The Performance Analysis of SIP-T Signaling System in Carrier Class VoIP Network

Jung-Shyr Wu * and Peir-Yuan Wang **
* Graduate Institute of Communication Engineering, NCU
**Department of Electrical Engineering, NCU
**Training Institute, Chunghwa Telecom Co., Ltd.

Abstract

This paper presents the performance modeling, analysis, and simulation of SIP-T (Session Initiation Protocol for Telephones) signaling system in carrier class VoIP(Voice over IP) network. The SIP-T signaling system defined in IETF (Internet Engineering Task Force) draft is a mechanism that uses SIP (Session Initiation Protocol) to facilitate the interconnection of PSTN with carrier class VoIP network. Based on IETF, the SIP-T signaling system not only promises scalability, flexibility, and interoperability with PSTN but also provides call control function of MGC (Media Gateway Controller) to set up, tear down, and manage VoIP calls in carrier class VoIP network. In this paper, we analyze the queueing size (i.e. buffer size), the mean of queueing delay, and the variance of queueing delay of SIP-T signaling system that are the major performance evaluation parameters for improving QoS (Quality of Service) and system performance of MGC in carrier class VoIP network focused on toll by-pass or tandem by-pass of PSTN. First, we assume a mathematical model of the M/G/1 queue with non-preemptive priority assignment to represent SIP-T signaling system. Second, we present the formulas of queueing size, queueing delay, and delay variation for the non-preemptive priority queue by queueing theory respectively. Besides, some numerical examples of queueing size, queueing delay, and delay variation are presented as well. Finally, the theoretical estimates are shown to be in excellent consistence with simulation results.

Key Words: SIP-T Signaling System, Carrier Class VoIP Network, Media Gateway Controller, Quality of Service, Performance.

1. Introduction

Until recently, one of the greatest challenges in the migration from existing PSTN (Public Switched Telephone Network) toward NGN (Next Generation Networks) is to build a carrier class VoIP (Voice over IP) network that preserves the ubiquity, quality, and reliability of PSTN services while allowing the greatest flexibility for use of new VoIP technology [1][2]. It shows the emerging carrier class VoIP technology will still dominate NGN in a few years. Carrier class packet telephony network will provide toll by-pass (i.e. Class 4 switch replacement) function of traditional PSTN. Furthermore, the replacement function of Class 5 switch (i.e. all IP solution) by the packet telephony technology will come true in the near future. Carrier class VoIP other than internet telephony and ISP (Internet Service Provider) class VoIP technology supports QoS (Quality of Service), security, and management functions which will present a tremendous opportunity to network/service providers of the world to offer both traditional services as well as a range of creative new services in the near future. However, there exist substantial challenge to be faced and overcome in supporting system performance, capability, reliability, and resource management that are significant for carrier class VoIP technology.

This paper presents the performance analysis and simulation of SIP-T (Session Initiation Protocol for Telephones) signaling system in carrier class VoIP network. The SIP-T signaling system defined in IETF draft is a mechanism that uses SIP (Session Initiation Protocol) to facilitate the interconnection of PSTN with carrier class VoIP network [13][16]. Based on IETF, the carrier class VoIP network includes three major system elements that are MGC (Media Gateway Controller), SG (Signaling Gateway), and MG (Media Gateway) respectively [9]-[11]. The system architecture of carrier class VoIP network shown in Figure 1 in this paper is focused on the toll by-pass of PSTN [12][14][16].

MGC, often referred to as the Softswitch or CA (Call Agent), is responsible for call control, service control, routing, charging, and other functions. SIP-T is the major inter-MGC signaling protocol used to seamlessly interwork with the SS7 (Signaling System No.7) signaling protocol of PSTN for a VoIP call [15][17]. For the future development of inter-MGC signaling protocol, SIP-T will be the major trend that not only promises scalability, flexibility, and interoperability with PSTN but also provides call control function of MGC to set up, tear down, and manage VoIP calls in carrier class VoIP network. MG is used to transform TDM (Time Division

Proceedings of the17 th International Conference on Advanced Information Networking and Applications (AINA'03)

Multiplexing) based bearer traffic into IP packets, which will be transported over the IP core network and vice versa. SG is used to interconnect SS7 signaling network with MGC.

