

10 기가비트 이더넷 기반 비압축 HDTV의 인터넷 전송을 위한 시스템 최적화 연구

조진용[†] · 석우진^{**} · 이민선^{**} · 변옥환^{***}

요약

네트워크 및 시스템에서 가용 대역폭의 확보와 전송 성능에 대한 고려는 대용량 데이터를 고속으로 처리하기 위한 중요한 요소 기술이다. 점보 프레임, 네트워크 버퍼 설정 등과 같은 시스템 튜닝을 통해 중단 간 전송 처리율이 향상되고 시스템 부하가 감소되어 결과적으로 시스템 효율성이 크게 증가된다. 이러한 시스템 튜닝은 대용량 데이터 전송을 필요로 하는 응용에 국한하여 연구되어 왔고, 비압축 HDTV와 같이 높은 데이터 집적도와 연산 처리 성능을 함께 요구하는 응용에 대해서는 충분히 고려되지 않고 있다. 본 논문에서는 다양한 시스템 튜닝 값들이 중단 간 전송 성능 및 시스템 부하에 미치는 영향을 10 기가비트 네트워크와 비압축 HDTV 응용 소프트웨어를 이용해 분석했다. 시스템 측면에서 전송 성능을 높이고 부하를 줄이기 위해서는 송/수신 버퍼 크기, 인터페이스 레지스터 값, 인터럽트 통합 등의 튜닝 과정이 중요하며, 실험을 통해 이들 튜닝 요소들이 비압축 HDTV의 재생 품질에 미치는 영향을 조사한다.

키워드 : 비압축 HDTV, 시스템 성능 향상, 시스템 튜닝, 광대역 네트워크 응용

Optimizing System Performance for Uncompressed HDTV over 10-Gigabit Ethernet

Jinyong Jo[†] · Woojin Seok^{**} · Minsun Lee^{**} · OkHwan Byeon^{***}

ABSTRACT

Guaranteeing network bandwidth and system performance becomes important technical matters to make seamless delivery of bulky data at high speed. Tuning system and network parameters including jumbo frame, kernel buffer size, and interrupt coalescence determines end-to-end transmission throughput. Additionally, fine-tuning of the parameters alleviates workload on high-end systems. Until now, many studies have concentrated on how to increase transmission throughput but rarely discussed how to mitigate system workload. In this paper, we have investigated various tuning parameters, which positively affect networking and processing performance of uncompressed HDTV system.

Key Words : Uncompressed HDTV, System Performance, System Tuning, Broadband Networked Application

1. 서론

그리드 컴퓨팅은 고성능 컴퓨터 자원과 첨단 네트워크를 하나로 연결해 고속연산 및 대용량 데이터 처리를 용이하게 하는 기술이다. 협업 환경을 지원하는 그리드 기술들은 현장의 생동감을 원거리로 전달하기 위해 고선명(high definition) 영상·가상현실(virtual reality) 등의 응용 기술에서 사용된다. 영상 압축 및 네트워크 전송 기술의 발달은 인터넷을 이용한 고선명 영상의 활용 범위를 IPTV나 원격의료와 같은 응용 분야로 확대시켰다[1, 2]. 또한, 그리드 네트워크 기술을 활용하는 비압

축 HDTV 기술들이 개발되고 있다[3-5].

압축 기술을 활용하는 고선명 영상 응용의 경우, 영상 프레임 간 종속성으로 인한 프레임 간 오류 전파, 부호화 및 복호화 과정에서 발생하는 프로세싱 지연, 압축 기술의 복잡도로 인한 시스템 부하의 증가 등 문제점이 있다. 이에 반해, 비압축 HDTV 전송 기술은 압축 기술의 사용으로 인한 문제점들이 발생하지 않으므로, 대륙 간 중계방송 시 전달 지연 감소, 방송 장비 및 영상 편집을 위한 구축비용 절감 등의 효과를 얻을 수 있다. 또한, 저지연, 고화질 특성으로 인해 원격 의료 등 실시간 응용에 활용 가능하다.

네트워크 기반 비압축 HDTV의 품질을 보장하기 위해서는 망 기반 구조 및 시스템 관점에서 각각 선결되어야 할 기술적 요소들이 존재한다. 먼저, 비압축 HDTV 신호의 네트워크 전송

[†] 정 회 원 : 한국과학기술정보연구원 슈퍼컴퓨팅센터 연구원
^{**} 정 회 원 : 한국과학기술정보연구원 슈퍼컴퓨팅센터 선임연구원
^{***} 중신회원 : 한국과학기술정보연구원 슈퍼컴퓨팅센터 책임연구원
논문접수 : 2006년 4월 13일, 심사완료 : 2006년 7월 24일

을 위해서는 기가급 대역폭이 안정적으로 확보 되어야 한다. 비압축 HDTV 시스템은 점포 프레임을 이용해 신호를 전송해야 하기 때문에 영상 트래픽이 대역폭을 선점하는 특성을 보인다 [6]. 대역폭 선점 특성은 표준 프레임을 사용하는 트래픽들에 부정적인 영향을 미치기 때문에 비압축 HDTV가 발생시키는 트래픽들은 외부 트래픽들로부터 격리되는 것이 바람직하다. 이와 같은 요건을 충족시키는 네트워킹 기술로 UCLP(user-controlled light path)가 있다[7].

UCLP는 그리드 기술을 활용하는 네트워크 응용으로 사용자가 중단 간 lightpath를 관리하고 대역폭을 할당하는 기술로써 CANARIE(Canada's advanced Internet development organization)에서 개발되었다. 국내에서는 GLORIAD-KR 및 KREONET 등에서 광 패스(light path)를 동적으로 할당하기 위한 기술로 사용되고 있다.

