qk}=\lambda_k/\{\mu_k(\mu_k–\lambda_k)\}=\rho_k/\{\mu_k(1–\rho_k)\}$.

• The expected message response time of a management command by the NMS user

$$W_{NMS}=W_{NI}+P_{EI0}W_{NO}+(1–P_{EI0})[W_{EIi}+P_{AIi0}(W_{EOi}+W_{NO}) + (1–P_{AIi0})(W_{AIij}+W_{AOij}+P_{Sij}W_{Sij}+W_{EOi}+W_{NO})].$$

The values of parameters used in this analysis are as follows:

• Arrival rates: $\lambda_{NMS}, \lambda_{EMSi}, \lambda_{SRij}, \lambda_{HMIij}, \lambda_{SOSij} = 0.07$.

• Branching Probabilities: $P_{NO}=0.99$, $P_{EI0}=0.1$, $P_{EIi}=(1–P_{EI0})/m$, $P_{EOi}=0.5$, $P_{EXi}=0.49$, $P_{AIi0}=0.1$, $P_{AIij}=(1–P_{AIi0})/n$, $P_{Sij}=0.5$, $P_{Fij}=0.9$.

• Service rates: $\mu_{NI}=2.9$, $\mu_{NO}=2.78$, $\mu_{EIi}=2.9$, $\mu_{EOi}=2.78$, $\mu_{AOij}=2.15$, $\mu_{AIij}=4.12$, $\mu_{Sij}=7.31$ (real data from reference papers [6], [9]).

1. Effect of λ_{NMS} on W_{NMS}

Figure 3 shows the effect of λ_{NMS} on W_{NMS} under the above conditions. The figure indicates that W_{NMS} increases drastically as λ_{NMS} increases and that it has the same trend regardless of the increase of n (the number of agents and switches) and m (the number of EMSs) (n,m=5, 10, 15, 20). A trend in which the graph increases drastically at the point of about λ_{NMS}=2.85 is revealed. At this point, ρ_{NI} is 1 (Fig. 4), the bottleneck of the system is queue NI (the NMS input queue), and the maximum throughput is about λ_{NI}=2.90, irrespective of the network size (the values of n and m).

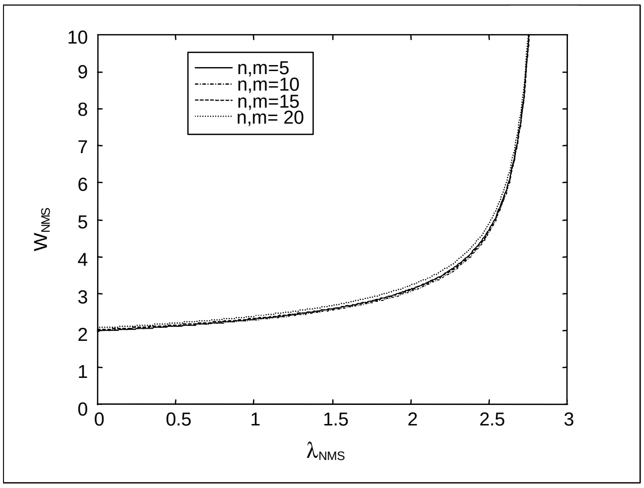

Fig. 3. Effect of λ_{NMS} on W_{NMS}.

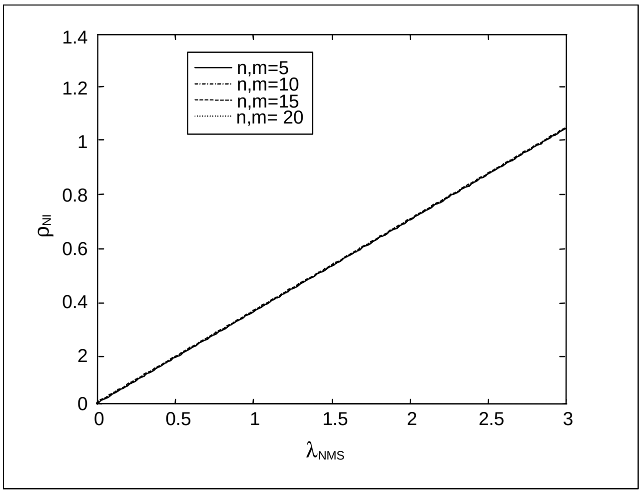

Fig. 4. Effect of λ_{NMS} on ρ_{NI}.