The SIP-T is a request-response protocol for initiating, maintaining, and terminating multimedia sessions. A VoIP call is considered as a kind of multimedia session in which voice is exchanged between two MGCs. Each SIP-T request is followed by one or more provisional responses, followed by one or more definitive responses. They are most commonly used for response to an INVITE request and provide information on call progress, such as trying (100), alerting (180), queueing (182), and session progress (183). SIP-T can run over UDP (User Datagram Protocol), TCP (Transmission Control Protocol), or SCTP (Stream Control Transmission Protocol) [9], but the message format is independent of the transport protocol. Therefore, the queueing model for SIP-T signaling system shall take account into the following features including (1)priority queue to meet system QoS and performance requirements, (2)SIP-T message with lower or higher priority, (3)fill-in message or synchronization message such as hello message for keeping alive purpose, (4)preemptive or non-preemptive queueing process. In this paper, non-preemptive priority queue is proposed for SIP-T signaling system.

Obviously, the performance analysis of SIP-T signaling system will play an essential role in optimizing network QoS provisioning. We will analyze the queueing size, the mean of queueing delay, and the variance of queueing delay that are the major performance evaluation parameters for improving QoS and system performance of MGC in carrier class VoIP network. The M/G/1 queue with non-preemptive priority assignment will be used to represent SIP-T signaling system that constitutes a control signaling channel for VoIP calls between inter-MGC communication. The non-preemptive priority assignment means the signaling message of SIP-T with unique priority communicating between two MGCs in carrier class VoIP network. We analyze the queueing size using imbedded Markov chain and Semi-Markov process and analyze the queueing delay and delay variation using LST (Laplace-Stieltjes Transform) from the distribution of queueing delay for a non-preemptive priority queue [3]-[8].

Besides, we present the numerical examples by simulation to show the accuracy of the performance analysis. The numerical examples show the excellent consistence for the comparison between theoretical estimates and simulation results. We will discuss how the results are useful for the system performance of MGC in carrier class VoIP network. Consequently, we can evaluate whether the ratio of cost to performance can meet the requirement of planning and design for carrier class VoIP network.

2. Performance Model

SIP-T is an application layer protocol that constitutes a control signaling channel for VoIP calls between inter-MGC communication. The queueing model proposed in this section will be illustrated under the following assumptions to model and analyze the SIP-T signaling system.

(1) M/G/1 queueing model: M means the Poisson distribution for the arrival process of SIP-T message (i.e.The inter-arrival time of SIP-T message is of exponential distribution), G stands for the general service distribution for service time of message, 1 stands for only one server in the queueing model.

(2) Nodal model for inter-MGC communication is analyzed for the reason of simplicity and cost-efficiency. Network model is not proposed in this paper since it shows too complicated to analyze SIP-T signaling system by queueing theory.

(3) Priority queue: Messages over the control signaling channel include SIP-T messages with arrival rate r and fill-in messages. Once the channel does not transport SIP-T messages, fill-in messages shall be inserted into the channel periodically or continually. H_1 (x) is the service time distribution (Cumulative Distribution Function-CDF) of SIP-T message and H_0 (x) is the service time distribution (CDF) of fill-in message. Denote $H_1(x) = \Pr(T_1 \le x)$, $H_0(x) = \Pr(T_0 \le x)$, where T_1 and T_0 are the service time of SIP-T message and fill-in message respectively.

(4) Service discipline adopts non-preemptive priority queue: Any new arrival with higher priority message is not allowed to transport over the channel until the end of current lower priority message being transported.

The queueing process of SIP-T signaling system in carrier class VoIP network is summarized in figure 2, which also exhibits non-preemptive priority queue.