시스템 측면에서 볼 때, 64 비트 133 Mhz를 갖는 확장형 주변 소자 상호연결 버스(PCI-X, peripheral component interconnect extended)의 이론상 최고 속도는 8 Gbps이다. 네트워크 대역폭이 10 Gbps에 이르고 있는 점을 감안할 때, 시스템 버스가 고성능 네트워킹에 있어 병목 지점으로 작용하는 것은 명백하다. 또한, 시스템의 메모리, 입출력 장치, 버스 대역폭 및 CPU 성능 등 하드웨어 특성 및 운영체제나 응용소프트웨어의 성능은 중단 간 전송 처리량(throughput)에 큰 영향을 주는 요소들로 데이터 전송의 효율성을 높이기 위해서는 반드시 시스템 튜닝이 필요하다[8-10].

시스템 튜닝을 필요로 하는 주요 요소들로는 OS 커널 변수(kernel parameters), 인터럽트 통합(interrupt coalescence), 전송 프레임의 크기 등이 포함 된다. 이러한 시스템 튜닝은 대용량 데이터의 효율적인 전송을 필요로 하는 분자 물리 그리드 및 고에너지 물리 등 다양한 데이터 그리드 프로젝트에서 활용되고 있다[11].

본 논문에서는 대용량 데이터 전송 및 높은 연산 처리 능력을 요구하는 비압축 HDTV 신호의 인터넷 전송을 위한 시스템 최적화 방안에 대해서 연구한다. 기존 연구들은 TCP를 전송 프로토콜로 하는 범용 응용의 대용량의 데이터 전송에 초점 [8,9]을 맞추거나 1 기가급 네트워크상에서 실험이 수행[10] 된 데 반해, 본 연구는 UDP를 전송 프로토콜로 하고 높은 연산 처리능력을 요구하는 비압축 HDTV 시스템의 성능을 최적화할 수 있는 방안에 대해서 조사했다. 또한, 기존 연구와 차별화된 하드웨어 구성 및 운영 체제를 이용했기 때문에 관련 연구의 다양성 측면에서 기여하고 국내에서 10 기가비트 이더넷을 활용한 선도 실험이라는 점에서 의미를 갖는다.

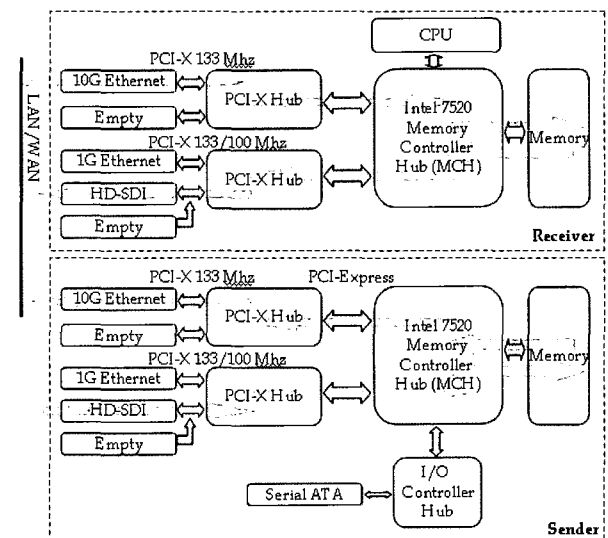
본 논문의 2장에서는 구축된 비압축 HDTV 시스템의 구성에 대해서 살펴보고, 3장에서 실험을 위한 환경 설정과 튜닝 요소들이 네트워킹 및 시스템 성능에 미치는 영향에 대해서 조사한다. 4장에서 시스템 튜닝 요소들을 응용 소프트웨어에 적용해 실험하고, 5장에서 결론을 맺는다.

2. IP기반 비압축 HDTV 시스템 구조

(그림 1)은 IP 기반 비압축 HDTV 송/수신 시스템의 구성도와 데이터의 흐름 방향을 보여 준다. 송신측은 HD 카메라로부터 입력되는 HD-SDI(high definition serial digital interface) 신호를 수집하고, 이를 패킷화해 10 기가비트 네트워크를 통해 전송하는 역할을 담당한다. 수신측에서는 네트워크 패킷 형태의 데이터를 음성 및 영상 프레임으로 재구성 한 후, HD-SDI 신호로 변환 출력한다. 송신측은 비월주사(interlace) 방식의 1080i 비디오(60/1,001 Hz)와 PCM(pulse code modulation) 방식의 음성(24-bit/48 KHz) 신호를 HD-SDI 인터페이스 장치에서 수집 한다. HD-SDI 신호의 총 해상도는 보조 데이터 공간인 HANC(horizontal ancillary data)와 VANC(vertical ancillary data) 등을 포함하는 2200 개의 총 샘플과 1125 개의 주사선으로 구성된다. 10 비트 4:2:2 샘플화된 HD-SDI 신호를 전송하기 위해서는 약 1.485 Gbps의 네트워크 대역폭이 필요하다. 현재 구성된 시스템은 1920*1080 크기의 확장 해상도만을 수집하기 때문에 영상 전송을 위해서 약 1.25 Gbps의 데이터 송신율(sending rate)을 갖게 한다.

