# 용량제한을 갖는 중첩형 대기행렬 네트워크의 성능 범위분석

# (A Boundness Analysis of Performance on the Nested Queueing Network with Population Constraint)

이 영 †

(Young Rhee)

**요 약** 각 계층 네트워크가 세마포에 의하여 용량제한을 갖는 중첩형 대기행렬 네트워크의 평균대기시간에 관하여 연구하였다. 중첩형 대기행렬 네트워크는 고객의 대기시간 관점에서 보다 간단한 대기행렬 네트워크로 변환될 수 있다. 이러한 중첩형 대기행렬 네트워크의 가장 중요한 특성은 하위계층의 흐름이 상위계층의 상태에 의존적이며, 제한을 받는다는 것이다. 이러한 형태의 대기행렬 네트워크는 성능분석을 정확히 할 수 없기 때문에, 변환된 중첩형 대기행렬 네트워크를 이용하여 평균대기시간에 대한 하한과 상한을 분석하고자 한다. 모의실험에 관한 측정은 도착분포가 단계형태 분포로서 포아송, 얼랑 그리고 초지수분포에 관하여 조사하였다. 이렇게 구한 범위는 추후에 좀 더 근접한 근사치를 구하는데 이용할 수 있다.

**키워드** : 중첩형 대기행렬 네트워크, 세마포, 용량

***Abstract*** In this study, we analyze the mean waiting time on the nested open queueing network, where the population within each subnetwork is controlled by a semaphore queue. The queueing network can be transformed into a simpler queueing network in terms of customers waiting time. A major characteristic of this model is that the lower layer flow is halted by the state of higher layer. Since this type of queueing network does not have exact solutions for performance measure, the lower bound and upper bound on the mean waiting time are checked by comparing them with the mean waiting time in the transformed nested queueing network. Simulation estimates are obtained assuming Poisson arrivals and other phase-type arrival process, i.e., Erlang and hyper-exponential distributions. The bounds obtained can be applied to get more close approximation using the suitable approach.

**Key words** : nested queueing network, semaphore, population

## 1. Introduction

In a typical queueing network, a customer may have traverse several layers of flow controlled mechanisms before it comes out from the network. In this paper, we present a model for analyzing the delays introduced by such nested flow mechanism. The number of customers in each subnetwork is controlled by a semaphore queue. Communication network protocols are designed a layered structure along the line of the OSI model. A common means of implementing the required sequencing, pacing, congestion control, and data integrity functions of protocol layer is a multiple window mechanism, say a sliding window mechanism. As it were, the window mechanism is a means for regulating the flow of active customers through the service nodes in distributed networks. This mechanism is widely applied in data transmission network, as it is the standard level 2 data link flow control, and for level 3 end-to-end flow control. Even if the remote process call and the buffer management schemes

† 정 회 원 : 계명대학교 경영공학과 교수
yrhee@kmu.ac.kr

논문접수 : 2009년 1월 12일
심사완료 : 2009년 5월 19일



정보과학회논문지: 시스템 및 이론 제36권 제4호(2009.8)

are conventionally encountered and well modeled as the window mechanism. The window mechanism has been extensively studied, especially in the context of communication networks. The performance of the mechanism under the nominal operation, and in the presence of stress conditions such as transmission errors or loss or delay of frames are well understood.

Reiser [1] modeled a computer communication system consisting of many virtual routes with end-to-end window flow control, as a closed multichain queueing network under the assumption of a loss system. Pennotti and Schwartz [2] and Schwartz [3] analyzed a virtual route as a closed tandem queueing network under the same assumption. Reiser [1] observed that in a real situations, packets that arrives to find a full window are not lost, but are queued in an input queue. Reiser [1] and Thomasian and Bay [4], use a flow equivalent server technique to model the sliding window link as a single server queue with state dependent service rate. In this approach, the effect of delays due to all sequence numbers in use is accounted for in the delivery service time of the equivalent server. Varghee, Chou and Nilsson [5] and Gihr and Kuehn [6], presented a similar approach to the above. Varghee, Chou and Nilsson [5] analyzed an open queueing network without an acknowledgment delay using the approximation method. Gihr and Kuehn [6] obtained the characteristics of the physical transmission process using hierarchical decomposition and aggregation methods. Rhee and Perros [7] modeled an open tandem queueing network with population constraint and constant service times. The total number of customers that may be presented in the network can not exceed a given value $k$. Customers arriving at the queueing network when there are more than $k$ customers are forced to wait in an external queue. The arrival process to the queueing network is assumed to be arbitrary.

