

# GPRS 망에서 GGSN 노드의 패킷 처리 향상을 위한 성능 모델 및 분석

곽용원 · 민재홍 · 정영식 · 박웅

한국전자통신연구원

## Performance Model and Analysis for Improving Efficient Packet Service of GGSN in GPRS Network

Yong-Won Kwak · Jae-Hong Min · Young-Sic Jeong · Wung Park

Electronics and Telecommunications Research Institute

E-mail : {yongwon, jhmin, jys, wungp}@etri.re.kr

### 요 약

비동기식 제 3세대 이동통신 시스템인 GPRS망은 기존의 제 2세대 통신인 GSM 방식에 GGSN과 SGSN 노드를 추가함으로써 패킷 교환 방식을 가능하게 하였다. 따라서 GPRS망과 인터넷 망을 연결시켜주는 관문(Gateway) 역할을 수행하는 GGSN 노드의 성능과 용량을 최적화 시키기 위한 패킷 트래픽 처리 방안에 대한 연구가 필요하다. 이를 위해 본 연구에서는 외부 인터넷 망으로부터 GGSN 노드에 도착되는 인터넷 패킷 트래픽의 모델을 정립하고, 이를 효율적으로 처리하기 위한 GGSN노드의 프로토콜 구조를 분석하여 처리 프로세서들의 트래픽 처리 능력을 분석하기 위한 성능모델을 만든다. 제안된 성능 모델의 시뮬레이션을 통해서 다양한 트래픽 부하 상태에서 처리 능력에 대해 분석한다. 입력 트래픽인 Speech, Video, WWW, E-mail에 대해 실시간 트래픽인 Speech와 Video의 QoS 요구사항을 만족시키기 위해 여러 스케줄링 알고리즘을 적용하여 각각의 기법에 대한 성능을 분석하였다.

### ABSTRACT

Asynchronous third generation mobile communication system is able to service Packet Switching through adding GPRS Network to the second generation system GSM. Therefore, it is necessary to study packet traffic service of GGSN node which is due to perform gateway role that GPRS Network is enable to inter-connect with Internet in order to optimize the capability and performance of GGSN

In this paper, the Internet packet traffic model that it is arrived to GGSN node from the Internet is studied and In order to process the Inter traffic efficiently, performance analysis model in GGSN is proposed to optimize packet processing capability of each processor. In order to guarantee QoS requirement of the real time traffic Speech and Video, several scheduling algorithm is applied to performance model and each mechanism is compared with several performance parameters

### 키워드

GPRS, GGSN, SGSN, GTP

### 1. 서 론

오늘날 다양한 무선 데이터 서비스에 대한 수요가 증가함에 따라 빠른 데이터 전송이 요구되고 있고, 그

에 따른 사용자의 요구를 충족시킬 수 있는 멀티미디어를 지원하기 위한 무선 데이터 서비스가 필요하게 되었다. 이러한 흐름에 따라, 북미에서는 ITU-T에서의 IMT-2000표준화 과정에 따라, 이를 만족할 수 있

는 차세대 이동 통신기술이 개발되고 있다. 또한, 유럽에서는 WCDMA(Wireband Code Division Multiple Access) 방식으로 기존의 2세대 TDMA(Time Division Multiple Access) 기술인 GSM(Global System for Mobile Communications)을 바탕으로 되어 있다. 2000년 초기에는 단지 일부분의 GSM(Global 가입자들만이 데이터 서비스를 이용했다. 왜냐하면 현재 존재하는 GSM은 회선교환(Circuit-Switching) 방식으로 멀티미디어 데이터 서비스를 위한 쉬운 액세스, 높은 데이터율을 제공하지 못하기 때문이다. GSM 사업자들은 이러한 요구사항을 충족시키기 위해 더 나은 서비스를 제공해야만 한다. 이것에 대한 해결책이 GPRS이다. GPRS는 종단간(end-to-end) 패킷 교환(Packet-Switched) 서비스를 제공하기 위해 GSM 구조를 재사용하고 새로운 노드인 GGSN(GPRS General Support Node)노드와 SGSN(Serving General Support Node)노드가 추가된다. 즉 GPRS는 기존의 GSM 사용자들이 멀티미디어 서비스를 위하여 데이터 패킷을 외부 인터넷망에 연결시켜주는 역할을 하게 된다. 여기서 외부 인터넷 망과의 게이트웨이 역할을 담당하는 것이 GGSN이다. GPRS를 위한 지원 노드인 GGSN은 하나의 PLMN(Public Land Mobile Network)에서 여러 개의 SGSN과 연결되어 외부 네트워크와의 접속점의 역할을 한다. GGSN은 PDP(Packet data protocol)에 의해 패킷 데이터망에 액세스 되는 노드이며, 접속된 GPRS 사용자들을 위한 라우팅 정보를 포함하고 있다. 라우팅 정보는 MS(Mobile Station)의 현재 접속점(예,SGSN)으로 PDU (Packet Data Unit) 들을 터널링하는 데 사용된다. [1,2,3]. GGSN은 GPRS를 지원하는 2세대, 3세대 PLMN과 함께 PDN의 상호 연결을 위한 첫번째 지점(Gi reference)으로 외부 망에서 들어온 데이터가 GTP(GPRS Tunneling Protocol)로 캡슐화(encapsulation)된 후 UDP, IP를 거쳐서 전송된다[4]. 이 때문에 GGSN에서는 클래스별 패킷 트래픽을 효율적으로 처리하기 위한 능력이 필요하고 클래스별 QoS 보장을 위한 정확한 성능 분석과 GGSN의 성능을 향상시키기 위한 연구가 필요하다. 이를 위하여 본 연구에서는 GGSN노드의 UDP, IP, GTP 등의 프로토콜을 분석하여 물리적 버퍼와 프로세서를 찾아 트래픽 처리 모델을 만들고, 구현된 GGSN 노드의 성능 지표를 찾기 위한 시뮬레이션과 분석을 한다. 인터넷 트래픽은 시간지연에 민감한 트래픽(Speech, Video)과 시간지연에 민감하지 않는 트래픽(WWW, E-mail)으로 나눈다. 본 논문에서는 이러한 트래픽들의 특성에 따른 각 프로세스별 버퍼 성능에 대해서 연구하고 FCFS(First Come First Serve), FQ(Fair Queueing), WFQ (Weighted Fair Queueing) 등의 스케줄링 알고리즘을

