

2종류의 heterogeneous 링크로 구성된 M/M/L 큐잉 시스템에서의  
지연시간 해석에 관한 연구

김 명 석      유 인 태      박 민 용      이 상 배  
연 세 대 학 교      전 자 공 학 과

Evaluation and Reduction of the Delay in a M/M/L Queueing  
System with 2 Class Heterogeneous Servers

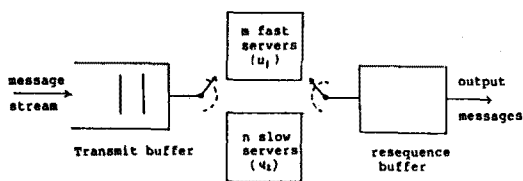
Kim, Myeong Seok    Ryoo, In Tae    Park, Mignon    Lee, Sang Bae  
Dept. of Electronic Eng, Yonsei Univ.

ABSTRACT

The resequence delay is an important performance measure in a store-and-forward computer network where messages go out of order while traversing logical channel. An analysis of the total delay including resequence delay is presented for M/M/L queueing system with 2 class heterogeneous servers and improvement of delay performance is investigated by simulation.

1. 개 요

근대 컴퓨터 통신망의 구조는 메시지를 전송하는 방법에 있어 virtual circuit를 빈번히 사용하며, 이 경우 통신 노드쌍들은 서로 다른 다중 링크로 구성된 토지컬 채널로 상호 연결되어 있다.<sup>[1]</sup> 이러한 통신망에서는 신뢰도 및 효율을 높이기 위하여, 메시지를 전송할 때 대역폭이 서로 다른 링크를 이용하여 전송된다. 이처럼 상호 다른 대역폭을 갖는 링크를 통하여 전송된 크기가 다른 메시지는 수신측 노드에서 송신측 노드에서 sequence된 순서대로 재조립 되어야 한다. 본 연구에서는 2 가지 서비스 속도로 서비스하는 heterogeneous 링크로 구성된 큐잉 시스템을 해석하고, 지연 시간을 감소하는 방법을 시뮬레이션을 통하여 고찰한다. 그림1에서는 2종류 heterogeneous server 시스템을 나타내었다.



(그림1) 2종류 heterogeneous 서버 시스템

2. M/M/L 큐잉 시스템의 모델링 및 해석

Logical 채널이 2가지 종류의 heterogeneous 링크 (총 L개)로 구성된 M/M/L 큐잉 시스템이 서비스 속도가  $u_1$ 인 m개의 fast 서버와  $u_2$ 인 n 개의 slow 서버로 구성되어 있다고 하자. 소스 노드에 도착하는 메시지는 평균도착을  $\lambda$ 를 갖는 포아손 프로세스이며, 메시지의 길이는 지수분포를 갖는다고 가정하였다.

A. 상태 확률 평형 방정식의 표현

상태 벡터를  $x = (i, j)$ 라 할때, i는 fast 서버를 이용하는 customer의 수, j를 slow 서버를 이용하는 customer의 수라 하자. 큐내에서 대기하는 customer의 수를 N이라 하고, 버퍼 크기를 B라 하고 평형 상태 확률  $P(N; x)$ 는

$$P(N; x) = \text{Prob.} \{N=n; (i, j)\}$$

로 표시하면, 상태 확률 평형 방정식의 표현은 다음과 같이 표현된다.

$$(i) \ N = 0(i+j < L) \text{인 경우}$$

$$P(0; x) = p(i, j) \text{라 표현하면,}$$

$$\begin{aligned} (\lambda + iu_1 + ju_2)p(i, j) &= \lambda p(i-1, j) \\ &+ (i+1)u_1 p(i+1, j) \\ &+ (j+1)u_2 p(i, j+1) \end{aligned} \quad (1)$$

$$(ii) \ N \geq 0(i+j = L) \text{인 경우}$$

$$P(N; (m, n)) = P(N) \text{로 표시하면,}$$

$$P(N) = (mu_1 + nu_2)P(N+1) \quad (2)$$

가 되며 utility  $\rho$ 를

$$\rho = \frac{\lambda}{mu_1 + nu_2}$$

라 하면,  $P(N) = \rho^N P(0)$

로 표현된다.

