

Short Bus Contention 방식의 Priority Output Queuing Model의 분석

정 용 주[†]

요 약

광역 ISDN에서 여러 패킷들은 그 사용 용도에 따라서 서버에 요구하는 속도를 달리하면 다른 결과를 보일 수 있을 것이다. 즉 일반적인 자료들은 보통의 속도에서 처리하여도 별 차이는 없을 것이나 그러나 어떤 자료들은 - 예를 들면 실시간 동적 영상자료나 목소리 자료들 혹은 인터넷 상거래에서 입찰을 위한 자료들은 - 일반적인 자료들에 비하여 보다 빠른 처리를 하는 것이 서비스의 질을 높일 것이다. 이것에 대한 해결책이 - 우선 순위 패킷 방식이 - 제안되었으나 그 분석이 아직 연구 중이다. 그래서 본 논문에서는 단일 서버의 출력 큐잉 스위칭을 가정하여 패킷들에게 우선 순위를 부여하고 또한 현실적인 상황을 모방하여 상관 관계를 패킷들 사이에 가정하여 분석하였다. 이를 위하여 이 패킷들은 세 가지 경우로 나뉘어 분석하였는데 첫째 상관 관계나 우선 순위가 모두 없는 경우, 둘째 단순히 상관 관계만을 가지는 경우와 또 마지막으로 우선 순위와 상관 관계를 함께 갖는 것을 가정하여 시뮬레이션을 통하여 분석하였다. 이 결과 상관 관계는 도착율이 일정 값 이하의 경우에는 대체로 영향을 미치지 않는 것으로 나타났고 또한 높은 우선 순위가 있는 패킷들의 지연 시간은 도착율과 상관 관계에 큰 관계없이 상당히 향상됨을 보였다. 여기서의 패킷들은 ATM에서와 같이 고정크기의 패킷들이고 출력 큐에서의 출력 방식은 근거리 버스 경쟁(short bus contention) 방식을 가정하여 각 패킷들의 평균 지연 시간, 또한 패킷들을 손실하지 않기 위한 최대 버퍼 길이 등을 분석하였다.

The Analysis of Priority Output Queuing Model by Short Bus Contention Method

Yong-Ju Chung[†]

ABSTRACT

In broadband ISDN every packet will show different result if it would be processed according to its usage by the server. That is, normal data won't show big differences if they would be processed at normal speed. But it will improve the quality of service to process some kinds of data - for example real time video or voice type data or some data for a bid to buy something through the internet - more fast than the normal type data. Solution for this problem was suggested - priority packets. But the analyses of them are under way. So in this paper a switching system for an output queuing model in a single server was assumed and some packets were given priorities and analysed. And correlation, simulating real life situation, was given too. These packets were analysed through three cases, first packets having no correlation, second packets having only correlation and finally packets having priority and correlation. The result showed that correlation doesn't affect the mean delay time and the high priority packets have improved mean delay time regardless of the arrival rate. Those packets were assumed to be fixed-sized like ATM fixed-sized cell and the contention strategy was assumed to be short bus contention method for the output queue, and the mean delay length and the maximum 버퍼 length not to lose any packets were analysed.

[†] 종신회원 : 단국대학교 전자계산학과 교수
논문접수 : 1998년 10월 30일, 심사완료 : 1998년 12월 29일

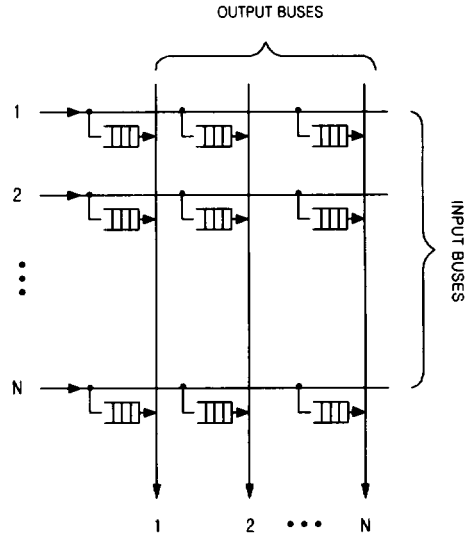
1. 서 론

광역 ISDN에서 통신의 양이 현재 점진적으로 증가하고 있다. 그리고 증가하는 통신의 양 뿐 만이 아니라 그 용도도 다양해지고 있다. 그런데 현재 기존의 많은 server에서 아직은 그 다양한 service를 충분히 만족시키지 못하는 경우가 종종 발생하는 것 같다. 현재의 이런 문제들을 해결하는데는 여러 가지 해결책이 있겠으나 가장 손쉬운 방법으로는 전반적인 통신망 전체를 변화시키는 방법이 있을 수 있을 것이다. 그러나 이것은 많은 비용이 든다. 또한 server가 전체적인 통신의 교통 양을 잘 소화한다 하더라도 통신의 용도가 다양해진 상황에서는 통신 자료의 용도에 따라서는 server의 기본 처리에 차등을 두는 것이 바람직 할 것이다. 즉 통신 할 자료의 용도에 따라 먼저 할 것은 먼저하고 다소 지연하여도 되는 자료에 대해서는 뒤에 처리하면 사용자에게는 보다 나은 환경을 제공할 수 있을 것이다.

