

# 비공유 저장장치를 가지는 클러스터 기반 시스템에서 다중 디스크립션 코딩을 이용한 스트리밍 방법

박유현<sup>0</sup>, 정진환, 이용주, 이상민, 김갑동, 김학영

한국전자통신연구원 디지털융연구단 인터넷서버그룹

## The Streaming Method using Multiple Description Coding for cluster-based server with shared-nothing storage

BAK, Yuhyeon<sup>0</sup> JEONG, Jinhwan LEE, Youngju WOO, Sangmin KIM, Gapdong KIM, Hagyoung

ETRI

[bakyh,jhjeong,youngju,smwoo,kdkim71,h0kim}@etri.re.kr](mailto:{bakyh,jhjeong,youngju,smwoo,kdkim71,h0kim}@etri.re.kr)

### 요 약

일반적인 비 공유 저장장치를 가진 시스템에서는 특정 콘텐츠에 대한 편향된(skewed) 요청 문제를 피할 수 없다. 본 논문에서는 비 공유 저장장치 모델로 설계/구현된 클러스터 기반의 스트리밍 서버에서 사용자의 편향된 콘텐츠 요청에 대처할 수 있고 스트리밍 서비스 중인 서버 또는 NS 카드에 문제가 생기더라도 스트리밍 서비스가 중단되지 않도록 다중 스트리밍 방법인 MDC(Multiple Description Coding)를 적용한 스트리밍 전송방법을 제안한다. 제안하는 방법은 클러스터 시스템에서 단일 세션으로 스트리밍 서비스를 제공하는 것이 아니라 MDC 를 활용하여 다중 세션을 통한 스트리밍 서비스를 제공함으로써 모든 서버의 부하가 같아지고, 특정 서버에 문제가 발생하더라도 스트리밍 서비스가 끊어지지 않으면서 MDC 의 특성인 적응형 서비스도 가능하게 되는 장점이 있다.

### 1. 서론

전통적인 클라이언트/서버 구조의 컴퓨터 시스템은 클라이언트가 데이터를 요청하면 서버가 서비스를 제공한다. 이러한 구조에서 데이터를 요청하는 사용자 수가 증가하면 단일 서버만으로는 모든 사용자의 요청을 처리할 수 없기 때문에 클러스터 기반의 다중 서버로 서비스를 제공하는 방법들이 제안되었다.

일반적인 클러스터 기반의 다중 서버를 가지는 시스템은 비 공유 저장장치를 가지는 구조이다. 따라서, 모든 데이터에 대해서 균등한 사용자 요청이 들어오면 최적의 성능을 내지만, VoD(Video On Demand) 서비스와 같이 인기 많은 데이터에 대한 요구와 인기 없는 데이터에 대한 요구가 현저하게 다르면, 인기 많은 데이터를 가지고 있는 서버는 사용자 요청을 제대로 처리하지 못하는 반면에, 인기 없는 데이터만을 가지는 서버는 시스템 자원을 낭비하는 단점을 가진다. 이러한 사용자 요청의 불균형으로 인해 시스템 자원의 효율적인 사용을 위한 부하분산에 관한 많은 연구가 이루어졌다.

이러한 시스템에서 부하 불균형 문제를 해결할 수 있는 대표적인 방법은 콘텐츠 복제와 스트라이핑(striping) 방법이 있으나 이에 따른 오버헤드가 존재한다.

본 논문에서는 클러스터 기반의 비 공유 저장장치를 가지는 스트리밍 서버에서 사용자의

편향된 데이터 요청에도 모든 서버의 부하를 동일하게 하면서, 특정 서버에서 에러가 발생하는 경우에도 스트리밍 서비스가 끊어지는 현상을 방지하기 위하여 MDC(Multiple Description Coding)를 활용하는 방법을 제안한다. 제안하는 방법은 또한 사용자 단말 및 네트워크 상황에도 동적으로 적응할 수 있는 장점을 가진다. 본 논문의 구성은 다음과 같다. 2 장에서는 스트리밍 서버의 유형, MDC 에 대해 살펴보고 3 장에서는 본 논문이 제안하는 시스템인 차세대 인터넷 서버에 대해서 설명한 후, 4 장에서 차세대 인터넷 서버로 구성된 클러스터 시스템에서 MDC 를 이용한 부하 분산, 서버의 고장 시 지속적인 스트리밍 서비스 방안 및 QoS 지원방안을 설명한다. 5 장에서는 구현 결과 및 시뮬레이션 결과를 살펴보고 6 장에서 결론을 맺는다.