For an analysis of nested layered communication network, the communication functions are partitioned into a vertical set of layers. Each layer performs a related subset of the functions required to communicate with another system. It relies on the next lower layer to perform more primitive functions and to conceal the details of those functions. It provides services to next higher layer. Mitchell and Lide [8] presented a general framework to model sliding window flow control from the closed queueing network models. Fdida, Perros and Wilk [9] presented a methodology for analyzing nested and tandem configurations of sliding window controlled networks. Each layer of sliding window control is reduced to a state dependent infinite server queue without acknowledgment using a flow-equivalence methodology. A single-hop OSI structured network with multiple layers of sliding window flow control and packet fragmentation between layers is analyzed by Shapiro and Perros [10]. They presented a hierarchical method to analyze nested sliding window flow controlled layers. Each layer with sliding window control is reduced to a single queue with state dependent service rate.

In this paper, we present a nested open tandem queueing network controlled by semaphore queue. This type of queueing networks have application in diverse area, such as pallet based production system, computer sharing and multiprogramming systems, communication network model and semaphore controlled software in an operating system. A major characteristic of this model is that the lower layer flow is halted by the state of higher layer. Rhee and Perros [11] analyzed a mean waiting time of an open tandem queueing network with population constraint assuming constant service times. 2 node queueing network which is equivalent to the original queueing network as far as a customer's waiting time is concerned, is obtained. Rhee [12] presented some properties that the inter-change of nodes does not make any difference to customer's waiting time in the nested queueing network under a certain condition. Using those properties, the dramatic reduction of network dimensionality is executed. It is also generalized that the reduction of network dimensionality can be extended to an n-layer open queueing network with population constraint and constant service times. This paper is a sequel to earlier two papers by Rhee and Perros [11] and Rhee [12].

This paper is organized as followed: Section 2 presents the model for a nested queueing network with semaphore queue. In section 3, some characteristics of a semaphore controlled queueing network are presented to analyze the customer's waiting time. In section 4, we demonstrate its value by comparing it against simulation estimates. Finally, the conclusion is presented at section 5.

We note that throughout this paper, we interpret the waiting time of a customer as the total time a customer waits in the queueing network, rather than the total time it takes to traverse the queueing network which also includes service times.

## 2. A nested queueing network under study

Let us consider a nested open tandem queueing network with population constraint and constant service times as shown in Figure 1. The population constraint of the queueing network is controlled by a semaphore. For presentation purposes, we shall refer to the outside network as the high layer, and also refer to the network with purely nested as the low layer. We assume that the queueing network consists of $n$ nodes for the low layer, $m$ nodes for the high layer respectively. The arrival process to the queueing network is assumed to be an arbitrary general distribution with a rate $\lambda$. The semaphore mechanism consists of a pool of $k_1$ tokens and an internal queue, $E_1$ for the low layer, and a pool of $k_2$ tokens and an external queue, $E_2$ for the high layer.

An arriving customer takes a token from the token pool, $P_2$ to enter the high layer queueing network. The customer holds the token until it leaves the high layer queueing network. The customer proceeds to the high layer queueing network until $E_1$. In order to enter the low layer queueing network, the customer needs another token from the low layer token pool, $P_1$. The customer is then subjected to the low layer window flow control. Upon service completion in the low layer queueing network, the customer returns its token immediately to $P_1$ and proceeds to the rest of the node of the high layer queueing network. Again upon service completion in the high layer queueing network, the token is returned to $P_2$ in zero time. Customers that arrive during the time when the corresponding token pool is empty, are forced to wait in either $E_1$ or $E_2$. The first arriving customer in the external queue, enters the queueing network as soon as a token is returned to its corresponding token pool.

Let us rearrange the low layered open queueing network into an open queueing network where the node with the longest service time is placed at the beginning of the queueing network. A customer's waiting time in either queueing network is the same [7]. Since there is no queue after the first node, the time a customer spends in the remaining nodes is the sum of the service times. In view of this, we can represent the queueing network in Figure 1 by a simpler 2 node queueing network as shown in Figure 2.