전송할 자료에 따라서는 자료의 성질이나 용도에 따라서 크게 두 가지로 나뉠 수 있다. 실 시간적 자료와 그렇지 않은 자료들이다. 보통의 text형 자료들은 server에서 컴퓨터로 전달되어 바로 처리된다. 그러나 동적 영상 자료나 목소리 자료들은 컴퓨터 내에서 재구성을 하고 사용자에게 전달된다. 이 경우 자료의 지연은 service의 질에 문제가 된다.[2, 16] 즉 TV 방송국을 internet의 작은 화면으로 본다던가 internet phone을 사용하는 경우들이다. 이러한 자료들은 자료의 지연에 상당히 민감하게 반응하는 것이다. 그리고 이외에도 입찰을 internet을 통해서 한다던가 하는 자료들도 실 시간적 자료로 분류될 수 있을 것이다. 즉 심한 경우 입찰이 결정 났는데도 자료를 보내는 경우가 발생할 수 있는 것이다.

본 논문에서는 이와 같이 빠른 처리를 요구하는 자료들에 대하여 단순 우선 순위 구조를 제안하고 분석한다. 이 우선 순위는 network의 스위치 내에 구성되며 전달되는 각 패킷들에 우선 순위를 구별 할 수 있음을 가정한다. 아마도 패킷들의 우선 순위의 구별은 ATM 셀 헤더 안의 PTI(Pay Load Indicator) 부분의 사용자용 bit나 사용자 안 bit를 사용할 수 있을 것이다. 그리고 이 연구를 위하여 본 논문에서는 큐잉 모델로 접근한다. 그리고 이 큐잉 모델은 패킷들의 우선 순위를 구별 할 수 있는 것으로 가정하고 ATM에서와

같이 고정 크기 셀을 처리 할 수 있도록 이산 시간 큐를 가정한다.



(그림 1) 출력 큐잉 모델
(Fig. 1) Output queuing model

스위치의 구성이나 분석에서 큐잉 모델로 많이 연구되었는데 특히 ATM에서와 같이 고정 크기의 셀 방식에서는 이산 시간 큐잉 모델(discrete-time queuing model)에 관한 연구가 많이 발표되었다.[3, 4, 5, 9, 11, 14, 15, 18] 이 연구들은 일반적인 이산 시간 큐잉 모델에 관한 문제들을 상당히 해결한 상태이다. 이 중 [9, 14]는 주기적이거나 연속된 큐잉 모델을 가정한 특수한 조건의 큐잉 모델에 관한 연구를 하였는데 최근에는 우선 순위를 부여한 큐잉 모델이 연구 발표되고 있다.[1, 8, 13, 17] 즉, 보통의 자료에는 하위 순위를, 동적 영상 자료나 목소리 자료들과 같이 전달 시간에 민감한 자료에는 높은 순위를 주어 연구한 것이다. 그러나 이들 연구는 출력 선에서 패킷의 선택에 관한 알고리즘을 선택하지 않고 큐잉 모델에서 행렬이나 생성함수(generating function)를 통하여 분석하였다. 그러나 본 논문에서는 출력 큐잉에서 각 패킷들에 우선 순위를 부여하고 또한 출력 방식에서 현실적으로 구체적인 알고리즘을 선정하여 분석한다. 즉 같은 출력 선에 여러 패킷들이 동시에 존재 할 때 처리하여야 할 패킷의 결정을 근거리 버스 경쟁 방식을 사용하여 시뮬레이션을 통하여 그 결과를 분석하여 본다. 이 방식은

각 버퍼에 패킷이 있는 경우 출력 에 탭(tap) 방식에 의하여 패킷이 선택이 되므로 패킷의 전달 시간에는 영향을 미치지 않는다.[2, 10]

