

IPTV QoS를 위한 비디오 코딩 기술

경희대학교 | 서덕영

1. 서론

IP는 통신을 대표하며 TV는 방송을 대표한다. 통신과 방송이 융합되어야 한다는 것은 기술의 발전상 극히 자연스러운 일이다. IP와 TV는 우수한 가문에서 태어난 선남선녀처럼 개념상은 주위의 축복을 받으면서 부부가 되었다. 그러나, 대부분의 결혼이 그렇듯 두 가문 사이의 마찰도 없지 않았고, 앞으로 어떠한 자식을 낳아 키우게 될지 사뭇 기대가 크다.

우리나라는 IP와 TV가 처음 나오던 40~50년 전에는 기술적으로 전혀 기여한바가 없다. 그러나, 우리나라는 현재 통신기기의 총아인 휴대폰 판매에 있어서 세계 1위이고, HDTV 판매에서도 소니를 제치고 있다. 그런데 그 둘이 결합된 IPTV에서는 기대와 달리 부진하다. 더더구나 IPTV 서비스와 관련된 미디어법이 정치적인 싸움의 유탄을 맞아 지지부진하게 진행되고 있다. 이를 빌미로 정치적인 이득을 보려는 세력들과 기술과 정치적인 것과 분리하지 않고 전체를 반대하는 세력들로 인하여 기술적인 발전이 지연되고 있다. 현재는 IPTV라는 키워드에 묻히고 있지만, multimedia over IP networks가 더 정확한 지향점이 아닌가 싶다. 앞으로 점점 국가간의 경계도 없어진다는 미래학자의 예측대로 TV가 가지는 법적, 정치적 울타리는 낮아질 것이다.

TV는 지상파, 위성, CATV망이라는 지정된(dedicated) 통신 채널을 통해 전송되었으나, IP망에서는 여러 가지 서비스가 혼재한다. 따라서, TV 서비스만의 요구조건을 맞추는 것은 어렵다. 이것을 맞추기 위해 필요한 기술이 QoS(Quality of Service)이다. 이를 위하여 우선 전송하려는 대상인 비디오 트래픽의 특성을 잘 알아야 한다.

2. 비디오 코딩 표준의 변천사

디지털 비디오 코딩의 표준화는 1990년 초에 시작된 이래 지속적인 진화를 하고 있다. 비디오 코덱이 사용되는 환경이 바뀌기 때문에 그렇지만, 비디오 코덱을 사용하는 이들은 비디오 코딩 표준이 너무 자주 바뀐다고 불평을 한다. 비디오 코딩 표준은 크게 2가지 방향으로 발전하고 있다.

첫째, 압축효율이 높아지는 쪽으로 발전하고 있다. 1994년경 표준화가 완료된 MPEG-2에 비해서, 2005년에 표준화된 MPEG-4 AVC, 즉, H.264는 압축률이 2배 이상 좋아졌다고 한다. HDTV를 MPEG-2로 압축하면 20Mbps인데, H.264로 압축하면 10Mbps가 된다는 것이다. 여러 가지 새로운 알고리즘이 도입되었으나, 기본적으로 압축에 필요한 계산량을 증가시키면서 압축률이 좋아졌다고 할 수 있다. 계산량이 커질 수 있었던 이유는 Moore의 법칙에서 찾을 수 있다. H.264도 MPEG-2 정도만의 계산량이 되도록 옵션을 선택하면 압축률이 큰 차이 나지 않는다.

둘째, 네트워크 친화성이 높아지는 방향으로 발전하고 있다. MPEG-1(1992)의 경우에는 CD 한장에 2시간짜리 영화를 저장하는 목적으로 표준화되었다. 그래서, 비트율로 CD 1배속에 맞추어 1.44Mbps로 고정되어 있다. MPEG-2는 안정된 방송망을 통해 전송되는 것을 목적으로 표준화되었다. MPEG-4는 시작할 때, 휴대단말기에 적합하도록 하는 것이 목표였다. 그래서, 압축율이 매우 높아야 하고, 채널에러에 강해야 하는 것이 요구조건이었다. 이후 SVC(Scalable Video Coding)은 혼재된 통신환경(heterogeneous networks)에 적합하게 만들어졌다고 할 수 있다.

