

고화질 VoD 서비스를 위한 멀티미디어 파일시스템 구현

김갑동^o 석성우 박유현 이용주 김창수 김학영

한국전자통신연구원

{kdkim71^o,swsok,bakyh,jongju,cskim7,h0kim}@etri.re.kr

Implementation of the Multimedia File System for HD VoD Service

Kap-Dong Kim^o, Song-Woo Sok, Yu-Hyeon Park, Yong-Ju Lee, Chang-Soo Kim, Hag-Young Kim

ETRI(Electronics and Telecommunications Research Institute)

요 약

고화질의 VoD 서비스를 위한 차세대 인터넷 서버는 디스크 컨트롤러, 네트워크 어댑터와 메모리를 하나의 하드웨어에 집적한 네트워크 스토리지 카드를 기본으로 리눅스 환경에서 개발되었다. 본 시스템에서 제공하는 입출력 경로를 리눅스에서 사용되고 있는 EXT3 파일시스템과 호환이 되도록 설계하고 구현하였다. 본 파일시스템은 기존의 EXT3 파일시스템의 기능을 그대로 사용하면서 대용량의 고화질 멀티미디어 데이터를 고속으로 처리하고, 시스템의 자원을 최소로 사용하면서 VoD 서비스에서 요구하는 데이터 입출력 처리 성능을 제공한다. 본 논문에서는 차세대 인터넷 서버 환경에서 개발된 멀티미디어 파일시스템의 설계, 구현 내용과 성능을 기술한다.

는 스토리지에서 네트워크로, 네트워크에서 스토리지로 고속의 데이터 전송을 가능하도록 하기 위하여 대용량의 버퍼를 스토리지와 네트워크 인터페이스 장치들이 위치하는 PCI 버스 세그먼트 내에 두는 구조로 대용량의 버퍼 메모리로 DDR SDRAM을 채용하고, 이를 제어하기 위하여 PCI Memory Controller를 사용하며, 디스크 저장장치의 인터페이스로는 Ultra320 SCSI를, 네트워크 인터페이스로는 Gigabit Ethernet을 사용한다. Zero-copy 메커니즘을 제공하기 위해서 PCI 버스를 통해서 접근 가능한 PCI 메모리를 제공하고, 고성능의 네트워크 기능을 제공하는 TOE는 PCI 메모리의 사용을 지원하면서 데이터 이동 경로를 최적화하여 PCI 메모리로부터 직접 네트워크로 스트림 데이터 전송이 가능하도록 하고, TCP/UDP/IP Checksum Offload와 Scatter/Gather 입출력 기능을 구현하였다.

1. 서 론

차세대 인터넷 서버는 지역망을 기반으로 하는 빌딩, 대학, 아파트 등과 같은 지역 서비스를 위해 20Mbps 통신망 환경에서 1000명의 사용자에게 HDTV급 고품질 멀티미디어 서비스를 지원하는 서버 기술을 사용한 고품질 VoD 서비스를 제공하는 종합적인 솔루션을 제공한다. 본 기술은 네트워킹 가속기술을 기본으로 채용하여 고품질의 실시간 콘텐츠 전송이 가능하며, 고품질 VoD 서비스 제공을 위한 시스템 및 각 요소 기술을 제공함으로써 서비스 확장 및 응용이 용이하다. 본 서버는 고품질 스트리밍 서비스 기반 인터넷 서비스를 제공하는 서비스 인프라, 아파트, 학교, 빌딩등과 같은 지역 커뮤니티 서비스 시스템, 고성능 멀티미디어 처리능력을 요구하는 광대역 통합망 기반 서비스 인프라 장비 등으로 활용가능하다. 그림 1은 차세대 인터넷 서버의 구성도를 나타낸다. 원활한 멀티미디어 서비스를 위해 차세대 인터넷 서버는 기존의 파일시스템이 지원하는 기능과 완전한 호환성을 가진 대용량의 고품질 멀티미디어 서비스를 제공할 수 있도록 개발된 새로운 파일시스템(이하 EXT3NS)을 적용하였다. 본 논문에서는 파일시스템의 개발환경에 대하여 설명하고, EXT3NS의 설계 내용과 성능에 대하여 기술하였다.