구성된 시스템은 3개의 133 Mhz와 2개의 100 Mhz PCI-X 슬롯을 가지고 있다. 모든 시스템 버스는 64 비트로 동작하기 때문에 133 Mhz PCI-X 버스를 사용하는 주변 장치의 최대 속도는 8.5 Gbps에 이른다. 2개 이상의 고성능 주변 장치가 하나의 PCI-X 허브를 공유할 경우 병목으로 인해 시스템 성능이 크게 저하된다. (그림 1)의 HD-SDI 인터페이스 장치는 66 Mhz로 동작하며 최대 속도는 4.2 Gbps이다. 위와 같은 구성을 가질 때, 100 Mhz 속도를 갖는 1 기가비트 이더넷 인터페이스는 66 Mhz로 클럭 레이트(clock rate)가 하향 된다.

서버 보드는 인텔 E7520 칩셋 구조로 구성되어 있다. E7520



(그림 1) 비압축 HDTV의 인터넷 전송을 위한 시스템 조감도 및 데이터 흐름

칩셋은 고성능 주변 장치들을 이용하는 데이터가 상단 MCH만 거처기 때문에 하단 브리지와의 사이에서 발생하는 병목을 최소화 시킨 구조를 갖는다. 중앙처리장치는 800 Mhz FSB(front-side bus)를 갖는 이중(dual) 구조이며 이론상 최대 대역폭은 51.2 Gbps이다. 메모리는 400 Mhz 클럭으로 동작하며 CPU와 동일한 최대 대역폭을 갖는다. 10 기가비트 네트워크 인터페이스는 Chelsio N210 모델[12]로 다중모드 광섬유(multimode optical fiber)를 이용해 스위치와 연동 된다. 네트워크 인터페이스는 64 비트 133 Mhz로 동작하기 때문에 이론상 최고 속도가 8.5 Gbps이다.

운영 체제는 Mandrake 리눅스(커널 2.6.11.3)이다. 패킷화된 비압축 HDTV 데이터의 송/수신을 위해서 UltraGrid의 전송 모듈을 이용했으며 실시간 영상/음성의 수집 및 재생을 위한 HD-SDI 인터페이스의 구현부가 개발되어 UltraGrid 소스에 추가 되었다. UltraGrid는 미국 남가주대학교에서 개발하고 오픈 소스화 한 비압축 HDTV용 응용 소프트웨어이다.

3. 실험 환경 및 평가 요소

3.1. 실험 환경

실험에 참가한 2대의 비압축 HDTV용 고성능 시스템들은 10 기가비트 Force10 [13] 스위치와 연동되어 있고 동일한 네트워크상에 존재한다. 10 기가비트 스위치의 MTU 크기는 9,180이며, 시스템의 네트워크 인터페이스의 이론상 최대 전송 속도는 1 GB/s 이다. Force10 스위치의 배경 트래픽(background traffic)은 수 Mbps이다.

설치된 리눅스 커널은 기본적으로 SMP(symmetrical multi-processing)가 지원된다. 단일프로세서(uniprocessor)용 커널에서 전송 처리율이 SMP 보다 약 10-25% 우수하다는 연구 결과가 있다[9]. 하지만, 연산 처리 능력과 전송 처리율을 함께 고려해야 하는 비압축 HDTV 응용의 경우 단일 프로세서용 커널에서 CPU 부하가 99%에 이르는 등 문제점이 발생했기 때문에 논의에서 제외한다.

본 연구의 목적은 처리율 및 CPU 부하의 최적화에 있다. 정확한 처리율 성능측정을 위해서 Iperf[14], Netperf[15] 및 Hpcbench[16]와 같은 벤치마크 응용들이 활용되었다. Iperf와 Netperf는 단위 시간 당 전송 처리율을 측정하고, Hpcbench는 두 종단 시스템 간의 라운드 트립 시간(round trip time)을 커널 계층에서 측정한다. 또한, CPU 부하의 측정을 위해 /proc/loadavg와 리눅스 top 명령이 사용되었다.

일반적으로 수신 버퍼의 최대 크기는 bandwidth-delay product의 값보다 커야 하며 라운드 트립 시간의 변화를 고려해 2*대역폭(바이트)*지연(ms)으로 설정한다. 10 기가비트 네트워크의 경우, 1.25(Mbytes)*라운드 트립 시간(ms)로 설정 한다 [17]. 실험 결과와 bandwidth-delay product를 바탕으로 커널 2.6.11.3에서의 버퍼 크기는 다음과 같이 최적화 설정했다.

```
echo 524288 > /proc/sys/net/core/optmem_max
echo 262144 > /proc/sys/net/core/wmem_default
```

```
echo 262144 > /proc/sys/net/core/rmem_default
echo 128388607 > /proc/sys/net/core/rmem_max
echo 128388607 > /proc/sys/net/core/wmem_max
echo 300000 > /proc/sys/net/core/netdev_max_backlog
```

```
/sbin/ifconfig eth2 txqueuelen 9000
```

하이퍼 쓰레딩(hyper threading)은 각 프로세서에게 쓰레드 계층에서의 병렬화 기법을 제공함으로써, 프로세서 자원을 보다 효율적으로 관리하는 기법이다. 하이퍼 쓰레딩을 이용하면 프로세싱 처리율 및 성능 향상을 기대할 수 있으며, 실험을 통해 비압축 HDTV 시스템에서 이용할 경우 시스템 부하를 3-7% 가량 낮출 수 있었다. 실험 환경의 기본 설정에서 하이퍼 쓰레딩은 적용하지 않는다.