These figures also indicate that λ_{NMS} has little effect on the determination of the optimal number of EMSs and Agents (n, m, network size).

2. Effect of λ_{SRij} on W_{NMS}

Figure 5 shows the effect of λ_{SRij} (messages from internal processors within switch j under the control of the EMSi may relate to fault event reports) on W_{NMS}. The figure indicates that W_{NMS} increases drastically as λ_{SRij} increases and that it also has a much quicker rising trend according to the increase of n and m (n,m=5, 10, 15, 20). The figure also reveals a trend wherein the graph increases drastically at about λ_{SRij}=0.24, 0.06, 0.025, 0.015. At this point, ρ_{NO} is 1 (Fig. 6), the bottleneck of the system is queue *NO*, and the maximum throughput is about λ_{NO}=2.78, irrespective of the network size (the values of *n* and *m*). That is, the bottleneck is the NMS output queue and waiting time increases with the fault event rate, which increases with the network size.

These figures also indicate that if λ_{SRij}=0.215, 0.055, 0.022, 0.013 and ρ_{NO} (the utilization factor of the bottleneck of systems) = 0.9, the optimal number of EMSs and Agents is *n*, *m*= 5, 10, 15, 20.

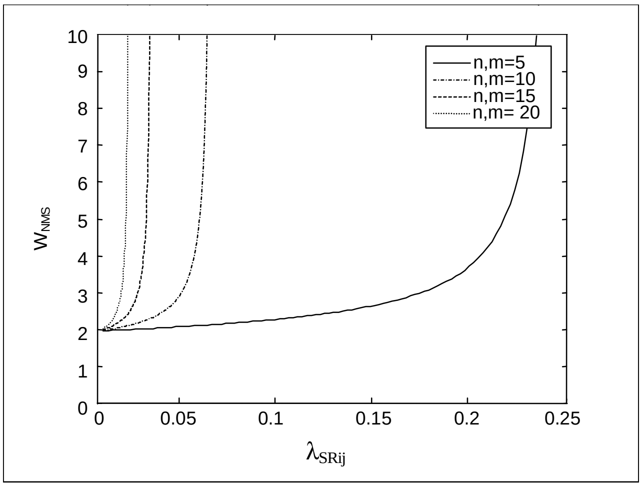

Fig. 5. Effect of λ_{SRij} on W_{NMS}.

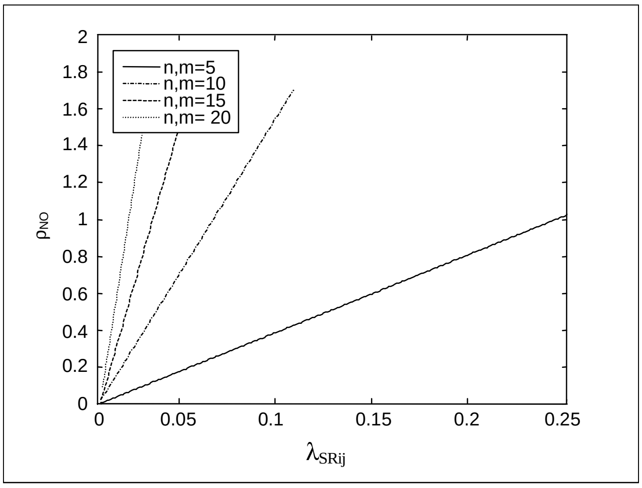

Fig. 6. Effect of λ_{SRij} on ρ_{NO}.

3. Effect of *n* on W_{NMS}

Figure 7 shows the effect of *n* on W_{NMS}. The figure indicates that W_{NMS} increases drastically as *n* increases, and that it has a much quicker rising trend due to the increase of λ_{SRij} (λ_{SRij}=0.10, 0.15, 0.20, 0.25). A trend wherein the graph increases drastically at about *n*=5, 6, 8, 12 is displayed. At these points, ρ_{NO} is 1 (Fig. 8),