비디오 코딩을 표준화하는 기구는 ISO/IEC의 MPEG과 ITU-T가 있다. MPEG에서는 방송서비스를 목적으로 MPEG-1, MPEG-2를 표준화하였고, ITU-T에서는 화상전화로 목적으로 H.261, H.263 등을 표준화하였다. 2000년대 초반에 이 두 기구가 JVT(Joint Video Team)을 만들어서 공동으로 표준화를 진행하였고, 그 결과로 H.264(MPEG-4 AVC), SVC, MVC(Multi-view

† 본 원고는 2008년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임 (No. R0A-2005-000-10061-0).

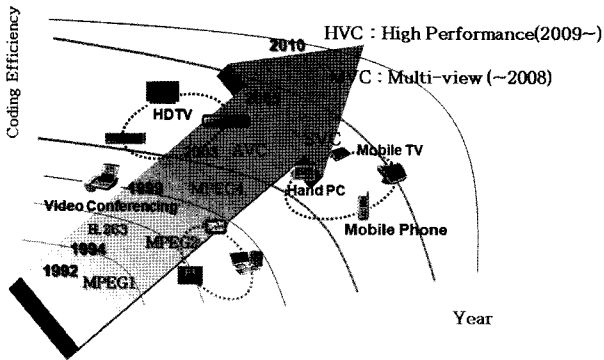


그림 1 비디오 코딩 표준의 변천사(건국대 임창훈교수 그림)

Video Coding)을 표준화하였다. 두 기구의 결합은 방송통신융합시대가 도래하였음을 상징적으로 보여준다. 2009년에는 HVC(High-performance Video Coding)이라는 이름으로 새로운 코덱 표준을 만들고 있다.

HVC의 목표는 크게 두 가지이다. 우선 8k×8k 정도의 고품질 비디오에 대한 코덱이 목표이고, 둘째로는 모바일과 같은 새로운 통신환경에 적합한 코덱을 만드는 것이다. 새로운 통신환경을 이해하려면 네트워크가 제공하는 QoS 프로토콜을 살펴보아야 한다.

3. 네트워크 QoS 프로토콜과 비디오 코덱

네트워크는 회선교환(circuit switching) 중심에서 패킷교환(packet switching) 중심으로 중심축이 이동하였고, 이제는 QoS를 보장하는 것이 중요한 화두가 되고 있다. 비디오 코덱에 대하여 논하더라도 네트워크의 프로토콜의 특성이 고려되어야 한다.

3.1 회선교환과 비디오 코딩

회선교환은 지정된 품질의 채널을 한 서비스에 보장한다. 전화서비스가 가장 중요한 예가 된다. 방송에서는 채널별로 대역폭을 고정적으로 할당하는 CATV나 DMB도 QoS 제어면에서는 회선교환의 범주에 넣을 수 있다. 같은 QoS(비트율, 지연)를 서비스 기간동안 고정적으로 제공한다는 면에서 실시간 멀티미디어 서비스에 적합하다. 회선교환을 이용하는 경우에는 비디오 비트율을 정해진 비트율보다 낮게 유지해야 한다.

그러나, 향후 여러 가지 성능의 단말기와 여러 가지 성능의 네트워크가 혼재된 환경에서 사용하기에는 유연성이 부족하다. ‘앞으로 음성전화서비스는 공짜가 될 것이다’라고한 시스코 사장의 말대로 회선교환은 이제 패킷교환망이 특별히 제공하는 한 방식으로 사라질 전망이다.