특정한 언급이 없는 한, 앞서 설정된 실험 환경 및 버퍼 크기가 계속 사용된다. 시스템 튜닝 요소는 최대 전송 단위(MTU), MMRBC(maximum memory read byte count), 인터럽트 통합(interrupt coalescence)등이 활용되었다. 이는, 선택된 튜닝 요소들이 수신 버퍼에 대한 튜닝과 더불어 가장 높은 성능 이득(performance gain)을 나타내기 때문이다.

3.2 시스템 튜닝 요소의 개요 및 성능 평가

IEEE 802.3[18] 이더넷 규약(ethernet protocol)은 최대 전송 단위를 1,500 바이트로 규정하며, TCP 패킷의 전형적인 MSS(maximum segment size)는 1,448 바이트이다. 다음 식 (1)에서 볼 수 있는 바와 같이 TCP의 상한 처리율은 MSS 크기에 비례한다. 즉, 점보 프레임(9,000 바이트 크기의 MTU)의 사용은 높은 전송 처리율을 보장한다.

$$TCP_{throughput} = \frac{constand \cdot MSS}{RTT \cdot \sqrt{packet\ loss}} \quad (1)$$

대부분의 기가비트 이더넷 인터페이스는 9,000 바이트 MTU를 지원한다. 점보 프레임의 사용은 높은 처리율을 보장하지만 규정 및 표준의 부재 등으로 상호 운용의 호환성 문제가 있다. 다음 (그림 2)는 10 기가비트 네트워크 상에서 측정된 MTU 크기에 따른 처리율의 변화를 보여 준다.

MMRBC(maximum memory read byte count)는 PCI-X의 세 부 사양으로 MMRBC 값이 증가하면 패킷들이 PCI-X 버스를 통과하는데 걸리는 시간을 단축시키기 때문에 전송 처리율을 향상시킬 수 있다. PCI-X를 이용하는 네트워크 어댑터의 MMRBC는 512, 1024, 2048 또는 4096 바이트로 설정 가능하다. 일반적으로 송/수신 시스템에서 4096 바이트 MMRBC를 가질 때 최대 성능을 기대할 수 있다[19].

인터럽트 통합(interrupt coalescence)은 네트워크 어댑터에서 개별 패킷의 송/수신 이벤트에 대해 인터럽트를 발생시키는 대신 단위 시간 당 발생시킴으로써 인터럽트를 처리하는데 필요한 프로세싱 부하를 감소시키는 기술이다. 인터럽트 통합기술의 사용은 버퍼링 효과로 인한 지터(jitter)를 발생시키고 네트워크 측정(network measurement) 시 부정적인 영향을 미치기도 한다[20].

성능 비교를 위해서 활용된 평가 요소들은 전송 처리율, 시스템 부하, 초당 프레임 재생율, 라운드 트립 시간 및 지터이고 각 요소의 평균값 또는 표준 편차가 측정되었다. 표준 편차는 각 평가 요소에 대한 시스템 또는 네트워킹의 안정성을 가늠하는 지표로 사용된다. 평가 요소 중 시스템 부하와 영상 프레임 재생율은 중단 시스템의 성능을 나타내고 전송 처리율, 라운드 트립 시간 및 지터는 네트워킹 성능을 보여준다. 네트워킹과 관련된 요소들은 중단 시스템의 성능에 영향을 줄 수 있다.

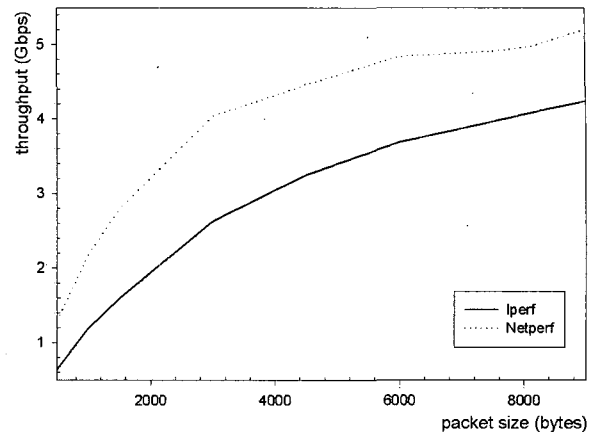
4. 실험 결과

4.1. 최대 전송 단위 (MTU)

(그림 2)는 중단 시스템에서 MTU 크기를 변화시키면서 10초간 측정된 평균 전송율이다. Force10 스위치의 MTU 크기가 9,180 바이트로 설정되었기 때문에 네트워크에서 단편화(fragmentation)가 발생하지 않도록 시스템의 MTU를 9,180 바이트 이하로 설정했다.

(그림 2)에서 같은 조건의 Iperf 실험과 Netperf 실험이 차이를 보이는 이유는 Iperf의 경우, acknowledged UDP 패킷을 사용하기 때문이다. 9,000 바이트 점보 프레임을 사용했을 때, 1,500 표준 프레임에 비해 약 86%의 전송 처리율이 향상되었으며 최대 5.1 Gbps의 중단간 처리율을 얻을 수 있었다. 일반 운영체제에서 메모리 블록을 4096바이트의 배수로 페이지징하기 때문에 8160바이트 MTU를 사용하게 되면, 9,000 바이트를 사용했을 때와 비교해서 메모리의 낭비를 최소화하고 전송 효율을 높일 수 있는 것으로 알려져 있다. 하지만, 실험 결과 8160 바이트 MTU를 사용했을 때 얻을 수 있는 성능 이득은 무시할 수준으로 평가된다.