적용하여 실시간 트래픽인 Speech, Video 트래픽의 QoS를 보장해 주고자 한다.

## II. GPRS 망 구조 및 패킷흐름도

### 2.1 GPRS망 구조 [2,3,6]

GPRS를 위한 지원 노드인 GGSN은 하나의 PLMN에서 여러 개의 SGSN과 연결되어 외부네트워크와의 접속점의 역할을 한다. GGSN은 PDP(Packet data protocol)에 의해 패킷 데이터망에 액세스 되는 노드이며, 접속된 GPRS 사용자들을 위한 라우팅 정보를 포함하고 있다. 라우팅 정보는 MS의 현재 접속점(예, SGSN)으로 PDU들을 터널링하는 데 사용된다. 그림 1에서 보듯이 GGSN은 GPRS를 지원하는 2세대, 3세대 PLMN과 함께 PDN의 상호 연결을 위한 첫번째 지점(Gi reference)이다. 외부망에서 들어온 데이터가 GTP(GPRS Tunneling Protocol)로 캡슐화(encapsulation)된 후 UDP/IP를 거쳐서 데이터링크 계층을 통해 SGSN으로 전송된다. UDP/IP는 라우팅 사용자 데이터와 제어 신호를 위해 사용되는 백본망 프로토콜로서, 한 GGSN에 연결되어 있는 SGSN들 간의 연결을 설정해주시기도 한다. SGSN은 Gb인터페이스를 통하여 MS를 지원하는 노드이다. GPRS 접속에서 SGSN은 MS에 대한 이동성과 보안정보를 포함하는 mobility management context를 설정한다. 사용자 데이터를 MS와 외부 데이터망 사이에서 투명하게 전송시키는 방법으로는 GTP(GPRS Tunneling Protocol)에 의한 SGSN과 GGSN간의 터널링 방법이 있다. 이를 위해 GPRS/UMTS PLMN은 외부 데이터 프로토콜을 해석하는 것이 요구되고, 미래에 추가적인 연동 프로토콜을 쉽게 받아들일 수 있어야 한다.

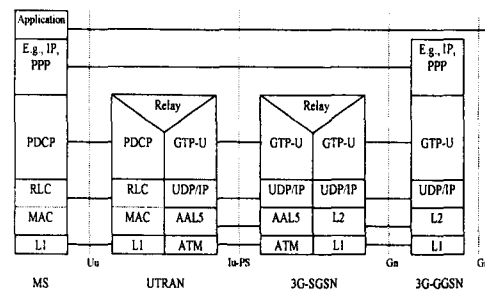


그림 2-1 GPRS/UMTS User Plane

### 2.2 PDP Context 설정 절차 및 패킷 흐름도[5]

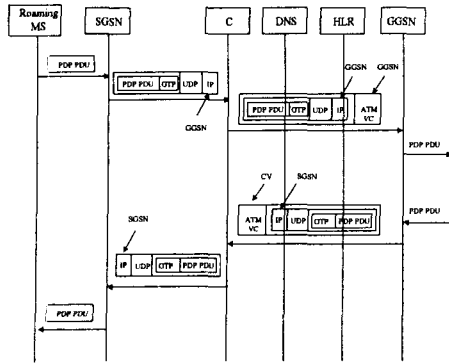


그림 2-2 GPRS망에서 패킷 흐름도

그림 2-2에서 CV는 전송 계층 변환기 (transport layer converter)를 의미한다. 이것은 노드 IP 주소를 ATM 주소에 매핑(mapping)시킨다. CV는 또한 ATM SVC링크에 대한 종단 점을 의미한다. CV는 IP 기반 터널링 망과 ATM SVC 터널링 망 사이에서 사용자 패킷을 전달하며 각각의 GTP 터널링에 대한 SVC 링크를 관리한다. 위의 그림 2-2는 GPRS망에서 패킷의 흐름도에 대해 나타낸다. MS가 SGSN에게 PDU를 보내면 SGSN은 이를 UDP 헤더와 GGSN의 주소를 갖고 있는 IP 헤더로 encapsulation한 후 핵심망(Core Network)으로 패킷을 보낸다.

여기서는 ATM-SVC가 설정되므로 GGSN의 ATM 주소를 갖고 있는 ATM-VC를 붙인 후 GGSN에게 전송된다.

### III. GGSN 노드의 구조

#### 3.1 GTP 프로세서[4]

GTP 프로토콜은 UMTS/GPRS 백본 망 내에 GGSN과 SGSN에 의해 구현된다. GTP는 여러 다양한 프로토콜의 패킷들이 SGSN과 GGSN사이에서 그리고 SGSN과 UTRAN사이에서 UMTS/GPRS 백본망을 통해 터널링이 가능하게 한다.