3. Performance Analysis

In this section, we briefly analyze the derivation results of the queueing size (i.e. buffer size) using imbedded Markov chain and Semi-Markov process and analyze the queueing delay and delay variation using LST from the distribution of queueing delay for a non-preemptive priority queue [3]-[8] owing to paper length constraint. Some definition and notation are introduced as follows:

(1) $h_r(k)$: the k-th moment of $H_r(\cdot)$, where k = 1,2,….N, r = 0,1.Here, $h_1(1) = h_1$ is the mean service time of SIP-T message when k = 1 and r = 1; $h_0(1) = h_0$ is the

mean service time of fill-in message when k = 1 and r = 0.

(2) $a = r\,h_1$: the mean traffic intensity of SIP-T message.

3.1.Queueing Size Derivation using Imbedded Markov Chain

Queueing size is defined as the number of SIP-T messages in the system. M/M/1 has the memory less property which also holds Markov property. So, the number of SIP-T messages in the system can be observed at any time. Unlike M/M/1, M/G/1 does not have the memory less property which shows the difficulty if we try to analyze the message numbers in the system at any time. Imbedded Markov chain is used to solve for the observation time changed from arbitrary instant to departure time of messages, which means we can observe the number of messages in the system just after the completion of message served for M/G/1 queueing model.

Let S(t) be the number of SIP-T messages in system at arbitrary instant t and let V_i be the i-th observation time just after the completion of message served, where i=1,2,…N. Then, $S_i = S(V_i)$ represents the number of SIP-T messages in the system for the i-th observation time just after the completion of message served.

Let M be the average number of SIP-T messages in the system (i.e. queueing size at departure instant) for the queueing model. Then, M can be derived as (1) from this queueing model.

$$M=\frac{(1-a)(r^2h_0^{(2)}+2rh_0)+r^3h_0h_1^{(2)}}{2(1-a)(rh_0+1-a)} \qquad (1)$$

3.2.Queueing Size Derivation using Semi-Markov Process

In section 3.1, the observation time is constrained at departure instant of message. In this section, Semi-Markov process is used to solve for the observation time changed from departure instant of message to arbitrary instant which means we can observe the number of messages in the system at any time for M/G/1 queueing model.

Let N be the average number of SIP-T messages in the system (i.e. queueing size at arbitrary instant) for the queueing model. Then, N can be derived as (2) from this queueing model.

$$N=a+\frac{rh_0}{2}+\frac{r^2h_1^{(2)}}{2(1-a)} \qquad (2)$$

3.3.Derivation of Queueing Delay and Delay Variation

In this section, we show the derivation result of the mean queueing delay and the delay variation using LST from the distribution of queueing delay for the queueing model.

Let $W(x)$ be the queueing delay distribution of SIP-T messages. Equation (3) holds in equilibrium state which means the probability with n SIP-T messages under the constraint of at least one SIP-T message in the system.

$$\frac{\pi_n}{1-\pi_0}=\int_0^\infty \frac{(rx)^{n-1}e^{-rx}}{(n-1)!}dW(x),\, n\ge 1 \qquad (3)$$

Where $\lim_{i\to\infty}\Pr[S_i=n]=\pi_n$

Let W_q be the mean queueing delay of SIP-T messages and V_q^2 be the variance of queueing delay of SIP-T message. Then W_q and V_q^2 can be derived as (4) and (5) from this queueing model respectively.

$$W_q=\frac{h_0}{2}+\frac{rh_1^{(2)}}{2(1-a)} \qquad (4)$$

$$V_q^2=\frac{h_0^2}{12}+\frac{1}{12(1-a)^2}[4rh_1^{(3)}(1-a)+3r^2(h_1^{(2)})^2] \qquad (5)$$

4. Numerical Analysis

In this section, we present some numerical examples by simulation in order to evaluate the accuracy of the performance analysis shown in section III. The traffic model of SIP-T message is proposed under the assumption using embedded method for encapsulation of SS7 ISUP. The parameters of proposed traffic model of SIP-T message, which excludes the header and other SDP without ISUP within SIP-T message, is shown in table 1. It shows the distribution of average length of SIP-T message obtained from a VoIP call.