MTU 크기가 증가하면 평균 처리율이 증가하지만, 8,000바이트 이상의 크기에서는 성능 이득이 줄어드는 것을 알 수 있다. 성능 이득이 줄어드는 이유는, 리눅스 커널의 메모리 할당



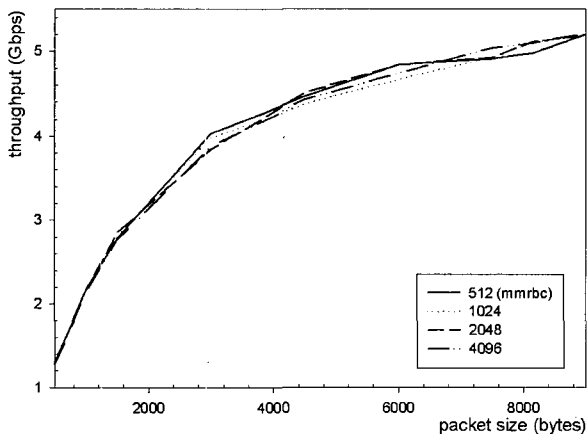
(그림 2) MTU 변화에 따른 처리율

시스템의 문제로, 대형 메모리 블록을 사용할 경우 인접한 페이지(pages)를 찾기 위한 시스템 부하가 커지기 때문이다[9].

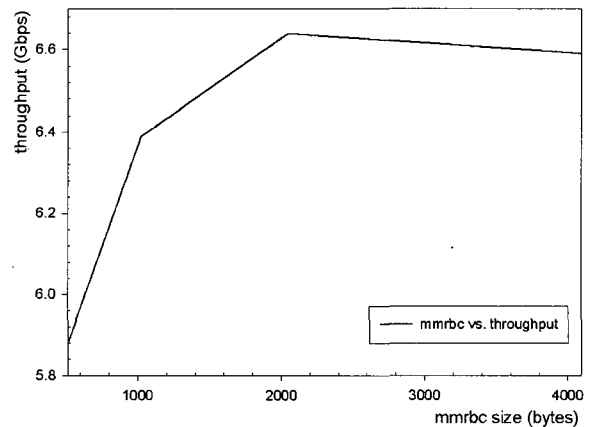
4.2. MMRBC와 점보프레임

(그림 3)과 (그림 4)는 MMRBC 값과 MTU 크기가 전송 처리율에 미치는 영향을 보여 준다. (그림 3)의 경우, netperf를 이용한 단일 스트림 전송 실험의 결과이다. Iperf를 이용해 테스트한 결과도 (그림 3)과 비슷하기 때문에 도시화하지 않았다. (그림 4)는 점보 프레임 설정 후, 다중 스트림(multiple stream) Iperf를 이용해 전송 시험한 결과이다.

1,500 바이트 표준 프레임과 MMRBC 값을 512 바이트로 설정했을 때, 약 3.3 Gbps의 전송 처리율을 보였고 4096바이트로 설정했을 경우에는 약 3.9 Gbps의 성능을 보임으로써 MMRBC의 증가가 약 15%의 성능 향상을 가져왔음을 확인했다. 하지만, (그림 3)에서 볼 수 있듯이 점보 프레임을 사용하는 단일 스트림 응용의 경우, MMRBC 값의 증가와 처리율의 향상이 비례관계에 있지 않음을 알 수 있다. 이는 MMRBC 값과 전송 처리율이 비



(그림 3) 단일 스트림 사용 시 MMRBC와 MTU 변화에 따른 처리율



(그림 4) 9,000 바이트 MTU와 다중 스트림 사용 시 MMRBC 변화에 따른 처리율

레한다는 기존 연구 결과 [9],[19]와 일치하지 않는 내용으로, 실험에서 사용된 Chelsio N210 네트워크 어댑터의 제품 특성으로 의심된다.

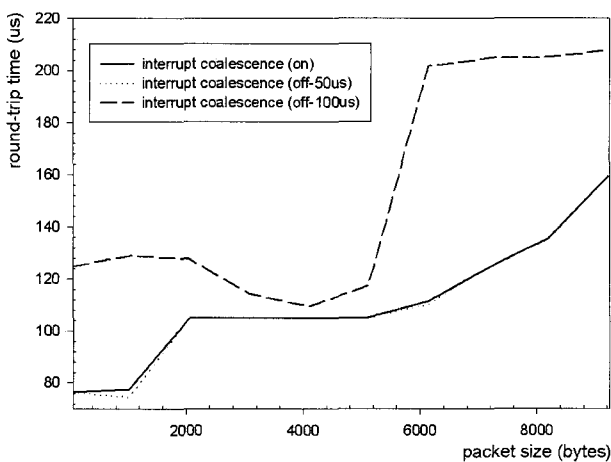
(그림 4)는 Iperf를 이용해 다중 스트림을 전송했을 때, MMRBC 값이 전송 처리율에 미치는 영향을 보여 준다. Chelsio N210 모델의 경우, MMRBC값이 2048 바이트 일 때 약 6.6 Gbps의 최대 성능을 보임을 알 수 있다. 이는 MMRBC의 기본 값인 512 바이트일 때 보다 약 12%의 처리율이 향상된 결과이다.

실험을 통해 MMRBC 값이 증가하면 약 12-15%정도의 전송 처리율 향상을 기대할 수 있고, 점보 프레임과 다중 스트림을 사용하는 경우에 MMRBC의 크기 증가가 처리율을 향상 시켰음을 확인할 수 있다.