3.2 패킷교환과 비디오 코딩

회선교환은 하나의 넓은 길을 여러 사람이 공유하

는 형태이다. 각자 사용하고 싶은만큼 패킷을 보낼 수 있다. 회선교환보다 매우 유연한 방식이다. 비디오의 경우, 여러 가지 비트율의 비디오를 예약이나 다른 특별한 절차없이 전송할 수 있으므로, 자유롭게 비디오를 압축하여 여러 가지 비트율로 전송할 수 있다. 자유를 갖는 댓가로 동시에 여러 사람이 하나의 길로 가려고 할 때 문제가 생긴다. 내게 할당된 대역폭이 따로 없으므로 라우터에 도착한 순서대로 버퍼에서 대기하다가 포워딩된다. 따라서, 혼재(congestion)상황이 되면 자기 차례가 올 때까지 지연된다. 버퍼가 넘치면 라우터는 무작위로 패킷을 버리게 된다. 비실시간 서비스에서는 손실된 것을 재전송할 수 있지만, 실시간 서비스나 멀티캐스트인 경우에는 재전송이 불가능하다. 이러한 기본적인 기능만 가지는 QoS프로토콜을 BE(Best Effort) 프로토콜이라고 한다. 특히, 무선망인 경우에는 채널의 조건도 시시각각으로 바뀔 수 있다.

3.3 QoS 프로토콜과 비디오 코딩

QoS 프로토콜은 per class와 per flow로 나눌 수 있다.

Per class 서비스는 패킷당 지불하는 돈이 많은 패킷은 특별히 취급하고 그렇지 않은 패킷은 상대적으로 낮은 서비스를 해주는 방식이다. 보통 클래스는 3~4개로 구분된다. 따라서 라우터는 몇 비트로 이루어진 class identifier만 보고, 누가 보내고 누가 받는지 상관없이 어떤 대우를 해야 할지 결정할 수 있다. 비디오 코덱에서는 항상 우선순위가 높은 데이터와 낮은 데이터가 존재한다. 따라서 하나의 비디오 서비스에서도 패킷별로 우선순위를 달리 할 수 있다.

Per flow 서비스는 누가 보내고(source address), 누가 받고(destination address), 어떤 서비스를 위한 것이며(source/destination port numbers), 어떤 프로토콜을 이용하는가, 이렇게 5개의 변수를 확인한 후 어떤

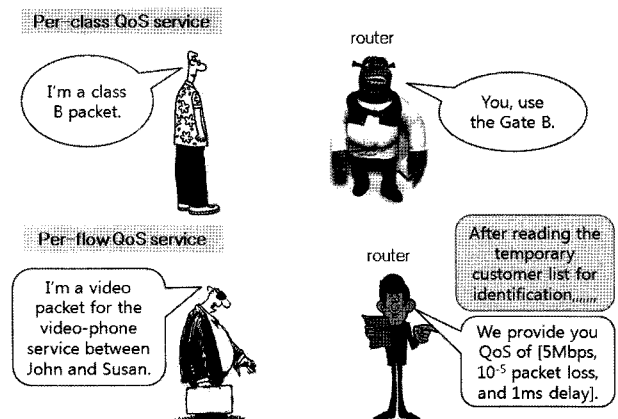


그림 2 Per-class와 per-flow QoS 프로토콜

대우를 할지 결정한다. 어떤 대우를 하는가는 서비스 시작할 때 미리 예약을 해놓아야 한다. 예약을 위하여 flow별로 원하는 보장비트율, 최고비트율, 버퍼크기 등을 tspec으로 정해놓는다. 기본적으로 leaky bucket model로 트래픽인자를 정한다. 비디오 서비스에서 flow란 오디오 스트림과 비디오 스트림이 독립적인 flow가 될 수도 있고, 둘을 합쳐서 하나의 스트림으로 할 수도 있다. IEEE802.16 WiBro(WiMAX)에서나, UMTS, LTE 등에서 per flow QoS 프로토콜을 지원한 예정이므로, 예약을 하고 예약한대로, 비디오 트래픽이 만들어 질 수 있도록(traffic shaping) 비디오 코덱이 제어되어야 한다.

4. IP망과 비디오 코딩

All IP 패러다임이 일반화되면서 최근에 표준화 되는 모든 통신프로토콜과 멀티미디어 서비스 표준에서는 IP망을 이용하는 것을 전제로 한다. 문서전달용으로 개발된 IP망은 실시간 멀티미디어 서비스를 원활하게 지원하기 위해 진화하고 있다. 이러한 변화를 효과적으로 이용하기 위해서 비디오 코덱도 같이 진화하여야 한다.

