

실시간 멀티미디어 데이터를 위한 RAID 구조의 실측 분석

전 상 훈[†] · 안 병 철^{††}

요 약

최근 MPEG등의 멀티미디어 데이터를 실시간으로 재생 서비스하는 멀티미디어 서버의 사용이 인터넷의 발전과 더불어 급격히 증가되고 있다. 이러한 멀티미디어 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 있다. 그러나, 특히 시스템의 환경과 비디오 데이터의 특성에 따른 스트라이핑 블록 크기의 관계와 같은 디스크 배열을 위한 다양한 특성변수들은 성능을 좌우하는 중요한 요인이다. 본 논문에서는 MPEG-1 파일을 제공하는 실제 멀티미디어 서버의 환경을 구축하고, 디스크배열을 위한 다양한 특성변수들을 여러 가지로 발생되어지는 작업부하에 적용해 봄으로써 최적의 저장시스템 구조를 평가하였다.

Experimental Analysis of RAID Architecture for Real-Time Multimedia Data

SangHoon Jeon[†] · ByoungChul Ahn^{††}

ABSTRACT

Recently, Usage of multimedia server is rapidly increased with expansion of internet that real-time displaying service for multimedia data, MPEG and etc. For these multimedia real-time applications, it is necessary to use a disk array as a storage system. For providing requested service to much more clients at the same time in a multimedia data storage server, considerable varied strategies in a several parts. But especially it is important feature to various parameters for disk array such as relation of striping block size as a system environment and characteristics of video data. In this paper, we implemented real environment of multimedia server which provides MPEG-1 files and evaluated the suitable storage system architectures by applying synthetically generated workloads in the various parameters for disk array.

키워드 : 실시간 멀티미디어 데이터(Real-Time Multimedia Data), 실측 분석(Experimental Analysis), 디스크 배열(Disk Array), 재건(Rebuild)

1. 서 론

정보통신기술의 발달은 대용량의 데이터를 빠른 시간 내에 처리하도록 요구하고 있다. 이러한 요구는 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 전송 속도를 각각 만족시켜야 한다. 그러므로 네트워크 상에서 멀티미디어 서비스를 지원하기 위해서는 QoS(Quality of Service) 유지 전송이 필수적이라 할 수 있다[2]. 이를 만족

시키기 위해서는 전처리 과정으로 멀티미디어 서버의 저장 시스템은 데이터의 저장 및 재생을 미디어 데이터의 전송 속도에 부합하도록 효과적으로 처리해줄 수 있어야 한다. 다수의 고객들의 요구에 응답하는 즉각적인 실시간 재생기능은 대단히 중요한 요소이다.

많은 양의 비디오 정보를 저장하고, 안정적이고 신뢰성 있는 서비스를 제공하기 위해서 고성능 저장장치로 RAID 시스템의 사용이 보편화되어 있다[3]. RAID시스템과 같은 디스크 배열은 다수의 디스크 상에 데이터를 분산 저장함으로써 대용량의 저장 공간을 얻을 수 있을 뿐 아니라, 다수의 디스크에 병렬로 접근함으로써 데이터 전송률도 증가하므로 멀티미디어 응용분야에 적합하다. 이러한 멀티미디어 응용에서 디스크 배열의 스트라이핑 단위, 입출력 요구의 크기, 디스크 배열내 디스크의 수, 디스크 배열 구성형태 등은 QoS를 만족시키는 척도가 된다[4].

Weikum은 디스크 배열을 구성하는 디스크 수를 미리 외부 환경 변수로 정의하였으며, 디스크들이 비동기라고 가정

[†] 정 회 원 : 경동정보대학 게임애니메이션과 교수

^{††} 종신회원 : 영남대학교 전자정보공학부 교수

논문접수 : 1999년 12월 7일, 심사완료 : 2001년 8월 30일

하고 요구 특성을 분석하였다. 요구 크기와 요구 도착률이 주어졌을 때 하나의 요구가 몇개로 나누어져 각 디스크에서 동시에 서비스를 받을 경우 I/O응답 시간이 최소가 되는 조건을 제시하였다[5, 6]. 그러나 이 연구는 디스크 배열을 구성하는 디스크 수를 미리 정한 것이 단점이다. Chen은 Weikum과 같은 가정에서 요구 크기는 한 블록 정도의 적은 요구 크기와 전체 디스크 수만큼의 크 요구 크기 두 가지만을 고려하였다. 디스크 서비스 시간과 요구 도착 시간이 지수 분포를 따를 때 디스크 요구에 따른 평균응답 시간을 계산하는 분석 모델을 제시했다[7].

