

VOD 시스템에서의 실시간 응용 프로그램을 위한 디스크 배열 구조의 성능 분석

정 현 식*

Analysis for Performance of RAID Architecture for Real-Time Application as VOD Storage System

Hyun-Sik Chung*

요 약

멀티미디어 화일 전송과 같은 실시간 응용에서 저장시스템으로서 디스크 배열의 사용은 필수적이다. 멀티미디어 데이터 저장 서버에서 동시에 보다 많은 고객들에게 요청되어지는 서비스를 제공하기 위해서는 여러 분야에서의 다양한 정책들이 고려될 수 있으나, 특히 시스템의 환경과 비디오 데이터의 특성에 따른 스트라이핑 블록 크기의 관계와 같은 여러 가지 디스크 배열 특성변수들은 성능을 좌우하는 중요한 요인이다. 본 논문에서는 실제 멀티미디어 서버의 환경에 저장시스템과 관련된 특성변수들을 다양한 작업부하에 적용해 봄으로써 최적의 저장시스템 구조를 평가하였다.

Abstract

Real-time applications as multimedia file transfer, it is necessary to use a disk array as a system. For providing requested service to much more clients at the same time in a multime storage server, considerable varied strategies in a several parts. But especially it is important to various parameters for disk array such as relation of striping block size as a system envir characteristics of video data. In this paper, we evaluated the suitable storage system archite implementing them applying synthetically generated workloads that varied in the param interest in real environment of multimedia server.

* 경도대학 컴퓨터정보과학과 전임강사

I. 서론

고속 네트워크의 발전 속도로 인해 문자정보, 정지화상, 동영상과 같은 멀티미디어 정보에 대한 온라인 서비스를 가능하게 하였다. 하지만, 멀티미디어 서버에서 제공되는 대부분의 멀티미디어 데이터들은 전송 지연에 민감하기 때문에 실시간 요구 조건(real-time constraints)을 충족시키면서 연속적으로 제공될 때만 의미 전달이 가능하다[1]. MPEG-1화일의 경우 초당 약 1.5Mbps, MPEG-2화일의 경우 약 6Mbps의 전송속도를 각각 만족시켜야 한다. 그러므로 네트워크 상에서 멀티미디어 서비스를 지원하기 위해서는 QoS(Quality of Service) 유지 전송이 필수적이라 할 수 있다[2]. 이를 만족시키기 위해서는 전처리 과정으로 멀티미디어 서버의 저장시스템은 데이터의 저장 및 재생을 미디어 데이터의 전송속도에 부합하도록 효과적으로 처리해줄 수 있어야한다. 다수의 고객들의 요구에 응답하는 즉각적인 실시간 재생기능은 대단히 중요한 요소이다. 본 논문에서는 MPEG 스트림의 효과적인 실시간 재생을 위해 실제 멀티미디어 서버의 환경 상에서 MPEG-1 화일을 데이터로 사용하여, 저장시스템의 작업부하에 따른 최적의 저장시스템 환경을 위한 특성변수들을 도출해 낸다.

본 논문의 구성은 다음과 같다. 2장에서 디스크 배열 구조에 대해 논의하고, 3장에서는 모델링 기법에 대해 나열하고, 4장에서는 RAID 구조 결정 및 평균 데이터 전송률을 분석한다. 5장에서는 결론을 나타낸다.

II. 디스크 배열 구조

멀티미디어 데이터는 연속성과 많은 저장 공간을 필요로 한다. 이러한 멀티미디어 데이터를 한 개의 디스크에 저장시킬 경우 읽기 시 접근 시간이 느려 몇 개의 클라이

언트밖에 사용할 수 없다. 이러한 단점은 여러 개의 디스크를 동시에 사용하여 입출력의 병렬성(parallel)과 병행성(concurrency)을 제공할 수 있는 디스크 배열을 이용하여 해결할 수 있다. 디스크 배열은 동시에 많은 수의 디스크를 접근하므로 대규모의 데이터를 저장하거나 데이터를 고속으로 전송할 수 있는 장점이 있다. 그러므로, 멀티미디어 서버를 위해 디스크 배열의 데이터 전송율과 관련된 특성변수들을 결정하는 것이 필요하다.