The error control scheme of SIP-T message uses SR (Selective Retransmission) method instead of GBN (Go-Back-N) method because of the advantage of the channel bandwidth efficiency. The comparison between mathematical and simulation results is shown the consistence including queueing size (i.e. buffer size), queueing delay, and delay variation for the non-preemptive priority queue.

4.1. Numerical Examples

In this section, we give some numerical examples in order to compare the mathematical and simulation results. By computer simulation, we observe the message numbers of one process (SIP-T message) every 100 seconds which hold the relationship of Poisson input for the queueing model. Then, we analyze the mathematical and simulation results for the queueing size (i.e. buffer size), mean queueing delay and standard deviation (i.e. delay variance) of queueing delay respectively. The error probability of SIP-T message is denoted by P_u. In this paper, three simulation cases of SIP-T message for P_u = 0.004 are presented respectively.

Table 2 shows the comparison data between mathematical and simulation results including buffer size (i.e. buffer size), mean queueing delay and standard deviation of queueing delay for Pu =0.004. Figure 3 shows the comparison curve of queueing size (i.e. buffer size) between mathematical and simulation results along the different message arrival rates respectively. Figure 4 shows the comparison curve of mean queueing delay between mathematical and simulation results along the different message arrival rates respectively. Figure 5 shows the comparison curve of standard deviation of queueing delay between mathematical and simulation results along the different message arrival rates respectively. From these figures, we can see that the difference between mathematical and simulation results almost remains the same and small whenever the message arrival rate varies. The curve of queueing size (i.e. buffer size), mean queueing delay and standard deviation of queueing delay vary slowly as arrival rate of SIP-T message less than 450. However, the curve of queueing size (i.e. buffer size), mean queueing delay and standard deviation of queueing delay increase abruptly as arrival rate of SIP-T message greater than 450. An intuitive and reasonable explanation for this phenomenon is that the SIP-T message arrival rate approaches the processing capability of system for heavy traffic intensity. The numerical examples show the excellent consistence for the comparison between theoretical estimates and simulation results.

4.2. Performance Evaluation

Based on the numerical examples, we discuss how these results are useful for carrier class VoIP network. From Figure 3 to 5, we can find out the value of 450 messages/sec, arrival rate of SIP-T message, stands for the system performance threshold for each SIP-T signaling channel (i.e. signaling link) in carrier class VoIP network. The system performance will be supported and guaranteed while arrival rate of SIP-T message is less than the threshold. Otherwise, the system performance may not be supported and guaranteed while arrival rate of SIP-T message is greater than the threshold. Therefore, the results can be used to evaluate the system performance of MGC in carrier class VoIP network. The call processing capability for each SIP-T signaling channel can be obtained to present how many SIP-T calls per hour are generated. Then, we not only decide how many SIP-T signaling channels are required for each MGC but also decide how many MGCs are required for whole network deployment. By the same way, we can obtain how many VoIP sessions handled by each SIP-T signaling channel in carrier class VoIP network. So, we decide how many VoIP sessions can be handled by each of MGC for whole network deployment.

We assume the average 3 of SIP-T messages per call and per direction to control the call processing between inter-MGC communications in carrier class VoIP network. Then, the call processing capability for each of SIP-T signaling channel is 150 average SIP-T calls per second. Therefore, the call processing capability for each of SIP-T signaling channel is 540000 average SIP-T calls per hour. If the requirement of call processing capability for each of MGC is designed to handle 1000000 BHCA (Busy Hour Call Attempt), two SIP-T signaling channels for each of MGC are required. If the requirement of call processing capability for the carrier class VoIP network is designed to handle 3000000 BHCA, three MGCs are then required for the carrier class VoIP network. Besides, we can decide the maximum of VoIP sessions per hour handled by each of SIP-T signaling channel and also decide the maximum of VoIP sessions per hour handled by each of MGC with two SIP-T signaling channels for the carrier class VoIP network deployment. The performance evaluation shows a very reasonable resource provisioning with load sharing and redundancy in carrier class VoIP network.