그리고 본 논문에서는 이와 같이 자료 전달 지연 시간에 민감한 자료들에 관하여 우선 순위 뿐 만이 아니라 다른 특성도 가정한다. 예를 들어 동적 화면 자료들은 그 자료의 특성상 처리가 빠를 뿐 아니라 자료의 크기가 큰 것이 또 다른 특징이다.[6, 7] 이것은 한 개의 패킷 안에 그 자료를 담을 수가 없어서 여러 개의 패킷들에 나뉘어 전달되게 되는데, 이러한 경우 이 패킷들은 같은 출발지에서 출발하여 같은 목적지를 가기 위하여 스위치에 연속적으로 도착하는 것으로, 본 논문에서는 그들에게 상관 관계, q 를 줌으로써 스위치 전체의 교통의 상태를 분석한다. 즉 본 논문에서는 기존의 스위치에 관한 연구 분석에 구체적인 알고리즘을 선정하고 또한 각 패킷들 사이에 상관 관계를 가정하여 상관 관계에 따른 변화를 연구 분석한다. 그리고 또한 본 논문에서는 버퍼가 출력 쪽에 있는 출력 큐잉 모델을 가정하였는데 출력 큐잉의 전체 버퍼 크기는 각 버퍼의 제곱에 비례하므로($N \times N$ 스위치에서 N^2 개의 버퍼) 패킷을 손실하지 않기 위한 버퍼의 크기도 분석한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 사용한 큐의 모델과 근거리 버스 경쟁 알고리즘에 관하여 설명하고 3절에서는 그 결과와 분석을 우선 순위나 상관 관계가 없는 경우부터 상관 관계와 우선 순위를 차례로 갖는 순서로 보인다. 4절에서는 결론과 향후 연구 방향에 대하여 논의한다.

2. 큐잉 모델

본 논문에서는 일반적 특징의 모델을 가정한다. 즉 이 처리 시스템은 단일 서버에 존재하고 입력과 출력 선들은 보통의 특정한 제약을 받지 않은 상태이다. 그리고 기술한대로 라우팅(routing)을 먼저하는 $N \times N$ 의 출력 버퍼의 출력 큐를 모델로 한다.

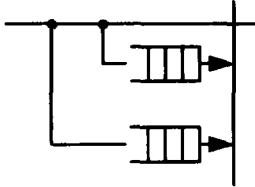
다음 본 논문 모델의 특성이다.

- (a) $N \times N$ 스위치이다.
- (b) 난수 값은 Unix의 drand48()을 사용하였고, 각각의 셀은 Poisson 분포를 한다. 즉 도착 시간은

지수 분포이다.

- (c) 서로 다른 출발지를 가진 여러 패킷들이 한 시간 구획(time slot)안에 동시에 서로 같은 목적지를 가는 것을 가정한다. 즉 어느 입력 선이든지 어느 임의의 출력 선으로 갈 수 있다.
- (d) 시간은 구획되어 각 구획 당 1개의 셀의 처리만 가능하고 그 시간은 단위 시간인 IT이다.
- (e) 각각의 시간 구획에서 큐에 셀이 있으면 바로 처리한다. 그리고 그 시간 구획이 끝나기 이전에 (끝나는 시각까지 포함하여) 도착하는 셀은 바로 그 다음 처리 대상이 된다.
- (f) 초기에는 각 셀에 correlation이 존재하지 않는다. 그리고 각 셀의 목적지는 균일 분포(uniform distribution)에 의하여 결정된다.
- (g) 도착 율(arrival rate), 즉 가중치 확률(loaded probability)은 p 로 가정한다. 즉 평균적으로 $1/p$ 시간구획마다 1개씩의 셀이 도착함은 가정한다. 그리고 기술한 대로 각 셀 사이의 시간 간격은 지수 분포에 의하지만, 처음 일정 개수의 난수 값은 취하지 않는다.
- (h) 출력 단계에서 각 버퍼에 패킷들이 존재한다면 출력 패킷의 선택은 근거리 버스 경쟁 방식을 택한다.[2] 즉, 각 버퍼에 존재하는 패킷들에 pseudo bit을 설정하여, 그 시간대에 동시에 존재하는 패킷들에게는 각 pseudo bit들을 set하고, 그 set되어진 패킷들을 모두 먼저 처리한다. 그 후 다시 - pseudo bit이 set된 것이 없는 상태에서 - 버퍼들을 조사하여 버퍼에 존재하는 패킷들의 pseudo bit들을 set하여 차례대로 처리한다. 3.1의 단순 패킷들의 분석은 (a)-(h)의 특성들만을 조건으로 처리한 결과이다.
- (i) 3.2 상관 관계가 있는 패킷들의 분석은 다음과 같은 가정을 택하였다. 어떠한 상관 관계 q 에 대하여 연속되는 패킷들이 서로 상관 관계가 있다면, 그 다음 패킷이 같은 목적지로 갈 확률은 q 이고, 또 그 다음 패킷이 역시 똑같은 목적지로 갈 확률은 $q \times q$ 로 설정하였다. 그리고 다른 목적지로 갈 확률은 각각 $1 - q$, $1 - q \times q$ 이다. 즉, r 개의 연속된 패킷이 동일한 목적지 i 로 가고 또 그 다음 패킷이 같은 목적지 i 로 갈 확률은 q^r 이고, r 회 이후 도착한 패킷이 다른 목적지로 갈 확률은 $1 - q^r$ 인데, 이 경우 또

다른 목적지 $j, j \neq i$ 는 확률 $1 - q^i$ 과 또 i 를 제외한 $N-1$ 개의 목적지들 중에서 균일 분포에 의하여 선택된다.



