

액세스 스위치단에서의 Triple Play Service 를 위한 멀티캐스팅 라우팅과 QoS 기능 구현에 관한 연구

이정옥*, 최지연, 범순균
(Lee Jeoung-Ok, Choi Jee-Yeoun and Beom Soon-Kyun)

Abstract : 현재 통신 시장은 보다 편리하고 다양한 서비스를 찾는 소비자의 욕구와 맞물려 TPS(Triple Play Service) 방향으로 급속히 나아가고 있다. 인터넷 전화의 보급은 이미 보편적 서비스가 된 초고속인터넷 서비스와 연계하여 빠르게 확산되고 있으며, 2007년 들어 방송서비스까지 결합한 TPS에 대한 통신사업자간 경쟁이 본격화되고 있는 추세이다.

트리플 플레이 서비스(Triple Play Service; 이하 TPS)는 초고속 인터넷과 인터넷 전화와 방송 서비스를 하나의 회선으로 제공하는 종합 인터넷 서비스로, 광대역 통합망(BcN; Broadband Convergence Network)의 핵심 서비스이다. 액세스 스위치(Access Switch) 단은 TPS로 인해 더욱 많은 역할을 담당할 것으로 전망되는데, 디지털 방송을 인터넷 프로토콜을 통해 전송하기 위한 멀티캐스트 기능과 혼재된 이종간의 디지털 신호를 각 서비스의 품질에 맞게 보장해 주는 QoS(Quality of Service) 기능이 무엇보다 중요하게 고려 되어야 한다. 본 논문에서는 TPS의 핵심일 수 있는 방송 서비스를 위해 액세스 스위치 단에서 제공해야 할 멀티캐스트와 QoS 기능 구현 방안에 대해 제시한다

Keywords: Triple Play Service, IP Multicast, QoS, BcN, VoIP, IPTV

I. 서론

TPS는 단방향의 방송망과 협대역의 음성 통신망과 초고속 인터넷을 한꺼번에 제공할 수 있는 양방향 광대역의 “방송 · 통신 융합망” 으로의 발전을 의미한다. 이는 또한 전화(음성), 디지털방송(비디오), 인터넷(데이터)의 각기 다른 신호가 인터넷 프로토콜(IP)로 통합됨을 의미한다.

실시간 서비스인 음성 통신과 방송 서비스를 제공하기 위하여 액세스 스위치 입장에서는 TPS를 지원하기 위해서는 멀티캐스트와 QoS 기능이 필요하고 장기적으로 IPv6와 end-to-end QoS를 위한 MPLS (Multiprotocol Label Switching)의 지원 등이 요구된다.

본 논문에서는 당장 제공되어야 할 멀티캐스트와 QoS를 위주로 기술을 설명하고, 2장에서는 TPS를 위해서 액세스 스위치에서 고려되어야 할 사항에 대해 정리했다. 3장에서는 액세스 스위치의 멀티캐스트 기능과 구현에 대해서 정리하고, 4장에서는 TPS를 위한 QoS 기능 및 구현 방식에 대해 정리한 후, 결론을 맺는다.

II. TPS에서 액세스 스위치의 요구사항

인터넷 전화, IPTV, 초고속 인터넷이라는 TPS를 제공하기 위하여는 각 트래픽의 등급 별 품질관리가 가장 핵심이다. 인터넷 전화 트래픽은 신호의 특성상 실시간성이 요구되는 음성 통신으로 최소의 패킷 손실/지터를 요구한다. VoIP는 IP 프로토콜을 사용하여 음성 신호를 전달하는 IP 전화 기술이다. 오늘날의 VoIP 기술은 충분한 대역폭과 QoS를 통해 전화

서비스의 품질 요구 사항을 만족시킨다.

IPTV는 “요구되는 수준의 QoS/QoE(Quality of Experience), 정보보호, 상호작용, 신뢰성 제공을 처리하는 IP 기반 네트워크상에서 전달되는 텔레비전/비디오/오디오/텍스트/데이터와 같은 멀티미디어 서비스”이다.[2] 인터넷과 텔레비전의 융합이라는 점에서 디지털 컨버전스의 한 유형이라고 할 수 있다. IPTV를 이용하기 위해서는 텔레비전 수상기와 셋톱박스과 인터넷 회선이 필요하며, IP 기반의 네트워크를 이용하여 가입자에게 다양한 콘텐츠 및 기존 TV 프로그램을 전용 수상기나 셋톱박스가 연결된 TV 수상기로 전송한다. IPTV는 유니캐스트와 멀티캐스트, 두 종류의 전송 방식을 사용한다. 영화, 드라마의 재방송, 원격 교육처럼 수신자들이 각각 다른 시간에 각각 원하는 콘텐츠를 시청할 수 있는 VoD 서비스는 서버와 클라이언트 간에 일대일 연결을 통한 유니캐스트로 전달되고, 뉴스나 스포츠 경기처럼 다수의 수신자들에게 동일한 시간에 동일한 콘텐츠를 제공하는 생방송 또는 실시간 지상파 방송 서비스는 동일한 데이터를 다수의 수신자들에게 전달할 수 있는 멀티캐스트로 전달된다.

차세대망이 추구하는 궁극적인 액세스 망은 디지털 홈의 다양한 멀티미디어 서비스 제공, 수용 거리의 한계 극복, 양방향성 보장, 보안과 품질 보장 등 대 국민의 사용 편리성 증진과 보편적 이용을 위해 FTTH 기반의 통신/ 방송 융합망이다[7].

결론적으로 액세스망은 단일 인프라를 이용하여 모든 디지털 서비스 제공을 목표로 물리적으로 50~100Mbps급 통합서비스를 유연하게 제공하는 광대역 네트워크이어야 하며, 국민의 사용 편의성을 추구하는 홈네트워크와 기간네트워크를 품질보장성

* 책임저자(Corresponding Author)

이정옥, 최지연, 범순균 :TIMO Technology 네트워크 연구소
(jolee@timo.co.kr, darby@timo.co.kr, skbeom@timo.co.kr)

광대역 전달을 통해 연계해주어야 한다. 이외에 사용자 보호기능으로 단계별 인증, VPN, 암호화 등을 담당하며, 정보화 역기능 차단기능으로 해킹, 바이러스, 비 건전성 정보 등의 유입을 방지함을 동시에 차별화된 서비스를 제공해야 한다[7].

여기서 TPS의 관점에서 추가적으로 액세스 스위치에 요구될 수 있는 중요한 기술적 요구 사항은 실시간 지상파 방송 수용을 위한 IP 멀티캐스팅 기능이다.

III. TPS 를 위한 멀티캐스트 구현

TPS 이전 인터넷 방송 서비스는 여러 수신자를 대상으로 하는 다자간 통신 서비스 임에도 불구하고, 현재 대부분의 인터넷 방송 전송시스템이 유니캐스트 방식에 의존하고 있다. 이는 네트워크에서의 대역폭 및 송신시스템 장비의 이용 측면에서 비효율적이며 동시 접속자 수 측면에서도 한계를 지닌다. 또한, 유니캐스트 트래픽의 중복 전송으로 인해 네트워크의 과부하 및 병목 현상을 유발할 수 있으며, 이로 인해 전체적인 서비스 품질 저하를 초래하고 있다[8]. 따라서 디지털 방송 서비스를 제공하는 TPS에서도 멀티캐스트 기술은 필수이다.

이러한 멀티캐스트 기술은 IETF (Internet Engineering Task Forum)의 중점 표준화 분야중의 하나로써 차세대 인터넷의 핵심 기술이다. IP 멀티캐스트 관련 프로토콜은 크게 멀티캐스트 트리를 만들기 위한 라우팅 프로토콜과 링크 상의 멀티캐스트 수신자를 관리하기 위한 멤버십 관리 프로토콜로 나눌 수 있다. 지금까지 제안되어 온 라우팅 프로토콜을 살펴보면 소스 중심 트리방식(Source Based Tree)인 DVMRP(Distance Vector Multicast Routing Protocol)[9], MOSPF(Multicast Extensions to OSPF)[10] 등이 있으며 중심점 트리 기반 방식(Core Based Tree)으로는 PIM-SM(Protocol Independent Multicast-Sparse Mode)[11] 및 CBT(Core Based Tree)[12] 등이 제안되고 있다. 이 중에서 PIM-SM은 실제로 서비스에 적용하기 위해 가장 많이 시도되고 있는 프로토콜이다.