일반적으로 디스크의 성능비교에 사용되는 특성변수들은 회전대기 시간(rotational delay), 명령 오버헤드(overhead), 디스크 드라이브의 전송률이다. 여러 개의 디스크를 사용하여 입출력 처리가 빈번한 응용에서는 회전대기시간이 디스크의 수에 비례하여 증가하며 고속 데이터 액세스 및 전송을 할 수 없다. 이것은 각 디스크가 근본적으로 회전동기가 되지 않는데 기인한다. 회전동기가 되지 않은 시스템에서의 전송률은 회전동기가 된 시스템에 비하여 요구율과 요구되는 데이터의 크기가 증가할수록 현저히 저하된다[8]. 그러므로 동기화된 디스크 배열은 한번의 접근으로 다수의 디스크로부터 많은 데이터를 동시에 입출력하여 클라이언트로 데이터를 전송하므로 짧은 전송시간에 많은 양의 데이터를 전송할 수 있다.

멀티미디어 서버 시스템의 하드웨어는 서버, 클라이언트 및 네트워크로 크게 나눌 수 있다. 본 논문에서 실측 모델링을 위해 구축한 서버의 하드웨어 환경은 (그림 1)과 같다.

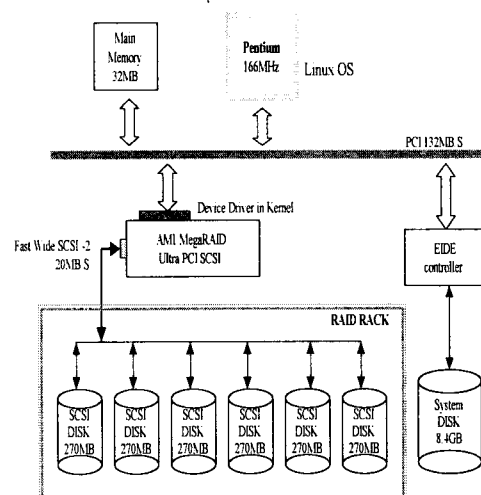


그림 1. 멀티미디어 서버의 하드웨어 환경
Fig 1. Hardware Environment of Multimedia Server

이 시스템은 RAID 컨트롤러로 American Megatrends 사의 PCI MegaRAID Ultra SCSI를 설치하였다. 이 컨트롤러는 3개의 채널을 지원하며, 1채널당 15개의 디스크를 장착할 수 있다. 8MB의 메모리를 가지며, 2KB에서 256KB사이의 스트라이핑 단위를 지원한다(9). 그리고 Fast Wide SCSI-2방식의 케이블로 연결된 RAID RACK내부에는 Quantum사에서 제조된 같은 용량의 LPS270S SCSI 하드디스크 6개가 연결되어 있다. Fast Wide SCSI-2 방식의 버스는 <표 1>에서 보듯이 최대 20MB의 초당 전송율을 가진다. 따라서 초당 2~4MB의 데이터 전송율을 갖는 SCSI 디스크로 디스크 배열을 구성할 경우 5~10개 이상의 디스크 장착시 SCSI 버스에 병목현상이 발생한다. 이 시스템은 RAID로 구성된 저장시스템의 환경에 대한 모델링을 위한 것이다. 저장시스템의 환경에 대한 특성변수들은 상대적인 결과로서 충분히 의미가 있고, 적은 디스크 용량은 매 실험마다 디스크의 초기화를 위한 시간을 단축시키기 위한 것이다. 또한 <표 2>는 본 논문에서 사용된 LPS270S SCSI 하드디스크의 물리적 특성을 나타내는 매개변수이다.

표 1. SCSI 버스의 종류
Table 1. A kind of SCSI bus

	SCSI-1	Fast SCSI-2	Fast Wide SCSI-2	Ultra SCSI (Fast-20)
Synchronous Transfer rates	5MB/S	10MB/S	20MB/S	40MB/S
Max Num of Devices	7	7	15	15
SCSI bus clock rate	5 MHz	10 MHz	10 MHz	20 MHz
Width of the bus	8 bits	8 bits	16 bits	16 bits

표 2. 디스크 매개변수
Table 2. Disk parameter values

Quantum LPS270S	
Cylinders per Disk	2,740
Tracks per Cylinder	4
Sector per Track	48
Bytes per Sector	512
Disk Capacity	270MB
Average Seek time	13ms
Single cylinder seek time	2.0ms
Max stroke seek time	25ms
Rotational Speed	4,500rpm
Average latency	6.67ms
Interface type	Fast Wide SCSI-2