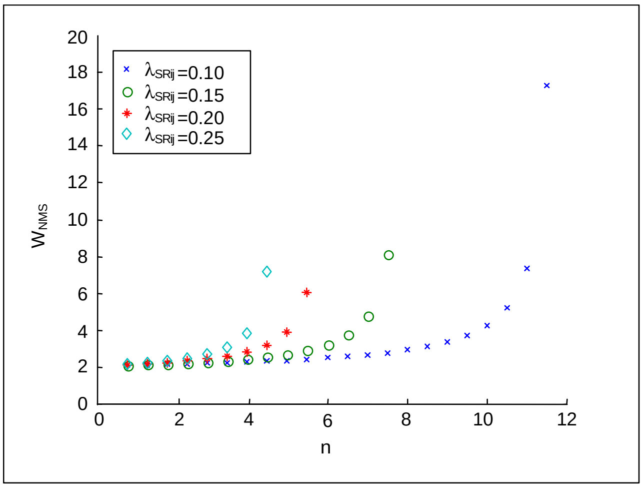

Fig. 7. Effect of *n* on W_{NMS}.

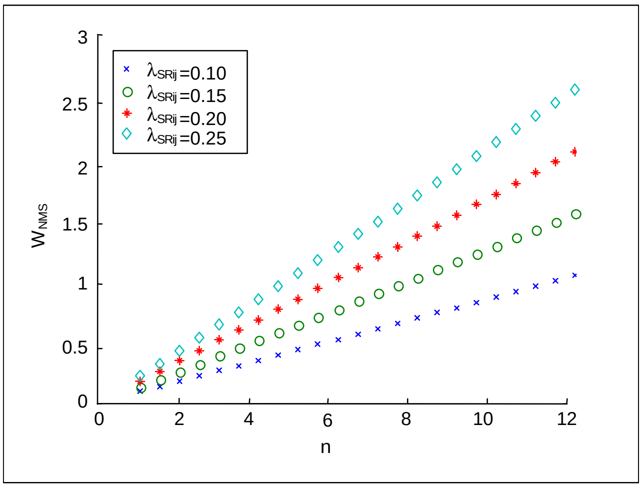

Fig. 8. Effect of *n* on ρ_{NO}.

the bottleneck of the system is queue *NO* (the NMS output queue), and the maximum throughput is about λ_{NO}=2.78 (regardless of the value of λ_{SRij}).

These figures also indicate that if λ_{SRij}=0.10, 0.15, 0.20, 0.25 and ρ_{NO} (the utilization factor of the bottleneck of systems) = 0.9, the optimal number of Agents is *n*=4, 5, 7, 10.

V. SIMULATION ANALYSIS

This section describes our simulation analysis of the performance of a TMN system composed of one NMS, *m* (*m*=5) EMSs, *n* (*n*=5) agents, and *n* (*n*=5) switches. AweSim (Visual SLAM [19]) is used as a simulation tool. The basic assumptions, the values of the parameters, and the formulae used in this analysis are the same as those used in the numerical analysis in section IV. The distribution of the inter-arrival time and the service time of the commands or messages in queue k are Exp ($1/\lambda_k$) and Exp ($1/\mu_k$), respectively. To obtain the results in a steady-state, we used one million times and two million times as run-times.

In the first simulation (Simulation I), the above conditions were used with the model in Fig. 2. However, in the second simulation (Simulation II), we used a different queuing network model (Fig. 9). For Simulation I, it was impossible to measure the mean message response time (W_{NMS}) because of the continuous circulation of messages by the branching probabilities ($1–P_{NO}$, $1–P_{EOi}–P_{EXi}$) within the gray circle in Fig. 9. In Simulation II those branching probabilities were changed as follows:

• Branching Probabilities: $1–P_{NO}=0$, $P_{NO}=1$, $1–P_{EOi}–P_{EXi}=0$, $P_{EOi}=0.5$, $P_{EXi}=0.5$.

We expected differences in the results of Simulations I and II. For the comparison of the results of the mean message response time (W_{NMS}) in section V.2, the results of Simulation II were used.

Table 2 shows the results of the two simulations: the results for the two run-times are nearly the same. Thus, the system is in a steady-state. The results of the two million run-time were used for the comparison with the analytic method proposed in this paper. However, there was a slight difference between the results of Simulation I and Simulation II because we used different queuing network models (Figs. 2 and 9).