멤버십 관리 프로토콜로는 IGMP(Internet Group Management Protocol)가 제안되어 사용되고 있다. 지금까지 IGMPv1[13], IGMPv2[14], IGMPv3[15]이 제안되어 있다. IPv6기반에서는 MLD(Multicast Listener Discovery) 프로토콜이 제안 되고 있으며 현재까지 MLDv1[16], MLDv2[17]가 제안되어 있다.

1. 멀티캐스트 라우팅 프로토콜 - PIM-SM

PIM-SM 프로토콜의 기본적인 기능은 송신자가 전송하는 멀티캐스트 패킷을 멀티캐스트 그룹의 멤버들에게 전달하기 위한 멀티캐스트 트리 경로를 설정하는 것이며, 타 멀티캐스트 라우팅 프로토콜(PIM-DM, DVMRP, MOSPF 등)에 비해서 멀티캐스트 그룹의 멤버들이 네트워크 상에 널리 간헐적으로 퍼져

있을 경우에 망 자원을 효율적으로 이용하도록 설계되었다. "Protocol Independent"는 RIP, OSPF 등 어떤 유니캐스트 라우팅 프로토콜이 제공하는 경로 정보도 MRIB를 통해 이용할 수 있다는 의미이다.

PIM-SM은 유니캐스트 라우팅 테이블을 직접 이용하거나, BGMP[18]와 같은 다른 멀티캐스트 라우팅 프로토콜이 제공하는 정보를 이용하여 MRIB를 구축한다. MRIB는 Join/Prune 메시지가 전달되는 인접 라우터에 대한 정보(RPF neighbor)를 제공하며, 실제 멀티캐스트 패킷은 Join 메시지의 역경로(reverse path)를 통해 전달된다. 다시 말해, 유니캐스트 RIB가 각 목적지 서브넷(subnet)으로 향하는 next-hop 정보를 제공함에 반해, MRIB는 멀티캐스트 패킷의 소스로 향하는 역경로 정보를 제공한다.

PIM-SM의 특징을 간략히 요약하면 다음과 같다. 송신 라우터는 멀티캐스트 그룹의 멤버십에 대한 아무런 정보나 신호절차 없이 패킷 전송을 시작하며, 해당 패킷을 수신하고자 하는 라우터는 해당 멀티캐스트 그룹에 대해 upstream방향 인접 라우터로 Join 메시지를 전송한다(Receiver-initiated Join). DVMRP, PIM-DM등 기존 멀티캐스트 라우팅 프로토콜과의 차이점은 명시적인 Join 메시지를 전송한 멤버로만 멀티캐스트 패킷을 전송한다는 점이다. 반면에 DVMRP 등은 멀티캐스트 패킷을 일단 모든 라우터로 플러딩하고, 수신을 원치 않은 라우터로부터 Prune 메시지를 수신하면 해당 라우터로 패킷 전달을 중단한다(Broadcasting and Pruning).

PIM-SM은 라우터-라우터간의 프로토콜이므로, 호스트는 아무런 업그레이드도 필요없으며, 라우터-호스트간에는 IGMP 프로토콜만 동작하면 된다.

PIM-SM은 소스별 트리(source-specific tree)와 공유 트리(shared tree)를 모두 지원한다. 공유 트리는 RP(Rendezvous Point)를 루트로 하는 멀티캐스트 트리를 의미하며, 모든 송신 라우터는 멀티캐스트 패킷을 RP로 전송하고, RP는 수신된 패킷을 해당 그룹의 각 멤버들에게 공유 트리를 통해 포워딩한다. 여기서 "공유"라 함은 특정 그룹으로 패킷을 전송하는 모든 송신 라우터들이 동일 멀티캐스트 트리를 사용한다는 의미이다. 소스별 트리는 송신 라우터 별로 독립적인 멀티캐스트 트리를 사용하며, 대부분 송신 라우터와 각 멤버 라우터를 연결하는 최단경로들로 구성된다. 멀티캐스트 그룹에 참여하고자 하는 PIM-SM 라우터는 RP를 통해 먼저 공유 트리인 RPT(RP Tree)에 연결된 후, 필요에 따라 (예를 들어, 소스로부터의 전송지연이 너무 클 경우) 리프 라우터(last hop router)는 소스별 트리인 SPT (Shortest Path Tree)에 참여하고 기존 RPT에서는 탈퇴할 수 있다.

네트워크 상황이나 동적인 그룹 멤버십 변화를 적절히 수용하기 위해 soft-state 기법을 사용한다. "soft-state"라 함은 PIM-SM 라우터의 멀티캐스트 상태 형상 정보는 일정 시간이 지나면 소멸되며, 따라서 해당

정보는 주기적으로 갱신(refresh)되어야 한다. Soft-state를 지원하기 위해서는 다수의 타이머들이 사용되어야 한다.

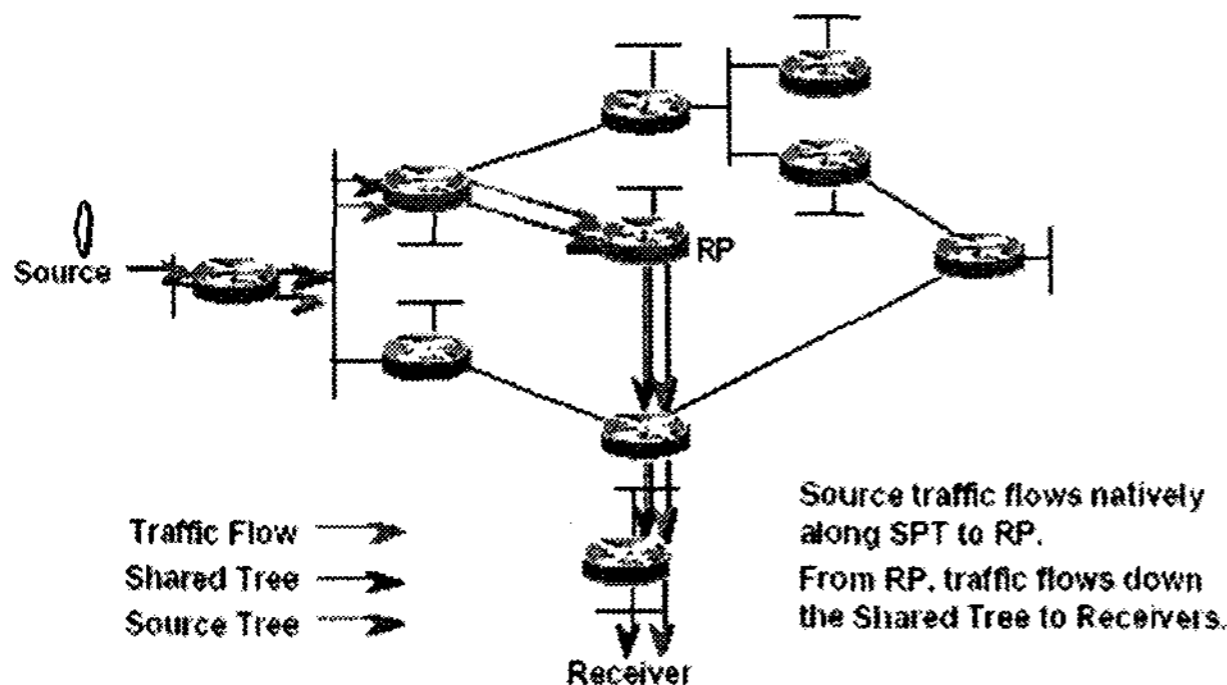


그림 1. 멀티캐스트 프로토콜 동작 방식
Fig 1. Multicast protocol operating method

2. 멀티캐스트 멤버십 관리 프로토콜 - IGMP

멀티캐스트는 IGMP (Internet Group Management Protocol)를 이용해 그룹의 멤버들을 관리한다. 멀티캐스트 그룹에 가입하고자 하는 수신 호스트는 멀티캐스트 주소를 토대로, IGMP 프로토콜을 이용하여 네트워크 라우터에 자신을 바인딩한다. 네트워크 라우터는 IGMP를 토대로 자신의 서브넷에 그룹 멤버가 있음을 알게된다. IGMP는 하나의 라우터와 여러 호스트로 구성되는 서브넷 안에서 라우터가 자신의 호스트들이 어떤 멀티캐스트 그룹에 관심이 있는지를 파악하기 위한 일종의 시그널링 (signaling) 프로토콜이다.