(그림 2) 보통 버퍼와 우선 순위 버퍼
(Fig. 2) normal buffer and priority buffer

- (j) 우선 순위는 두 가지 상태만을 가정하였다. 높은 순위 자료 패킷과 낮은 순위 자료 패킷이다. 각 우선 순위는 각각의 버퍼를 가지는 것을 가정하므로, 전체적으로 $2 \times N^2$ 의 버퍼가 있다.
- (k) 우선 순위가 있는 패킷들은 낮은 상관 관계든 높은 상관 관계든 상관 관계를 가지는 것으로 한다.
- (l) 우선 순위를 가진 패킷들도 역시 출력 선택 방식으로 근거리 버스 경쟁 방식을 사용한다. 그러나 우선 순위가 있는 버퍼에 패킷이 존재하는 경우에는 낮은 순위의 버퍼에 pseudo bit이 set 되어 있어도 우선 순위 패킷을 먼저 처리한다.
- (m) 패킷의 손실이 없기 위한 최대 버퍼의 크기는 전체 시행 횟수 중 대기하는 패킷들의 최대 길이로 한다. 그러나 패킷들의 대기 길이가 그 최대 길이로 도달한 횟수가 0.005%(즉 100,000회 중 5회) 이상인 경우만을 최대 길이로 가정한다.

3. 결과와 분석

이 절에서는 각 패킷의 평균 대기 시간과 $N \times N$ 의 버퍼들에서 버퍼의 최대길이를 분석한다. 패킷들의 평균 개수에 대하여는 다음과 같은 식이 유도된다.[3]

$$E(n) = p + \frac{p^2(1 - \frac{1}{N})}{2(1-p)} \quad \dots \text{식 (1)}$$

$E(n)$: 시스템 내에서의 패킷 평균 숫자
 p : 도착 율
 N : 입력선의 수

$E(n) = \lambda D$ 이고, (Little의 공식, λ 는 평균 도착 율, D 는 평균 지연 시간) $\lambda = p/T$ 이므로 (T 는 단위 시간)

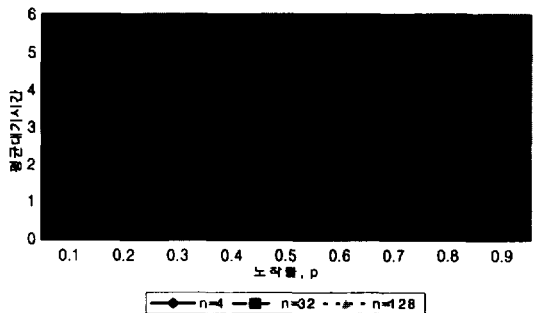
$$D = T \left[1 + \frac{p(1 - \frac{1}{N})}{2(1-p)} \right] \quad \dots \text{식 (2)}$$

로 된다.

따라서 평균 지연 시간 D 는 위의 식에 의하면 p 의 값이 1에 접근함에 따라 D 는 커지나 입력 큐잉(input queuing)때와 같이 head-of-block 문제는 발생하지 않는다. (N 은 일정이상의 값에서는 D 값에 큰 영향이 없다.) 시뮬레이션은 C^{++} 로 작성하여 난수 값의 오류를 방지하기 위하여 각 입력선 별로 처음 10000개의 도착 패킷들은(난수 값은) 버리고, 시간 구획 0부터 100,000까지 시험하였다.

3.1 단순출력 큐잉 모델

$N=1$ 일 때는 $D=T$ 이므로 전혀 교통에 밀림이 없다. 즉 입력과 출력이 하나이므로 들어오는 즉시 처리된다. 그리고 p 가 1에 접근할 때는, 예상했던 대로, 대기 시간이 급격히 증가함을 보이고 입력 수 N 은 크게 작용되지 않음이 보인다. $p=0.8$ 일 때는 N 이 아무리 증가를 하여도 평균 지연 시간은 $3.0T$ 부분에서 수렴함을 알 수 있다. (그림 3)은 도착 율에 대한 평균 대기 시간이다. 그리고 $N \times N$ 버퍼들 중에서 버퍼의 최대 크기는 <표 1>에 나타나 있다. 즉, 입력 선이 128($N=128$)이고, 도착 율이 0.8일 때 총 도착 횟수는 10235022회 ($128 \times 100000 \times 0.8 = 10240000$)이고 시스템 내에서의 평균 시간은 $2.98T$ ($2.98T$, 식 (2)), 최대 버퍼 크기가 4이었던 횟수는 42회이었다.