For presentation purposes we shall refer to these two nodes as the first node and the second node. $s^*$ represents the longest service time in the network and $\bar{s}$ is the sum of the remaining service times. The number of parallel servers at the second node is infinite. A customer's waiting time in the two-node queueing network is the same as in the low layered queueing network. The procedures of reducing the network dimensionality are stated for

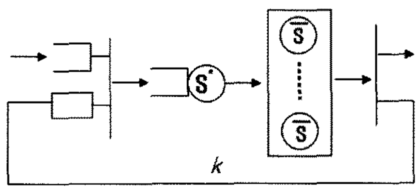


Figure 2 two node queueing network

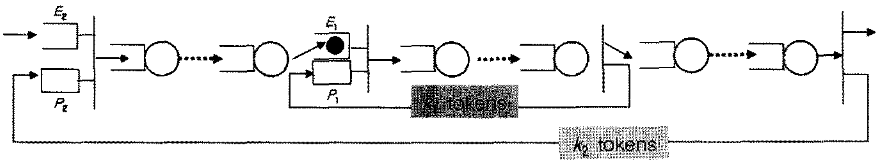


Figure 1 A nested queueing network

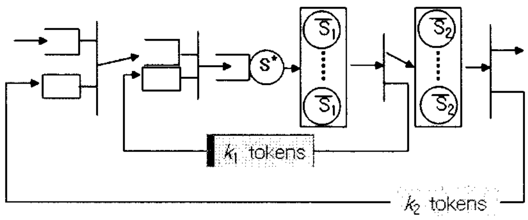


Figure 3 node queueing network

the nested queueing network under in Figure 1. By applying Proposition 1, 2 and 3 [12], a simple and equivalent queueing network can be generated as far as a customer's waiting time is concerned. Letting $\bar{s}_1$ and $\bar{s}_2$ be the sum of the remaining service time in each layer, the nested queueing network can be transformed into only 3 node queueing network as shown in Figure 3.

One of the main issues in the semaphore controlled queueing network is how to organize the number of tokens. That is, the number of tokens in the network, i.e., the size of the window, may influence where customers wait internally in the network.

However, a customer's waiting time in the network is independent of the number of tokens in the low layer network, when $k_1 s^* \geq T_1$, where $T_1$ is the sum of service times in the high layered queueing network. And, the number of tokens in the high layer queueing network has no influence on the customer's waiting time, when $k_2 s^* \geq T$, where $T$ is the sum of service time in the whole nested queueing network. Although the above described model has only two layers in the queueing network, the reduction of network dimensionality can be extended to an $n$ layered open queueing network with population constraint and constant service times. It is evident that in order to be able to tackle any number of multi-layers of window flow control we need to be able to construct a simpler queueing network.

## 3. Some basic characteristics

In this section, we present a lower and an upper bound of the mean waiting time in the queueing network assuming that $ks^* < T$, where $T = s^* + \bar{s}$, in Figure 2. And we show that these are true bounds. In the case where the token pool consists of one token, since the network allows only one customer, there is no waiting after the external queue. 2 node queueing network becomes $GI/D/1$ queue with a service time equal to $T$. The subject for the mean waiting time on $GI/D/1$ is not considered in this paper, but quoted to its previous well-known researches. If the arrival process to the queueing network is Poisson, then we obtain an $M/D/1$ queue, and the mean waiting time, $W$ is given by Khinchin-Pollaczek formula, i.e.,

$$W = \frac{\rho T}{2(1-\rho)} \tag{1}$$

where $\rho = \lambda T$ is traffic intensity. However, for the $GI/G/1$ queue, there is no exact expression available for the mean waiting time. Marshall [13] and Marchal [14] give the following bounds for the $GI/G/1$.

$$\max\left\{0, \frac{\lambda^2 \sigma_B^2 + \rho(\rho-2)}{2\lambda(1-\rho)}\right\} \leq W \leq \frac{\lambda(\sigma_A^2 + \sigma_B^2)}{2(1-\rho)} \tag{2}$$

where $\sigma_A^2$ and $\sigma_B^2$ are the variances of the interarrival distribution and service time distribution respectively. For the case of the constant service time, the variance $\sigma_B^2$ is zero. Therefore, the lower bound is always zero since $\rho < 1$. Marchal [14] obtained the following approximation for the lower bound of the mean waiting time in a $GI/G/1$ queue;

$$\frac{\lambda^2(\sigma_A^2 + \sigma_B^2)}{2\lambda(1-\rho)} - \frac{1+\rho}{2\lambda} \tag{3}$$