신호단계에서 GTP는 SGSN이 MS(Mobile Station)에 대한 패킷 데이터망 접근이 가능하도록 터널조정과 관리 프로토콜(GTP-C)을 상세화한다. 전송단계에서 GTP는 사용자 데이터 패킷을 전송하기 위한 서비스를 제공하기 위해 터널링 기법(GTP-U)를 사용한다. 어느 경로를 선택할지는 터널링 되어야 할 패킷들이 신뢰성 있는 링크를 요구하는지 아닌지에 달려있다.

GTP 헤더의 최소 길이는 8옥텟(Octet)으로 IP 스택에서 내려온 IP 데이터그램에 추가된다.

#### 3.2 UDP 프로세서[7]

GGSN 스택 내에서 UDP는 GTP 신호 메시지를 전달하기 위해 정의된 경로(Path) 프로토콜이다. UDP/IP는 비연결형 T-PDU(Packet Data Unit)들을 터널링하기 위한 비 연결형 경로로 사용된다. 연결형 경로로는 TCP가 사용될 수 있지만 TCP는 신호의 오버헤드가 크고 연결을 설정하는데 시간이 걸린다는 단점으로 신속하고 빠르게 보낼 수 있는 UDP가 선호된다. UDP 처리 과정에서 UDP 헤더인 8byte가 GTP에서 전송된 T-PDU에 추가된다.

#### 3.3 IP 프로세서[7]

IP 프로세서에서는 상위 계층인 UDP로부터 입력되는 패킷들을 헤더 추가모듈(Header-Adding Module)에서 IP 헤더 20바이트가 더해진다. 이 IP 헤더는 GGSN과 터널링 하기 위한 SGSN의 IP주소를 포함하고 있다. 처리모듈(Processing module)에서는 목적지 주소가 루프백 패킷 인지를 검사하고 라우팅 모듈로 패킷을 보낸다. 라우팅 모듈은 다음 보내져야 할 노드의 주소와 패킷이 보내져야 하는 인터페이스의 번호를 찾는다. 그리고 이 정보와 함께 패킷을 단편화 모듈(Fragmentation Module)로 보낸다. 단편화 모듈에서는 MTU 테이블을 참조하여 데이터링크 계층으로 패킷을 보내기 위해 해당 크기로 단편화 과정을 수행한다.

### IV. GGSN 노드의 성능 분석 모델 및 패킷 스케줄링 알고리즘

#### 4.1 패킷 스케줄링 알고리즘

##### 4.1.1 FCFS(First Come First Serve)[8]

FCFS Queueing의 원리는 패킷들이 도착한 순서대로 서비스 되는 것이다. FCFS 방법은 도착순서대로 패킷을 공통 버퍼에 저장하였다가 도착 순서대로 패킷을 전송한다. 이것은 인터넷 망에서 가장 널리 쓰이고 있는 기법이다. 그러나 FCFS 방법은 각 사용자 흐름(flow)의 독립성을 유지하지 못하고 공평성을 보장하지 못한다. 즉, FCFS는 패킷을 무조건 많이 보내는 흐름에게 망 자원을 더 많이 할당한다는 단점을 가지고 있다. 그러므로, 스케줄링 방법을 이용하여 QoS를 지원하는 문제외에 패킷 교환망에서 흐름간의 간섭을 없애고 독립성을 유지하는것도 중요한 문제이다.

##### 4.1.2 WFQ(Weighted Fair Queueing)[8, 9]

WFQ는 CBQ(Class Based Queueing)상에서 공정성과 성능 보장면에서 상당한 향상을 가져오는 기법이다. 네트워크상에서 전송시간이 큰 패킷이 같은 순서로

전송이 될 경우, 여기에 적용되는 다양한 서비스 클래스 사이에는 공정성과 자원 할당이 이루어져야 한다.

이에 스케줄러는 다양한 서비스 알고리즘을 사용해야만 하고, 낮은 우선 순위의 서비스 클래스 큐보다는 좀 더 자주 이용되는 서비스 클래스 큐를 이용하게 된다. 다시 말해서 서비스의 정확성과 타당한 수행결과를 이루기 위해서 높은 우선 순위의 트래픽이 낮은 우선 순위의 트래픽보다 자주 서비스되어야 한다.

#### 4.2 GGSN 노드의 성능 분석 모델

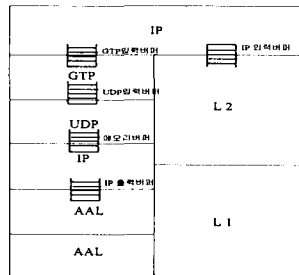


그림 4-1 GGSN 노드의 물리적 모델

그림 4-1은 GGSN 노드의 실제 물리적인 모델이다. 입력 트래픽들은 L1, L2의 데이터링크 계층을 거쳐서 IP의 입력버퍼에 저장된다. IP프로세스를 거친 후 GTP 계층으로 전송이 되는데 GTP는 응용프로그램이기 때문에 트래픽들은 GTP 응용프로그램상의 버퍼에 임시 저장된다. 다음 GTP에서 UDP로 처리되기 위해 UDP입력버퍼에 저장되는데 이는 실제 큐가 아닌 메모리 버퍼(buffer)이다. 역시 UDP에서 다시 IP로 처리되기 위해 메모리 버퍼에 저장된 후 IP를 거쳐 ATM으로 전송된다.