IGMP 프로토콜은 IGMP Query와 IGMP Report 두 가지의 제어메시지를 기반으로 동작한다.

라우터는 주기적으로 서브넷 호스트들에게 IGMP Query 메시지를 전송한다. IGMP Query 메시지에 응답하여, 호스트는 자신이 가입하고자 하는 그룹주소 G 정보를 Report 메시지를 통해 라우터에게 알려준다. 호스트는 Query 메시지 없이도 먼저 Report 메시지를 라우터에게 전송할 수 있다. [그림 2]은 위의 두 가지 메시지를 이용하는 IGMP 동작 메커니즘을 보여준다.

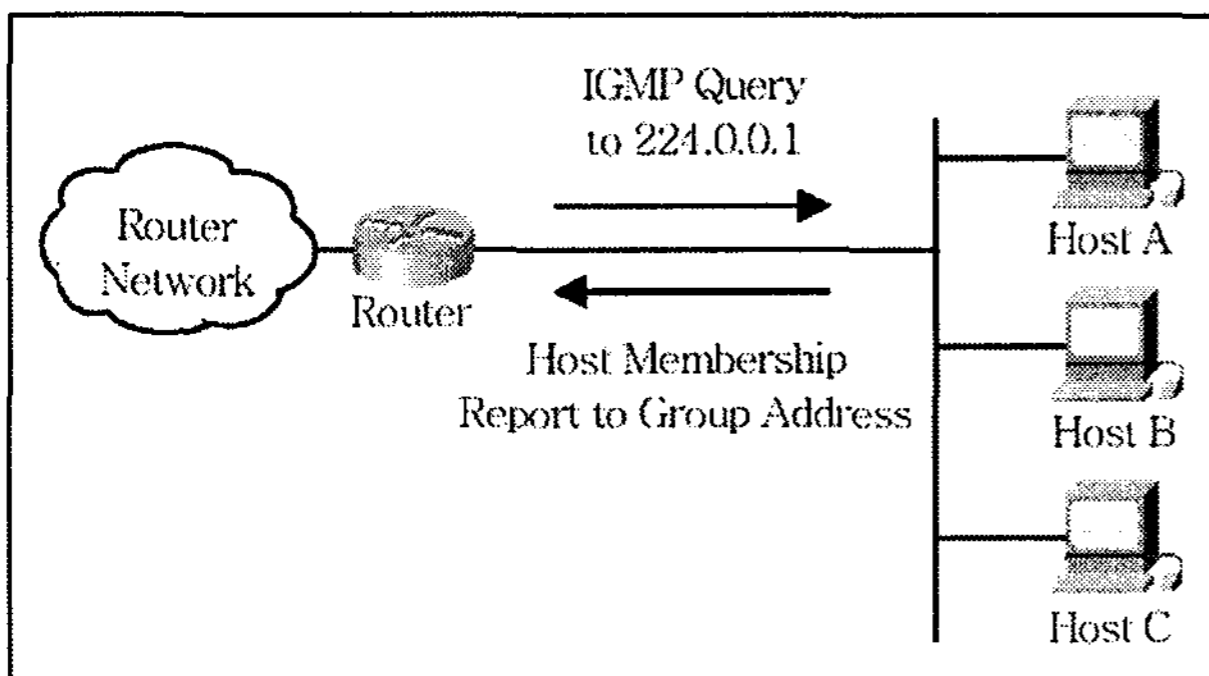


그림 2. IGMP 동작 방식
Fig 2. IGMP operating method

IGMP에서 주목할 것은 라우터의 입장에서 그룹 G에 대한 호스트 목록을 유지할 필요가 없다는 점이다. 즉, 해당 서브넷이 그룹 G에 대한 멀티캐스트 수신을 원하는 지 여부만 판단하고 구체적으로 어떤 호스트들이 그룹 G에 가입했는지는 알 필요가 없다. 그룹 G의 데이터 수신을 원하는 경우, 라우터는 네트워크 트리를 통해 유입되는 멀티캐스트 데이터를 자신의 서브넷으로 포워딩한다.

IGMPv1 및 IGMPv2와는 달리, IGMPv3에서는 그룹주소 G뿐만 아니라 송신자의 IP 주소 'S'를 IGMP 메시지에 포함시킨다. 이를 통해, 소위 'Source Filtering' 기능을 제공하며, 각 수신자는 같은 그룹주소 G에 대해서도 특정 Sources가 보내는 데이터만을 수신할 수 있다.

3. 액세스 스위치에 구현된 멀티캐스트 기술

일반적으로 고속의 Ethernet Switching Chip을 기반으로 개발되고 있는 액세스 스위치에서 멀티캐스트 기능은 우선 고려해야 할 것들이 있다. 이는 멀티캐스트 프로토콜들의 표준화 작업이 소프트웨어에 의해 패킷이 처리되는 것을 가정하고 있기 때문이다. 또한 이들 프로토콜들은 라우터를 기반으로 동작하도록 설계되어있어 Layer 3 계위의 인터페이스 개념만 존재할 뿐, 이더넷 스위치에서 사용되고 있는 VLAN(Virtual Local Area Network) 등 Layer 2 계위의 여러 개념들은 고려하고 있지 않기 때문이다.

따라서, 액세스 스위치에서 멀티캐스트 기능을 구현하기 위해서는 사용되는 Ethernet Switching Chip이 지원하는 멀티캐스트 기능과 Layer 2 계위의 프로토콜들을 잘 파악하여 설계하여야 한다. 액세스 스위치에서 멀티캐스트 기능을 구현할 때 고려해야 할 사항들은 다음과 같다