The network scalability and network topology can also be decided for the carrier class VoIP network, which not only provides system performance guarantees but also provides cost-efficiency. Consequently, we can evaluate whether the ratio of cost to performance can meet the requirement of planning and design for carrier class VoIP network.

5. Conclusion

In this paper, a non-preemptive priority queue is proposed and analyzed for the performance evaluation of SIP-T signaling system in carrier class VoIP network. We present the derivation formulas of the queueing size, the mean queueing delay, and the variance of queueing delay that are the major performance evaluation parameters for improving QoS and system performance of MGC in carrier class VoIP network. After the comparison between mathematical and simulation results, the theoretical performance analysis is shown the accuracy from Figure 3 to 5 including queueing size, queueing delay, and delay variation for the queueing model. The theoretical performance analysis of SIP-T signaling system shows the robustness regardless of the effect of traffic intensity and error probability of SIP-T message for the non-preemptive priority queue. Then, we can determine how much the ratio of cost to performance can tolerate, how the planning and design is needed so as to meet the requirements of carrier class VoIP network.

6.Acknowledgements

The authors wish to thank for the financial support under the grant NSC 91-2219-E008-002 and MOTC-STAO-91-24 respectively.

7.References

[1] Abdi R. Modarressi and Seshadri Mohan "Control and Management in Next–Generation Networks: Challenges and Opportunities", IEEE Comm., pp.94-102, Oct. 2000.

[2] Paul G. A. Sijben, John P. L. Segers, Louise F. A. Spergel, and J. Kozik, "Building the Bridge: Devising an Architecture to Migrate Voice-Band Calls to Packet Transport and Multimedia Services", Bell Labs Tech. J., pp. 166-185, July-Sep. 2000.

[3] R. B. Cooper, "Queues Served in Cyclic Order: Waiting Times", The Bell System Technical Journal, Vol.49, No.3, pp.399-413, Mar. 1970.

[4] M. Scholl and L. Kleinrock, "On the M/G/1 Queue with Rest Periods and Certain Service-Independent Queueing Disciplines", Operations Research, Vol.31, No.4, pp.705-719, 1983.

[5] H. Takagi, "Queueing Analysis: A Foundation of Performance Evaluation, Volume 1", North-Holland, Amsterdam, 1991.

[6] D. Gross and C. M. Harris, " Fundamentals of Queueing Theory", 3 rd. Edition, John Wiley & Sons Inc., 1998.

[7] L. Kleinrock, " Queueing System Volume 1: Theory", John Wiley & Sons Inc., 1974.

[8] A. J. Fabens, " The Solution of Queueing and Inventory Models by Semi-Markov Process", J. Roy. Statis. SOC., Ser. B, 23, pp. 113-127, 1961.

[9] L. Ong et al., " Framework Architecture for Signaling Transport", IETF RFC 2719, Oct.1999.

[10] F. Cuervo et al., " Megaco Protocol Version 1.0", IETF RFC3015, Nov. 2000.

[11] I. Elliott, C. Huitema, and S. Pickett , "Media Gateway Control Protocol (MGCP) Version 1.0", IETF RFC 2705, Oct. 1999.

[12] ITU-T, "Bearer Independent Call Control Protocol", ITU-T Recommendation Q.1901, June 2000.