버퍼 크기 B가 유한인 경우는 해가 항상 존재하며, 무한한 경우 해가 존재할 조건은  $\rho < 1$  이어야 한다.

B. M/M/L 큐잉 시스템의 해석

의 에서 표현한 상태확률 평형 방정식으로 부터 상태 확률에 관한 1차 연립 방정식을 얻을 수 있다. 또한 보존정리 (모든 상태 확률의 합은 1)로 부터

$$\sum_{\substack{N=0 \\ \text{모든 상태}}} P(N; x) = \sum_{i+j < L} P(0; x) + \sum_{N=0}^B P(N) \quad (3)$$

$$= \sum_{i+j < L} p(i, j) + \frac{1-\rho^{B+1}}{1-\rho} P(0) = 1$$

따라서

$$P(0) = -\frac{1-\rho}{1-\rho^{B+1}} [1 - \sum_{m+n < L} p(i, j)] \quad (4)$$

이 때, 평균 큐의 크기는

$$E[n] = \sum_{N=0}^B NP(N)$$

$$= P(0) \rho \left[ \frac{1-(B+1)\rho^B + B\rho^{B+1}}{(1-\rho)^2} \right] \quad (5)$$

이며, 평균 큐잉 시간은

$$E[T_q] = \frac{E[n]}{\lambda(1-P_b)} \quad (6)$$

로 표시 된다.

여기서  $P_b$  는 브로킹 확률 이다.

또한 Little's formular 에 의해 서비스 시간은

$$E[T_s] = \frac{L_s}{\lambda(1-P_b)} \quad (7)$$

로 부터 구할 수 있다.

여기서  $L_s$  은

$$L_s = \sum_{i+j < L} (i+j)p(i, j) + (m+n) \sum_{N=0}^B P(N) \quad (8)$$

로 표현 된다. [2]

C. 재조립 시간 지연 해석

재조립 시간지연은 대역폭이 다른 heterogeneous 링크를 통하여 서비스를 받는 variable 메세지 길이에 기인한다. 여기서는 sequence된 메세지의 수가 링크의 수보다 충분히 크다고 가정하였다. 즉, 재조립 시간 지연은 어떤 메세지가 전송되는 동안 bypass하는 모든 메세지를 기다리는 시간으로 정의하였다. 전송을 시작하려는 sequenced 메세지는 우선 링크의 상태를 먼저 확인하여 이용할 수 있는 속도가 빠른 링크에 할당된다. 이 때, busy한 링크의 집합을  $S = (p, q)$ 라 하고 idle한 링크의 집합을  $I = (m-p, n-q)$ 가 된다. sequenced 메세지는 집합  $U = (r, s)$ 를 bypass한다고 가정하자. 전송시간의 memoryless 성질 때문에 링크 k 에서 전송되고 있는 메세지가 t[sec] 경과한 후에도 전송되고 있을 확률은  $e^{-\mu_k t}$ 이다. 메세지 전송 시간은 모두 상호독립인 랜덤 변수이므로 시간 t가 경과한 후 sequenced 메세지가 U 에 속하는 메세지를 bypass할 확률은

$$P_{bp}(U|t, S) = \left( \prod_{k \in U} e^{-\mu_k t} \right) \left( \prod_{k \in S-U} (1 - e^{-\mu_k t}) \right) \quad (9)$$

이며, t에관한 조건을 제거하면

$$P_{bp}(U|S) = \int_{t=0}^{\infty} P_{bp}(U|t, S) dF_{T_k}(t) \quad (10)$$

여기서  $F_{T_k}(t)$ 는 링크 k내에서의 sequenced 메세지의 전송 시간 분포함수이다.  $W_k$ 를 링크k에 있는 메세지를 bypass한 sequenced 메세지를 기다리는 대기시간 이라하면 W의 분포함수는 모든 메세지가 지수 분포이므로