Peter와 Patterson은 16개의 디스크로 구성된 RAID-0를 대상으로 시뮬레이션을 통해 스트라이핑 단위와 병행성의 이율배반성(tradeoff)을 조사하였으며, 스트라이핑 단위의 크기는 작업 부하의 평균 크기보다 작업부하의 병행성에 밀접하게 관련되어 있음을 보였다[8].

Lee와 Katz는 비중복(non-redundant) 디스크 배열의 분석 모델(analytic model)을 제시하여 최적의 스트라이핑 단위를 구하는 방정식을 유도하고, 이 방정식을 이용하여 스트라이핑 단위는 디스크 성능 곱(performance product)과 병행성에 종속적임을 보였다. 그러나 이러한 종속성도 5% 이내의 성능 손실을 감안하면 무시할 수도 있음을 Peter와 Patterson은 보였다[9, 10].

Park은 VOD서버를 위한 디스크 배열 특성변수를 실측과 시뮬레이션 상에서의 성능 비교를 통해 다소 제한적인 결과를 보였다. 이 모델에서의 제한적인 실측환경과 결과를 보다 확장하고, 더욱 일반적인 상황에서 디스크 배열의 특성변수를 결정하는 것이 필요하다[11].

본 논문에서는 MPEG 스트림의 효과적인 실시간 재생을 위해 실제 멀티미디어 서버의 환경 상에서 MPEG-1 화일을 실측 데이터로 사용하여, 저장시스템의 다양한 작업부하에 따른 최적의 저장시스템 환경을 위한 특성변수들을 도출해 낸다.

본 논문의 구성은 다음과 같다. 2장에서 실험에 사용될 RAID레벨 구조와 멀티미디어 서버의 구현에 대해 논의하고, 3장에서는 실측 모델링 기법에 대해 그리고 4장에서는 성능 측정 및 결과를 비교 분석한다. 5장에서는 결론을 도출한다.

2. 멀티미디어 서버의 구현

2.1 디스크 배열

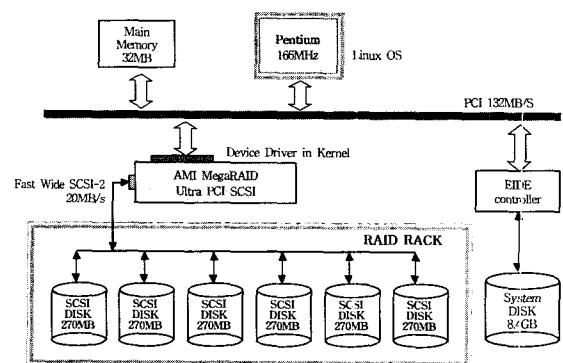
멀티미디어 데이터는 연속성과 많은 저장 공간을 필요로 한다. 이러한 멀티미디어 데이터를 한 개의 디스크에 저장시킬 경우 읽기 시 접근 시간이 느려 몇 개의 클라이언트밖에 사용할 수 없다. 이러한 단점은 여러 개의 디스크를 동시에 사용하여 입출력의 병렬성(parallel)과 병행성(concurrency)

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 요구률과 요구되는 데이터의 크기가 증가할수록 현저히 저하된다[12]. Kim의 연구에 의하면 4KByte이상의 데이터가 요구되어질 경우 이러한 성능저하는 증가함을 보인다. 그러므로 동기화된 디스크배열은 한번의 접근으로 다수의 디스크로부터 많은 데이터를 동시에 입출력하여 클라이언트로 데이터를 전송하므로 짧은 전송시간에 많은 양의 데이터를 전송할 수 있다.