### 2. 관련 연구

#### 2.1 스트리밍 서버의 유형

초기의 스트리밍 서버에 관한 연구는 주로 대용량 병렬 컴퓨터로 구성하였다. 하지만 경제성과 확장성을 위해서 클러스터 기반의 스트리밍 서버에 관한 연구들이 진행되고 있다[1]. 클러스터 기반의 스트리밍 서버는 비 공유/공유 저장장치를 가지는 스트리밍으로 구분할 수 있다.

대용량 병렬 스트리밍 서버는 대용량 병렬 컴퓨터와 대용량 저장 시스템 기술을 이용하여 구성하고, 중앙 집중식 처리 방식이기 때문에 시스템

구성 및 구현이 용이한 장점을 가지나 저장 공간의 확장이나 처리 용량 증대에 융통성이 적고 단일점 고장이 가능하기 때문에 가용성이 적으며 구현 비용이 높은 단점을 가진다.

비 공유 저장장치를 가지는 스트리밍 서버는 다수의 중저가형 시스템을 상호 연결하여 구성하며, 관리 서버는 사용자 요청을 받아 최적의 서버로 분배하는 방법이다. 저장 공간 및 처리 용량의 확장성이 있으나 저장공간을 공유하지 않기 때문에 특정 콘텐츠에 대한 접근이 불가능할 경우가 존재한다.

공유 저장장치를 가지는 스트리밍 서버는 초기 연구에서 주로 RAID 시스템으로 구현하였다. 이러한 시스템은 관리 및 구성하기 쉽고, 저장장치 비용이 비교적 저렴하였으나, 확장성과 신뢰성에 제약이 있으며, 부가 관리 비용과 고장 가능성이 높은 단점을 가졌다. 최근에는 다수의 중저가형 시스템을 연결하고 SAN(Storage Area Network)과 같은 고속의 네트워크로 연결된 저장장치로 구성하는 방법이 제안되고 있다. 이러한 구성은 저장 공간 및 처리 용량의 확장성이 있고, 전체 서버에 문제가 생기지 않는 한 어떤 콘텐츠라도 접근이 가능하다. 하지만 이 방법 역시 SAN 을 지원하기 위한 고가의 장비가 추가적으로 필요한 단점을 가진다.

## 2.2 MDC (Multiple Description Coding)

MDC 는 벨 연구소에서 최초로 개발된 방법이다[2]. 그 후에 소스 신호에 대한 MDC 는 정보 이론 분야와 신호 처리 분야에서 널리 연구되어 왔다[3],[4]. 신호 처리 분야에서 MDC 의 개념은 음성 압축에 차례로 적용되었으며[5] 그 후에 영상과 동영상 압축에 차례로 적용되었다[6],[7]. MDC 는 하나의 스트림을 여러 개의 분리된 비트 스트림으로 코딩 하는 것을 의미한다. 즉, 하나의 스트림을 여러 개의 스트림으로 분할하고 각 스트림에는 자신의 정보와 다른 스트림에 관한 부가 정보를 추가하여 어느 하나의 스트림을 잃더라도 나머지 스트림으로 복구할 수 있고 모두를 수신하게 되면 송신 시와 똑 같은 품질의 콘텐츠를 생성할 수 있는 코딩 기술이다[8][9][10]. 이것은 Layered Coding 과는 달리 각 비트 스트림은 서로 독립적이며 어느 것이 더 중요한 것이 없이 각각 똑 같은 중요도를 가지며 어떤 스트림을 수신하든 재생이 가능하고 수신되는 스트림이 많을수록 재생 품질이 좋아지는 장점이 있다[11].