본 논문의 구성은 먼저 2장에서 파일시스템이 개발된 환경인 차세대 인터넷 서버의 환경에 대하여 설명하고, 3장에서는 EXT3NS 파일시스템의 설계 내용과 성능에 대하여 기술하고, 끝으로 4장에서 결론을 맺는다.

2. 차세대 인터넷 서버

차세대 인터넷 서버는 다음과 같은 하드웨어와 시스템 소프트웨어, 스트리밍 어플리케이션으로 구성이 된다.

2.1 하드웨어

하드웨어 플랫폼의 주요 구성요소인 네트워크 스토리지 카드

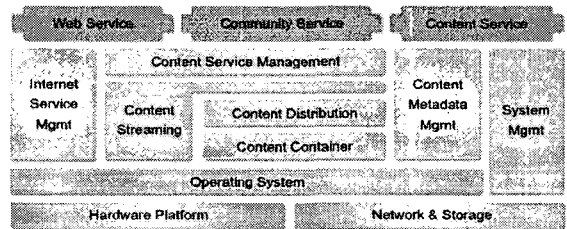


그림 1 차세대 인터넷 서버 구성도

2.2 운영체제

운영체제는 멀티미디어 콘텐츠 분배와 네트워크 처리가 강화된 특성화된 리눅스 커널로 하드웨어 자원을 관리하고 효과적 인 시스템 서비스를 제공한다. 커널은 리눅스 커널 2.6.10을 기반으로 한 운영체제를 사용하고, 각 노드를 구성하는 서버는 하드웨어가 2개의 Intel Xeon 프로세서, 최대 16GB의 메인 메모리, 그리고 최대 4개의 네트워크 스토리지 카드를 지원한다. 라이브러리 및 명령어는 사용자에게 응용프로그램 인터페이스를 제공하고 리눅스 사용자 라이브러리, 표준 리눅스 명령어로 구성된다.