4.1 VBR과 CBR

IP망은 패킷망이므로 기본적으로 VBR(Variable BitRate) 전송이 가능하다. 비디오는 VBR이 자연스럽다. 동영상의 품질을 똑같이 유지하려면 동영상의 엔트로피(무질서도)에 따라 비트율이 변화한다. ‘까마귀 우는 밤’ 장면에서는 비트율이 낮아지고, ‘풋볼경기’ 장면에서는 비트율이 올라간다. DVD는 읽는 속도가 3Mbps에서 9Mbps까지 가능하므로 VBR로 인코딩하여 저장한다. 불법 DVD인 경우 대개 품질이 떨어지는 이유가 CBR로 인코딩하기 때문이 아닐까 추정한다.

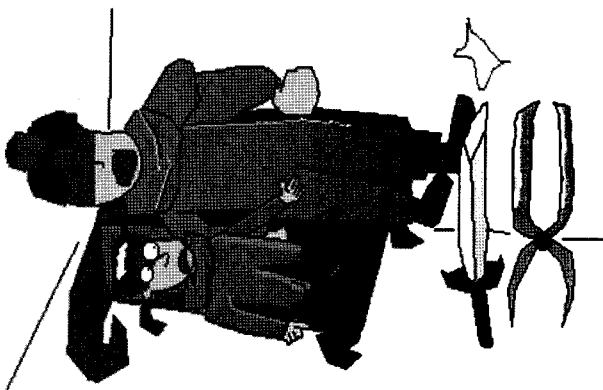


그림 3 VBR 비디오를 CBR 채널로 전송하는 것은 Procrustes의 침대와 같다(서재형 그림)

HDTV 표준에서도 비트율이 20Mbps로 정해져있었지만 아직까지 CBR(Constant BitRate)가 일반적으로 통용되고 있다. 이것은 비트율을 정해서 대화하는 것이 엔지니어간 의사전달이 편리하기 때문인 것 같다. CBR의 경우 울제어(rate control) 기술을 이용하여 비트율을 일정하게 유지한다. 대개 한 GOP(Group Of Pictures, 약 0.5초)간 전체 비트수를 제어한다. 따라서, Intra, Predictive, Bi-directional 프레임들간 비트율 차이는 존재한다. 동영상마다 큰 차이가 있지만 대표적인 예는 I:P:B=8:3:2 정도이다. 이 차이를 평탄화하여 CBR 채널을 통해 전송하기 위해서는 버퍼를 이용해야하는데, 버퍼지연이 약 200ms 정도까지 될 수 있다. 이는 화상전화와 같은 대화형 서비스에서는 허용할 수 없는 지연이다.

모든 GOP의 비트수를 일정하게 유지하기 위해서는 복잡한 장면(풋볼경기)에서는 화질을 떨어뜨려야 하고, 단순한 장면(까마귀 우는 밤)에서는 화질을 필요 이상으로 올리게 된다. 둘다 문제이다. 풋볼경기장면은 경계가 흐릿하게 되고, ‘까마귀 우는 밤’을 보통 화질로 인코딩하고 남는 비트율을 다른 서비스에 사용할 수 있는 여지가 없어진다.

4.2 SVC와 적응적 비디오 코딩

SVC(Scalable Video Coding)는 저평가된 우량주라고 하고 싶다. SVC를 사용함으로써 얻을 수 얻는 이익은 막대한데 소문이 잘못나서 사용이 잘 안되고 있다. 본고에서는 SVC에 대한 몇 가지 오해에 대한 변명을 하고자 한다.

SVC는 하나의 동영상을 우선 순위별을 나누어 여러 개의 비트스트림으로 나누어 인코딩하는 것을 말한다. 기본계층(base layer)는 가장 기본적인 품질의 동영상을 재생하는데 사용하며, 여기에 향상계층(enhancement layer)가 더해지면 더욱 좋은 품질의 동영상을 재생할 수 있다. 계층수는 자유롭게 정할 수 있으며 MPEG/JVT에서 표준화를 할 때에는 64개 계층까지 실험하기도 했다.