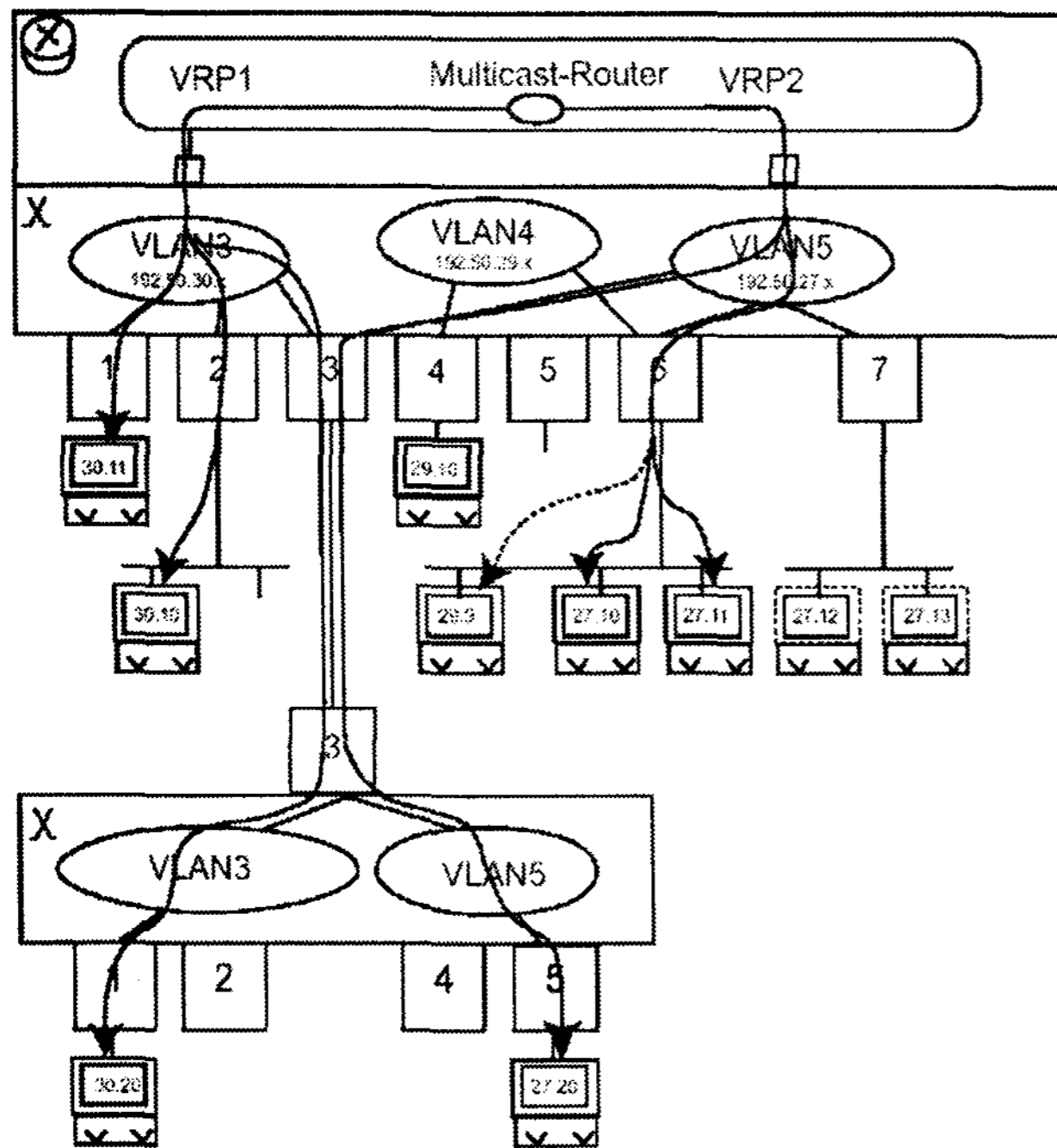


그림 3. VLAN과 IP 멀티캐스트 동작 방식
Fig 3. VLAN and IP multicast operating method

먼저, IGMP Snooping 기능이 고려되어야 한다. 멀티캐스트 그룹의 멤버십 관리를 맡고 있는 IGMP 프로토콜은 앞서 살펴본 바와 같이 여러 물리적인 포트가 하나의 인터페이스로 구성되는 VLAN 환경에서는 적절하지 않다. 이를 보완하기 위하여 IGMP Snooping이 필요하다. IGMP Snooping은 VLAN내의 모든 멤버 포트에게 멀티캐스트 트래픽을 전송하지 않고, 멀티캐스트 트래픽이 필요한 포트들에게만 전달될 수 있도록 멀티캐스트 그룹을 동적으로 추가/삭제함으로써 모든 포트에 멀티캐스트 트래픽이 Flooding 되는 것을 막는 기능을 수행한다. 이를 위해 IGMP Snooping은 포트를 기반으로 하는 멀티캐스트 그룹 테이블을 만들어 관리한다. 특정 멀티캐스트 그룹에 대한 IGMP Join 메시지를 받으면, 관련된 멀티캐스트 포워딩 테이블 엔트리에 그 호스트가 연결된 포트를 더하고, 호스트로부터 IGMP Leave 메시지를 받으면 반대로 그 호스트가 연결된 포트를 테이블 엔트리에서 제거한다. 또한, 멀티캐스트 라우터로부터의 IGMP Query를 VLAN 내의 포트들로 릴레이 한 후, IGMP Join 메시지를 받지 못한 포트들은 삭제한다.

다음으로, Ethernet Switching Chip의 사용에 따르는 몇가지 고려사항을 알아보자. 고속으로 하드웨어 스위칭이 진행되고 있는 상황에서, 제어 모듈의 프로토콜 기능 수행을 위해 Forwarding되고 있는 모든 패킷을 CPU로 전달할 수는 없다. 이는 수 Gbps ~ 수십 Gbps의 패킷을 CPU가 처리할 수 있는 능력을 갖추어야 된다는 의미이다. 따라서 멀티캐스트 패킷에 대해 와이어 스피드를 제공하기 위한 하드웨어 스위칭과 CPU의 부하를 최소화하는 제어 모듈의

프로토콜 수행이 동시에 만족할 수 있도록 제어하는 기술이 중요하다. 이러한 기술은 프로토콜들이 꼭 전달 받아야 하는 패킷과 시기를 최소화하도록 제어하는 것과 CPU로 전달되는 패킷의 양을 QoS 관리 기술을 통해 CPU가 처리할 수 있는 적당한 양으로 조절하는 기술을 통해 구현될 수 있다.

IGMP나 PIM-SM의 제어 메시지가 스위칭 도중에 혼잡으로 인해 버려지는 것을 방지하여야 한다. 제어 메시지의 손실은 정상적인 패킷 전달이 이루어지지 않는 결과를 낳기 때문이다. 이를 위해서 IGMP나 PIM-SM 제어 패킷에 높은 우선순위를 설정하여 송수신을 보장 한다.

PIM-SM 프로토콜이 멀티캐스트 패킷을 전달 받아야 하는 경우는 직접 접속된 소스의 멀티캐스트 패킷이나, 라우팅 테이블에 부합하는 엔트리가 존재하지 않는 멀티캐스트 패킷, (S, G)에 대한 라우팅 엔트리가 생성된 후 등록 절차가 종료될 때까지의 멀티캐스트 패킷 등이다.

PIM-SM 구현을 위해 추가적으로 고려하여야 할 사항은 (S, G)에 대한 라우팅 엔트리가 멀티캐스트 패킷 Forwarding에 의해 최초 참조되는 시점을 감지하거나, 지속적으로 라우팅 엔트리가 참조되고 있음을 모니터링하는 방법에 대한 고려이다. 그리고 (S, G)에 대해 RPT에서 SPT로 전환할 수 있는 방법도 고려하여야 한다. 즉, Ethernet Switching Chip은 (*, G)의 라우팅 엔트리에 의해 Forwarding되는 패킷의 양을 (S, G)별로 모니터링하여 PIM-SM 프로토콜이 SPT 전환을 위해 설정한 임계치와 비교할 수 있도록 하는 방법이 고려되어야 한다는 의미이다.

그리고, 멀티캐스트 패킷의 중복된 전송을 방지하기 위하여 Outgoing Interface를 통해 해당 멀티캐스트 패킷이 수신되는 경우를 감지할 수 있는 방법도 고려하여야 한다.

IV. TPS를 위한 QoS 구현

TPS를 위한 액세스 스위치는 음성 트래픽, 비디오 트래픽, 인터넷 트래픽을 분류하고 각 서비스 별로 요구되는 품질(QoS)을 보장해주어야 된다.

서비스의 품질인 QoS(Quality of Service)는 대역폭, 지연, 지터 등에 대한 지정된 값을 보장하는 end-to-end 서비스이며, 트래픽이 전달되는 모든 장비에 품질 보장을 위한 QoS 기능이 구현되어 있어야 된다. 본 논문에서는 TPS 가입자 트래픽을 입력 받아 망으로 전달하는 액세스 스위치에서 QoS를 보장하기 위하여 구현될 기능에 대하여 설명한다.