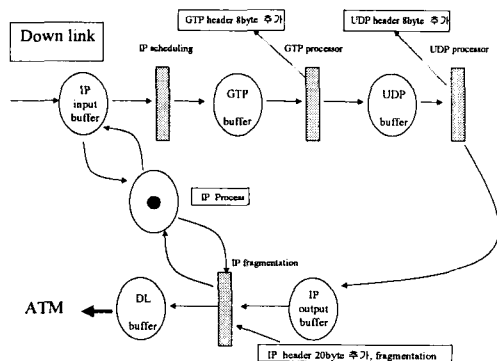


그림 4-2 GGSN 노드의 성능 분석을 위한 논리적 모델

그림 4-2는 GGSN 노드의 성능을 분석하기 위한 논리적인 모델이다. 본 연구에서는 GGSN노드의 하향

링크만을 고려한다.

외부 망에서 데이터 링크를 통해 여러 가지 트래픽 (Voice, Video, E-mail, Web)들이 IP 입력 버퍼에 도착한 후 IP Process에 의해 QoS에 따른 스케줄링 과정을 거치게 된다. 여기서 IP 입력 버퍼는 클래스별로 Speech Queue, Video Queue, WWW Queue, E-mail Queue등 네 가지로 구분한다. GGSN과 SGSN사이의 터널링(tunneling)을 위하여 IP 데이터그램에 GTP (GTP Tunneling Protocol) 헤더가 붙여진다. GTP로 캡슐화(encapsulated)된 IP 데이터그램을 해당 SGSN에게 보내기 위한 절차로 UDP(Usage Datagram Protocol) 헤더와 IP 헤더가 첨가된다.

Path 프로토콜로 UDP가 쓰이는 것은 시그널링의 오버헤드가 적고 빠르게 전송할 수 있기 때문이다. 그 다음에 IP에서는 하부 계층인 Data link의 최대전송단위인 MTU(Maximum Transfer Unit)에 따라 패킷을 단편화(fragmentation)시키는 작업이 필요하다. 본 연구에서는 Data link 계층을 ATM으로 가정하여 각각의 패킷들에 대해 53바이트씩 단편화하여 전송하게 된다. GGSN의 물리적인 구조상에서 실질적인 버퍼는 IP의 입력버퍼와 출력 버퍼뿐이다. GTP, UDP 버퍼는 실제 버퍼가 아닌 커널(Kernel)상의 Mbuffer(Memory buffer)이다.

#### V. 입력 트래픽 모델[11]

지금까지 대부분의 인터넷 트래픽 연구는 백본 망에서 트래픽 모델링을 위한 연구이다. 이들 연구는 주로 백본망에서 웹 트래픽의 모델로서 이미 이전부터 패킷 데이터 트래픽을 모델링하는 데 제안되어온 Poisson Process와 Markov Modulated Poisson Process(MMPP) 같은 Poisson모델의 부적합성을 보이고 있다. 최근의 연구에 의하면 이러한 Poisson 모델이 패킷 데이터 트래픽의 몇 가지 중요한 특성을 간과하는 경우가 있는데[10], 예를 들어 exponential distribution의 memoryless 특성을 나타내지 않는 heavy-tail distributed가 ON과 OFF기간동안 발생되는 것이 그것이다. ON과 OFF의 heavy-tail distribution은 aggregated된 traffic의 self-similar behavior를 이끌어낸다. self-similarity는 패킷 데이터 트래픽의 arrival pattern을 예상 가능케 하고, 네트워크 리소스를 올바르게 준비할 수 있도록 하는데 매우 중요한 기술로 이용된다. 본 논문에서는 GPRS에서 입력 트래픽 모델-간헐적/연속적으로 발생하는 클래스별 인터넷 트래픽(Conversational class(voice), Streaming class(video), Interactive class(www), Background

class(e-mail))을 모델링 함으로써 GPRS에서 GGSN 시스템의 성능분석을 하고자 한다.

### 5.1 Speech Traffic Model

On-Off 모델이라고 가정하면 일반적으로 voice on=1초, voice off=1.35초 Speech on, off length가 지수함수를 따른다고 가정한다. Speech ON 평균 시간을 1초로 하면  $k_{von}=1$  이고 exponential한  $t_{on}$ 의 결과를 얻어낸다. 이 상태에서  $t_{on}$ 은 active period(active time)를 나타낸다. U는 0~1까지 uniform한 distribution이다.

$$e^{-(t_{on}/k_{von})} = U(0 \sim 1) \quad t_{on} = -k_{von} \log(u)$$

Speech OFF 평균 시간을 1.35초로 하면  $k_{voff}=1.35$  이고, 이 상태에서 exponential한 시간  $t_{off}$ 의 결과를 얻어낸다.

$$e^{-(t_{off}/k_{voff})} = U(0 \sim 1) \quad t_{off} = -k_{voff} \times \log(u)$$

$t_{off}$ 는 inactive period(inactive time)를 나타낸다.

### 5.2 Video Traffic Model

각각의 video 사용자에서 AR(Autoregressive) 모델을 이용하여 트래픽을 분석하였다.

$$\lambda_v(n) = a\lambda_v(n-1) + bw(n)$$

$\lambda_v(n)$ : new video 패킷의 생성율[packets/sec]

$\lambda_v(n-1)$ : old video 패킷의 생성율[packets/sec]

$\lambda_v(0) = 0$ 이라고 놓는다.