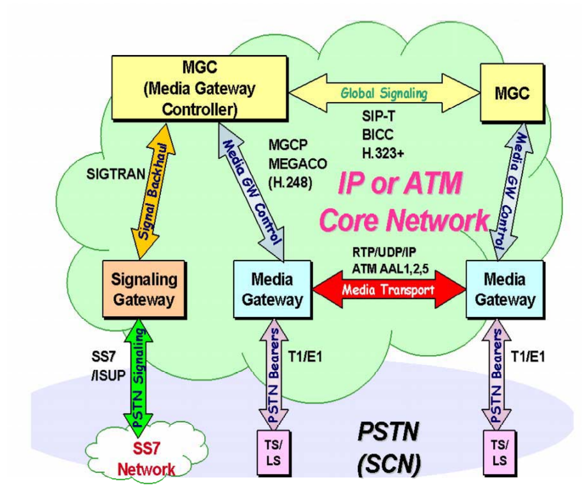
[13] M. Handley, H.Schulzrinne, E.Schooler, and J.Rosenberg, "SIP: Session Initiation Protocol", IETF RFC 2543, Mar.1999.

[14] ITU-T, "Packet Based Multimedia Communication Systems", ITU-T Recommendation H.323, Version 3, Geneva, Switzerland, Sep. 1999.

[15] ITU-T, "Signaling System No. 7 - ISDN User Part Functional Description", ITU-T Recommendation Q.761- Q.764, Sep. 1997.

[16] A. Vemuri and J. Peterson, " SIP for Telephones (SIP-T): Context and Architectures", < draft-vemuri-sip-t-context-02.txt >, IETF Internet Draft, Feb. 2001, Work in Progress

[17] G. Camarillo et al., " ISUP to SIP Mapping, "<draft-ietf-sip-isup-01.txt>, IETF Internet Draft, May 2001, Work in Progress



TS: Toll Switch
LS: Local Switch
SCN: Switched Circuit Network

Figure 1. The system architecture of carrier class VoIP network

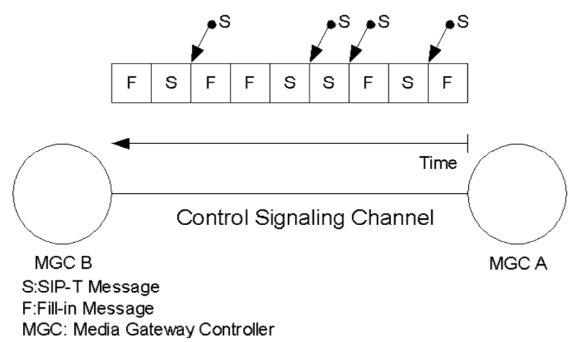


Figure 2. Queueing process of SIP-T signaling system

Table 1. Parameters of proposed traffic model of SIP-T message

Message Length (Bits)	168	120	112	Total
Number of Message per call	1	1	4.9	6.9
Percentage (%)	14.5	14.5	71	100
Average Message Length (Bits)	121.28			

Table 2. Comparison between mathematical and simulation results including buffer size, mean queueing delay and standard deviation for the non-preemptive priority queue (P_U=0.004)

Mathematical				Simulated				
SIP-T message (messages/s)	Mean (ms)	Std. Dev. (ms)	Buffer Size	SIP-T message (messages/s)	Mean and 95th percent (ms)	Std. Dev. (ms)	Buffer Size	Sample Size(SIP-T message)
50	0.49	0.49	0.12	59.2	0.49±0.01	0.47	0.12	4925
100	0.62	0.68	0.23	100.7	0.61±0.01	0.65	0.25	10074
150	0.78	0.89	0.37	150.4	0.79±0.01	0.88	0.39	15044
200	1.00	1.14	0.52	201.8	0.98±0.02	1.10	0.54	20178
250	1.29	1.46	0.72	251.4	1.27±0.02	1.40	0.76	25138
300	1.71	1.90	0.97	301.1	1.70±0.02	1.91	1.03	30115
350	2.37	2.57	1.33	348.9	2.36±0.03	2.57	1.40	34891
400	3.57	3.75	1.92	399.5	3.62±0.04	3.98	1.97	39955
450	6.39	6.41	3.17	450.7	6.49±0.06	6.67	3.49	45067
500	25.00	18.57	8.20	502.5	21.08±0.18	20.12	8.85	50251