Further, the mean waiting time in a $GI/D/1$ queue can be approximated by the following expression due to Kramer and Langenbach-Belz [15];

$$W = \frac{\rho^2(c_a^2 + c_s^2)}{2\lambda(1-\rho)} g(c_a^2, c_s^2, \rho) \tag{4}$$

where $c_a^2$ and $c_s^2$ are the squared coefficient of variation of the interarrival time and service times respectively, and

$$g(c_a^2, c_s^2, \rho) = \begin{cases} \exp\left(-2(1-\rho)\left(\frac{(1-c_a^2)^2}{3\rho(c_a^2 + c_s^2)}\right)\right) & \text{if } c_a^2 < 1 \\ \exp\left(-(1-\rho)\frac{(1-c_a^2)^2}{(c_a^2 + 4c_s^2)}\right) & otherwise \end{cases}$$

### 3.1 A Lower Bound

Let us consider the equivalent queueing network shown in Figure 2. We note that tokens are used in the order in which they arrive at the token pool. For presentation purpose, let us number the tokens from 1 to $k$. Then, since service times are all constant, token $i$ will always be behind token $(i-1)$. In view of this, every $k_{th}$ arriving customer will use the same token. If we regard each token as a separate server, then the queueing network can be represented by $k$ queues in parallel. Each queue will consist of customers waiting to use the same token. The service time at each queue is the time it takes for a token to traverse the two nodes inside the semaphore controlled queueing network. Obviously, this service time depends on how many other tokens are being used at the same time. In other words, the service time in a queue depends on the state of the remaining $(k-1)$ queues.

A lower bound on the mean waiting time can be easily obtained by setting the service time of each of these $k$ queues equal to $T$, i.e., independent of the state of the other queues. If the arrival process to the original queueing network is a general arrival process with arrival rate $\lambda$, then the arrival process to each of the $k$ queues is the convolution of $k$ such general arrival processes each with an arrival rate $\frac{\lambda}{k}$. Thus, each queue can be analyzed as a $GI \otimes GI \otimes \cdots \otimes GI/D/1$ queue, where $GI \otimes GI \otimes \cdots \otimes GI$ is the convolution of the $k$ arrival processes, and the service time is equal to $T$. When the arrival process to the queueing network is Poisson with an arrival rate $\lambda$, the arrival process to each queue becomes an Erlang distribution with $k$ phases and an arrival rate $\lambda$ for each phase. For Poisson and non-Poisson arrivals, the mean waiting time is calculated using (2) or (4).

### 3.2 An Upper Bound

Let us consider the queueing network under study assuming that the external queue is saturated. That is, there is always at least one customer waiting in the external queue. In this case, all $k$ tokens are continuously used. Let us consider the case where $\frac{T}{k} > s^*$. Since the external queue is always saturated, sooner or later there will be no token left in the token pool. The interdeparture time from the first node is larger than or equal to $s^*$, which means that the interarrival time of a token to the pool is larger than or equal to $s^*$. Thus, a token arriving to the first node always finds the node empty. The time it takes for a token to return to the token pool is $T = s^* + \bar{s}$ and the average departure time for customer is $\frac{T}{k}$. Thus, when the external queue is saturated, the throughput is $\frac{k}{T}$. We conjecture from this, that the throughput of 2 node queueing network lies between $\frac{1}{s^*}$ and $\frac{k}{T}$. Therefore, an upper bound on the mean waiting time can be obtained by representing a $GI/D/1$ queue with a service time equal to $\max\left\{s^*, \frac{T}{k}\right\}$. The mean waiting time in this queue can be obtained using (1) if the arrival process is Poisson. For a non-Poisson arrival process, we use (2) or (4).

The above lower and upper bounds are true bounds is shown for the mean waiting time in the 2 node queueing network [11].

## 4. Numerical examples

The lower bound and upper bound given in section 3, were checked by comparing them with the mean waiting time in 3 node nested queueing network.

A lower bound on the mean waiting time can be obtained by removing the level 1 semaphore queue. One way of removing the level semaphore queue is to set $k_1$ to $k_2$. Then, the customer never experiences queueing in the internal queue $E_1$. Let $\bar{s} = T - s^*$, the lower bound on the mean waiting time is shown in Figure 4.

Obviously an upper bound on the mean waiting time also can be obtained by removing the level

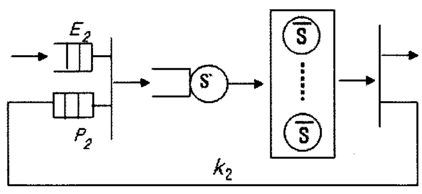


Figure 4 A lower bound of the nested queueing network

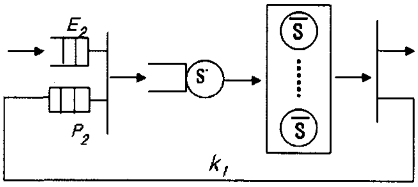


Figure 5 An upper bound of the nested queueing network

semaphore queue. If we set $k_2$ to $k_1$, then the queueing network provides an upper bound on the mean waiting time. This is because customers never experience any queueing in the internal queue $E_1$ in the 3 node queueing network. the upper bound on the mean waiting time is shown in Figure 5.