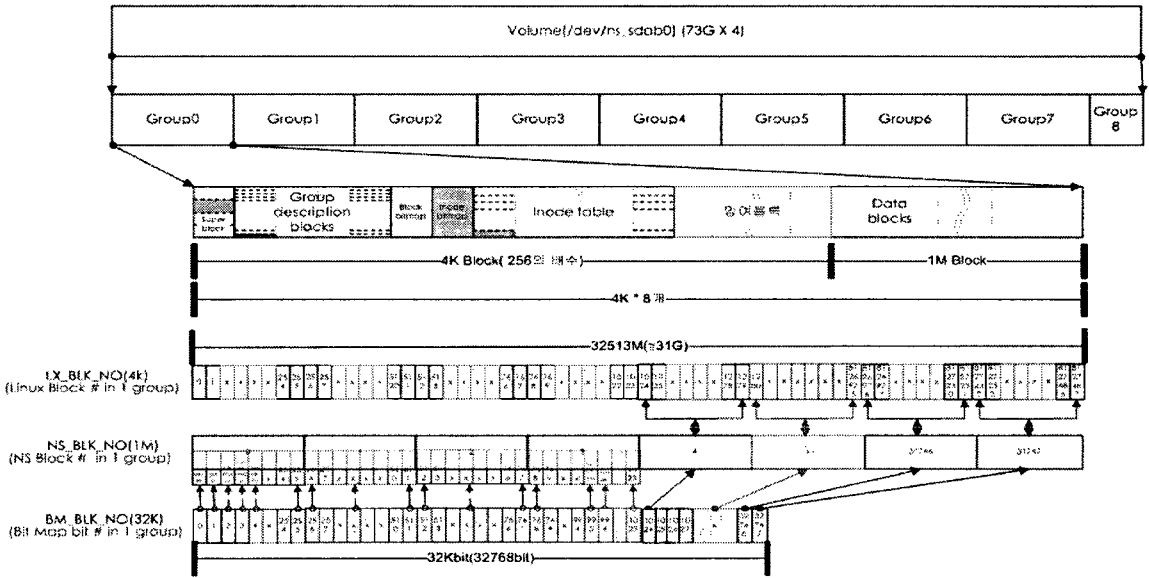


그림 2 EXT3NS 파일시스템 레이아웃

2.3 스트리밍 어플리케이션

스트리밍 어플리케이션은 콘텐츠 스트리밍을 지원하는데 필요한 프로토콜 엔진과 스트리밍 엔진으로 구성이 된다. 프로토콜 엔진은 서버와 클라이언트 사이에 제어 메시지를 전송하는 스트리밍 제어 프로토콜, 실제 콘텐츠를 전송하는데 필요한 스트리밍 전송 프로토콜, 스트리밍 엔진과의 인터페이스 역할을 하는 메인으로 구성된다. 이것은 클라이언트와 스트리밍 서버 사이에서 전송되는 RTSP 제어 메시지를 네트워크를 통하여 실제로 송수신하는 역할, 스트리밍 서버에서 클라이언트로 스트리밍하는 멀티미디어 데이터의 전송 및 멀티미디어 데이터 송수신 관련 제어 메시지를 네트워크를 통하여 전송하는 역할, 프로토콜 엔진 서비스의 전체적인 기능을 관리하고 제어하는 역할을 수행한다. 스트리밍 엔진은 콘텐츠와 프로토콜, 전송 방식에 따라 콘텐츠의 데이터를 클라이언트에 전송하는 일을 담당하며, MPEG-1/2 PS와 MPEG-2 TS, MPEG-4 isma를 기본으로 콘텐츠를 제공한다.

3. EXT3NS 파일시스템

기존의 응용프로그램과 호환성을 높이기 위하여 표준 파일 시스템 지원을 기존의 리눅스 커널에 포함되어 있는 EXT3 파일 시스템을 차세대 인터넷 서버에 장착된 네트워크 스토리지 카드에 맞도록 확장하여 구현한 EXT3NS 파일시스템을 설계 구현하였다. 기존에 표준 파일 시스템으로서의 인터페이스를 갖추지 못한 멀티미디어 전용 파일시스템의 경우 기존의 응용프로그램과 호환성이 없어 시스템의 관리 비용이 많이 들고 병용성이 떨어지는 단점을 가지고 있었다. EXT3NS 파일시스템은 EXT3 파일시스템에서 채택하고 있는 디스크 레이아웃을 네트워크 스토리지 카드가 지원하는 큰 블록을 지원하도록 확장하였다. 그림 3에서 보는 것과 같이 리눅스 커널 2.6의 파일시스템 구조에 EXT3NS를 추가하였다.

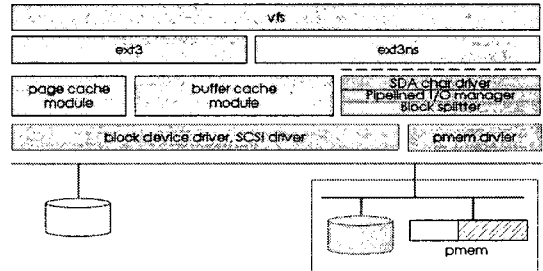


그림 3 리눅스 커널 2.6 EXT3NS 파일시스템 구조

3.1. 파일시스템 설계

기존의 EXT3 파일시스템은 1K, 2K, 4K의 블록 사이클을 지원하지만 네트워크 스토리지 카드는 256K, 512K, 1M, 2M의 상대적으로 큰 블록 사이클을 지원한다. 확장된 디스크 레이아웃을 정확하고 효율적으로 관리하기 위해 EXT3NS 파일시스템의 데이터 블록의 할당 및 해제 루틴을 수정하였다. 또한, 기존의 파일시스템과 달리 EXT3NS 파일시스템은 네트워크 스토리지 카드에 포함된 PCI 메모리를 활용하여 높은 성능을 달성하기 위해 read 및 write 연산을 수정하였다. 즉, read 및 write 연산의 대상이 되는 메모리 영역이 일반 메모리일 경우 기존의 EXT3와 동일한 형식으로 리눅스 커널의 캐시 영역을 통한 입출력을 수행하지만, 연산 대상 메모리가 PCI 메모리인 경우 높은 성능을 달성하기 위해 리눅스 커널의 캐시 영역을 건너뛰어 NS 카드에서 제공하는 Character 디바이스를 활용한다. 그림 2는 리눅스의 기본 블록 크기를 4K로 하고 네트워크 스토리지 카드가 지원하는 블록의 크기를 1024K로 설정한 EXT3NS의 디스크 레이아웃의 예를 보여준다. 그림 2에서 보는 것과 같이 기존의 EXT3 디스크 레이아웃의 블록 비트맵에서 사용하는 기본 단위를 이원화하여 시스템 영역의 디스크 할당과 비트맵의 관계는 기존의 EXT3 파일시스템과 동일하게 유지