1. QoS 관련 기술

액세스 스위치는 L2 스위치와 L3라우터의 기능을 모두 가지고 있으며 L2 스위치에서 지원하는 802.1p 기능과 L3 스위치 기능에서 지원하는 DiffServ 기능을 모두 지원한다.

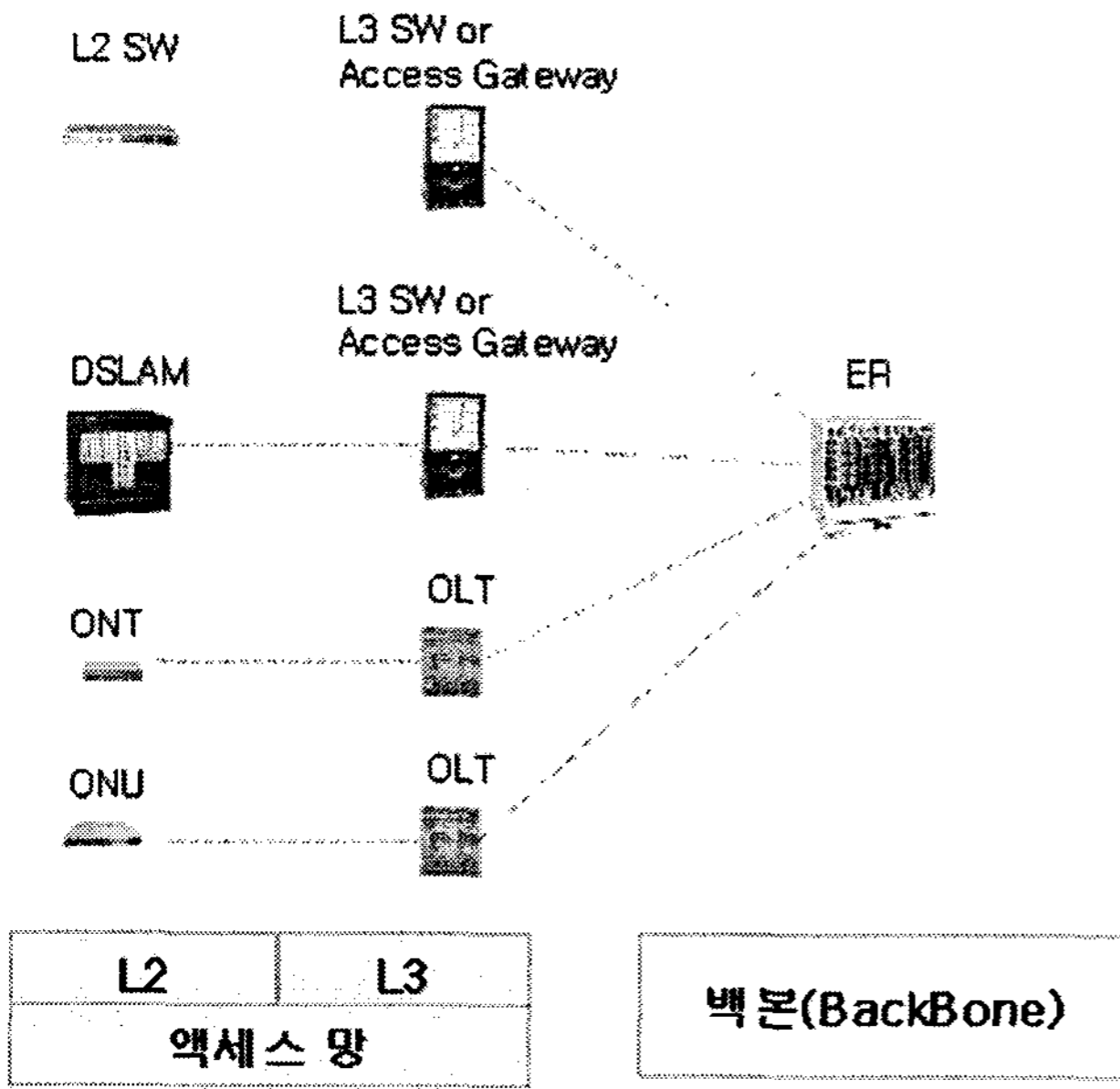


그림 4. 액세스 망
Fig 4. Access Network

1) Over Provisioning

액세스 스위치의 대역폭을 서비스에서 필요로 하는 최대 대역폭으로 할당함으로써, 모든 서비스에 최고의 품질을 제공할 수 있다. 이는 액세스 스위치에 QoS 기능의 구현 없이 서비스 품질을 만족시킬 수 있지만, 대역폭과 스위치 장비가 낭비되며, 무한하게 대역폭을 증가시키지 않는 한 대용량의 데이터 트래픽에 의하여 음성과 비디오 트래픽의 품질 저하가 발생할 수 있다.

2) CoS

CoS는 IEEE 802.1p의 우선순위에 따라 트래픽을 클래스로 분류하고 클래스별 처리순서를 관리한다. CoS는 L2 계층에서 지원하는 기능으로, 이더넷 헤더에 있는 VLAN 태그(tag)의 우선순위(priority) 필드를 이용한다. 우선순위 필드는 3비트(bit)로 0~7의 값을 가진다. 0이 우선순위가 가장 낮고 7이 우선순위가 가장 높다. CoS는 트래픽의 처리순서에만 관여하므로 대역폭이나 전달에 대한 보장은 할 수 없다.



그림 5. VLAN 태그
Fig 5. VLAN Tag

우선순위는 가입자 단에서 추가한 우선순위 필드에 설정된 값에 따라 정해질 수 있다. 가입자 장비에서 트래픽의 종류에 따라 우선순위를 다르게 설정하여 전달하는 경우 사용될 수 있지만, 모든 가입자 장비에 우선순위 필드를 추가하는 기능이 있는 것은 아니다.

가입자 단에서 우선순위 필드가 없는 패킷을 전달하는 경우는 액세스 스위치에서 우선순위 필드를 설정할 수 있다. 액세스 스위치는 트래픽이 입력되는 포트 별로 우선순위를 설정하거나 L2 헤더의 MAC 주소를 이용하여 우선순위를 설정할 수 있다. 이 경우 TPS를 위한 트래픽 분류가 되지 않을 수 있다.

3) QoS

QoS 관련 프로토콜 및 표준에는 MPLS, RSVP[3], IntServ, DiffServ[4] 등이 있다.

IntServ는 RSVP signaling을 이용하여 end-to-end 대역폭 및 자원을 미리 할당하는 방식으로 모든 서비스 트래픽의 품질을 보장할 수 있는 방법이지만 확장성에 제한이 있다.

DiffServ는 IP 헤더의 ToS 필드를 이용하여(IPv6는 traffic class 필드) 8비트의 ToS 필드를 6비트의 DSCP[5] 필드와 2비트의 CU 필드로 재정의한다. 여기서 DSCP 필드는 다시 3비트의 CSCP(Class Selector Code Point) 필드를 가지는데, 이 부분이 전송 우선 순위를 결정하는 부분이다. CSCP는 ToS를 위한 하위 호환성을 유지하기 위해 TOS 필드의 precedence와 위치가 같게 설계되었다. 또한, 폐기 우선 순위를 가지는 2비트 필드도 가지고 있어, 총 32개의 서비스 클래스를 가질 수 있다. 현재 DiffServ는 14개의 서비스 클래스가 정의되어 있다.