Simulation estimates were obtained assuming Poisson arrivals and phase-type arrival process such as Erlang and hyper-exponential distributions. This is because the coefficient of variation can be controlled variously in the phase-type arrival process. And it is well known that the mean waiting time in a queue depends on the coefficient of variation its arrival process.

For each example below, the table 1 through 3 give the lower and upper bounds and the simulated mean waiting time as a function of $k_2$, the number of tokens in the level 2 layer, assuming that $k_1$ is fixed. And, table 4 through 6 show the lower and upper bounds and the simulated mean waiting time as a function of $k_1$, the number of tokens in the level 1 layer, but $k_2$ is fixed($k_1 < k_2$). In particular, it gives a) the simulated mean waiting time, and b) the upper and lower bound by simulating the lower and upper bound queueing models. A lower bound is chosen from the maximum value of (4). And, an upper bound is chosen from the minimum value of (4) for the better tightness.

As can be seen the above table 1 to 3 when $k_1$ is fixed, the simulated mean waiting time is close to the upper bound of the mean waiting time in the beginning, and later it tends to approach to its lower bound. This is because, the token in the first

Table 1 Poisson ($\lambda = \frac{1}{7}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_1 = 5$

| number of tokens($k_2$) | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|
| upper bound | 13.82 | 10.85 | 10.85 | 10.85 | 10.85 |
| simulated mean waiting time | 11.13 | 3.31 | 1.47 | 0.94 | 0.77 |
| lower bound | 10.85 | 3.16 | 1.37 | 0.94 | 0.77 |

Table 2 Erlang 2 ($\lambda = \frac{1}{3.5}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_1 = 5$

| number of tokens($k_2$) | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|
| upper bound | 6.43 | 4.11 | 4.11 | 4.11 | 4.11 |
| simulated mean waiting time | 4.37 | 0.98 | 0.35 | 0.35 | 0.35 |
| lower bound | 4.11 | 0.93 | 0.35 | 0.35 | 0.35 |

Table 3 Hyper ($p_1 = \frac{1}{3}, p_2 = \frac{2}{3}, \lambda_1 = \frac{1}{15}, \lambda_2 = \frac{1}{3}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_1 = 5$, $c_a^2 = 2.31$

| number of tokens($k_2$) | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|
| upper bound | 32.48 | 11.4 | 11.4 | 11.4 | 11.4 |
| simulated mean waiting time | 28.87 | 9.42 | 4.93 | 2.94 | 2.21 |
| lower bound | 11.4 | 4.25 | 3.25 | 2.59 | 1.87 |

Table 4 Poisson ($\lambda = \frac{1}{7}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_2 = 7$

| number of tokens($k_1$) | 4 | 5 | 6 | 7 |
|---|---|---|---|---|
| upper bound | ∞ | 10.85 | 3.16 | 1.48 |
| simulated mean waiting time | 1.81 | 1.47 | 1.47 | 1.47 |
| lower bound | 1.71 | 1.37 | 0.70 | 0.70 |

Table 5 Erlang 2 ($\lambda = \frac{1}{3.5}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_2 = 7$

| number of tokens($k_1$) | 4 | 5 | 6 | 7 |
|---|---|---|---|---|
| upper bound | ∞ | 4.11 | 0.93 | 0.36 |
| simulated mean waiting time | 0.45 | 0.35 | 0.35 | 0.35 |
| lower bound | 0.36 | 0.35 | 0.20 | 0.20 |

Table 6 Hyper ($p_1 = \frac{1}{3}, p_2 = \frac{2}{3}, \lambda_1 = \frac{1}{15}, \lambda_2 = \frac{1}{3}, s^* = 2.5, \overline{s_1} = 12.5, \overline{s_2} = 14$) when $k_2 = 7$, $c_a^2 = 2.31$

| number of tokens($k_1$) | 4 | 5 | 6 | 7 |
|---|---|---|---|---|
| upper bound | ∞ | 11.4 | 6.30 | 4.25 |
| simulated mean waiting time | 5.30 | 4.93 | 4.68 | 4.25 |
| lower bound | 3.70 | 3.25 | 1.45 | 1.45 |

layered queueing network has influenced much for the flow dynamics. However, the simulated mean waiting time is close to the lower bound of the mean waiting time in the beginning, and later it tends to approach to its upper bound when $k_2$ is fixed. This result is similar to that of 2 node queueing network.