$\lambda_v(n) = a\lambda_v(n-1) + bw(n)$ 에서

$w(n)$  = white gaussian noise

$w(n)$ : 독립적인 Gaussain 확률변수  $N(\eta, \sigma_w^2) \sim$

n: video 프레임 인덱스

a, b: AR 모델의 파라미터

한 video 사용자에게서 나오는 패킷의 평균 도착율 :

$$\bar{\lambda} = \frac{b}{1-a} \eta$$

한 video 사용자에게서 나오는 패킷의 도착율의 편차 :

$$\sigma_v^2 = \frac{b^2}{1-a^2} \sigma_w^2$$

a = 0.8781 b = 0.1108, video traffic은 바로 전의 영상과 차이만큼만 보내면 되므로 auto-regressive model을 사용하면 거의 비슷하다고 볼 수 있다. Traffic의 변화는 33ms마다 발생한다고 볼 수 있다.

Voice나 다른 traffic 과 틀리게 off 기간이 없다는 특징이 있다.

### 5.3 E-Mail Traffic Model

사용자가 메시지를 전송 받을 경우, 그 메시지는 전용 E-mail server에 저장되고 E-mail server는 모든 메시지들을 사용자가 네트워크에 로그인 해서 메일을 열어볼 때까지 mailbox에 안전하게 보존한다. 사용자가 E-mail application을 실행시킬 때 메시지의 헤더 부분은 E-mail server로부터 컴퓨터로 다운로드 되는데 각각의 메시지가 다운로드된 후, 사용자는 메시지를 읽을 수 있고, sender에게 응답을 보낸다. mailbox에 도착하는 메시지들은 Poisson process에 의해서 접근될 수 있고, 각각의 E-mail 사용자는 다음 그림5-1 처럼 E-mail session이 일어나는 동안 ON/OFF traffic pattern을 발생한다.

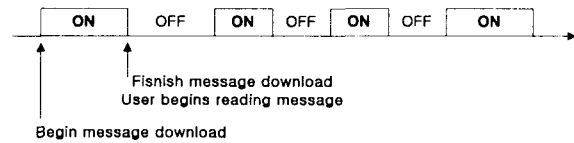


그림5-1 Email : ON/OFF Traffic model

ON period는 메시지가 다운로드 되는 시간을 나타내고, OFF period는 두개의 연속적인 메시지의 시작과 끝 사이의 간격을 나타낸다. 즉, OFF period는 사용자가 메시지를 읽는 reading time을 가리킨다.  $x_e$ 와  $t_e$ 는 각각 E-mail message 크기와 OFF period의 길이를 나타내는 iid(independent and identically distributed) random variables이고 다음수식은  $x_e$ 의 probability distribution function 이다.

$$F(x_e) = \begin{cases} 1 - e^{-e^{-k_1} x_e^{\eta}} & \text{if } F(x_e) \leq 0.5 \\ 1 - e^{-e^{k_2} x_e^{\eta}} & \text{if } F(x_e) > 0.5 \end{cases} \quad F(x_e) = U(0 \sim 1)$$

위에 식은 CDF이므로 이렇게 식을 이용할 수 있다.

$x_e$ 의 값을 구하면 channel throughput을 이용하여 ON 지속 시간을 구할 수 있다.

message length 인

$$x_e = ((-e^{k_1}) \times \log(1 - U))^{(1/c_1)}$$

$k_1 = 18.36 \quad c_1 = 2.04$

$k_2 = 3.33 \quad c_2 = 0.37$

$U = \text{uniform distribution } (0 \sim 1)$

Off period는 heavy-tailed 분포라고 가정하면 일반적으로 Pareto PDF를 사용한다.

$$\Gamma_e(t_e) = 1 - \left(\frac{k_e}{t_e}\right)^{a_e}$$

$k_e = 30$  second, OFF period의 최소 시간

$a_e = 1.5$  랜덤 변수의 평균과 편차를 결정하는 값

$$1 - \left(\frac{k_e}{t_e}\right)^{a_e} = U(0 \sim 1) \quad t_e = \frac{k_e}{(1-u)^{1/a_e}}$$

$t_e$  = off period의 시간을 말한다.

### 5.4 WWW Traffic Model

사용자가 URL 요청을 했을 때, WWW server는 사용자에게 응답을 발생시키는 client program을 실행시키고, URL에 관계된 모든 요청이 완료될 때, 사용자는 다음 요청을 시작하기 전에 정보를 읽는 시간이 필요할 것이다. 따라서 WWW traffic을 active와 inactive periods로 구성하여 모델링한다. WWW traffic의 arrival pattern은 다음 그림5-2와 같다.

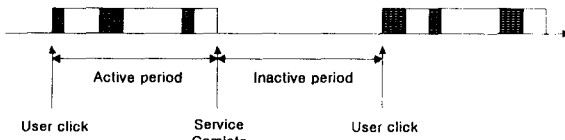


그림5-2 WWW Traffic model

Active ON(active period)은 사용자의 요청시간부터 서비스가 완료되는 시간을 말하며, 각 파일들의 전송시간과 두개의 파일을 수신할 때의 짧은 간격인 active OFF time으로 구성된다. Inactive OFF (inactive period)는 사용자가 수신한 정보를 읽는 시간(서비스 완료와 다음 서비스 요청의 시작 시간)을 말하는데, 다시 말해서, ON period는 downlink상에서 파일의 전송시간을 말하고, OFF period는 데이터 전송이 없는 시간을 말한다. 다음 수식은 ON time에서의 파일 크기와 유용한 downlink bandwidth의 함수, 즉, Pareto distribution CDF를 나타낸다.

$$F_w(x_w) = 1 - \left(\frac{k_w}{x_w}\right)^{a_w} = U(0 \sim 1) \quad t_w = \frac{k_w}{(1-u)^{1/a_w}}$$

$x_w$ 는 WWW 파일 크기를 나타내는 iid (independent and identically distributed) random variable을 가리킨다.