(그림 3) $N=4, 32, 128$ 의 평균 대기 시간
(Fig. 3) Average time in system when $N=4, 32$ and 128

<표 1> 상관 관계나 우선 순위가 없는 상태에서의 근거리 버스 방식을 사용한 결과. 시뮬레이션은 N=4, 8, 16, 32, 64, 128에서 각각 실험되었다.

<Table 1> The result of short bus contention algorithm with no priority and no correlation. The simulation was tested when N = 4, 8, 16, 32, 64, 128.

N	p	total input	aver. time in sys.	times of maximum 버퍼 length, time slot=100000										
				0	1	2	3	4	5	6	7	8	9	10
4	0.2	80124	1.09	92698	7254	48								
	0.4	159700	1.25	67461	31291	1196	50	2						
	0.6	240090	1.58	28965	59008	10343	1406	230	44	4				
	0.8	319690	2.59	2816	37831	33094	15417	6339	2567	1057	514	241	87	36
128	0.2	2559050	1.12	4867	95063	70								
	0.4	5121066	1.33	0	97886	2106	8							
	0.6	7682629	1.74	0	78594	21294	112							
	0.8	10235022	2.98	0	6434	88908	4616	42						

이 결과에서 평균 대기 시간은 식 (2)의 예상 결과와 거의 정확히 일치하였으나, N의 값이 작고 p의 값이 커질 경우 약간의 오차가 생김을 보인다.(N이 4, p가 0.5 이상에서 오차 범위 0.1 이하.) 그러나 N이 32 이상에서는 p값이 0.8 이상이라 하여도 오차는 0.01이 하였다.

3.2 상관 관계가 있는 패킷들에서의 큐잉 모델

상관 관계가 0.3이라 하면 어느 한 패킷이 임의의 목적지에 도착했을 때 같은 입력 선으로 온 그 다음 패킷이 같은 목적지에 도착할 확률은 0.3 이다. 그러나 다음 세 번째 패킷이 그 앞의 패킷과 연결된 자료일 확률은, 2. 큐잉 모델에서 가정되었듯이, 0.09이다. 이것은 본 시뮬레이션에서 입력 선에 도착하는 패킷의 난수 값이 주어진 상관 관계 비율 q 보다 작으면 같은 목적지로, 크면 다른 목적지로 가는 것으로 하였다. 따라서 상관 관계 확률이 q라하면 원래의 파일의 크기, M, 는 다음과 같다.

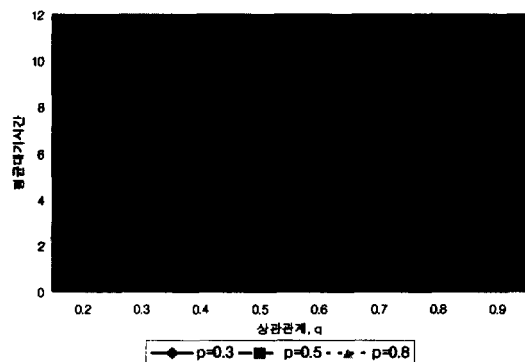
$$\begin{aligned}
 M &= (1+E) \cdot (\text{cell size}) \\
 E &= (1 \cdot q + 2 \cdot q^2 + 3 \cdot q^3 \dots) \\
 &= \lim_{n \rightarrow \infty} q \frac{1 - q^{n+1}}{(1 - q)^2} \\
 &= \frac{q}{(1 - q)^2}
 \end{aligned}$$

(E : 어느 목적지에서 한 개의 패킷을 받은 후, 더

늘어날 파일의 크기 조건부 기대값)

즉, q=0.5이라면 원래의 파일의 크기는 단위 패킷 크기의 3(=1+2)배이다.

일반적 예상과 같이 도착율이 낮을 때는 평균 지연 시간이 상관 관계에 상관없이 상관 관계가 없는 경우와 비슷함을 알 수 있으나 도착율이 커질수록, 상관 관계가 높을수록 특정 출력 선에 패킷들이 몰리는 특성으로, 평균 대기 시간과 최대 패킷 대기 길이가 늘어 나는 것이 (그림 4)에 나타나 있다.