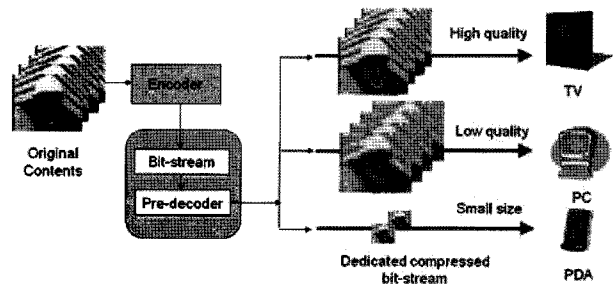


그림 4 SVC를 이용한 OSMU(ETRI 최해철 박사 그림)

SVC를 사용하여 얻는 장점으로는 첫째, OSMU(One Source Multi Use)에 활용할 수 있다는 것이다. 서버가 HDTV급(6Mbps), SD급(3Mbps), 4CIF(600kbps), CIF(300kbps), QCIF(100kbps)급으로 합이 10Mbps가 되도록 SVC 인코딩한 후 단말기의 능력, 네트워크의 대역폭, 사용자의 요구에 맞추어서 서비스할 수 있다. 유선망에서 무선망으로 넘어갈 때에도 트랜스코딩이 필요없고 맞는 계층만 선택해주면 된다. 둘째, 가변적 채널에 적응할 수 있다. 시간에 따라 가용대역폭이 100 kbps에서 10Mbps 사이에서 변화할 때도 끊기지 않고 전송할 수 있다. 그러나, 만일 단일 계층으로 인코딩하였을 때는, 10Mbps보다 내려가면 서비스가 끊기게 된다. 셋째, 예러나 손실에 강하다. 기본계층만 예러나 손실에 강하게 해놓으면 채널품질이 나빠지더라도 끊길 염려가 없다. 넷째, 이외에도 계층적 암호화, PIP(Picture In Picture), 재핑지연 줄이기 등등 많은 새로운 응용방법이 있을 수 있다.

여기서 SVC에 대한 오해를 몇 가지 풀기로 한다. 첫째, 비트율이 올라간다는 오해가 있다. SNR 계층화를 사용하는 경우에는 이 말이 맞다. 그러나, 공간적(spatial) 계층화, 시간적(temporal) 계층화를 사용하면 같은 품질의 단일 계층부호화와 비트율 차이가 거의 없다. 둘째, 복잡도가 증가한다는 오해가 있다. 2계층이라고 해서 2배 복잡도가 증가하는 것이 아니다. SNR 계층화와 시간적계층화에서는 복잡도 증가가 거의 없고, 공간적 계층화의 경우에서 25%증가하는 정도이며, 이마저도 약간의 아이디어만 가지고 설계하면 10%이하로 떨어뜨릴 수 있다. 셋째, 네트워크에서 활용하는데 있어서 복잡해진다는 오해가 있다. 복잡해지는 것은 사실이다. 그러나, 어떤 방송이 한국어, 영어, 중국어, 일본어 방송을 하는데, 어떠한 단말기에서는, 또는 어떠한 네트워크로는 이중 하나만 허용할 수 있는 기능이 있다면 SVC도 똑같은 처리방식으로 구현할 수 있다. 넷째, '1993년에 MPEG-2 SVC가 표준화되었는데 뭔가 안좋은 게 있으니까 그동안 사용 안했겠지'라는 입장이 있다. 그러나, 그동안 단일계층 서비스하는 시스템 겨우 만드는데 까지만 기술이 발전했기 때문에 사용하지 않은 것이지 SVC 기술 자체가 나빠서 그런 것이 아니다.

SVC에 대한 오해를 버리고, SVC의 장점을 살리는 서비스나 제품을 개발하는 회사가 향후 멀티미디어 시장을 석권할 것이다.