4.3. Interrupt Coalescence

(그림 5)는 다양한 패킷 크기 환경 하에서 인터럽트 발생 주기를 변화시키면서 측정한 라운드 트립 시간의 결과이다. MMRBC 값은 2048바이트로 설정되었고 라운드 트립 시간의 측정을 위해 Hpcbench를 사용했다.

실험 결과, 동적 인터럽트 통합 기술을 이용했을 경우(on), 송신 시스템의 CPU 부하는 약 33.6%로 측정됐고 정적(static) 인터럽트 발생 주기(이하, 단위시간)를 50 μ s로 설정했을 경우(off-50 μ s)에는 약 34.6%를 나타냈다. (그림 5)에서 보듯이 두 경우에 해당하는 라운드 트립 시간의 결과 그래프는 매우 비슷했으나 인터럽트 통합으로 인해 약 1 %의 CPU 부하가 감소되었다. 측정된 라운드 트립 시간은 인터럽트 발생을 위해 인위적으로 설정된 단위 시간(50 μ s 또는 100 μ s)과 전달 지연 및 네트워크 엔티티(entity)에서의 패킷 프로세싱 시간이 더해진 것으로 볼 수 있다. (그림 5)를 통해 5120 바이트 이상의 패킷이 사용될 경우, 정적 인터럽트 발생을 위해 설정한 단위 시간의 값 보다는 패킷 프로세싱 지연이 전체 라운드 트립 시간에 큰 영향을 줌을 확인할 수 있다.



(그림 5) 인터럽트 통합 vs. 라운드 트립 시간

5. 응용 소프트웨어 적용 실험 및 결과

본 장에서는 앞서 실험한 시스템 튜닝값들이 실제 비압축 HDTV 응용에 미치는 영향을 조사한다. 중단 시스템은 실시간 라이브 콘텐츠를 전송 및 재생한다. 이를 위해, HD-SDI 미디어 인터페이스가 수정 구현된 UltraGrid 응용 소프트웨어가 활용되었다. UltraGrid는 RTP/RTCP(real-time transport protocol, RTP control protocol)를 이용해 콘텐츠를 전송하는데 RTCP는 라운드 트립 시간을 예측하는데 활용 가능하다[21]. 시스템 부하를 측정하기 위해서 리눅스 top 명령이 사용되었다. 성능 측정을 위해 5분 간 비압축 HDTV 콘텐츠를 전송했다. 이는 9,000 바이트 크기의 점보프레임을 사용했을 경우, 초당 약 17,300 패킷들이 전송된 것 또는 57.8 μ s 당 1개의 패킷이 송신 된 것과 동일한 결과이다.

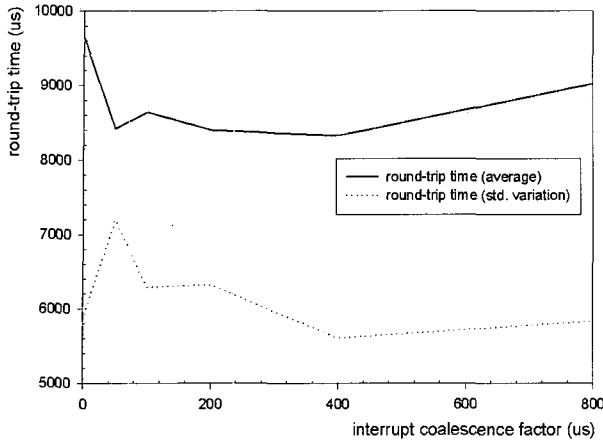
먼저, 시스템 튜닝이 없고 인터럽트 통합만 적용했을 때의 성능을 수신측 시스템에서 측정했다. 이때, 5분간 측정된 평균 CPU 부하는 약 80.175%이고 표준편차는 0.90372를 보였다. 패킷 손실율은 3-4%를 나타냈는데 이는 중단 시스템에서 수신 버퍼의 크기가 작게 설정되어 발생한 결과이다. 시스템 튜닝과 정(버퍼 크기, MTU, MMRBC 등의 최적화)을 거치게 되면 중단 시스템 버퍼에서의 패킷 손실을 막을 수 있고 시스템 부하는 약 1.2% 감소한다.

<표 1>과 <표 2>는 동적 인터럽트 통합 또는 정적 인터럽트 발생을 위한 단위 시간 사용 시 측정된, 수신측 시스템 부하 및 초당 프레임 재생율(fps)이다. 설정된 단위 시간 값이 200 μ s를 넘을 때 얻는 성능 이득은 시스템 부하 및 초당 프레임 재생율의 관점에서 상대적으로 낮아 보인다. 실험 결과, 비압축 HDTV 신호의 송신을 위해서 약 2-4개의 패킷 도착 시간(100-200 μ s)을 단위 시간 값으로 설정하는 것이 중단 시스템의 안정성(시스템 부하, 초당 프레임 재생 및 표준편차)을 향상시켰다.

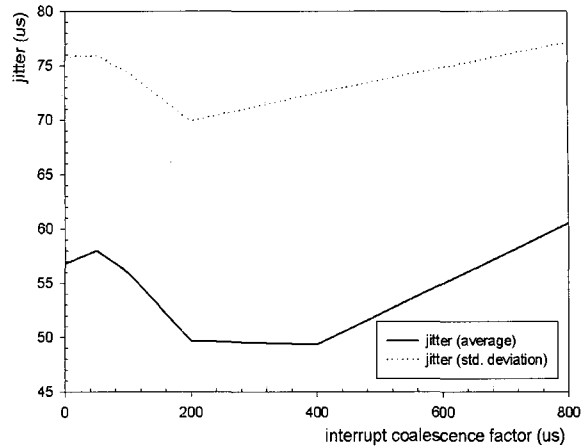
개발된 비압축 HDTV 시스템이 29.97 psf(progressive segmented frame)를 사용하기 때문에 <표 2>의 초당 프레임 재생율(fps)은 외부 요인이 없는 환경 하에서 29.97을 갖는다. 정적 인터럽트 값이 증가함에도 프레임 재생율이 29.96 fps를 유지하는 것은 개발된 시스템이 2개의 프레임 버퍼를 가지고 있어서 지터를 흡수하기 때문이다.