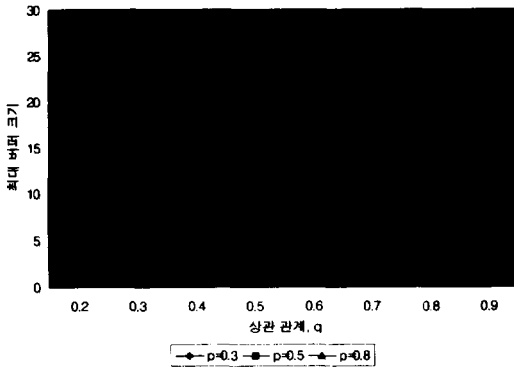
(그림 4) N=64이고 p=0.3, p=0.5 p=0.8 각각의 상관 관계에 따른 평균 대기 시간의 변화

(Fig. 4) Average times in system of N=64 with p=0.3, 0.5 and 0.8

(그림 4)에서 평균 대기 시간은 p=0.3일 때와 p=0.5일

때는 큰 차이는 없으나 $p=0.8$ 인 경우 그 차이는 $q=0.5$ 에 비해 훨씬 크다. 특히 $p=0.3$ 이하일 때는 (그림 3)과 (그림 4)를 비교하면 시스템 내의 평균 대기 시간은 q 에 거의 무관함을 알 수 있으나 $p=0.8$ 일 때는 많은 차이를 보인다. 실지로 p 가 0.6 이상일 때부터 평균 대기 시간은 상관 관계에 많은 영향을 받는 것으로 나타난다.

(그림 5)에서 64×64 출력 버퍼들 중에서 대기 패킷들의 최대 길이도 평균 대기 시간과 비슷하게 p 가 낮을 때는 적으나 p 가 커짐에 따라($p=0.8$) q 의 증가 정도에 따라 패킷의 대기 길이도 커진다.



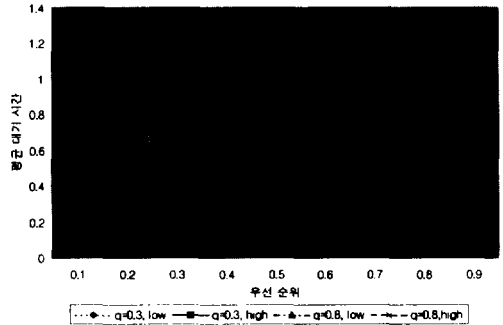
(그림 5) $N=64$ 에서 p 가 각각 0.3, 0.5, 0.8일 때 상관 관계에 따른 최대 버퍼 크기의 변화
(Fig. 5) The variations of the maximum buffer sizes according to $p=0.3, 0.5$ and 0.8 . $N=64$

3.3 상관 관계와 우선 순위가 있는 큐잉 모델

$k\%$ 의 우선 순위가 q 비율로 상관 관계가 있다고 하면 입력되는 패킷들 중 $k\%$ 가 우선 순위가 있고 또 그 패킷들은 q 비율만큼 서로 상관 관계가 있다. 즉 30% 우선 순위가 0.5 상관 관계는 입력되는 패킷들 중 30%가 우선 순위가 있고, 그들은 다시 0.5의 상관 관계를 가진다.

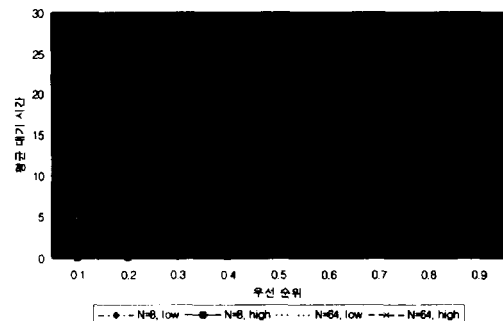
우선 순위가 높은 패킷들은 전체 도착율이 0.5에서는 높은 우선 순위의 패킷의 퍼센트가 커진다 하더라도 평균 대기 시간이 0.4 이하로 양호한 결과를 나타낸다. (그림 6) 또한 상관 관계 자체도 0.3이거나 0.8이라 하여도 높은 우선 순위와 낮은 우선 순위의 패킷들의 평균 대기 시간에는 거의 영향을 미치지 않는 것으로 나타난다. 그리고 낮은 우선 순위의 패킷들의 평균 대기 시간은 높은 순위의 패킷의 퍼센트가 90%라 하여도 2.29T로 나타났다. ((그림 3)과 (그림 4)는 시스템

에서의 처리 시간을 포함한 전체 시간이고, (그림 6)과 (그림 7)은 버퍼 내에서 만의 대기 시간임. 그 값은 처리 시간 1T만큼 차이 남.)