Table 3 shows the results of the comparison of the analytic method and simulation method; the results are almost the same. For the comparison, we used the results of Simulation I except for the mean message response time, for which we used the result of Simulation II. Thus, there was a little difference between the results of the mean message response time (W_{NMS}).

In accordance with the compared results of the analytic and the simulation method, there was no significant difference between the results of the two methods. Hence, we can conclude that the analytic method of the performance analysis proposed in this paper is suitable.

VI. CONCLUSION

We devised a queuing network model and an analytic model for performance analysis using Jackson's Theorem and then performed numerical and simulation analyses for specific cases.

The numerical analysis clearly showed that the number of subordinate subsystems and the quantity of traffic within the system had substantial effects on the performance of the system. In addition, by determining the optimal number of subordinate subsystems in specific cases we showed how to design an appropriate TMN system and to evaluate its performance efficiently.

This paper presented a general model for performance analysis

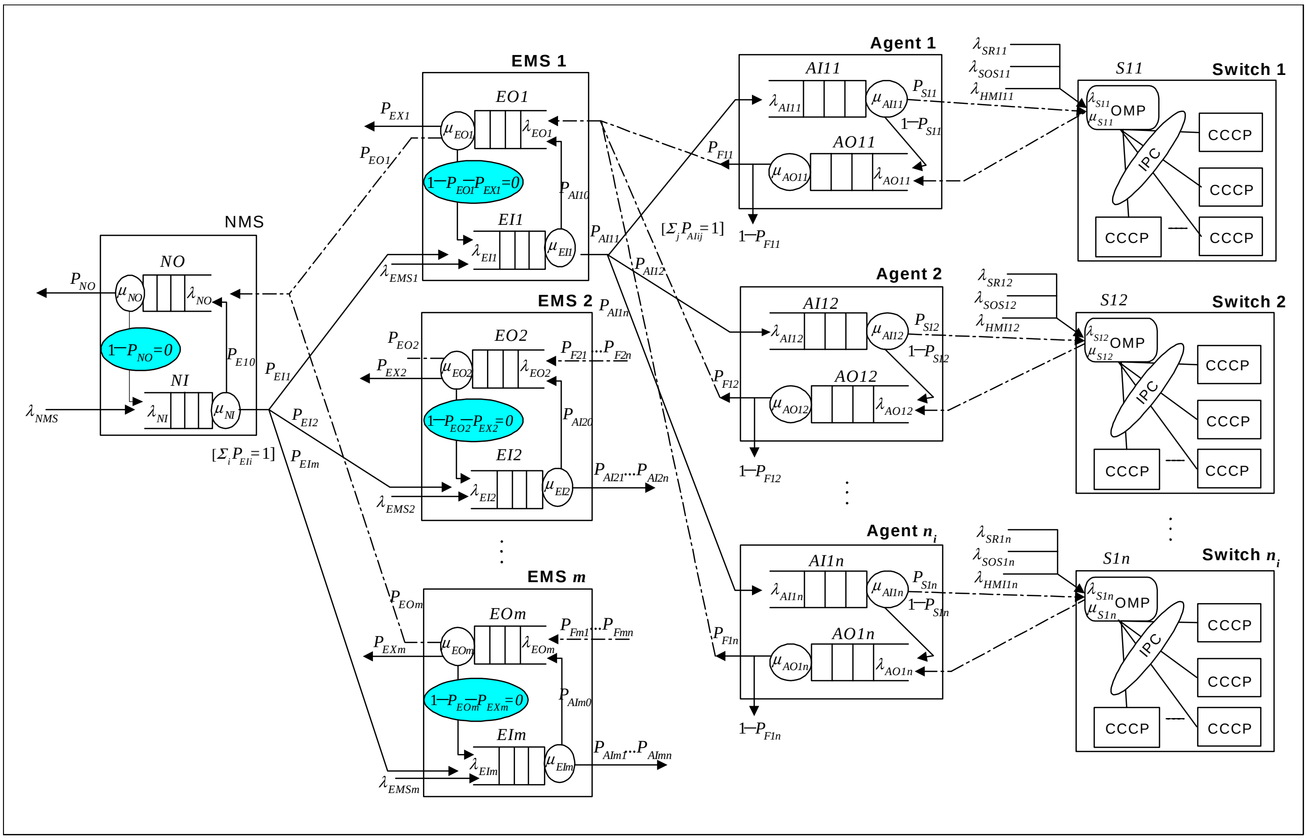

Fig. 9. Queuing network model used in simulation II.