$$F_{W_k}(t) = 1 - e^{-\mu_k t} \quad (11)$$

로 표현된다. sequenced 메세지가 U로 주어지는 링크의 집합을 bypass 하였다면 sequenced 메세지의 대기시간은

$$W_{u,s} = \max_{k \in U} W_k \quad (12)$$

로 표현된다. 또한  $W_k$ 는 상호독립 하므로

$$\text{Prob}[W_{u,s} \leq t] = \prod_{k \in U} (1 - e^{-\mu_k t}) \quad (13)$$

이며, 여기서 U와 공집합을 포함한 모든 부분집합이며 S에 대한 조건을 제거하면 대기 시간 분포함수는

$$F_w(t) = \sum_{U \subseteq U_s} \text{Prob}[W_U \leq t] Q[s] \quad (14)$$

이며,  $U_s$  은  $L = \{1, 2, \dots, L\}$ 의 모든 진부분 집합 즉, 공집합은 포함하나 L은 포함하지 않는 집합을 말한다.  $Q[s]$ 는 sequenced 메세지가 S내의 다른 메세지와 결합될 확률이며 다음과 같이 상태확률과 관련된다.

$$O[S] = \begin{cases} \text{(i) 2개 이상의 링크가 idle한 경우} \\ \quad m C_i n C_j p(i, j) \\ \text{(ii) 단지 하나만의 링크가 idle한 경우} \\ \quad L C_1 \left[ \sum_{N=0}^B P(N) \left( \frac{-\mu_k}{\mu_1 + n \mu_2} \right) + p(i, j) \right] \\ \text{(iii) 그 외의 경우} \\ \quad 0 \end{cases}$$

moment generating 함수를 이용하면, 평균 재조립 시간지연은

$$W = \sum_{S \subseteq U_s} Q[S] \sum_{U \in U_s} P_{bp}(U|S) \bar{W}_{u,s} \quad (15)$$

이며, 여기서

$$\bar{W}_{u,s} = \int_{t=0}^{\infty} [1 - \prod_{j \in U} (1 - e^{-\mu_j t})] dt \quad (16)$$

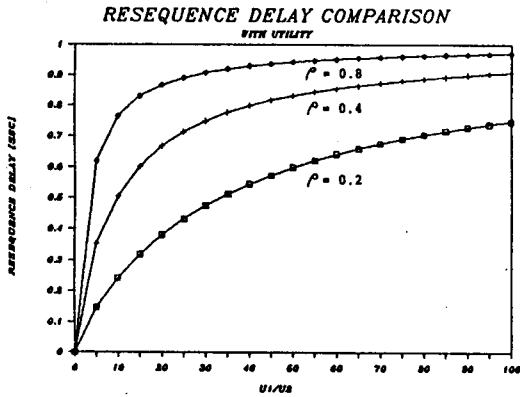
는 조건 평균 대기 시간이다.

평형상태에서 재조립 버퍼로의 평균 도착율은  $\lambda(1-P_b)$ 와 같다. 따라서 재조립 버퍼의 평균 큐길이는 Little's formular에 의해  $L_R = \bar{W} \lambda(1-P_b)$ 로 나타낸다. [4][5]

3. M/M/2 큐잉 모델에서의 시간 지연 및 지연 시간 감소 방법

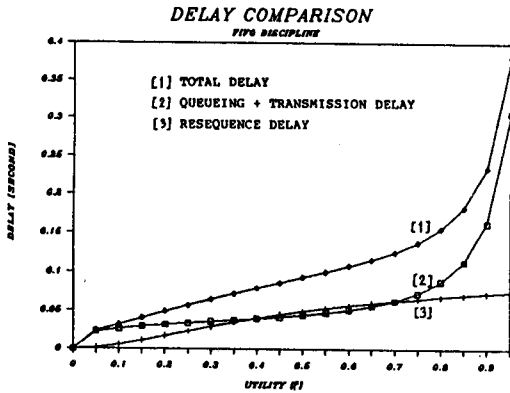
이 장에서는 2장에서 기술한 큐잉 해석에 의하여 간단한 2개의 heterogeneous 링크로 구성된 큐잉 시스템에 대한 큐잉 지연과 재조립 시간 지연 및 지연 시간을 감소시키는 방법에 대하여 알아보기로 한다. 그림2에서는 slow 서버의 서비스율이 1[Kbps]이고 fast 서버의 서비스율이 2에서 100[Kbps]까지 변화 할 때의 재조립

시간 지연을 3가지 값에 대하여 나타내었다. 이로부터 값이 클수록 그리고 2개의 서비스율의 차가 클수록 재조립 시간 지연이 증가함을 알 수 있다.