III. 모델링 기법

멀티미디어 서버 시스템과 같은 실시간 응용 환경에서 고객으로부터의 요구는 일반적으로 쓰기 요구보다 읽기 요구의 비율이 대부분이다. 본 논문에서는 이러한 요구 특성을 고려하여 고객들로부터의 모든 실시간 데이터 요구를 읽기 요구로 가정하였으며, 이를 위한 측정 프로그램은 멀티미디어 클라이언트에 적용되듯이 쓰레드를 이용해 작성하였다. 재생률이 1.5Mbps인 MPEG-1 파일의 실시간 재생을 보장기 위해서는 초당 185KB의 데이터가 저장시스템으로부터 서비스되어야 한다. 연속 매체 파일 시스템의 데이터 요구 블록 크기로 256KB로 결정하였는데, 265KB블록을 이용하매초 185KB의 데이터를 제공하기 위해서는 3초당 2번씩 데이터 요구 블록을 공급하여야 한다. 최대한의 고객들에게 이러한 요구를 만족시키는 최적의 특성변수를 구하기 위하여 실측을 위한 멀티미디어 서버 시스템의 처리 절차를 매 실험마다 반복한다. 일반적으로 입출력 시스템의 성능 비교시 단위 시간당 평균 데이터 전달률이 높을수록, 응답시간이 짧을수록 성능이 우수하다고 평가한다. 디스크 시스템의 응답 시간은 서비스 시간과 큐 대기 시간의 합으로 정의되는데, 탐색 시간과 회전지연시간 등 물리적인 요인들이 응답 시간을 결정하는 주요한 요인들이다. 따라서, 동일한 상황에서 디스크 입출력 횟수를 줄이고 한 번의 디스크 접근으로 전달되는 데이터 양을 늘려주면 단위 시간당 데이터 전달률이 높아지고, 응답 시간이 감소되어 디스크 시스템의 입출력 성능이 향상된다.

IV. RAID 구조 결정 및 평균 데이터 전송률 분석

하나의 요구를 모든 디스크가 동기화되어 처리하는 비

트 인터리브 방식의 RAID레벨보다는 여러 사용자의 요구를 각 디스크가 독립적으로 서비스하는 스트라이핑 방식의 RAID레벨이 더욱 효율적이다. 병행성을 이용한 RAID 레벨로는 RAID-0, 1, 4, 5가 있다. RAID-0는 중복검사를 위한 패리티와 관련된 오버헤드가 없기 때문에 구성 중 가장 간단한 RAID 모델이며 패리티 정보를 갱신할 필요가 없으므로 읽기 성능 뿐 만 아니라 쓰기 성능도 우수하다. 그러나 RAID-0는 패리티 정보가 없기 때문에 저장 데이터의 신뢰성은 가장 뒤떨어진다. RAID-1은 반사형(mirroring) RAID 레벨로서 하나의 RAID를 구성하는 디스크의 수가 n개일 경우, 2n개의 디스크가 소요된다. 비용면에서 볼때 100%의 오버헤드를 지나는 반면 신뢰성면에서는 가장 우수하다고 할 수 있다. RAID-4는 RAID-0에 비해 신뢰성도 우수하고 비용면에서도 하나의 패리티 정보를 위한 디스크가 존재하므로 RAID를 구성하는 디스크의 수가 n개일 경우, 100/(n+1)%의 오버헤드만이 존재한다. 그러나 패리티 정보 하나의 디스크에 집중되어 있으므로 쓰기 작업시 병목현상이 초래된다. RAID-5는 패리티를 모든 디스크에 균등하게 분산시켜 RAID-4의 단점인 패리티 디스크의 병목현상을 피할 수 있다. 모든 디스크가 읽기 요구에 기여할 수 있으며, 중복검사를 수행하는 다른 RAID 레벨에 비해 가장 우수한 성능을 나타내고 있다(6,7). 입출력 시스템은 멀티미디어 서버의 성능에 가장 결정적인 역할을 하므로 여기에 가장 적합한 RAID 레벨의 결정은 중요하다.

그러므로 본 논문에서는 병행성을 제공해주는 4가지 RAID 레벨 중 상대적으로 가장 우수한 데이터 전송률을 제공 할 수 있는 RAID-0와 RAID-5의 두 가지를 선택하여 적용하였다. 그러므로 본 논문에서는 병행성을 제공해주는 4가지 RAID 레벨 중 상대적으로 가장 우수한 데이터 전송률을 제공 할 수 있는 RAID-0와 RAID-5의 두 가지를 선택하여 실측에 적용하였다. 각 특성변수에 대한 실측 결과를 비교한 후 이 두 가지 중에서 사용할 멀티미디어 서버의 저장 시스템에 가장 적합한 RAID 레벨을 결정하고자 한다. 그리고 이용하여 스트라이핑 단위에 대한 데이터 요구의 크기와 작업부하의 병행도에 따른 최대 처리률을 분석한 후 멀티미디어 서버에 가장 적합한 디스크 배열의 특성 변수들을 도출하고자 한다.