2.2 멀티미디어 서버 시스템

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

이 시스템은 일반적인 Pentium PC에 운영체제로 Linux를 설치하였다. RAID 컨트롤러로는 American Megatrends사의 PCI MegaRAID Ultra SCSI를 설치하였다. 이 컨트롤러는 3개의 채널을 지원하며, 1채널당 15개의 디스크를 장착할 수 있다. 8MB의 메모리를 가지며, 2KB에서 256KB사이의 스트라이핑 단위를 지원한다[13]. 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

보통이 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 저장 시스템의 환경에 대한 실측 모델링을 위한 것이다. 저장 시스템의 환경에 대한 특성변수들은 상대적인 결과로서 충분히 의미가 있으므로, 서버의 성능은 크게 중요하지 않다. 또한 적은 디스크 용량은 매 실험마다 디스크의 초기화를 위한 시간을 단축시키기 위함이다. 본 실험에서는 시스템 디스크를 따로 두어 운영체제와 실험에 사용될 멀티미디어 데이터를 저장하였다. 매 실험마다의 결과는 신뢰성을 위해 3번 수행한 후 평균값을 이용하였다.

<표 1> SCSI 버스의 종류

	SCSI-1	Fast SCSI-2	Fast Wide SCSI-2	Ultra SCSI (Fast-20)
Synchronous Transfer rates	5MB/S	10MB/S	20MB/S	40MB/S
Man 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>는 본 논문에서 사용된 LPS270S SCSI 하드디스크의 물리적 특성을 나타내는 매개변수이다.

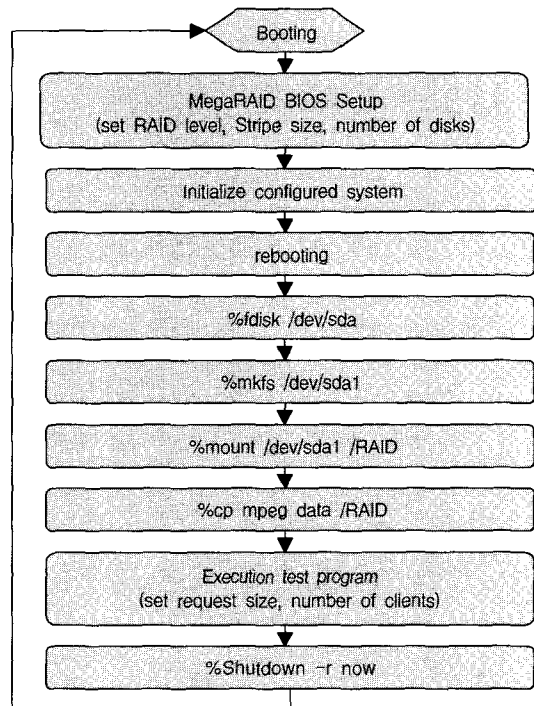
<표 2> 디스크 매개변수

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

3. 실측 모델링 기법

멀티미디어 서버 시스템과 같은 실시간 응용 환경에서 고객들로부터의 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 적용되었던 쓰레드를 이용해 C언어로 작성

하였다. 재생률이 1.5Mbps인 MPEG-1 파일의 실시간 재생을 보장하기 위해서는 초당 185KB의 데이터가 저장시스템으로부터 서비스되어야 한다. 연속 매체 파일 시스템의 데이터 요구 블록 크기로 256KB로 결정하였는데, 265KB블록을 이용하여 매초 185KB의 데이터를 제공하기 위해서는 3초당 2번씩 데이터 요구 블록을 공급하여야 한다. 최대한의 고객들에게 이러한 요구를 만족시키는 최적의 특성변수를 구하기 위하여 (그림 2)와 같은 실측을 위한 멀티미디어 서버 시스템의 처리 절차를 매 실험마다 반복한다.



(그림 2) 실측을 위한 멀티미디어 서버 시스템의 반복절차

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 상황하에서도 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.

<표 3>은 단일 디스크에 1.5Mbps MPEG-1 파일을 저장한 멀티미디어 서버 환경에서 데이터 요구 블록 크기에 따른 연속 매체 파일 시스템의 입출력 성능 측정 결과를 나타낸다. 성능 평가 기준은 연속 매체 파일 시스템이 1, 10, 30, 50, 70개의 활성화된 스트림들에게 MPEG 데이터를 제공하는 상황에서 데이터 요구 블록 크기를 8KB~512KB가

지 변경하면서 평균 데이터 전달률을 측정하였다.