$k_w = 2730.67$  bytes · minimum 파일 크기

$a_w = 1.2$  랜덤변수의 평균과 편차를 결정하는 값.

Active OFF는 active와 inactive OFF times의 두 가지 segments로 구성이 되고, 다음수식과 같이

Weibull distribution Pdf를 따른다.

$$\Gamma_w(t_w) = 1 - e^{-\left(\frac{t_w}{a}\right)^b} = U(0 \sim 1)$$

$t_w$  = active off 의 시간을 나타낸다.

$a=0.382$   $b=1.46$

Inactive OFF는 Pareto distribution Pdf를 따른다.

$$\Gamma_w(t'_w) = 1 - \left(\frac{k'_w}{t'_w}\right)^{a'_w} = U(0 \sim 1)$$

$t'_w = \frac{k'_w}{(1-u)^{1/a'_w}}$   $t_w$  = active off period의 시간을 말한다.

$k'_w = 1$  second, minimum inactive OFF time

$a'_w = 1.5$  랜덤 변수의 평균과 편차를 결정하는 값.

Active on에서 off로 가는데 이때 off 의 종류가 두 개이므로 사용자 click 시점에서 파일의 개수를 구해야 service complete시점을 구할 수 있다.

Asymptotically Parato distribution이라고 가정한다.

$$F_f(x_f) = 1 - \left(\frac{k_f}{x_f}\right)^{a_f} = U(0 \sim 1)$$

$$x_f = \frac{k_f}{(1-u)^{1/a_f}}$$

$k_f$  = minimum file number

$a_f = 2.43$  랜덤 변수의 평균과 편차를 결정하는 값.

## VI. 시뮬레이션 분석 및 결과

본 장에서는 GGSN 노드의 성능 모델을 구현하여 전반적인 GGSN 노드의 성능을 분석한 후, 입력된 IP 트래픽에 대해 FCFS, FQ, WFQ등의 스케줄링 알고리즘을 적용하였다.

분석내용은 첫번째 각 클래스 별 IP 입력큐에서의 평균 지연 시간, 두 번째 실시간 서비스인 Speech와 Video 트래픽에 대해 QoS를 보장하기 위한 방법의 관점에서 분석을 하였다. 즉, WFQ 알고리즘에서 다양하게 weigh값을 조정함으로써 Speech 와 Video 트래픽의 QoS(Quality of Service)를 보장해주기 위해 시뮬레이션을 하였다.

\* 시뮬레이션 파라미터

- IP Process rate = 10Mbps

- GTP, UDP Process rate = 20Mbps

- Speech transfer speed (Kbps) : 64, Video transfer speed(Kbps) : 1500

- WWW transfer speed(Kbps) : 384, Email transfer speed(Kbps) : 384

모든 입력 버퍼와 프로세스의 버퍼는 4Mbit이다. 데이터 링크 계층을 ATM으로 가정하여 IP fragmentation 모듈에서 각각의 패킷을 53byte 크기의 cell로 자른다. IP 스케줄링 알고리즘은 FCFS와 WFQ로 비교하였다.

6.1 각 클래스별 패킷의 평균 큐 지연 시간 분석  
본 절의 시뮬레이션은 패킷의 discard를 고려하지 않은 상황에서 평균 큐 지연시간을 분석하였다.

각각의 클래스별 트래픽 사용량은 Speech 30%, Video 50%, WWW 20%, Email 10%로 가정하였다. 시뮬레이션에 입력되는 사용자 호의 수는 다음과 같다.

표 6-1 : 각 트래픽 별 사용자 호의 수

	7.6M	8.0M	8.4M	8.8M	9.2M	9.6
Speech	57	60	63	66	68	70
Video	21	22	24	25	26	27
Www	39	41	43	45	47	49
Email	1945	2048	2150	2252	2355	2457

1) FCFS

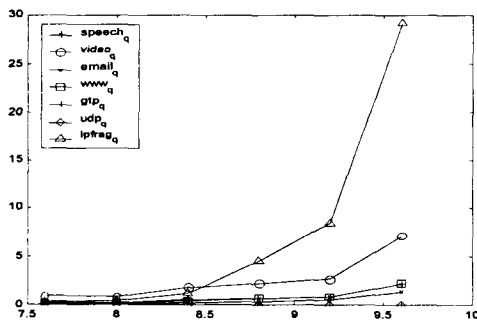


그림 6-2 FCFS일때 평균 큐 크기

그림 6-2는 FCFS일 때 평균 큐의 크기를 분석한 결과이다. 입력 트래픽 중 Video의 큐 크기가 가장 크며 나머지 입력 트래픽의 평균 큐 크기는 비슷한 양상을 보이고 있다. IP fragmentation 모드에서 패킷들을 셀(cell)로 자르기 때문에 트래픽의 부하가 증가함에 따라 버퍼의 크기가 급격히 커지는 현상을 나타내고 있다.