4.1 RAID 구조 결정

본 절에서는 병행도에 따른 최적의 RAID 구조를 결정

하고자 한다. 병행성을 제공해주는 RAID 레벨 중 상대적으로 가장 우수한 데이터 전송률을 제공 할 수 있는 RAID-0와 RAID-5에 대해 디스크 수가 6개인 경우, 고객 수 변화에 따른 세 가지의 스트라이핑 단위에서의 성능을 비교하였다. (그림 2)에서 보듯이, 점선은 RAID-0를 나타내고, 실선은 RAID-5를 나타낸다. 전반적으로 레벨 0의 데이터 전송률이 약간 우수하지만, 고객 수가 많아질수록 거의 차이가 없음을 알 수 있다. 저장시스템 내 멀티미디어 데이터의 신뢰성을 고려한다면, 데이터 전송률의 차이가 거의 없는 레벨 5가 가장 이상적이다..

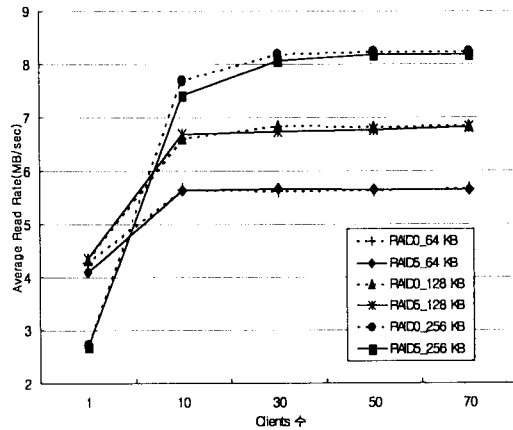


그림 2. RAID 레벨에 따른 평균 데이터 전달률
Fig 2. Average data transfer rate according to RAID Level

4.2 단일 디스크 평균 데이터 전송률 비교

단일디스크에 대해 내고장성을 가지는 RAID5구조에 있어, 단일디스크 고장시 새로운 디스크로의 교체 후, 고장난 디스크의 데이터를 재건하기까지의 기간인 MTTR(Mean Time To Repair)이 길어질수록 상대적으로 RAID5의 성능은 고장난 디스크상의 데이터를 (그림 3)과 같이 매번 재건(rebuild)시켜 요구에 응답하여야 하므로 그에 따른 부하로 인해 성능은 급격하게 저하된다(4). 이러한 RAID 시스템의 단일디스크 고장은 디스크의 수가 증가할수록 자주 발생할 가능성이 매우 크며, 이때 고장난 디스크상의 데이터를 즉시 복구시키는 것은 입출력 성능에 매우 중요한 영향을 끼친다. 대부분의 멀티미디어 서버 시스템의 서비스는 읽기 요구에 대한 응답이므로 요청작업의 수가 많아질 수록 디스크 배열내의 고장난 디스크로의 부하가 많아져 2배 가까운 성능의 저하가 발생할 수 있다(5). 본 절에서는 단일 디스크 고장시 다양한 디스크

크 배열 특성변수에 따른 성능을 분석해 보았다. (그림3)에서 보듯이 실선은 정상모드에서의 성능이며, 점선은 실패모드에서의 성능이다. 고객의 수를 늘려가면서 스트라이핑 단위 64KB, 128KB, 256KB에 대해 비교해 보았다. 실험 결과 고객의 수가 증가할 수록 작업부하가 커짐에 따라 성능의 저하가 두드러짐을 알 수 있다. 또한 스트라이핑 단위가 증가할 수록 데이터 전송률의 증가에 따른 상대적인 성능 저하가 역시 발생한다.