〈표 3〉 데이터 요구 블록 크기에 따른 파일 시스템의 평균 데이터 전달률

데이터 요구 블록 크기	평균 데이터 전달률(MBS)				
	1 Stream	10 Streams	30 Streams	50 Streams	70 Streams
8 KB	0.328	0.332	0.331	0.332	0.330
16 KB	0.556	0.555	0.555	0.560	0.559
32 KB	0.854	0.850	0.848	0.849	0.847
64 KB	1.056	1.067	1.070	1.074	1.074
128 KB	1.276	1.275	1.274	1.274	1.276
256 KB	1.400	1.410	1.411	1.410	1.409
512 KB	1.568	1.548	1.548	1.547	1.547

성능 측정 결과 512KB일때 8KB 보다 5배에 가까운 가장 평균 데이터 전달률을 보였으며, 스트림의 수는 단일 디스크에서의 측정이었기 때문에 큰 영향이 없었다. 그러나, 가장 우수한 성능을 나타내는 512KB 크기의 데이터 요구 블록을 사용할 경우 메모리 관리의 어려움이 따른다. 즉, 인덱스 블록이나 디스크로부터 읽어온 데이터를 보관하기 위해 512KB 크기의 버퍼 공간을 할당해야 하는데, 연속 매체 파일 시스템이 동작하는 동안 지속적으로 512KB 크기의 메모리 영역 할당이 어렵다. 따라서 512KB에서의 평균 데이터 전달률에 가장 근접한 256KB를 연속 매체 파일 시스템의 데이터 요구 블록 크기로 결정하였다.

4. 성능측정 및 분석

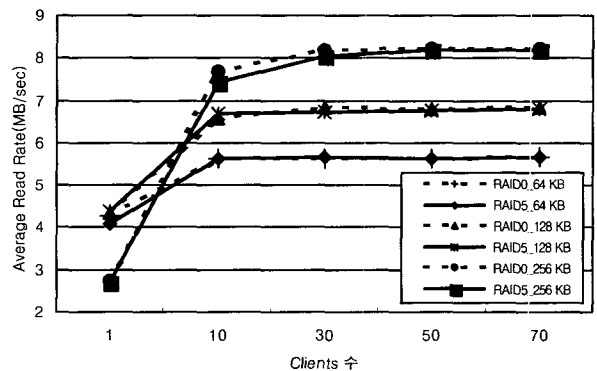
멀티미디어 저장시스템으로서의 디스크 배열은 RAID-0에서 RAID-5까지 여러 가지 레벨중에서 대용량의 멀티미디어 데이터를 저장하여 사용자들의 요구에 실시간으로 공급해줄 수 있는 것이 선택되어야 한다. 그러므로 하나의 요구를 모든 디스크가 동기화되어 처리하는 비트 인터리브 방식의 RAID레벨보다는 여러 사용자의 요구를 각 디스크가 독립적으로 서비스하는 스트라이핑 방식의 RAID레벨이 더욱 효율적이다. 여기서 전자는 병렬성(parallelism)을 이용한 구성이 되고, 후자는 병행성(concurrency)을 이용한 구성이라 할 수 있다. 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 성능뿐 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 지니는 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 [14, 15]. 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 제공할 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 실측 결과를 이용하여 스트라이핑 단위에 대한 데이터 요구의 크기와 작업부하의 병행도에 따른 최대 처리량을 분석한 후 멀티미디어 서버에 가장 적합한 디스크 배열의 여러 가지 특성 변수들을 도출하고자 한다.

4.1 RAID 구조 결정

본 절에서는 병행도에 따른 최적의 RAID 구조를 결정하고자 한다. 병행성을 제공해주는 RAID 레벨 중 상대적으로 가장 우수한 데이터 전송률을 제공할 수 있는 RAID-0와 RAID-5에 대해 디스크 수가 6개인 경우, 고객의 수 변화에 따른 세 가지의 스트라이핑 단위에서의 성능을 비교하였다. (그림 3)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 나으리라 본다.