4.3 MVC와 3D

할리우드에서 이제 모든 영화는 3D로 찍는다고 한

다. 3D로 본 내용은 2D로 본 것에 비해 30% 기억에 더 남는다고 하기도 하고, 3D에 맞들인 관객들이 2D를 멍멍하게 생각하게 될 것이라고도 한다. 국내에서는 3D 디스플레이 휴대폰을 2007년에 출시한 바 있다. 3D 디스플레이를 이용한 광고기기는 국내에서 편의점에 수십만대 보급되었다고 한다.

3D 동영상의 압축표준인 MVC(Multiview Video Coding)는 2008년에 JVT에서 표준화되었다. 하나의 대상을 여러 각도 카메라로 찍어서 입체로 볼 수 있게 하는 것이다. 사용자가 입체 디스플레이가 없다면 하나의 시점(視點, viewpoint)만 선택하여 볼 수 있게 할 수도 있다. 현재까지는 2시점, 즉, 스테레오스코픽 동영상이 많이 쓰이고 있지만, 8시점 디스플레이는 이미 2006년부터 시판되고 있다.

여러 시각에서 취득된 동영상을 각각 압축할 수도 있지만 서로의 상관성을 이용하여 압축하면 효율이 높아진다. 먼저 하나의 시점의 동영상을 압축하여 이를 참조영상으로 하여 다른 시점을 압축한다. 양방향으로 참조하기도 한다. 보는 시각에 따라 영상이 약간씩만 차이가 나므로 다른 시점을 참조하면 압축효율이 높아진다. 기존의 ME/MC(motion estimation and motion compensation)과 유사하지만, 차이점은 카메라의 하드웨어적인 특성의 차이에 대한 보상도 필요하다는 것이다. 즉, 노출이나 스케일링에서 차이가 있을 수 있다. 시점별로 참조하므로 시점별로 우선순위가 달라지며 시점중심의 SVC가 가능하다.

2009년부터 표준화가 시작된 HVC(High performance Video Coding)에는 깊이정보(depth information, 카메라에서 물체까지 거리)을 이용한 3D 압축방식이 포함되어 있다.

4.4 QoS 보장 방법

IEEE, 3GPP, 그리고 IETF의 QoS 프로토콜을 잘 활용해야 할 것이다. IEEE는 무선랜이나 WIBRO관련, 3GPP는 휴대폰관련, IETF는 인터넷관련 표준화를 하고 있다.

Per class 프로토콜인 경우에는 우선순위에 따라 IP 헤더에 중요도를 표시만 하면 된다. 실시간 서비스는 비실시간 서비스보다 중요하고, 통신비를 더 지불하는 서비스가 더 중요하며, 오디오는 대개 비디오보다 더 중요하고, SVC에서 기본계층이 상위계층보다 중요하다. 이렇게 SVC는 per class 프로토콜을 활용하는데 매우 적합한 예가 될 수 있다. 그러나 아직까지는 하나의 멀티미디어 서비스에 대하여 여러 가지 트래픽 클래스를 지정하여 서비스하는 예는 없다.

Per flow 프로토콜을 이용하는 경우에는 좀 복잡해진다.

◆ Double Leaky Bucket

- Peak bitrate (p, bp)
- Guaranteed bitrate (r, b)

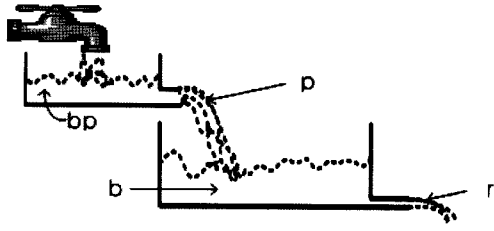


그림 5 VBR 트래픽의 리키버킷 모델

CBR인 경우에는 비트율만 정하면 되고, 그것도 회선교환 서비스에서는 대개 고정적으로 정해있지만, VBR의 경우에는 리키버킷(leaky bucket) 형태로 예약해야 한다. 기본적으로 예약에 대한 개념은 1990년대 초반 IP와 초고속통신망의 기본 프로토콜을 놓고 일전을 겨루다 장렬히 산화한 ATM(Asynchronous Transfer Mode)의 철학을 따른다. ATM에서는 실시간 서비스는 KTX처럼 라우터(ATM 스위치)를 먼저 빠져나가고(leak rate), 비실시간 서비스 패킷은 무궁화호처럼 대기한다. 이때 대기에서 사용되는 버퍼의 크기(buffer)가 어느 정도 크기 이상을 넘지 않도록 해야 한다. 빨리 빠져나갈수록 버퍼가 작아도 된다.