Table 2. Results of the simulation.

Measures	Measures for each queue	Simulation I		Simulation II	
		Run-time		Run-time	
		1 million time	2 million time	1 million time	2 million time
Utilization Factors (ρ_k)	ρ_{NI}	0.033	0.033	0.024	0.024
	ρ_{NO}	0.932	0.933	0.920	0.919
	ρ_{EIi}	0.034	0.034	0.028	0.028
	ρ_{EOi}	0.372	0.372	0.368	0.367
	ρ_{AOij}	0.106	0.106	0.105	0.105
	ρ_{AIij}	0.004	0.004	0.004	0.004
	ρ_{Sij}	0.030	0.030	0.030	0.030
Queue Lengths (excluding messages being served) (L_{qk})	L_{qNI}	0.001	0.001	0.001	0.001
	L_{qNO}	12.798	12.846	10.411	10.348
	L_{qEIi}	0.001	0.001	0.001	0.001
	L_{qEOi}	0.221	0.220	0.214	0.212
	L_{qAOij}	0.012	0.012	0.012	0.012
	L_{qAIij}	0.000	0.000	0.000	0.000
	L_{qSij}	0.001	0.001	0.001	0.001
Waiting Times (excluding service time) (W_{qk})	W_{qNI}	0.011	0.011	0.008	0.008
	W_{qNO}	4.937	4.955	4.070	4.048
	W_{qEIi}	0.012	0.012	0.010	0.010
	W_{qEOi}	0.214	0.213	0.209	0.208
	W_{qAIij}	0.001	0.001	0.001	0.001
	W_{qAOij}	0.054	0.054	0.055	0.055
	W_{qSij}	0.004	0.004	0.004	0.004
Mean Message Response Time	W_{NMS}	-	-	6.124	6.123

Table 3. Results of analytic and simulation methods.

Measures	Measures of each queue	Results	
		Analytic method	Simulation method
Utilization Factors (ρ_k)	ρ_{NI}	0.0332	0.033
	ρ_{NO}	0.9403	0.933
	ρ_{EIi}	0.0337	0.034
	ρ_{EOi}	0.3747	0.372
	ρ_{AOij}	0.1059	0.106
	ρ_{AIij}	0.0043	0.004
	ρ_{Sij}	0.0299	0.030
Queue Lengths (excluding messages being served) (L_{qk})	L_{qNI}	0.0011	0.001
	L_{qNO}	14.8044	12.846
	L_{qEIi}	0.0012	0.001
	L_{qEOi}	0.2246	0.220
	L_{qAOij}	0.0125	0.012
	L_{qAIij}	0.000018306	0.000
	L_{qSij}	0.0009221	0.001
Waiting Times (excluding service time) (W_{qk})	W_{qNI}	0.0118	0.011
	W_{qNO}	5.6636	4.955
	W_{qEIi}	0.0120	0.012
	W_{qEOi}	0.2156	0.213
	W_{qAIij}	0.0010	0.001
	W_{qAOij}	0.0551	0.054
	W_{qSij}	0.0042	0.004
Mean Message Response Time	W_{NMS}	7.8947	6.123

of TMN systems. With the values of any parameter for a specific system design (for example, n, m, λ, μ), a performance analysis for the specific design is possible, and the results are useful in designing an appropriate system.

As a further study, with an analysis of the structure of the communication protocol stack of a real TMN system, we will be able to consider many kinds of operational processes and management messages (e.g., an extra ordinal fault situation and burst type event reports) occurring within systems. In addition, by measuring actual traffic and obtaining empirical data (including the number of Network Elements and the speed of lines), an analysis of the performance of TMN systems considering real user demand and non-Poisson traffic could be the focus of significant research.

REFERENCES

[1] ITU-T Recommendation M.3010, "Principles for the Telecommunications Management Network," Feb. 2000.

[2] ITU-T Recommendation X.710, "Common Management Information Service," Oct. 1997.

[3] ITU-T Recommendation X.711, "Common Management Information Protocol," Oct. 1997.