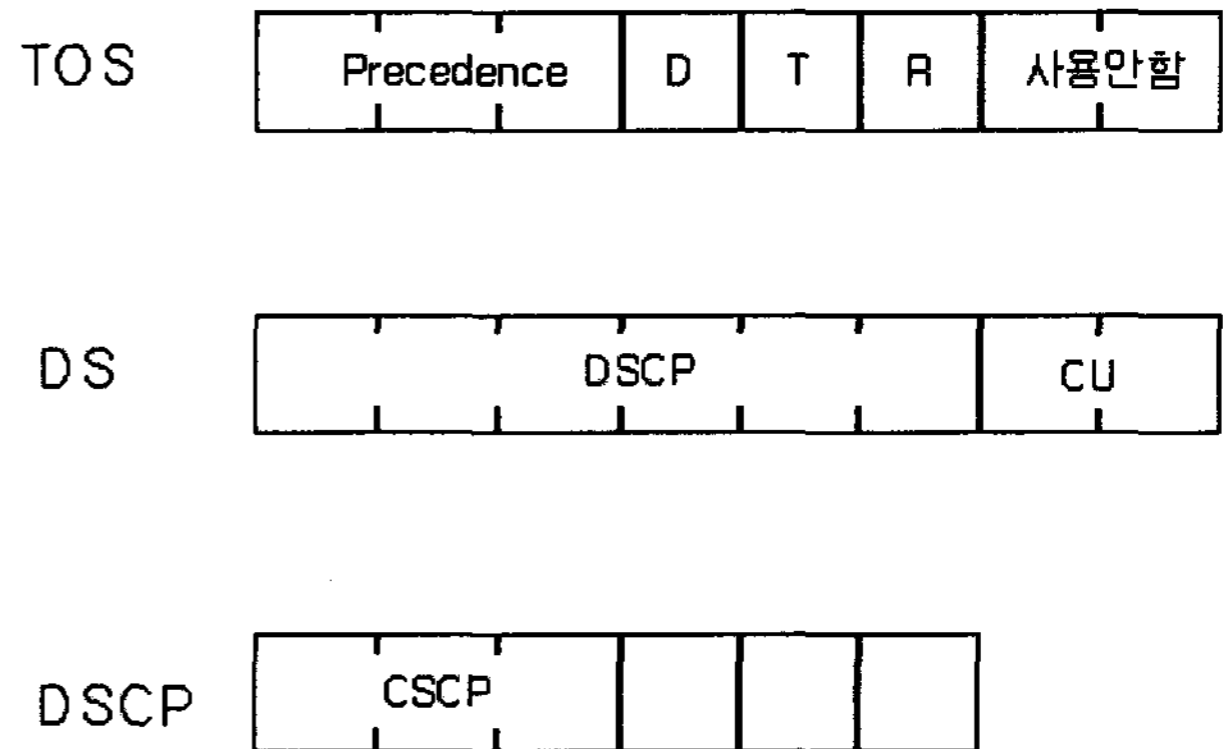


그림 6. ToS와 DSCP의 관계
Fig 6. The relationship between ToS and DSCP

IntServ가 확장성의 문제를 가지고 있어, 현재는 DiffServ가 사용되는 추세에 있다. 그러나, DiffServ만으로는 end-to-end QoS 보장에 취약점을 가지고 있기 때문에, 앞으로의 망 진화는 IntServ의 end-to-end QoS 보장 개념이 추가되거나, MPLS 기반의 QoS가 제공될 것이다

2. 액세스 스위치에 구현된 QoS 기술

현재 액세스 스위치에 구현된 QoS 기술은 IEEE 802.1p와 DiffServ를 지원하기 위한 기능으로, 우선순위 할당 및 우선순위에 따른 처리 기능과 패킷 분류 기능, 트래픽 조절 기능, 스케줄링 기능이 있다.

1) 우선 순위 설정 및 처리

액세스 스위치에 입력되는 모든 패킷은 우선순위 설정 규칙에 따라 우선순위가 설정되고, 스위칭이나 라우팅에 의하여 출력 포트가 결정된다. 출력 패킷은 우선순위에 따라 출력 포트의 큐에 각각 입력되고, 큐의 우선순위에 따라 처리순서가 정해진다. 액세스 스위치는 포트별 최대 8개의 큐가 있고, 0번 큐가 우선순위가 가장 낮고 7번 큐가 우선순위가 가장 높다. 즉 우선순위가 높은 7번 큐에 입력된 패킷이 먼저 처리된다. 출력되는 패킷에 우선순위 필드를 추가하는 경우, 출력 큐의 번호가 우선순위 필드에 설정된다.

액세스 스위치에 입력되는 패킷에 우선순위 필드가 있는 경우는 우선순위 필드의 값이 그 패킷의 우선순위가 된다. 우선순위 필드가 없는 패킷이 입력되는 경우는 3가지 방법으로 패킷의 우선순위를 결정한다. 첫째는 패킷이 입력되는 포트에 우선순위를 할당하여 그 포트에 입력되는 모든 패킷에 동일한 우선순위를 설정한다. 둘째는 맥 주소에 우선순위를 할당하고, 목적지 맥주소가 일치하는 패킷에 그 우선순위를 설정한다. 셋째는 DSCP 필드에 우선순위를 할당하고 패킷의 DSCP 필드 값에 따라 해당 우선순위를 설정한다.

2) 패킷 분류(Classification)

패킷 분류 기능은 트래픽 조절 기능의 전처리 단계이며, 패킷을 특성별로 분류하는 기능이다. 일반적으로 입출력 포트, L2, L3 및 L4의 헤더 정보에 따라 서비스 및 사용자 별로 패킷을 분류한다. 분류된 패킷은 새로운 우선순위나 DSCP 값이 할당된다.

패킷 분류 시 참조하는 L2 헤더의 필드는 목적지 맥 주소, 출발지 맥 주소, 이더넷 타입, VLAN ID, 우선순위 필드 등이다. 패킷 분류 시 참조하는 L3 헤더 필드는 목적지 IP 주소, 출발지 IP 주소, DSCP, 프로토콜 타입 등이다. 패킷 분류 시 참조하는 L4 헤더 필드는 목적지 TCP/UDP 포트, 출발지 TCP/UDP 포트, TCP flag 등이다.

3) 트래픽 조절(Traffic Conditioning)

트래픽 조절기는 액세스 스위치에 입력되는 트래픽이 사전에 정의된 트래픽 프로파일에 따르는 지를 확인하고, 위반 트래픽에 대하여 다른 QoS 처리를 한다.

패킷 분류기에 의하여 분류된 패킷은 미터 (meter)에서 트래픽의 양이 측정되어 정상 트래픽과 위반 트래픽으로 나누어진다. 위반 트래픽은 마커(marker)에서 DSCP 값이 재설정되어 혼잡 시 우선 폐기되도록 한다. 셰이퍼(shaper)와 드라퍼(dropper)로 위반 트래픽을 스위칭이나 라우팅하지 않고 바로 폐기할 수 있다.