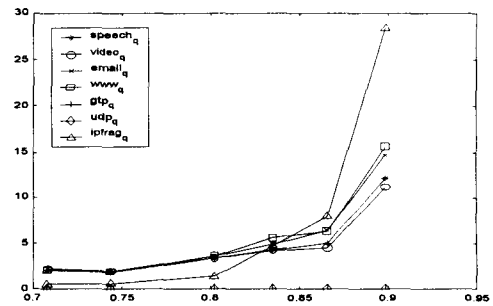


그림 6-3 FCFS일때 평균 큐 지연 시간

그림 6-3은 FCFS인 경우 IP 사용율을 증가시켰을 때 각각의 버퍼에 대해 평균 큐 지연 시간을 보여주는 결과이다. Video 트래픽의 양이 전체 트래픽의 50%를 차지하고 speech는 20%차지하기 때문에 비디오 호의 평균 큐 지연 시간이 Speech 트래픽에 비해 약간 낮은 값을 보여주고 있으며, 비 실시간 트래픽인 WWW과 E-mail 트래픽은 상대적으로 지연시간이 큰 현상을 보여주고 있다. 그림 6-2에서의 경우와 마찬가지로 IP fragmentation에서 평균 큐 지연 시간이 급격히 커짐을 알 수 있다. Voice와 Video 트래픽의 평균 큐 지연 시간이 WWW, E-mail 트래픽에 비해서는 상대적으로 낮지만 트래픽 부하가 9.5M 이상일 경우에는 보장형 서비스의 QoS 요구조건에 대해서는 보장해줄 수 없는 것으로 볼 수 있다.

2) FCFS와 FQ의 비교

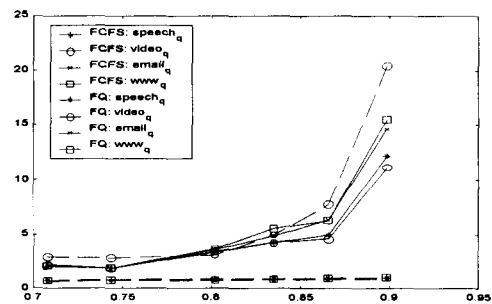


그림 6-4 FCFS와 FQ의 평균 큐 지연 시간 비교

그림 6-4는 FCFS와 FQ의 평균 큐 지연시간을 비교한 결과이다. FQ를 적용했을 때 Speech와 WWW, E-Mail의 성능은 눈에 띄게 향상되지만 실시간 트래픽인 Video의 평균 큐 지연 시간이 FCFS에 비해 급격히 나빠짐을 볼 수 있다.

이는 Video 트래픽의 양이 다른 트래픽들에 비해 많기 때문에 FQ 알고리즘을 적용할 경우, 성능이 저하되는 현상을 나타낸다.

3) FQ와 WFQ(5:5:2:1)비교

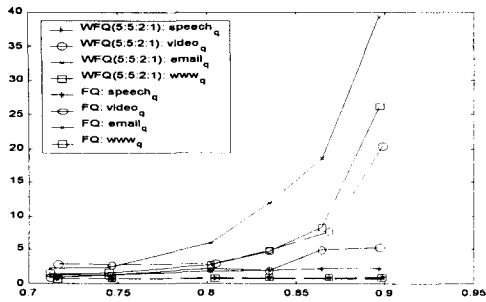


그림 6-5 FQ와 WFQ(5:5:2:1)의 평균 큐 지연 시간 비교

그림 6-5는 WFQ를 적용한 경우와 FQ를 적용한 경우의 평균 큐 지연 시간을 비교한 결과를 나타낸다. Speech, Video, WWW, E-mail 트래픽의 각각에 대한 weight는 5:5:2:1을 적용하였다. WFQ를 적용했을 때 Video 트래픽의 평균 큐 지연 시간이 FQ에 비해 상대적으로 낮아지는 현상을 보이고 있다. weight가 작은 WWW과 weight를 주지 않은 E-mail 트래픽의 평균 큐 지연 시간은 급격히 커짐을 알 수 있다. WWW과 E-mail은 비 실시간 트래픽으로 평균 큐 지연 시간에 민감하게 반응 할 필요가 없으므로 크게 고려하지 않는다.

6.2. IP 패킷의 Precedence를 고려한 weight값의 결정 방법 및 분석

IP 패킷의 Precedence를 고려한 weight값의 결정 방법과 그에 따른 분석을 하였다. 실제 참고문헌[12]에서 사용되는 IP 패킷의 Precedence를 고려했을 때의 weight 값과 Precedence에 따라 각 트래픽에 대한 bandwidth portion을 본 연구의 성능모델에 적용하였다. 각각의 클래스별 입력 트래픽의 IP Precedence와 그에 따른 weight는 다음과 같다.

표 6-3 각각의 클래스별 입력 트래픽의 IP Precedence와 weight

	precedence	Bandwidth	FQ의 weight	WFQ의 weight
Speech	5	6/14	68	405
Video	5	6/14	450	2700
WWW	0	1/14	100	100
E-Mail	0	1/14	1	1

- Bandwidth portion : 각각의 패킷은 (P+1)의 bandwidth proportion을 얻는다. 여기서 P는 패킷의 Precedence 값을 나타낸다. Precedence 값 5를 갖는 실시간 트래픽 Speech와 Video가 없는 bandwidth portion은 (5+1) / (6+6+1+1)이 된다. 비 실시간 트래

픽인 WWW과 E-Mail 트래픽은 (0+1) / (6+6+1+1)의 bandwidth를 할당받게 된다.

- FQ의 weight : FQ와 WFQ의 차이점은 패킷의 Precedence를 고려하는지에 달려있다. 대략적으로 하나의 Speech 호가 차지하는 bandwidth는 27K, 하나의 Video호가 차지하는 크기는 180K, 하나의 WWW호가 차지하는 크기는 40K, 하나의 E-Mail호가 차지하는 크기는 0.4K이다. 이때 전체적으로 차지하는 bandwidth가 가장 적은 E-Mail호의 크기로 나머지 트래픽을 나누게 되면 위와 같은 weight를 구할 수 있다.

- WFQ의 weight : WFQ에서는 각 패킷의 Precedence가 고려된다. FQ에서 구한 weight에 각 패킷의 Precedence를 곱하여 weight를 얻게 된다.