그림 5와 같이 수도꼭지에서 가변적으로 물이 나오더라도 새는 정도(leak rate)와 버킷 크기를 잘 정하면 넘치지 않게 할 수 있다. 3GPP나 IEEE 모두 이러한 변수의 정의는 IETF의 RSVP 표준에서 나온 tspec을 따른다[1]. 하나의 VBR 트래픽은 기본적으로 피크 비트율에 대한 리키 버킷 인자 2개(p, bp)와 보장 비트율에 대한 리키 버킷 인자 2개(r, b)로 모델링할 수 있다.

Per flow 프로토콜이 제대로 작동하려면 서비스 시작전에 자원의 예약 절차(CAC, Call Admission Control)와 함께 서비스 도중에 전송단말기측이 약속을 지키려는 노력인 트래픽 셰이핑(traffic shaping)과 망쪽에서 예약을 지키는지 감시하는 트래픽 폴리싱(traffic policing) 기능이 필요하다.

5. 결론

QoS 제어의 종류로는 서버 응용과 클라이언트 응용 시스템이 참여하는 응용 및 트랜스포트 계층 제어가 있으며, 이를 위해 비디오 인코더와 디코더, RTP 및 UDP 프로토콜과 FEC(Forward Error Correction) 등이 사용된다. 이어서 라우터가 참여하는 네트워크 계층

QoS 제어에서는 diffServ 또는 intServ와 3GPP의 IMS 및 NGN(Next Generation Networks) 등이 이용될 수 있다. MAC 계층 QoS 제어로는 IEEE 802 시리즈의 per-class 및 per flow 프로토콜이 이용될 수 있고, UMTS와 LTE 역시 per-class 및 per flow 프로토콜을 지원하고 있다.

여러 계층의 QoS 프로토콜을 결합하여 서비스 또는 제품을 계획할 때에는 각각의 표준의 time-to-market을 정확하게 예측해야 한다. 응용계층과 트랜스포트 계층, 또한 MAC 계층 표준은 표준화에서 상용화까지 시간이 짧은 편이다. 예를 들어, MPEG-4 AVC/H.264의 경우에는 2003년에 표준화되고 2006년에 DMB에 상용화되어 사용되었다. 그러나, 네트워크 계층은 매우 상용화가 느린 편이며, 호환성을 갖추기가 어렵다. 예를 들어, KT의 Megapass의 IPTV망에서 per class 프로토콜인 diffServ[2]를 사용한다고 되어 있으나, 망자체가 IPTV와 그외 중요한 서비스를 하기 위해 따로 가설한 Premium망이므로 per class 프로토콜의 효과는 크지 않다고 생각된다. 또한 KT와 다른 통신사간 통신에서 diffServ가 언제 사용될 수 있는지도 미지수이다. 따라서, 2007년에 표준화된 SVC와 어떤 네트워크 계층 QoS 표준과 묶어서 하나의 서비스나 제품을 만드는 것이 좋을지에 대해서는 신중한 고려가 필요하다.

참고문헌

- [1] RFC 2212, "Specification of Guaranteed Quality of Service," ietf.org, 1997.
- [2] RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," 1998.



서덕영

1980 서울대 핵공학과 학사
 1986 미국 Georgia Tech. 핵공학 석사
 1990 미국 Georgia Tech. 전기및컴퓨터공학과 박사
 1990~1992 상공부 생산기술 연구원, HDTV 연구개발단 선임연구원
 현재 경희대학교 전자정보학부 교수

관심분야 : networked video
 E-mail : suh@khu.ac.kr