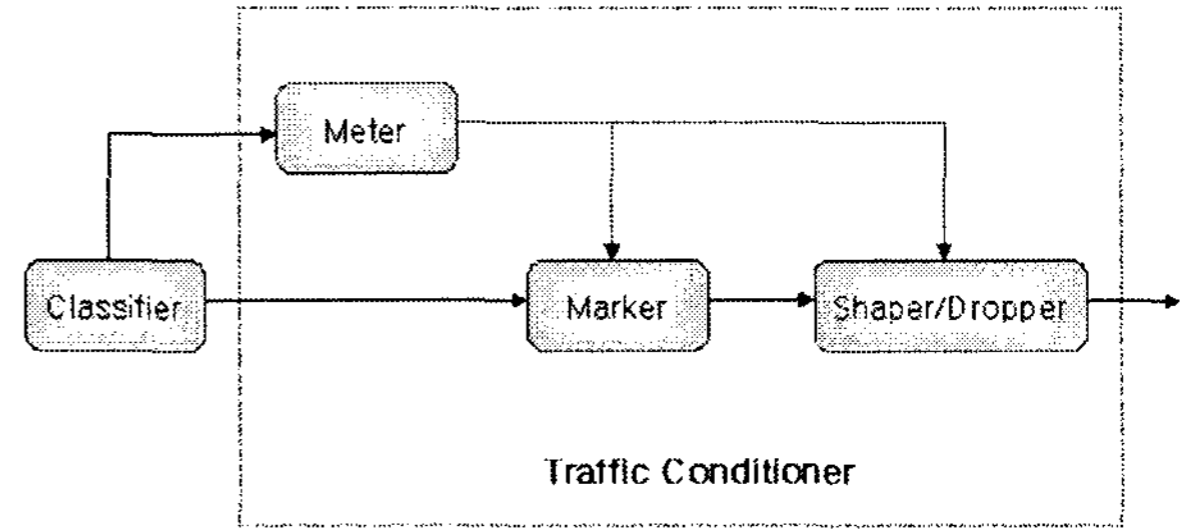


그림 7. 트래픽 조절기
Fig 7. Traffic Conditioner

셰이퍼는 버퍼링을 하여 일정 크기의 버스트 트래픽을 수용할 수 있다. 드라퍼는 버퍼링을 하지 않는 레이트 리미트(rate limit) 기능이 주로 사용된다. 셰이퍼를 이용하면 패킷 손실을 줄일 수 있다는 장점이 있는 반면, 패킷의 지연이 생겨 실시간 트래픽에는 적합하지 않다. 현재는 패킷 손실을 방지하기 위해 약간의 Over Provisioning을 레이트 리미팅에 접목시켜 사용하는 것이 일반적이다.

4) RED/WRED에 의한 혼잡 제어

RED(Random Early Detection/ Random Early Discard)는 TCP 동작 특성을 이용한 대표적인 혼잡 제어 기법으로, 혼잡 발생을 미리 예상하여 패킷의 폐기를 시작한다. 이는 TCP 플로우로 하여금 전송 속도를 줄이게 한다.

기본적인 동작은 두 개의 쓰레시홀드 TH_min과 TH_max를 두고 세 구간에서 서로 다른 드랍 확률을 적용하는 것이다. 평균 큐 크기가 TH_min 보다 작으면 모든 패킷을 전달하고, 큐 크기가 TH_max 보다 큰 경우는 입력되는 모든 패킷을 버린다. 큐 크기가 TH_min과 TH_max 사이에 있는 경우는 확률 함수 값에 따라 패킷을 버린다. 즉, 혼잡이 심해질수록 많은 패킷을 버림으로써 입력되는 트래픽의 양을 줄이는 방법이다.

WRED는 하나의 큐에 여러 개의 RED를 적용하여 동일한 큐에 할당된 트래픽이 서로 다른 RED에 따라 혼잡제어가 되도록 하는 기능이다.

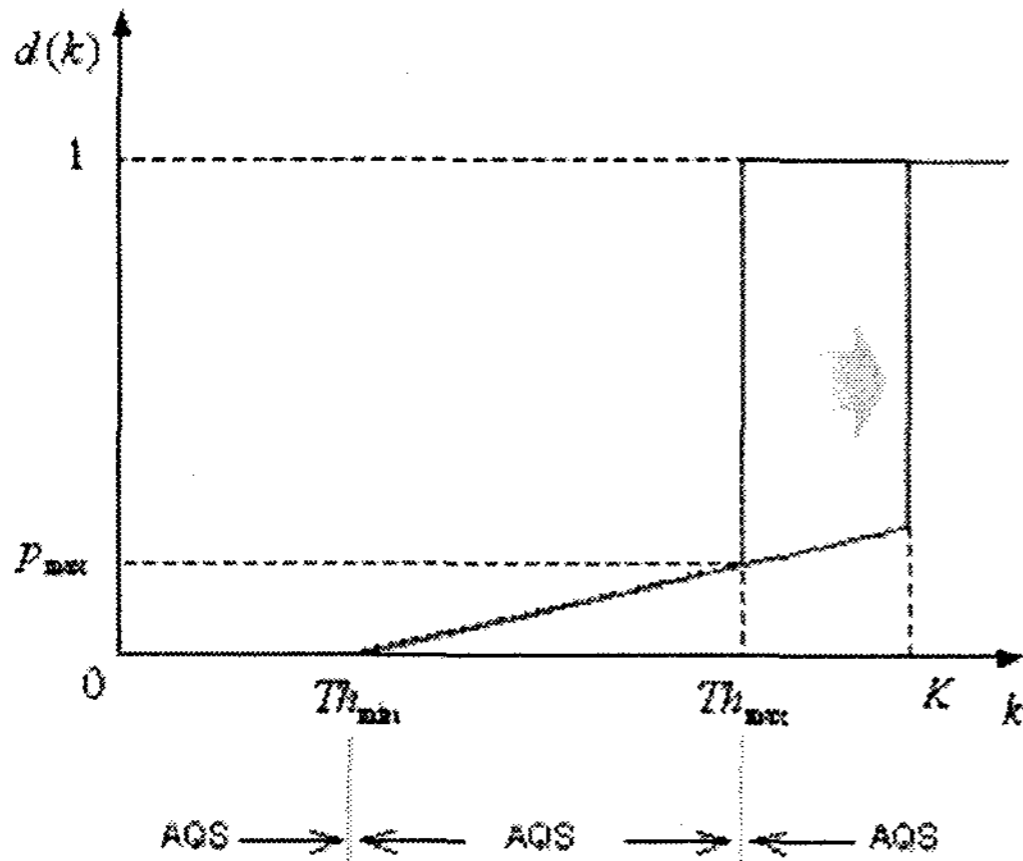


그림 8. RED 패킷 드랍 확률 함수
Fig 8. RED Packet Drop probability function

액세스 스위치의 출력 포트 별 8개 큐에 각각 RED나 WRED를 설정하여 혼잡이 예상되는 경우 특정 트래픽 플로우를 우선 폐기할 수 있다.

5) 스케줄링(Scheduling), 큐잉(Queueing)

출력 패킷을 우선 순위에 따른 처리를 위해 출력 포트에서 다수의 FIFO 큐를 이용한다. FIFO 큐가 많아지면 우선 순위에 따라 세분화된 스케줄링이 가능하다는 장점이 있는 반면 스케줄링이 복잡해진다는 단점이 있다. 출력 포트 별 FIFO 큐가 많아질 수록 스케줄링의 복잡도는 기하급수적으로 커지게 된다. 액세스 스위치는 다양한 종류의 서비스를 처리하기 위하여 최대 8개의 큐를 사용할 수 있다.

스케줄링중 우선 순위 적용을 위한 가장 간단한 방법이 SPQ(Strict Priority Queueing)이다. 이 방법은 우선 순위가 높은 패킷이 우선 순위가 낮은 패킷보다 먼저 처리되는 방식이다. 우선 순위가 높은 패킷은 전달이 확실히 보장되지만, 우선 순위가 높은 패킷이 대역폭을 모두 사용할 경우 우선 순위가 낮은 패킷은 전혀 전송되지 못한다는 단점이 있다. 이 단점을 보완하기 위해 레이트 리미팅 기능과 조합되어 사용될 수도 있다.

그 외의 대표적 방법으로 WFQ(Weighted Fair Queueing)이나, WRR(Weighted Round Robin)등이 있다. 이들 방법은 SPQ의 단점을 보완하는 방식으로 우선 순위가 높은 패킷이 많이 들어오더라도 우선 순위가 낮은 패킷이 어느 정도는 전송될 수 있도록 한 것이다.

3. TPS를 위한 QoS 적용 방안

액세스 스위치에서 제공하는 TPS의 QoS 관리를 위하여 네트워크 인터페이스의 용량에 따라 서비스별 용량을 제한하고, 가입자 트래픽을 제한된 용량 이내로 입력되도록 제어한다.

액세스 스위치에서는 입력된 트래픽을 분류하여 DSCP와 우선순위를 설정한다. 미터에서 각 트래픽을 측정하여 계약된 용량을 준수하는 지 관리하고 계약 위반 트래픽은 DSCP를 재설정하고, 폐기 우선순위를 높여 혼잡 시 우선 폐기되도록 한다.