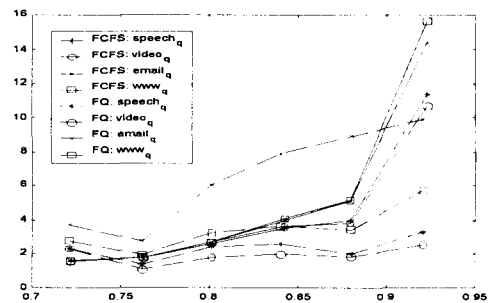


그림 6-6 IP 패킷의 Precedence를 고려했을 때 FCFS와 FQ의 평균 큐 지연시간

그림 6-6은 각각의 입력 트래픽의 Precedence를 고려 했을 때, FCFS와 FQ에 대해 각각의 입력 트래픽의 평균 큐 지연시간을 비교한 결과이다. FCFS인 경우, IP 사용율이 85%이상을 넘어서는 과부하 상태가 되며 실시간 트래픽인 Speech, Video 트래픽의 평균 큐 지연시간이 상당히 높게 나오는 현상을 볼 수 있다. FQ를 적용했을 경우 E-Mail의 지연시간이 조금 높아졌을 뿐, 나머지 트래픽들은 평균 큐 지연 시간이 상당히 낮아진 결과를 볼 수 있다.

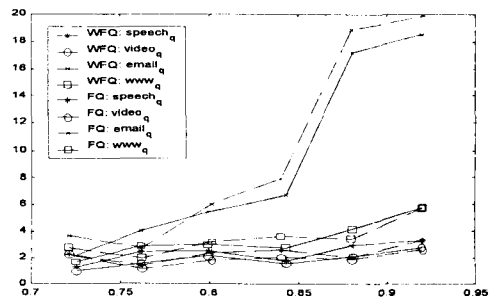


그림 6-7 IP 패킷의 Precedence를 고려했을 때 WFQ와 FQ의 평균 큐 지연시간



그림 6-7은 각각의 입력 트래픽의 Precedence를 고려했을 때, WFQ와 FQ에 대해 각각의 입력 트래픽의 평균 큐 지연시간을 비교한 결과이다. E-Mail을 제외한 나머지 트래픽의 경우 WFQ와 FQ에 대한 평균 큐 지연시간의 거의 비슷하게 나오는 결과를 볼 수 있다. 이는 Cisco사에서 적용되는 FQ의 경우 weight가 적용되기 때문이다.

## VII. 결 론

본 논문에서는 GPRS/UMTS망에서의 클래스별 패킷 트래픽 처리의 성능향상을 위해서 망의 게이트웨이 역할을 담당하는 GGSN의 IP버퍼에 대한 클래스별 트래픽 제어에 대해서 살펴보았다.

(1) 각 클래스별 평균 큐 지연시간 분석 : FCFS, FQ, WFQ스케줄링 기법을 이용하여 평균 큐 지연시간에 비교 분석하였다. FCFS일 경우에는 실시간 트래픽에 대해서 트래픽 부하가 심할 경우에 QoS 요구사항을 만족시키지 못하고, FQ일 경우에는 Speech는 월등히 QoS 요구사항을 만족시키면서 좋아지지만 Video는 반대로 현저히 평균 큐 지연시간이 증가하는 것을 보았다. WFQ일 경우, 실제 환경에서는 실시간 트래픽의 사용량의 변화에 따라 적절히 weight값을 변경 시켜야 할 것이다.

(2) IP 패킷의 Precedence를 고려한 weight값의 결정 방법 및 분석 : 실제 Cisco사에서 사용되는 IP 패킷의 Precedence를 고려했을 때의 weight 값의 결정과 Precedence에 따라 각 트래픽에 대한 bandwidth portion을 본 연구의 성능모델에 적용하였다.

## 참고논문

[1] Harri Holma and Antti Tosksla, WCDMA FOR UMTS, WILEY, 2000, pp 9-23  
 [2] 3G TS 22.060, General Packet Radio Service (GPRS); Service description, Stage 1, 2001.  
 [3] 3G TS 29.060, General Packet Radio Service (GPRS); Service description, Stage 2, 2001  
 [4] 3G TS 29.060, General Packet Radio Service (GPRS); GPRS Tunneling Protocol(GTP) across the Gn and Gp Interface, 2001  
 [5] DTR/SMG-UMTS 23.925, Special Mobile Group(SMG) Universal Mobile Telecommunication System (UMTS) UMTS Core Network based on ATM Transport, 2001

[6] 3G TS 23.107, QoS Concept and Architecture, 2001.  
 [7] Behrouz A. Forouzan, TCP/IP, McGraw-Hill, 2000, pp 166-171, pp 248-252  
 [8] Geoff Huston, Internet Performance Survival Guide, WILEY, 1999, pp191-209  
 [9] Farooq M. Anjum, and Leandros Tassiulas, Fair Bandwidth Sharing among Adaptive and Non-Adaptive Flows in the Internet, IEEE INFOCOM99, Vol.3, 1999, pp1412~1420.  
 [10] Hyekeun Park, Choongho Cho, Hyongwoo Lee, Internet-Traffic modeling in the access network 정보과학회지 17권 3호, 1999  
 [11] Joseph Ho, Yixin Zhu, and Seshu Madhavapeddy, Throughput and Buffer Analysis for GSM General Packet Radio Service(GPRS) WCNC, 1999  
 [12] Srinivas Vegesna, IP Quality of Service, Cisco Systems, 2001, pp 77-83