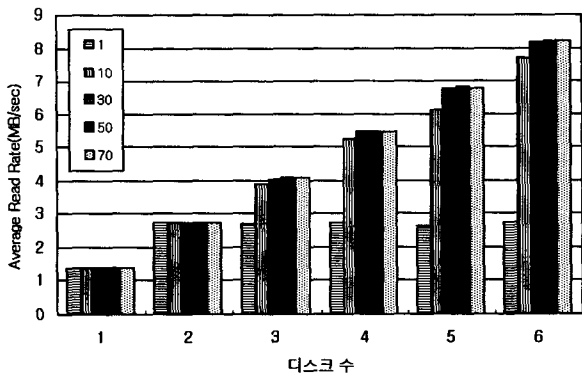
(그림 3) RAID 레벨에 따른 평균 데이터 전달률

4.2 디스크 수에 따른 평균 데이터 전달률

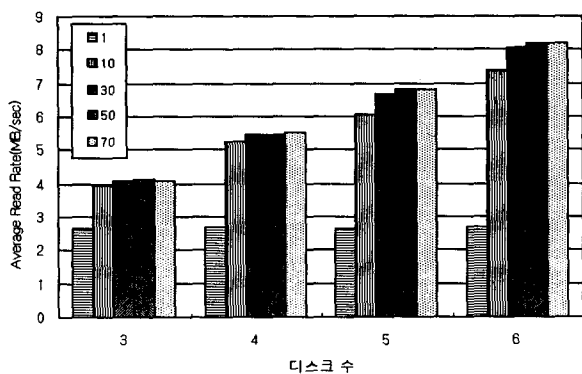
256KB의 데이터 요구 블록 크기에서 스트라이핑 단위의 크기를 256KB로 고정시킨 다음 RAID-0와 RAID-5에 대해 디스크 수에 따른 성능 비교를 하였다. 실측을 위한 멀티미디어 데이터 저장 시스템 내에 하드디스크 수가 6개이므로, 레벨 0인 경우는 디스크 수를 1에서 6까지, 레벨 5인 경우는 3에서 6까지 변경하여 실험해 보았다. 레벨 0에서 1개의 디스크인 경우는 단일 디스크상에서의 성능과 일치한다. 그리고 레벨 5인 경우 최소 3개의 디스크가 있어야지만 패리티 블록을 포함한 하나의 패리티 그룹을 형성할 수 있다. (그림 4)에서 보듯이 RAID-0와 5 모두 디스크 수에 비례하여 평균 데이터 전송률이 증가한다. 하지만, 디스크 수가 계속 증가하더라도 평균 데이터 전송률은 SCSI 버스 대역폭의 한계를 넘을 수는 없다. 결과에서 각 디스크 수마다 고객의 수가 30이 넘을 경우 더 이상의 작업부하에도 전송률은 증가하지 않는다. 즉, 각 디스크마다 전송률의 한계가 있으므로 최적의 디스크 수는 아래와 같은 수식에 의해 결정할 수 있다.