스위칭이나 라우팅에 의하여 출력포트가 결정된 트래픽은 우선순위에 따라 각각의 큐에 입력된다. 입력되는 트래픽이 출력되는 트래픽 보다 많아 큐의 크기가 커지면서 혼잡이 예상되면, WRED에 의하여 폐기 우선순위가 높은 패킷이 먼저 폐기된다.

각 큐에 입력된 트래픽은 스케줄러에 의하여 처리 순서가 결정되어 출력된다. 스케줄러는 SPQ가 사용되고 실시간 트래픽인 음성 트래픽에 가장 높은 우선순위를 설정하고 대용량 버스트 인터넷 트래픽에 가장 낮은 우선순위를 설정한다. 우선순위가 높은 트래픽에 의하여 우선순위가 낮은 트래픽의 전달이 안되는 SPQ의 단점에 의하여 비디오 트래픽이 전달되지 않는 경우가 발생하지 않도록 음성 트래픽이 처리되는 큐에 셰이핑을 설정한다. 계약된 용량 이내의 음성 트래픽이 셰이핑에 의하여 폐기되지 않도록 over-provisioning한다

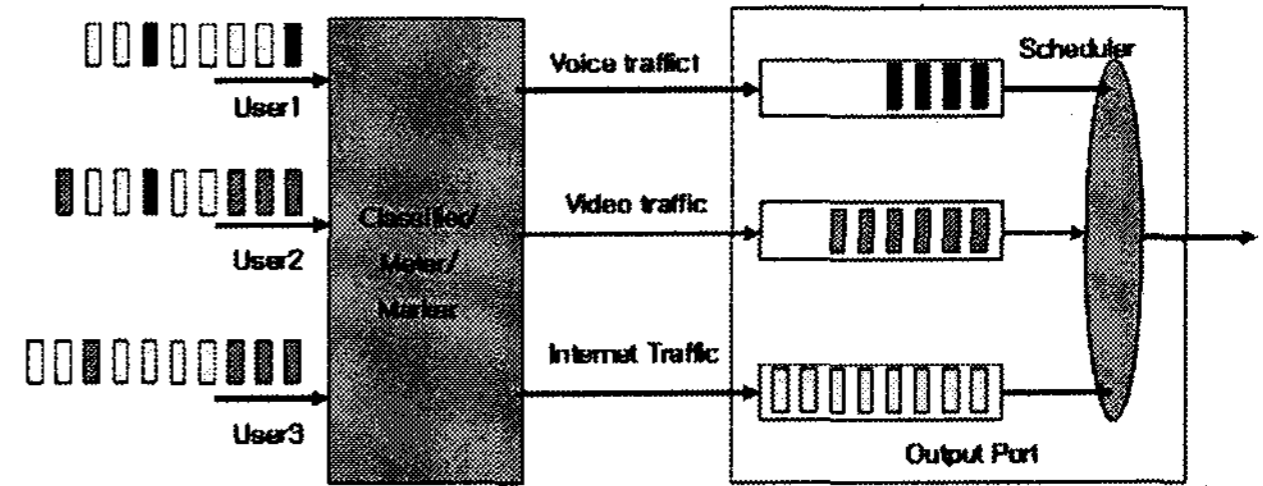


그림 9. TPS를 위한 QoS 기능
Fig 9. QoS functions for TPS

DiffServ와 IEEE 802.1p의 적용과 적절한 수준의 over-provisioning으로 상당한 수준의 QoS 제공이 가능하나 가입자 집선비와 네트워크 인터페이스의 용량과 서비스별 용량 산출 및 준수 관리 방안 등의 고려사항이 있다.

V. 결론

TPS에서 방송 서비스의 경우 구현을 위해서는 멀티캐스트와 QoS와 같이 서비스 품질을 보장할 수 있는 기술들의 적절한 적용이 무엇보다 중요하다.

앞으로 코어망은 QoS만 보장할 수 있는 간단한 전송 망이 될 것으로 전망되는 반면, 액세스 망의 입장에서 TPS를 비롯한 많은 신규 서비스들이 창출됨으로써, 액세스단 스위치의 기능은 더욱 고도화된 기능이 요구된다.

본 논문에서는 TPS 초기 적용 단계에서의 지원을 위한 액세스 스위치 구현 방법을 주로 다루었다. 하지만 HDTV와 같은 고품질 방송, 인터넷 전화, 초고속 인터넷이 서로 융합된 서비스들의 지원을 위해서는 궁극적 단계에서는 IPv6 및 MPLS의 QoS

제어 기능과 같은 기술들도 액세스 스위치 안에 적용되어야 하고, 악의적인 코드 방지를 위한 보안(Security) 기능과 인증 시스템과의 연계 기능 등도 점차 강화되어야 할 것이다.

참고문헌

- [1] 김미애 “2 단계 BcN 구축 추진 전략” TTA Journal No. 109, pp. 22 - 27, 2007
- [2] 권수갑 “IPTV 세부 기술동향” 산업자원부 보고서 2007. 7
- [3] Barden, R., Zhang, L., Berson, S., Herzog, S. and Jamin S., “Resource Reservation Protocol (RSVP) Version 1 Functional Specification”, RFC 2205, Proposed Standard September 1997
- [4] S.Blake, et. al. “An Architecture for Differentiated Service”, RFC 2475, December 1998.
- [5] Nichols, Kathleen, et al., “Definition of the Differentiated Services Field(DS Field) in the IPv4 and IPv6 Headers”, RFC 2474, December 1998.
- [6] 남일성, ”가입자망 관점의 BcN 분석 및 전망”, Telecommunications review 특집부록 통권제89호 pp.117-133 2004.
- [7] 고석주 외, “인터넷방송을 위한 멀티캐스트 기술 동향” 전자통신동향분석, 제17권, 제3호, pp. 1 - 14. 2002. 6.
- [8] D. Waitzman, C. Partridge, S.E. Deering., “Distance Vector Multicast Routing Protocol”, RFC 1075, Nov.1988.
- [9] J. Moy, “MOSPF: Analysis and Experience”, RFC 1585, March.1994.
- [10] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, L. Wei, “Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification”, RFC 2362, June.1998.
- [11] A. Ballardie, “Core Based Trees (CBT version 2) Multicast Routing -- Protocol Specification”, RFC 2189, September.1997.
- [12] B. Fenner, “IANA Considerations for IPv4 Internet Group Management Protocol (IGMP)“, RFC 3228, February.2002.
- [13] W. Fenner, “Internet Group Management Protocol, Version 2”, RFC 2236, November.1997.
- [14] B. Cain, S. Deering, I. Kouvelas, B. Fenner, A. Thyagarajan, “Internet Group Management Protocol, Version 3”, RFC 3376, October.2002.
- [15] S. Deering, W. Fenner, B. Haberman, “Multicast Listener Discovery (MLD) for IPv6”, RFC 2710, October.1999.
- [16] R. Vida, Ed., L. Costa, Ed., ” Multicast Listener Discovery Version 2 (MLDv2) for IPv6”, RFC 3810, June.2004.
- [17] D. Thaler, “Border Gateway Multicast Protocol (BGMP): Protocol Specification”, RFC 3913, September.2004.



이 정 옥

1987년 이화여자대학교 전자계산공학.
1987년~2007년6월 LG-Nortel R&D Center. 2007년6월~현재 (주) 티모테크 놀로지 연구위원. 관심분야는 Security, Qos, 멀티미디어 등임.



최 지 연

1991년 경북대학교 전자공학과.
1996년 경북대학교 전자공학과 석사.
2000년~2007년 6월 LG-Nortel R&D Center. 2007년6월~현재 (주) 티모테크 놀로지 수석연구원. 관심분야는 QoS, BcN 등임.



범 순 균

1985년 중앙대학교 전자계산학과.
1985년~2007년6월 LG-Nortel R&D Center. 2007년6월~현재 (주) 티모테크 놀로지 네트워크 연구소장. 관심분야는 BcN, IPTV, 홈네트워크 등임.