## 3. 시스템 구성[12] 및 시스템 제한점

차세대 인터넷 서버는 고성능 인터넷 환경, 고품질 무선통신 환경, 높은 인터넷 사용율을 보유한 한국에서 HDTV 급 고품질, 실시간 스트리밍 서비스를 제공하는 고성능 인터넷 서버이다.

차세대 인터넷 서버는 두 개의 프로세서(processor)와 1 개 이상의 NS (Network/Storage) 카드로 구성된다.

NS 카드[13],[14][15]는 SCSI 저장장치, 이더넷 네트워크 컨트롤러, 메모리를 하나의 카드에 장착하여 PCI 인터페이스를 통하여 호스트와 접속되는 PCI 카드이다. NS 카드는 저장장치로부터 읽은 데이터를 메인 메모리에 로드시키지 않고 PMEM 이라고 하는 NS 카드 내에 장착된 메모리로 데이터를 읽었다가 네트워크 인터페이스를 통해 바로 네트워크로 전송하는 기능을 가진다. NS 카드에서는 TCP/UDP 스택을 처리하는 TOE(TCP Offload Engine) 칩을 사용하고 있으며, 이로 인하여 불필요한 메모리 복사 없이 데이터 스트림을 저장장치에서 네트워크로 전송시킬 수 있다. 따라서 NS 카드는 저장장치와 네트워크 성능에서 기존의 I/O 가속 장치보다 적은 CPU 사용율과 높은 성능을 가진다.

차세대 인터넷 서버는 다음과 같은 소프트웨어 컴포넌트[16]를 가진다.

- 멀티미디어 파일 시스템[17]
- 콘텐츠 스트리밍
- 콘텐츠 분배 모듈
- 운영체제
- 메타데이터 관리기

차세대 인터넷 서버의 지역서버 클러스터는 파일서버, 디스패처와 다수의 스트리밍 서버들로 구성된다. 각 스트리밍 서버는 1~3 개의 NS 카드를 장착하여 스트리밍 서비스를 제공한다.

차세대 인터넷 서버는 저비용의 서버 구성을 위해 비 공유 저장장치를 가지는 모델로 개발되었다. 따라서 다음과 같은 문제점이 발생할 수 있다.

- 특정 콘텐츠에 대한 집중적인 요청이 발생하면 해당 콘텐츠를 저장하고 있는 특정 NS 카드 또는 특정 서버에 부하가 집중된다.
- 스트리밍 서비스를 하는 중에 서비스를 제공하는 NS 카드 또는 서버가 문제가 발생하면 사용자 측에서는 스트리밍 서비스가 중단된다.

이러한 문제점을 해결하기 위해서 미리 콘텐츠를 복수 개의 NS 카드에 중복해서 저장하거나, 동적으로 콘텐츠를 복제(replication) 방법이 사용되었다. 하지만, 콘텐츠 중복으로 인한 저장공간의 낭비, 동적 복제의 경우 정확한 시점을 선택하는 오버헤드와 복제로 인한 서비스의 지연 문제와 콘텐츠를 모든 NS 카드에 복제하지 않을 경우에는 콘텐츠 배치 문제가 발생하게 된다.

## 4. MDC 를 이용한 차세대 인터넷 서버 설계

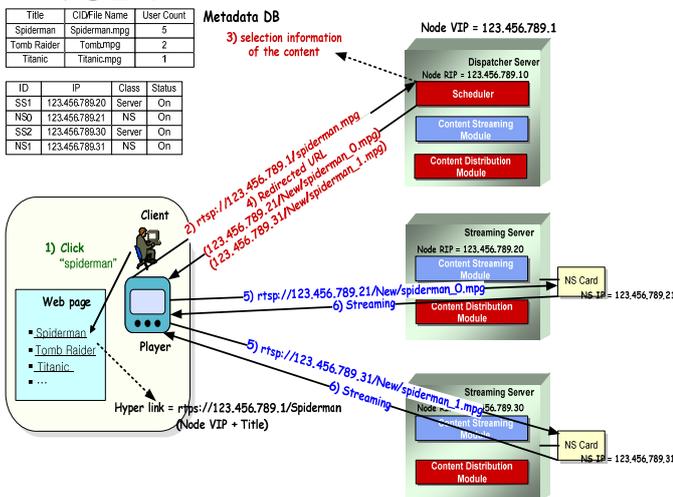
3.5 에서 언급한 시스템 제한점을 해결하기 위해 MDC(Multiple Description Coding)를 적용하여 설계하였다. 차세대 인터넷 서버와 같이 비 공유 저장장치를 가진 시스템에 MDC 를 적용하면 3.5 에서 언급된 시스템의 제한점을 해결할 수 있다. 또한 MDC 를 활용하게 되면 단말기, 네트워크 상태, 서버 또는 클라이언트의 부하 상태에 따라 적응적으로