하는 반면, 데이터 영역은 대용량을 처리하는 크기의 블록으로 할당을 하고 비트맵에서도 대용량 블록과 매핑을 시키도록 하였다. 시스템 영역의 블록할당의 경우는 대용량의 데이터 영역의 할당과 다르므로 그림 2에서 잉여블록으로 표시된 것과 같이 데이터 블록의 크기에 맞도록 추가적인 영역을 두었다.

3.2. 입출력 메커니즘

네트워크 스토리지 카드의 드라이버가 제공하는 SDA(Stream Disk Array)는 최악의 경우의 블록 액세스 패턴에서도 적어도 1Gbps 대역폭으로 전송할 수 있는 초고속의 디스크 입출력 성능을 제공한다. 비디오/오디오 데이터는 블록 분할 방식에 의하여 디스크 어레이를 구성하는 모든 디스크들이 동일하게 병렬적으로 동작하여 아주 높은 성능을 제공하며, 블록 분할 방식과 더불어서 파이프라인 방식을 개발하여 어떠한 경우에도 디스크의 사용률이 어느 디스크 어레이 방식보다 월등하며 디스크 접근시간을 줄여준다. SDA는 Linux Character Device Driver, Stream Disk Array Interface, Pipelined I/O manager, Block Splitter, Disk Array Configuration, Linux Block Device Driver로 구성되어 있다. 스트림 처리를 위한 빠른 경로는 Stream Disk Array Interface로 입력되고, 일반적인 속도의 입출력은 Linux Block Device Driver를 통한다. 각 네트워크 스토리지 카드마다 독립적인 Request queue가 있고 일단 입출력 요구를 하면 이 큐에 들어가게 된다. Pipelined I/O manager가 커널 쓰레드로 동작하며, Request queue에 있는 입출력 요구를 라운드 로빈 방식으로 가져와서 디스크 어레이로 보낼 명령들을 큐잉한다. 이 명령들은 Block Splitter를 통하여 각각의 디스크 명령으로 변환된다. Block Splitter는 Linux block I/O driver와 정합되어 있고, Linux block I/O driver는 raw 블록 입출력 접근을 가능하게 한다. 그리하여 저수준 블록 입출력을 사용하는 stream disk array는 멀티미디어 스트림 서비스를 위한 파일시스템의 새로운 디스크 어레이를 제공한다. EXT3NS 파일시스템은 기존의 read(), write() 인터페이스를 그대로 사용하여 사용자 메모리를 사용한 입출력의 요구는 기존의 리눅스 파일시스템 경로를 통하여 처리하여 주고, PCI 메모리를 사용한 입출력의 경우 SDA 경로를 통하여 큰 블록(256K, 512K, 1M, 2M)의 데이터를 처리하도록 설계, 구현되었다.