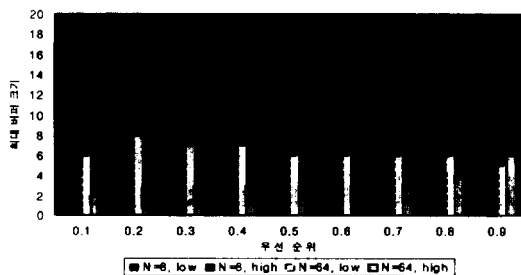
(그림 6) $N=16$, 도착 율, p 가 0.5일 때 $q=0.3, q=0.8$ 에서의 높은 우선 순위 패킷의 양에 따른 높은 우선 순위와 낮은 우선 순위 패킷들의 버퍼내의 평균 대기 시간
(Fig. 6) Average waiting times of low priority packets and high priority packets according to the amount of the high priority packets with $q=0.3$ and 0.8 respectively. $N=16$ and $p=0.5$

(그림 7)과 (그림 8)은 통신 양이 많은 경우의($p=0.9$) 패킷들의 버퍼 내 평균 대기 시간과 패킷 손실이 없기 위한 최대 버퍼의 크기이다. 이 결과는 전체 도착율이 커진다 하더라도, $p=0.9$, 높은 우선 순위의 패킷들의 도착 퍼센트가 50%이하인 경우에는 높은 우선 순위의 패킷들의 버퍼 내 평균 대기 시간은 0.5T 이하로 양호한 결과를 보였다.(그림 7)



(그림 7) $N=8, N=64$ 일 때 burst 경우($p=0.9$)의 높은 우선 순위의 패킷들과 낮은 우선 순위의 패킷들의 버퍼 내에서의 평균 대기 시간. 높은 우선 순위의 패킷들의 상관 관계, q ,는 0.5로 가정
(Fig. 7) Average waiting times of low and high priority packets when $p=0.9$, $N=8$ and 64 respectively. The correlation of the high priority packets is assumed 0.5

그리고 높은 우선 순위의 패킷들이 90%라 하더라도 높은 우선 순위 패킷들의 버퍼 내 평균 대기 시간은 2.11T(N=64인 경우)로 나타났다. 그러나 낮은 순위 패킷들의 버퍼 내 평균 대기 시간은 주어진 조건이 다소 다르나(그림 7)에서는 $p=0.9$, (그림 4)에서는 $p=0.8$ (그림 4)에서 $p=0.8$, $q=0.5$ 의 값(3.91T)과 비교하여 보면 높은 순위 패킷들의 퍼센트가 커짐에 따라 처음 1T 미만의 차이에서 최대 25T(높은 순위 패킷이 90%인 경우)까지 나타났다. 그리고 $p=0.9$ 에서의 최대 버퍼 길이는 입력 선 수, N,의 차이에 따라 큰 차이가 났는데 N의 값이 클수록 최대 낮은 순위 버퍼와 최대 높은 순위 버퍼의 크기는 각각 비교적 고르게 나타났다. (그림 8) N=64인 경우 높은 순위 패킷이 90% 이하에서는 낮은 순위 버퍼의 길이는 최대 8이었고 필요한 높은 순위 버퍼 최대 크기는 6이었는데 (그림 5)의 $p=0.8$, $q=0.5$ 의 값(9)과 비교하여 보면 전체적으로 5의 증가이나 이것은 출력 스위치 모델에서 버퍼의 수가 N^2 임을 고려하면 전체적으로는 큰 증가이다.



(그림 8) N=8, N=64에서 $p=0.9$ 인 경우의 최대 버퍼 크기. 높은 우선 순위의 패킷들의 상관 관계 $q=0.5$ 로 가정 (Fig. 8) The maximum buffer sizes when $p=0.9$, N=8 and 64 respectively. The correlation of the high priority packets is assumed 0.5

4. 결 론

본 논문은 스위치의 교통량의 분석에 관한 것으로 입력되는 패킷들에 우선 순위가 있고 또 그들 사이에 상관 관계가 존재하는 것을 가정하여 분석했다. 이들은 출력 큐잉 모델에서 근거리 버스 경쟁 방식으로 버퍼의 패킷들이 선택됨을 가정하여, 우선 순위가 높은 패킷들도 같은 알고리즘으로, 평균 대기 시간과 패킷을 손실하지 않기 위한 버퍼의 최대 길이를 분석했다. 실험 결과에서 도착율, p 가 0.3 이하에서는 상관

관계 자체는 평균 대기 시간이나 버퍼 길이에 영향을 미치지 않는 것으로 나타났고 p 가 0.3과 0.5사이 일 때는 비교적 작은 영향을, 그러나 p 가 0.6 이상에서는 상관 관계가 클수록 큰 영향을 미치는 것을 보였다. 그리고 우선 순위를 부여한 경우에도 상관 관계는, 상관 관계가 0.5 이하인 경우는, 높은 우선 순위의 평균 대기 시간에는 큰 영향을 미치지 않음을 보였다. 그리고 통신 양이 많은 경우라 하더라도, $p=0.9$, 입력되는 패킷들 중 높은 우선 순위의 패킷의 양이 50% 이하이면 평균 대기 시간은 0.5T 이하인 양호한 결과를 보인다. 그러나 이 경우 패킷의 손실이 없기 위한 낮은 우선 순위의 버퍼의 크기는 스위치 입력선, N의 크기가 작을수록 현격히 증가함을 보인다.