## 5. Conclusions

In this paper, we analyzed the mean waiting time on the nested open queueing network, where the population within each subnetwork is controlled by a semaphore queue. The queueing network can be transformed into a simpler queueing network. A major characteristic of this model is that the lower layer flow is halted by the state of higher layer. The lower bound and upper bound were checked by comparing them with the mean waiting time in 3 node nested queueing network. Simulation estimates were obtained assuming Poisson arrivals and other phase-type arrival process, i.e., Erlang and hyper-exponential distributions. The bounds obtained can be applied to get more exact approximation using the suitable approach.

An important problem that is yet to be considered, is the characterization of the departure process from the external queue, which is the arrival process to the first node. Once we provide the characterization of the arrival process to the first node, then the restriction of the semaphore queue can be relaxed. Also it would be interesting to further study the queueing network to obtain performance characteristics such as approximation of the mean waiting time and the variance of the interdeparture time.

## References

[1] M. Reiser, A queueing network analysis of computer communication networks with window flow control, *IEEE Trans. on Comm.*, 27, pp.1199-1209, 1979.

[2] M. Pennotti, M. Schwartz, Congestion control in store and forward tandem link, *IEEE Trans. on Comm.*, 23, pp.1434-1443, 1975.

[3] M, Schwartz, Performance analysis of the SNA virtual routing pacing control, *IEEE Trans. on Comm.*, 30, pp.172-184, 1982.

[4] A. Thomasian, P. Bay, Analysis of Queueing network models with population size constraint and delayed blocked customers, *ACM SIGMETRICS conference*, pp.202-216, 1984.

[5] G. Varghee, W. Chou and A. Nilsson, Queueing delays on virtual circuits using a sliding window flow control scheme, *ACM SIGMETRICS conf. Minneapolis*, pp.275-281, 1983.

[6] O. Gihr, P. Kuehn, Comparison of communication services with connection-oriented and connectionless data transmission, *Processing of the International Seminar on Computer Networking and Performance Evaluation*, Tokyo Japan, pp.431-444, Sep. 1985.

[7] Y. Rhee and H. Perros, Analysis of an Open Tandem Queueing Network with Population Constraint and Constant Service Times, *European J. Of Operations Research*, vol.92, pp.99-111, 1996.

[8] L. Mitchell, and D. Lide, End-to-end performance modeling of local area networks, *IEEE J. Select. Comm.*, vol.SAC-4, pp.305-320, 1986.

[9] S. Fdida, H. Perros, A. Wilk, Semaphore Queue; modeling multilayered window flow control mechanism, *IEEE Trans. on Comm.*, 83(3), pp.309-317, 1990.

[10] G.W. Shapiro and H.G. Perros, Nested sliding

window protocols with packet fragmentation, *IEEE Trans. Comm.*, vol.41, no.1, pp.99-109, 1993.
[11] Y. Rhee and H. Perros, On the Mean Waiting Time of a Population Constrained Open Tandem Queueing Network with Constant Service Times, *IIE Trans. on Design and Manufacturing*, vol.30-10, pp.973-979, 1998.
[12] Y. Rhee, Some notes on the reduction of the network dimensionality in the nested open queueing network, *European J. Of Operations Research*, vol. 174, pp.124-131, 2006.
[13] K.T. Marshall, Some inequalities in queueing, *Operations Research*, vol.16, pp.651-665, 1968.
[14] W.G. Marchal, Some simpler bounds on the mean queueing time, *Operations Research*, vol.26, no.6, pp.1083-1088, 1978.
[15] Kramer and Langenbach-Belz, Approximate formulae for the delay in the queueing system GI/G/1. *8 Int. Tele. Con.* Melbourne, 235/1-8, 1976.

이 영
1983년 고려대학교 산업공학과(공학사) 1885년 고려대학교 산업공학과(공학석사) 1990년 University of Oklahoma 산업공학과(MS). 1994년 North Carolina State University OR/CS(Ph.D). 1994년 삼성 SDS 정보기술연구소 수석연구원 1998년~현재 계명대학교 산업시스템공학과 교수. 관심분야는 Network 모델링과 Performance 분석, 웹 서버 capacity planning