3.3. 파일 시스템 성능

EXT3NS 파일시스템의 성능을 측정하기 위해서 먼저 차세대 인터넷 서버의 하드웨어와 디바이스 드라이버가 제공하는 입출력 성능을 측정하여 보았다. 네트워크 스토리지 카드가 제공하는 디스크의 입출력 성능은 15K RPM 73GB SCSI 디스크 4개를 연결한 경우, 블록크기에 따라 다음과 같은 성능을 보인다.
 2MB : 1156.8Mbps, 1MB : 1168.2Mbps,
 512K: 900.7Mbps, 256K: 641.2Mbps
 디스크의 기본 입출력 성능을 바탕으로 EXT3NS 파일시스템에서 5GB를 순차적으로 읽는 경우의 평균 시간을 측정해 보면 사용자 메모리로 읽는 일반 리눅스 입출력 경로를 통하여 177.49초, PCI 메모리를 통하여 사용자 메모리로 이동하는 경로는 64.09초, PCI 메모리로 직접 이동하는 경로는 49.28초이다. PCI 메모리로 읽는 경우가 사용자 메모리로 읽는 경우보다 2.76배나 빠르면서 CPU의 부하도 PCI 메모리로 읽는 경우가 사용자 메모리로 읽는 경우의 40% 정도밖에 차지하지 않는다. 이렇게 일반 파일시스템에 비하여 월등한 입출력 성능을 가진 파일시스템을 이용하여 앞에서 기술한 스트리밍 어플리케이션

으로 차세대 인터넷 서버가 서비스 할 수 있는 스트림의 수를 측정하여 보았다. 측정한 환경은 3.2GHz Xeon 프로세서가 두 개 장착되었고, 네트워크 스토리지 카드는 1장에서 3장까지 사용하였다. 사용된 컨텐츠는 MPEG-2 TS 10M이다. 스트리밍 성능으로 카드 1장은 179 스트림, 2장은 370 스트림, 3장은 525 스트림의 성능을 보이며, CPU의 자원도 여유가 있어 시스템이 안정적으로 운영이 된다.

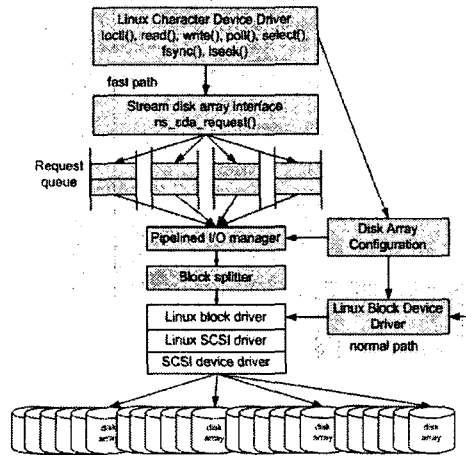


그림 4 SDA 구조

4. 결론

본 논문에서는 차세대 인터넷 서버 환경에서 네트워크 스토리지 카드를 기반으로 리눅스 커널 2.6의 가상파일시스템의 표준 인터페이스를 통하여 고속으로 큰 블록의 입출력을 처리해주는 EXT3NS 파일시스템을 설계하고 개발을 완료하였다. 본 논문에서 기술된 파일시스템은 표준 인터페이스를 사용하였으므로 일반 VoD 응용 프로그램들을 EXT3NS 파일시스템을 사용하여 성능을 향상시킬 수 있다. 또한, 표준 파일시스템 인터페이스를 제공하므로 VoD 뿐만이 아닌 고속의 입출력을 필요로 하는 다양한 분야에 응용이 가능하다.

참고문헌

- [1]석성우, 김갑중, 김학영, "차세대 인터넷 서버 (SMART)에서 PMEM을 사용한 캐싱", 제9회 차세대 통신소프트웨어 학술대회, 2005.12
- [2]백성훈, 김성훈, "주변메모리장치 및 스트림 디스크 어레이를 이용한 네트워크 스토리지 가속장치의 소프트웨어 구조 및 성능평가" Conference on Computer Software, 2004.10
- [3]김명준, 임기욱, "차세대 인터넷 서버 (SMART 서버) 기술 개발", 한국컨텐츠학회지, 제1권, 제1호, 2003.06
- [4]김성훈, 김명준, 김보관, "차세대 인터넷 서버를 위한 스트리밍 가속장치", 전자공학회 추계학술대회, pp. 49-52, 2003
- [5]백성훈, 김성훈, "스트리밍 디스크 어레이: 매우 큰 블록 크기를 가지는 블록 분할 디스크 인터리빙", Conference on Electronics and Information Communication, pp.20-23, 2003