<표 1>과 <표 2>가 시스템의 응용 계층에서 측정된 반면, (그림 6)과 (그림 7)은 패킷 단위에서 측정된 결과 값들로 비압축 HDTV 시스템의 프레임 버퍼에서 보상(compensation)되는 요소들(지연 흡수 등)이 배제된 결과이다. (그림 6)과 (그림 7)은 수신측에서 측정된 값들을 바탕으로 송신측에서 획득한 자료이다. 두 그림의 x 축 시작점(0)은 동적 인터럽트 통합이 적용됐을 때의 결과이다.

(그림 6)의 라운드 트립 시간은 RTCP 보고(Receiver Report) 패킷을 이용해서 측정했다. 라운드 트립 시간은 ARQ(automatic repeat request)와 같은 오류 복구 방식에서 복구 성공률을 가늠할 수 있는 중요한 지표이다. 측정 결과, 정적 인터럽트 값을 100-400 μ s로 설정했을 때 전송 계층에



(그림 6) 인터럽트 통합 vs. 라운드 트립 시간



(그림 7) 인터럽트 vs. 지터

<표 1> 인터럽트 통합 (Interrupt coalescence) 값에 대한 시스템부하와 표준편차

	on	50	100	200	400	800
시스템부하	79.0	78.9	78.4	77.5	77.4	77.5
표준편차	0.48	0.44	0.36	0.56	0.50	0.41

<표 2> 인터럽트 통합 (Interrupt coalescence) 값에 대한 fps 평균과 표준편차

	on	50	100	200	400	800
fps	29.95	29.94	29.97	29.96	29.96	29.96
표준편차	0.15	0.16	0.08	0.11	0.11	0.11

서의 라운드 트립 시간 및 표준 편차가 상대적으로 낮음을 알 수 있다.

(그림 7)은 송신 시스템에서 수집한 수신단의 패킷 지터로써 인터럽트 통합과 지터와의 관계를 보여준다. 송신 시스템은 RTCP 보고 패킷을 통해 수신 시스템에 측정된 값을 전달받는다. 중단 시스템들이 Force10 스위치에 직접 연결되어 있으므로 배경 트래픽에 의한 지연 변이는 무시할 수준이다. 구축된 비압축 HDTV는 패킷의 비순차 도착(out-of-ordered arrival)으로 인한 재생 품질 저하와 동기화 문제를 막기 위해 소프트웨어적인 재순서화(re-ordering)과정을 거치게 된다. 네트워크 지터가 없을 경우, 패킷의 수신 간격이 약 57.8 us 이므로 재순서화는 us 단위의 지터에 영향을 받게 된다.

측정된 지터 값에서 볼 수 있듯이, 정적 인터럽트 발생을 위한 단위 시간 값의 증가가 지터를 크게 했으나, 지터의 크기가 단일 패킷의 송신 간격(spacing)인 약 57.8 us 수준이어서 응용 계층의 시스템 성능에는 큰 영향을 미치지 못함을 알 수 있다.

정적 인터럽트 값이 200-400 us로 설정될 때 지터가 낮게 발생했다.

본 장에서는 앞서 조사된 시스템 튜닝값들을 광대역 네트워크 응용인 비압축 HDTV 시스템에 적용했고 인터럽트 통합이 시스템 부하에 미치는 영향을 분석했다. 시스템 튜닝을 통해 중단 호스트에서 버퍼 자원 고갈로 인한 패킷 손실을 막을 수 있고 시스템 부하가 감소되었다. 또한, 정적인 인터럽트 통합을 통해 지터의 감소, 안정화된 프레임 재생율, 및 시스템 부하의 경감 등 성능 향상을 얻을 수 있었다. 비압축 HDTV 신호의 네트워크 전송을 위해 권장되는 정적 인터럽트 통합의 단위 시간은 100-200 us이다.

6. 결 론

본 논문에서는 시스템 튜닝 값들이 대용량 전송과 높은 연산 처리 능력을 요구하는 광대역 응용(비압축 HDTV)에 미치는 영향에 대해서 조사했다. 커널 버퍼 크기의 조절을 통해 시스템 버퍼에서의 패킷 손실을 막을 수 있다. 점보 프레임 및 MMRBC의 사용이 중단 간 전송 처리율을 증가시켰으며 시스템 부하를 감소시켰다. 또한, 정적으로 최적화된 인터럽트 통합은 패킷 간 지터를 줄이고 미디어 재생을 위해 필요한 시간 간격을 균일화시킴을 확인했다.

참 고 문 헌

[1] B. Alfonsi, "I want My IPTV: Internet Protocol Television Predicted a 'Winner,'" IEEE Distributed Systems Online, vol. 6, no.2, 2005.
 [2] S. Shea, J. Starren, and R.S. Weinstock et al., "Columbia university's informatics for diabetes education and

telemedicine (IDEATel) project: rationale and design," Journal of Am Med Inform Assoc., pp.49-62, 2002.