[4] J.Y. Lee, C.J. Hwang, G.H. Lee, and W.D. Woo, "Implementation and Performance Analysis of a TMN System for Public ATM Networks," *The 23rd Int'l Conf. on Computers & Industrial Engineering*, 1998, Computers & Industrial Engineering.

[5] H.S. Hwang, B.Y. Lee, G.H. Lee, and W.D. Woo, "Implementation of TMN Agent for ATM Switch: Considering Integration of Agent into ATM Switch," *KICS J.*, vol. 23, no. 5, May 1998, pp. 1360-1371.

[6] H.S. Hwang, "Performance Evaluation of HANbit ACE64 TMN EMS," *The 3rd Conf. on Next Generation Communication Software*, 1999, Korean Institute of Communication Sciences.

[7] Y.B. Kim, S.S. Lee, C.H. Oh, Y.S. Kim, C. Han, and C.H. Yim, "An Architecture of Scalable ATM Switching System and Its Call Processing Capacity Estimation," *ETRI J.*, vol. 18, no. 3, Oct. 1996, pp. 107-125.

[8] S.S. Lee, J.G. Lee, Y.B. Kim, and Y.S. Kim, "An Architecture of Scalable ATM Switching System and Its Performance," *Telecommunication Systems*, vol. 14, 2000, pp. 269-290.

[9] S.J. Kim et al, "Estimation of Call Setup Delay Time in an ATM Switching System," *Technical Memo* (*TM*97-*KET*34572), Electronics and Telecommunication Research Institute, Aug. 1997.

[10] G.M. Lee et al., "Flow-Based Admission Control Algorithm in the DiffServ-Aware ATM-Based MPLS Network," *ETRI J.*, vol. 24, no. 1, Feb. 2002, pp. 43-55.

[11] B.-H. Choi et al., "Rate Proportional SCFQ (RP-SCFQ) Algorithm for High-Speed Packet-Switched Networks," *ETRI J.*, vol. 22, no. 3, Sept. 2000, pp. 1-9.

[12] Y.M. Jang et al., "Estimation and Prediction-Based Connection Admission Control in Broadband Satellite Systems," *ETRI J.*, vol. 22, no. 4, Dec. 2000, pp. 40-50.

[13] D.W. Hong et al., "An ATM Network Management System for Point-to-Multipoint Reservation Service Provision," *ETRI J.*, vol. 24, no. 4, Aug. 2002, pp. 299-310.

[14] J.R. Jackson, "Networks of Waiting Lines," *Operations Research*, vol. 5, no. 4, 1957, pp. 518-521.

[15] J.R. Jackson, "Jobshop-Like Queuing Systems," *Management Science*, vol. 10, no. 1, 1963, pp. 131-142.

[16] Frederick S. Hiller and Gerald J. Lieberman, *Introduction to Operations Research*, McGraw-Hill, New York, 6th ed., 1995.

[17] Kishor Shridharbhai and Trivedi, *Probability and Statistics with Reliability, Queuing and Computer Science Applications*, Prentice-Hall, 1982

[18] Divakara K. Udupa, *TMN Essentials*, McGraw-Hill, 1999.

[19] A. Alan, B. Pritsker, and Jean J. O'Reilly, *Simulation with Visual SLAM and AweSim*, Wiley, 1999.

Young Ha Hwang received MS degree in Industrial engineering from Chonnam National University, Korea, in 2001. He is a Senior Member of Engineering Staff at ETRI, Korea. His research areas are system performance analysis and product quality assurance.

Sang Wook Chung received his PhD in Industrial engineering from KAIST (Korea Advanced Institute of Science and Technology) in 1992. He has been an Associate Professor of Industrial Engineering at Chonnam National University since 1999. His research areas are system performance analysis, reliability analysis, product quality assurance and optimal design of accelerated life tests.

Gil-Haeng Lee is a Principal Member of Engineering Staff and Team Leader at ETRI, Korea. He received MS and PhD degrees in computer science from KAIST in 1986 and 1996, respectively. His research interests are load balancing and distributed processing, network management, speech recognition, and real time DBMS.

Young Il Kim is a Principle Member of Engineering Staff and Director at KT, Korea. He received his MS in computer science from Hankuk University of Foreign Studies in 1996 and PhD in computer science from Chungbuk University in 1999. His current research interests are IP network planning, IP network management, and internet traffic.