스트리밍 서비스를 제공할 수 있다.

### 4.1 스트리밍 서비스 단계

MDC 를 이용한 스트리밍 서비스는 다음과 같은 단계를 거쳐 제공된다.

- 1) 사용자는 서비스 받을 콘텐츠를 선택한다.
- 2) 사용자의 요청은 디스패처 서버로 전송된다.
- 3) 디스패처 서버는 사용자의 요청에 해당하는 description 의 이름과 위치를 사용자에게 전송한다.
- 4) 사용자는 디스패처 서버로부터 전달받은 NS 정보를 바탕으로 각 description 에 대해 해당 위치로 rtsp 요청을 수행한다.
- 5) 요청 받은 NS 카드들은 스트리밍 서비스를 수행한다.



[그림 1] MDC 를 이용한 스트리밍 서비스 단계

### 4.2 부하 분산 및 서버 에러 대응 방법

이와 같이 차세대 인터넷 서버에 MDC 를 적용하게 되면 요청 받은 콘텐츠에 대한 서비스를 모든 NS 카드가 제공하게 되어 모든 NS 카드의 부하가 균일하게 된다. n 개의 NS 카드로 구성된 스트리밍 시스템은 사용자에게 동시에 n 개의 스트리밍을 제공하고 사용자는 이를 받아서 디코딩 한 후 플레이 한다.

만일 스트리밍 서비스를 하고 있는 중에 특정 NS 카드 혹은 특정 서버에 문제가 생겨서 더 이상 스트리밍 서비스를 제공할 수 없는 경우에도 사용자 측에서는 스트리밍 서비스가 중단되는 것이 아니라 처음 서비스를 받을 때 비해서는 콘텐츠의 질은 떨어지지만 정상적으로 서비스를 제공하는 NS 카드들로부터 전송 받는 description 을 통해 서비스를 계속 받을 수 있다. 또한 정상적인 NS 카드들은 비정상적인 NS 카드들이 저장하고 있는 description 을 파일 서버로부터 전송 받은 후에는 사용자에게 고품질의 스트리밍 서비스를 제공할 수 있다. MDC 를 적용하지 않았을 때에 비해서 파일서버로부터 전송 받아야 하는 콘텐츠의 크기가 작아지기 때문에 사용자의 대기시간은 짧아지게 된다.

### 4.3 단말 적응 및 네트워크 적응 스트리밍 서비스

사용자의 스트리밍 요청 시 단말기의 성능이 떨어지거나 네트워크의 상태가 불량 또는 서버의 부하로 인해 요청 콘텐츠의 description 모두를 보낼 수 없거나 보낼 필요가 없는 경우가 발생한다. 이러한 경우에는 그 상태에 맞는 적당한 description 의 수를 결정하여 스트리밍 서비스를 하며 이는 스트리밍 서비스 중에서도 동적으로 변화될 수 있다. 서버의 부하 정보를 제외한 모든 서비스 적응에 관한 부분은 클라이언트가 판단한다. 즉, 클라이언트는 초기에 서버로부터 n 개의 description 을 받고 있다가 전송 지연이 생기거나 클라이언트의 부하로 인해 n 개의 description 을 재생시키지 못하게 되면 디스패처 서버에 description 을 줄이는 명령을 보낸다. 디스패처 서버는 n 개의 NS 카드 중에서 부하가 가장 큰 NS 카드를 선택하여 해당 클라이언트에 대해 더 이상 스트리밍 서비스를 수행하지 않도록 명령한다. 이때 해당 클라이언트에 대해 스트리밍 서비스를 수행하는 NS 카드는 n-1 개가 된다.