[3] L. Gharai, C. Perkins, and A. Saurin, "UltraGrid: A high definition collaboratory," USC/ISI, Sept., 2005. <http://ultragrid.east.isi.edu/>

[4] NTT Innovation Lab., "Uncompressed HDTV transmission system over the Internet," NTT Press Release, Oct., 2001.

[5] ResearchChannel and AARNet, "First multi-gigabit interactive video transmission between Australia and the US," AARNet news, Nov. 2004. <http://www.aarnet.edu.au/news>

[6] 조진용, 광재승, 변옥환, "대형 프레임이 중단간 전송 성능에 미치는 영향 분석," 한국정보처리학회논문지, 11-C권, 6호, pp. 787-789, 2004.

[7] J. Wu, S. Campbell, J. M. Savoie, H. Zhang, G. V. Bochmann, and B. St. Arnaud, "User-managed end-to-end lightpath provisioning over CA*net 4," in Proc. National Fiber Optic Engineers Conference (NFOEC), pp.275-282, 2003.

[8] T. Dunigan, M. Mathis, and B. Tierney, "A TCP Tuning Daemon," SC2002, 2002.

[9] W. Feng, J. Hurwitz, H. Newman, S. Ravot, R. L. Cottrell, O. Martin, and F. Coccetti, C. Jin, X. Wei, "Optimizing 10-Gigabit Ethernet for Networks of Workstations, Cluster, and Grids: A Case Study," SC2003, Nov., 2003.

[10] J. Jo, J. Kong, and O. Byeon "Optimizing System Performance for Uncompressed HDTV over Gigabit Ethernet: A Case Study," IEEE ICACT 2006, Feb., 2006.

[11] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and S. Tuecke, "The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets," Network and Computer Applications, vol.23, pp.187-200, 1999.

[12] VeriTest, "Chelsio Communications: 10-gigabit Network Adapter Performance Study," Aug., 2004.

[13] Force10 Networks, "<http://www.force10networks.com/>"

[14] "Iperf 1.7.0 - The TCP/UDP Bandwidth Measurement Tool," <http://dast.nlanr.net/Projects/Iperf/>

[15] "Netperf: Public Netperf Homepage," <http://www.netperf.org/>

[16] B. Huang, M. Bauer, and M. Katchabaw, "Hpcbench - A Linux-based Network Benchmark for High Performance Networks," in Proc. of the 19th International Symposium on High Performance Computing Systems and Applications, May, 2005.

[17] W. R. Stevens, "TCP/IP Illustrated, vol.1, the protocols," p.289.

[18] IEEE, "Carrier Sense Multiple Access with Collision Detection(CSMA/CD) Access Method and Physical Layer Specifications", ANSI/IEEE Std. 802.3-1985, 1985.

[19] H.J. Richard, C. Peter, and D. Steven, "Performance of Gigabit and 10 Gigabit Ethernet NICs with Server Quality Motherboards" Grid Edition of Future Generation Computer Systems, vol.21, pp.469-488, April, 2005.

[20] R. Prasad, M. Jain and C. Dovrolis, "Effects of Interrupt Coalescence on Network Measurements," in Proc. of Passive and Active Measurement (PAM) Workshop, April, 2004,

[21] P. Karn and C. Partridge, "Improving round-trip time estimates in reliable transport protocols," Computer Communications Rev., vol.17, no.5, pp.2-7, Aug., 1987.

조진용



e-mail : jiny92@kisti.re.kr

1999년 전남대학교 컴퓨터공학과(학사)

2002년 광주과학기술원 정보통신공학과
(공학석사)

2003년~현재 한국과학기술정보연구원
연구원

관심분야: 멀티미디어 스트리밍, Interactive 멀티미디어 응용기술,
멀티캐스트 모니터링, 고성능 네트워킹 등.

석우진



e-mail : wjseok@kisti.re.kr

1996년 경북대학교 컴퓨터공학과(학사)

2002년 노스캐롤라이나 주립대학교
(이학석사)

1998년~현재 한국과학기술정보연구원
선임연구원

관심분야: 고속통신망, 무선통신망, TCP 성능향상 등

이민선



e-mail : mleeoh@kisti.re.kr

1989년 숙명여자대학교 물리학과(학사)

1997년 Univ. of Nebraska-Lincoln, 전기
공학과(공학석사)

2000년 Univ. of Nebraska-Lincoln, 전기
공학과(공학박사과정수료)

1997년~2000년 Univ. of Nebraska-Lincoln, 전기공학과 강의
및 연구 조교

2005년~현재 한국과학기술정보연구원 선임연구원

관심분야: 데이터압축기술, 고성능 네트워크 기술 및 네트워크
관리등.



변 옥 환

e-mail : ohbyeon@kisti.re.kr

1979년 한국항공대학교 통신정보공학

1993년 경희대학교 전자공공학과

(공학박사)

1984년 미국 OSM Corp 객원연구원

1997년 미국 NCSA/UIUC 객원연구원

1978년~1995년 KIST 시스템공학연구소 책임연구원, 연구전산
망개발실장, 슈퍼컴퓨팅응용실장

1995년~1999년 ETRI 슈퍼컴퓨팅센터 책임연구원, 고성능망연
구실장, 슈퍼컴퓨팅연구실장

1999년~현재 KISTI 슈퍼컴퓨팅센터 책임연구원, 초고속연구망
부장, 슈퍼컴퓨팅인프라개발실장

관심분야: 고성능망 관리 및 보안, 그리드 네트워킹 및 협업 액
세스 그리드, 유비쿼터스 네트워킹 및 모바일 그리드 등