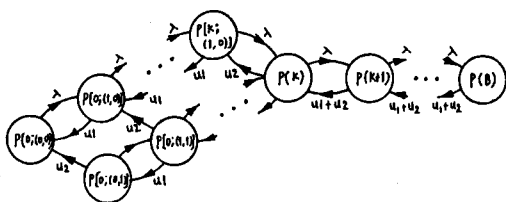
(그림 2) 서비스율의 비에 따른 재조립 지연 시간

그림 3은 노드쌍을 상호 연결하는 2개의 heterogeneous 링크 시스템의 용량이 각각  $u_1 = 56$  [Kbps] 와  $u_2 = 9.6$  [Kbps] 인 때, 지연 시간을 시스템의 이용률( $\rho$ )의 함수로 나타낸 것이다.



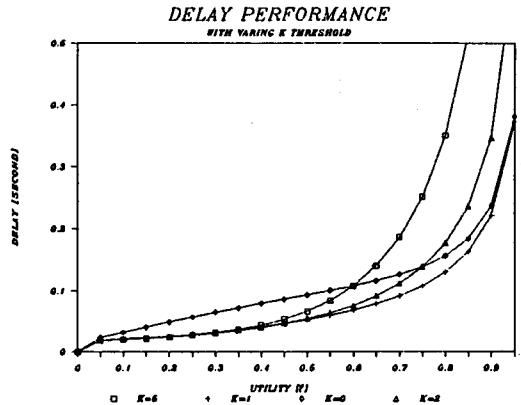
(그림 3) M/M/2 시스템에서 이용률에 따른 지연시간

다중 heterogeneous 링크 시스템의 시간 지연 감소를 시키기 위하여, fast 서버에 대하여 threshold  $k$  를 정의하여  $k$  이하의 customer가 큐내에 대기하는 경우 slow 서버가 idle한 경우에도 서비스를 받지 않도록 모델링할 수 있다. 그림 4에서  $k$ 를 고려한 M/M/2 큐잉 시스템의 상태 천이 diagram을 나타내었다.



(그림 4)  $k$ 를 고려한 M/M/2 큐잉 시스템의 상태 천이 diagram

$k$ 를 고려한 위의 큐잉 모델에 대한 지연 시간 해석은 2장에서 기술한 지연 해석을 통하여 구할 수 있다. 이 경우 재조립 지연 시간은  $u_1/u_2$ 의 최대 정수값을 취하여 구할 수 있으나 이는 오히려 input 버퍼내 큐를 크게 증가시켜 총지연 시간을 증가시킨다. 그러므로 적당한  $k$ 값을 선택함으로써 총지연 시간을 감소하도록 해야 한다. 그림 5에서  $k$ 값에 따른 총지연 시간을 나타내었다.



(그림 5)  $k$ 값에 따른 총 지연 시간

#### 4. 결 론

본 연구에서는 2 class heterogeneous 링크로 구성된 M/M/L 큐잉 시스템에서의 시간 지연을 해석 하였다. 또한 지연 performance를 향상 시키기 위하여 threshold를 적절히 선택하여 총 지연시간이 감소함을 알 수 있었다. 여기서는 간단히 M/M/2 모델에서 살펴 보았으나 3개 이상의 링크로 구성된 모델에서도 마찬가지로 해석할 수 있다.

#### 참 고 문 헌

- [1] M. Schwarz, Telecommunication Networks: Protocols, Modeling and Analysis, Addison-Wesley Publishing Company, 1987
- [2] L. Kleinrock, Queueing Systems, Volume 1 : Theory, John Wiley & Sons, New York, 1975
- [3] A.S. Tanenbaum, Computer Networks, Prentice-Hall, Inc. 1981
- [4] T.P. Yum and T. Ngai, "Resequencing of Messages in Communication Networks," IEEE Trans. on Com., vol. COM-34, no.2, Feb. 1986
- [5] Y.c. Luke Lien, "Evaluation of the resequence Delay in a Poisson Queueing System with 2 Heterogeneous Servers," Computer Networking and Performance Evaluation, IFIP, 1986