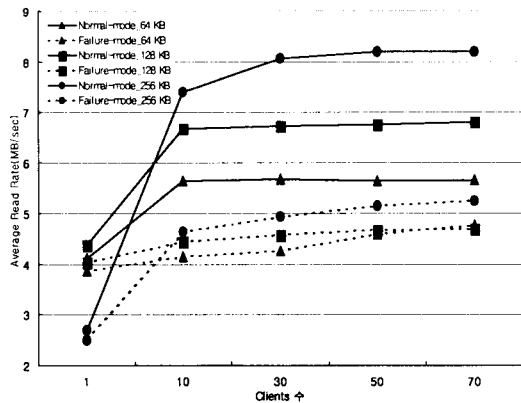


그림 3. 단일 디스크의 성능 분석
Fig 3. Performance analysis on single disk

V. 결론

멀티미디어 서버 시스템과 같은 환경을 구축하고, 단일 디스크 고장시 다양한 디스크 배열 특성변수에 따른 성능을 분석해 보았다. 최근 가장 많이 사용되는 동영상 매체로서 MPEG-1 데이터를 실시간으로 재생하기 위해서는 1초동안 약 185KB의 데이터간 전송되어야 한다. 그리고 실시간 응용을 위한 저장시스템으로써 RAID 구조의 결정에 있어 레벨 0와 레벨 5가 전송률은 거의 비슷함을 보였다. 이러한 이유는 대부분 멀티미디어 서버의 요구가 읽기 요구, 즉 재생 서비스 위주이기 때문이다. 하지만 데이터의 오류나 디스크 장애시에 대비하여 RAID 레벨 5를 사용하는 것이 신뢰도면에서 효율적일 것이다. 디스크 수에 따른 평균 데이터 전달률은 디스크 수에 비례하여 증가함을 알 수 있었으나, 최적의 디스크

수는 SCSI 버스 대역폭의 한계에 반영된다는 것을 알 수 있다.

채널 수에 따른 성능 분석에서는 작업 부하가 적을 경우에는 채널간의 통신 오버헤드로 인해 단일 채널이, 그리고 작업 부하가 커질수록 복수 채널이 우수한 데이터 전송률을 보임을 확인하였다. 그러므로 단일 디스크 고장시 평균데이터 전송률의 비교에서 고객의 수가 증가할수록 작업부하가 커짐에 따라 성능의 저하가 두드러짐을 알 수 있었으며, 스트라이핑 단위가 증가할 수록 데이터 전송률의 증가에 따른 상대적인 성능 저하가 발생하였다.

본 논문에서 실제 멀티미디어 서버 시스템상의 동화상 데이터를 사용하여 얻어낸 여러 가지 특성 변수의 값들은 다양한 실시간 멀티미디어 데이터 저장시스템으로서의 RAID 시스템 환경 구성 즉, 컨트롤러 BIOS 설정시 RAID 구조, 디스크 수, 스트라이핑 단위, 디스크 블럭, 채널 수 등의 결정에 중요한 자료로 활용되어 디스크 배열 구조에 따른 처리 효율에 도움이 될 것이다.

참고문헌

- [1] Jenwei Hsieh, Mengiou Lin, Jonathan C. L. Liu, and David H. C. Du, "Performance of A Mass Storage for Video-On-Demand," Journal of Parallel and Distributed Computing, 1994.
- [2] Gerhard Weikum, Peter Zabback, Peter Scheuermann, "Dynamic File Allocation in Disk Arrays," Proc. ACM SIGMOD Conference, Denver, May, 1991.
- [3] Gerhard Weikum, Peter Zabback, "Tuning of Striping Units in Disk Array-Based File Systems," Proc. 2nd IEEE International Workshop on Research Issues in Data Engineering, Feb. 1991.
- [4] Shenze Chen, Don Towsley, "A Queuing Analysis of RAID Architectures," Department of Computer and Information Science.

University of Massachusetts, 1992.

[4] Peter M. Chen and David A. Patterson, "Maximizing Performance in a Striped Disk Array," Proc. of the 1990 International Symposium on Computer Architecture, pp. 322-331, May, 1990.

[5] Edward K. Lee and Randy H. Katz, "An Analytic Performance Model of Disk Arrays and its Application," Technical Report UCB/CSD 91/660, Berkeley CA, 1991.

[6] Edward K. Lee and Randy H. Katz, "An Analytic Performance Model of Disk Arrays," Proc. of the 1993 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, pp. 90-109, May, 1993.

[7] American Megatrends, "MegaRAID and MegaRAID Ultra PCI SCSI Disk Array Controller Hardware Guide," Jun. 1998.

[8] Edward K. Lee, "Performance Modeling and Analysis of Disk Arrays," Ph. D Thesis, Carnegie Mellon University, 1994.

[9] D. A. Patterson, G. A. Gibson, and R. H. Katz, "A Case for Redundant Arrays of Inexpensive Disks(RAID)," International Conference on Management of Data(SIGMOD), pp. 109-116, June, 1988.

저자 소개



정 현 식
 1987년 2월 경일대학교 전산공학과 졸업(공학사)
 1990년 2월 영남대학교 대학원 전자공학과 졸업(공학석사)
 1997년 2월 영남대학교 대학원 전산공학과 박사수료
 1999년 3월~현재 경도대학 컴퓨터정보과학과 전임강사
 관심분야 : 이동통신데이터베이스, 분산처리시스템, 컴퓨터 구조, 멀티미디어