### 5. 성능 평가

본 논문에서는 차세대 인터넷 서버에서 MDC 를 활용했을 경우의 성능 측정을 위해서 아리조나 대학의 MDC 자료 [18]를 참조하여 시뮬레이션 하였다. 참조한 비디오의 이름은 mobile 이며, 파일 크기는 약 12.9MB 이고, 상영시간은 12 초, 평균 bit rate 는 8.36Mbps 이다. [표 1]은 mobile 비디오의 자세한 속성을 나타낸다. [그림 2]은 이 데이터를 기반으로 description 의 수에 따른 대역폭의 변화를 나타낸다. 1 개의 description 의 대역폭은 8.6Mbps 이지만, 20 개의 description 으로 인코딩 한 경우에는 각 description 1 개의 대역폭은 0.9Mbps, 20 개 전체의 대역폭 합은 약 18.3Mbps 가 된다.

[표 1] mobile(CIF Video Sequence)

Video name	Mobile	Number of frames	300
File size(Byte)	13,503,432	Min Frame size(Byte)	16,360
Time(s)	12	Max Frame size(Byte)	196,944
Mean bit rate	9,002,288	peak bit rate	63,022,080

따라서, 4.3 에서 설명한 서비스 중인 NS 카드 중에서 하나가 문제를 발생한 경우, 1 개의 description 인 경우는 최소 8.6Mbps 의 속도 이상으로 파일 서버로부터 데이터를 받아와야 에러복구가 가능하지만, 20 개의 description 의 경우는 최소 0.9Mbps 의 속도만 보장되어도 파일 서버로부터 콘텐츠를 받아오면서 스트리밍 서비스를 할 수 있기 때문에 네트워크 자원을 효율적으로 사용할 수 있는 장점을 가진다.

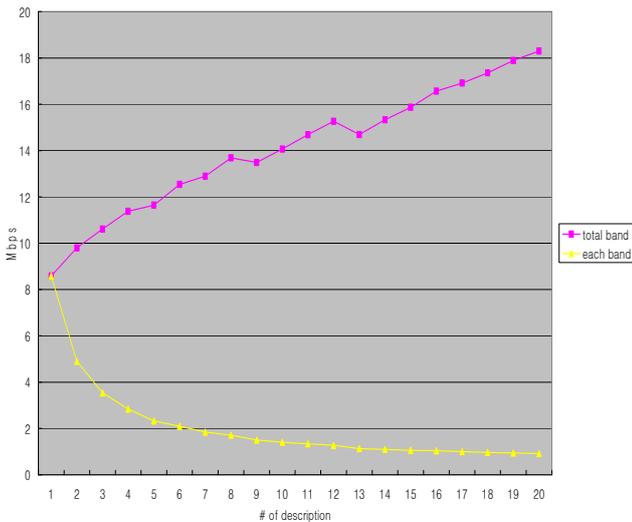


그림 2] description 에 따른 bandwidth 의 변화

## 6. 결론

차세대 인터넷 서버는 비 공유 저장장치를 가지는 서버이다. 각 서버는 디스크-네트워크 전송의 효율을 위한 NS 카드가 장착되어 있으며, 이 카드에는 2 개의 기가비트 네트워크 포트가 있으며 여러 개의 SCSI 디스크가 연결되어 있다. 따라서, 차세대 인터넷 서버는 NS 카드의 도움으로 일반 서버에 비해 적은 자원을 사용하면서도 많은 사용자에게 스트리밍 서비스를 제공할 수 있다. 하지만, 비 공유 저장장치 구조를 가지는 시스템이 가지는 부하 불균형 문제를 피할 수 없다. 부하 불균형을 해결하기 위해서 콘텐츠 복제를 사용할 수 있지만 복제로 인한 저장장치의 낭비되고 이를 위한 오버헤드가 존재한다. 따라서 본 논문에서는 MDC 를 이용하여 비 공유 저장장치 구조를 가지는 시스템에서 부하분산을 효율적으로 수행하면서도 적응적으로 스트리밍 서비스를 제공할 수 있는 방법을 제안하였다. 또한 스트리밍 서비스 중에 이를 제공하는 서버 및 NS 카드의 문제가 발생하더라도 스트리밍 서비스가 중단되지 않고 QoS 에 맞게 스트리밍 서비스를 계속할 수 있다. 문제 발생 이후에 파일 서버로부터 문제가 되는 description 을 정상 NS 카드에 복구시키면 스트리밍 서비스 중에 양질의 스트리밍 서비스를 계속해서 제공할 수 있다.