- 최적의 디스크 수 : N_{disk}
- 디스크의 한계 전송률 : LTR_{disk}
- SCSI 버스의 대역폭 : BW_{scsi}

$$N_{disk} = \lceil \frac{BW_{scsi}}{LTR_{disk}} \rceil$$



(a) RAID Level 0



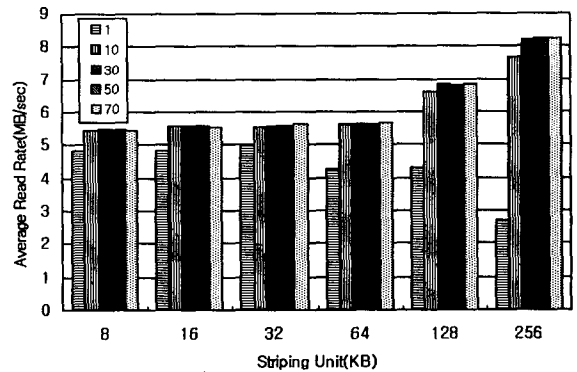
(b) RAID Level 5

(그림 4) 디스크 수에 따른 평균 데이터 전송률

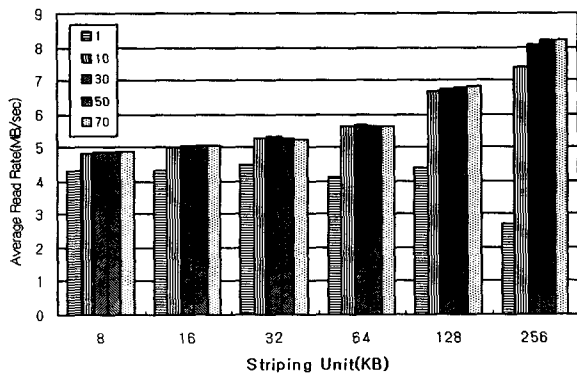
4.3 디스크 배열의 스트라이핑 단위

스트라이핑 단위는 디스크 배열의 성능을 결정하는 중요한 요소이므로 디스크 배열의 성능을 극대화하기 위해서는 최적의 스트라이핑 단위가 선택되어야 한다. 대부분 작업부하가 많이 걸리는 디스크 배열에서의 서비스 시간은 연속 매체 데이터 블록 크기에 의존적이다. 스트라이핑 블록 크기가 작은 경우, 작업부하는 디스크들간에 분산되어 부하의 불균형이 줄어들기 때문에, 저장시스템에 입출력요구가 받아들여지는 동안 접근되어질 블록의 수는 증가한다. 그러나, 이로 인한 탐색시간과 회전지연시간의 증가가 오버헤드이다. 반대로, 스트라이핑 블록 크기가 큰 경우, 접근되어질 블록의 수가 감소하여 탐색시간과 회전지연시간에 대한 오버헤드는 줄지만, 작업부하가 적절히 분산되지 못해 불균형이 생긴다. 이러한 이유 때문에 저장시스템 서버는 반드시 작업부하에 대한 불균형을 줄이고, 배열 내 디스크의 서비스 시간을 최소화할 수 있는 연속 매체 데이터의 스트라이핑 블록 크기를 결정하여야 한다[16].

멀티미디어 서버에서 가장 필요한 것은 쓰기 요구보다 읽기 요구가 주된 서비스이므로 실측에서는 읽기 요구에 중점을 두었다. MPEG-1 데이터인 경우 1초 동안 필요한 요구의 크기는 185KB이다. 구현된 멀티미디어 서버 시스템 환경에서 최적의 스트라이핑 단위를 선정하기 위해 6개로



(a) RAID Level 0



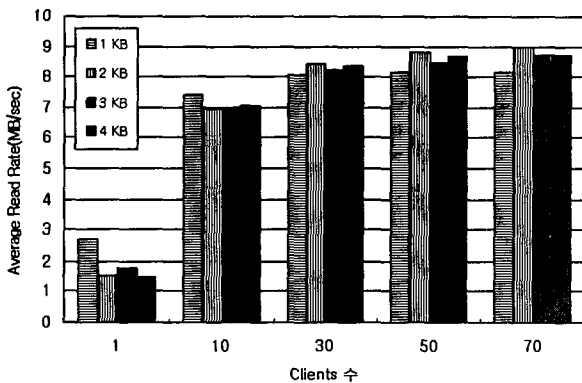
(b) RAID Level 5

(그림 5) 스트라이핑 단위에 따른 평균 데이터 전송률

구성된 RAID0와 RAID 5에서 데이터 요구 블록의 크기가 256KB일 때 스트라이핑 단위의 변화에 따른 디스크 배열의 평균 데이터 전송률을 측정하였다. 각 실험에서 고객의 수가 1에서 70까지 증가할 경우 다양한 결과를 (그림 5)에서 보인다. 가장 좋은 전송률을 보여주는 스트라이핑 단위의 크기는 256KB인 것을 알 수 있다. 단, 고객의 수가 1일 때 전송률이 나쁜 이유는 한 번의 데이터 요구 블록 크기인 256KB와 크기가 일치하기 때문에 하나의 디스크의 접근만으로 충분히 서비스가 행해지므로 디스크간의 병행성이 떨어지기 때문이다. 64KB에서 256KB까지 비례적으로 증가하나, 더 이상의 스트라이핑 단위가 RAID 컨트롤러에서 지원되더라도 더 이상의 전송률 증가를 각 디스크의 병행성 저하로 이루어지지 않을 것이다. 또한 스트라이핑 단위가 상대적으로 너무 커지면 컨트롤러 자체내의 버퍼유지 관리에 문제가 따른다. 또한 스트라이핑 단위가 커질수록 RAID 레벨 0와 5의 전송률은 거의 같아짐을 알 수 있다.