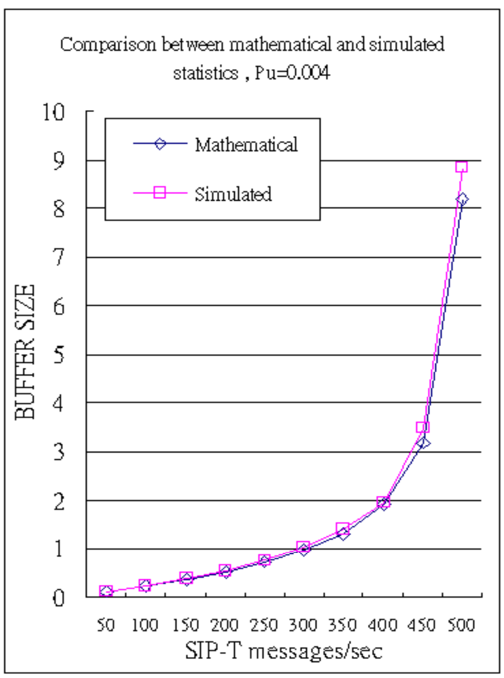


Figure 3. Comparison of buffer size for the non-preemptive priority queue (P_U=0.004)

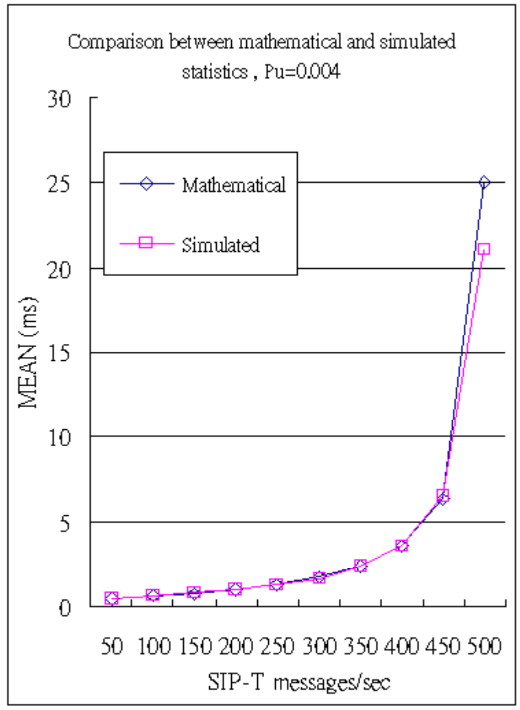


Figure 4. Comparison of mean queueing delay for the non-preemptive priority queue (P_U=0.004)

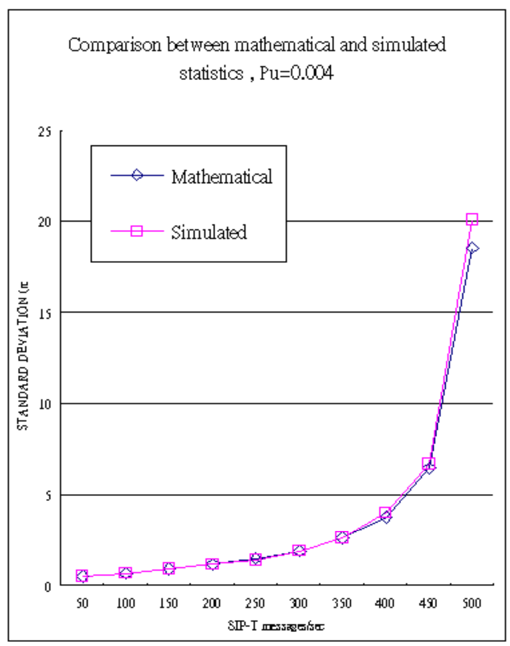


Figure 5. Comparison of standard deviation of queueing delay for the non-preemptive priority queue (P_U=0.004)

Proceedings of the17 th International Conference on Advanced Information Networking and Applications (AINA'03)
0-7695-1906-7/03 $17.00 © 2003 **IEEE**