그러나 본 논문에서는 전송할 파일의 크기와 전송량을 현실적인 상황을 가정하였다. 이것을 위하여 실제 통신에서의 전송 상황을 통계적 샘플링 방법으로 연구함이 필요하고, 우선 순위를 부여한 경우 전체적인 낮은 순위 버퍼의 크기가 기하급수적으로 증가함으로 이에 대한 새로운 알고리즘이나 대체 방법의 연구가 필요하다.

참 고 문 헌

- [1] A. Khamisy and M. Si, "Discrete-time priority queuing system with two state Markov modulated arrival process," *INFOCOM-91*, pp.1456-1463, 1991.
- [2] Anthony S. Acampora, "An Introduction to Broadband Networks: LANS, MANS, ATMS, B-ISDN and Optical Networks for Integrated Multimedia Telecommunications," Plenum, 1994.
- [3] B. Sengupta, "Sojourn time distributions for the M/M/1 queue in a Markovian environment" *European J. of Operational Research*, Vol.32, pp. 140-149, 1987.
- [4] D. Gross and C. M. Harris, *Fundamentals of Queuing Theory*, New York, Wiley, 1974.
- [5] D. M. Lucantoni, "New results in the single serverqueue with a batch Markovian arrival process," *Stochastic Models*, Vol.7, pp.1-48, 1991.
- [6] D. S. Lee, S. Q. Li and K. H. Tzou, "Control and Analysis in video packet loss in ATM Net-

works," *GLOBECOM-90*, pp.857-861, 1990.

[7] G. Ramamurthy and B. Sengupta, "Modeling and Analysis of a variable bit rate video multiplexor," *INFOCOM '92*, Florence, Italy, pp.817-827, 1990.

[8] J. A. Morrison, "Two discrete-time queues in tandem," *IEEE Trans. Comm.*, COM-27, pp.563-573, 1979.

[9] J. C. W. Van Ommereen, "The discrete-time single-server queue," *QUESTA*, Vol.8, pp.279-294, 1991.

[10] J.S. Turner and L.F. Wyatt, "A packet network architecture for integrated services," in *GLOBECOM'83 Conf. Rec.*, pp.45-50, Nov., 1983.

[11] M.F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and their applications*, Marcel Dekker, New York, 1989.

[12] Mark J. Karol, Michael G. Hluchyj, Samuel P. Morgan, "Input Versus Output Queuing on a Space-Division Packet Switch," *IEEE Trans. Comm.*, Vol.COM-35, No.12, pp.1447-1356, 1987.

[13] O. Hashida and Y. Takahashi, "A discrete-time priority queue with switched batch Bernoulli process inputs and constant service time," *Teletraffic and Datatrafic in a Period of Change*(A. Jensen and V. B. Iversen, eds.), pp.521-526, North Holland, Amsterdam, 1991.

[14] O. J. Boxma and W. P. Groenendijk, "Waiting times in discrete-time cyclic service systems," *IEEE Trans. Comm.*, COM-36, pp.164-170, 1988.

[15] S. Q. Li, "A general solution technique for discrete queuing analysis of multimedia traffic on ATM," *IEEE Trans. in Comm.* Vol.COM-39, No. 7, pp.1115-1132, 1991.

[16] T. Meisling, "Discrete-time queuing theory," *Operations Research*, Vol.6, pp.96-105, 1958.

[17] T. Takine and B. Sengupta, "An analysis of a discrete-time queue for broadband ISDN with priorities among traffic classes," *IEEE Trans. Comm.* Vol.42, pp.1837-1845, Feb., 1993.

[18] U. Briem, T.H. Theimer and H. Kroner, "A general discrete-time queuing model: Analysis and applications," *Teletraffic and Datatrafic in a Period of Change*(A. Jensen and V.B. Iversen, eds.), pp.13-20, North Holland, Amsterdam, 1991.

정용주

e-mail : chungy@cs.dankook.ac.kr

1981년 서울대학교 계산통계학과 졸업(학사)

1983년 한국과학기술원 전자계산학과(공학석사)

1983년~1984년 단국대학교 부설 통계문제연구소 연구원

1984년~1995년 단국대학교 전자계산학과 부교수

1995년 8월~1998년 2월 R.P.I. 박사과정

1998년 3월~현재 단국대학교 전자계산학과 부교수, R.P.I. 박사과정 휴학중

관심분야 : 프로그래밍언어, 자연어 처리, 네트워크, 병렬처리