## 7. 참고 문헌

[1] W. Bolosky, J. Barrera, R. Draves, G. Fitzgerald, Gibson, M. Jones, S. Levi, N. Myhrvold, and Rashid R, "The Tiger video fileserver", In Proc. NOSSDAV'96, 1996

[2] Michael Zink, Andreas Mauthe, "P2P Streaming using Multiple Description Coded Video", Proceedings of EuroMicro2004, Sep. 2004

[3] L. Ozarow, "On a source coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, pp. 1909-1921, Dec. 1980.

[4] V. A. Vaishampayan, "Design of multiple description scalar quantizer," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821-834, May 1993.

[5] A. Ingle and V. A. Vaishampayan, DPCM system design for diversity systems with applications to packetized speech," *IEEE Trans. Speech and Audio Processing*, vol. 3, pp. 48-57, Jan. 1995

[6] M. Orchard, Y. Wang, V. A. Vaishampayan, and A. Reibman, "Redundancy rate distortion analysis of multiple description image coding using pairwise correlating transforms," *Proc. Int. Conf. Image Processing*, Oct. 1997, pp. 608-611

[7] Y. Wang, M. Orchard, V. A. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Processing*, vol. 10, pp. 351-366, Mar. 2001.

[8] Apostolopoulos, J.G., "Error-resilient Video Compression via Multiple State Streams", In Proceedings of VLBV'99. Oct. 1999.

[9] 김일구, 조남익, "다중 표현을 이용한 에러에 강인한 동영상 부호화 방법", 방송공회논문지 2004년 제 9권 제 1호.

[10] 김일구, 조남익, "비트율-왜곡 최적화된 DCT 계수 부호화를 이용한 다중 표현 동영상 압축 방법", 방송공학회 논문지, Vol.9, No. 1, 2004.3

[11] 김선호, 송병호, "멀티미디어 콘텐츠의 QoS를 개선한 전송 매커니즘", 정보처리학회 논문지 B, 제 12-B권 제 2호, 2005.4

[12] "Next Generation Internet Server Requirement Specifications V1.0", ETRI, 2002.9

[13] Chong-Won Park, Myung-Joon Kim, Jin-Won Park, "Design of NGIS: The Next Generation Internet Server for Future E-society", GCC2003, Shanghai, China, 2004

[14] Byung Kwon Jung, June Kim, Myung Joon Kim, Sung-Ho Hahm, "Implementation of high speed multimedia streaming card using the zero-copy mechanism for streaming server", Proceedings of International Technical Conference on Circuits/Systems, Computer and Communications(ITC-CSCC 2005), Vol.2, 2005

[15] Byung-Kwon Jung, Sung Woo-Seok, Seong-Woon Kim, Myung-Joon Kim, "Performance analysis of high speed multimedia streaming card for streaming server", International Symposium of Signals, Circuits and Systems(ISSCS 2005), Vol.2, 2005

[16] Chang-Soo Kim, Hag-Young Kim, Myung-Joon Kim, Jae-Soo Yoo, "Design of System for Multimedia Streaming Service", EUC05, Nagasaki, Japan, 2005.

[17] Baik-Song Ahn, Sung-Hoon Sohn, Chei-Yol Kim, Gyu-Il Cha, Yun-Cheol Baek, Sung-In Jung, Myung-Joon Kim, "Implementation and Evaluation of EXT3NS Multimedia File System", Proceedings of the 12<sup>th</sup> annual ACM international conference on Multimedia, New York, USA, Oct. 2004

[18] Video Traces Research Group, <http://trace.eas.asu.edu/tracemain.html>