4.4 디스크 블록 크기에 따른 평균 데이터 전달률

본 절에서는 디스크 입출력의 최소단위인 디스크 블록 크기에 따른 평균 데이터 전송률을 측정한다. 앞에서의 실험결과에서 나온 최적의 데이터 요구 블록 크기와 연속 매체 스트라이핑 블록 크기를 각각 256KB로 고정시킨 후, 1KB에서 4KB까지 디스크 블록 크기를 변경해 보았다. 디스크 배열 내 디스크 수는 6개이며, RAID-5 구조에서의 측정이다. (그림 6)에서 보듯이, 작업부하가 적을 경우에는 디스크 블록 크기가 1KB에서 우수한 전송률을 보였으나, 작업부하가 증가할 수록 디스크 블록 크기가 2KB일 때 가장 우수한 전송률을 보임을 알 수 있다.

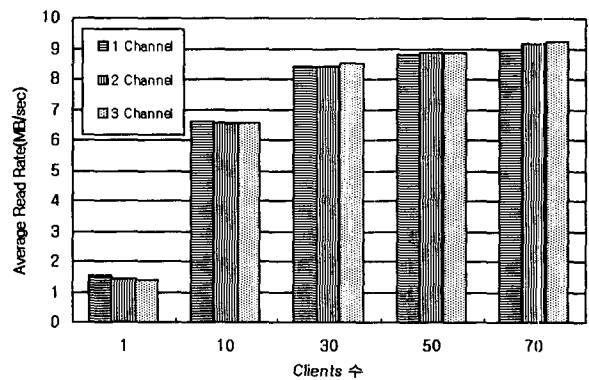


(그림 6) 디스크 블록 크기에 따른 평균 데이터 전송률

4.5 채널 수에 따른 성능 분석

실험을 위한 멀티미디어 서버 시스템에 사용된 RAID 제어기는 3개까지의 채널을 지원한다[13]. 앞에서의 실험결과에서 나온 최적의 데이터 요구 블록 크기와 연속 매체 스트라이핑 블록 크기를 각각 256KB로 고정시킨 후, 1개 채널에서 3개까지 저장시스템내의 디스크 배열 구조를 변경해 보았다.

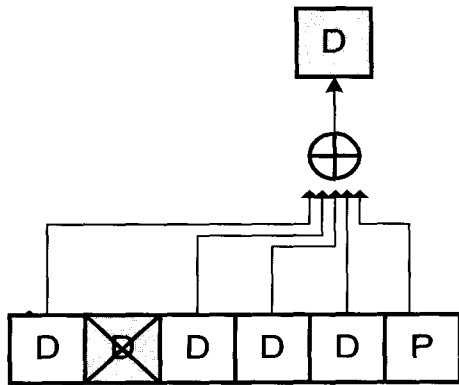
디스크 배열 내 전체 디스크 수는 6개이며, RAID-5 구조에서의 측정이다. 이 경우 1채널인 경우 채널 당 6개, 2채널인 경우 채널 당 3개, 3채널인 경우 채널 당 2개의 디스크가 각각 할당된다. (그림 7)에서 보듯이, 작업부하가 적을 경우에는 채널간의 통신 오버헤드로 인해 단일 채널일 때가 가장 우수한 데이터 전송률을 보이나, 작업부하가 커질수록 채널간의 통신 오버헤드보다는 채널 내에서의 부하가 다수의 채널로 분산됨으로써 복수 채널일 때가 가장 우수한 데이터 전송률을 보인다.



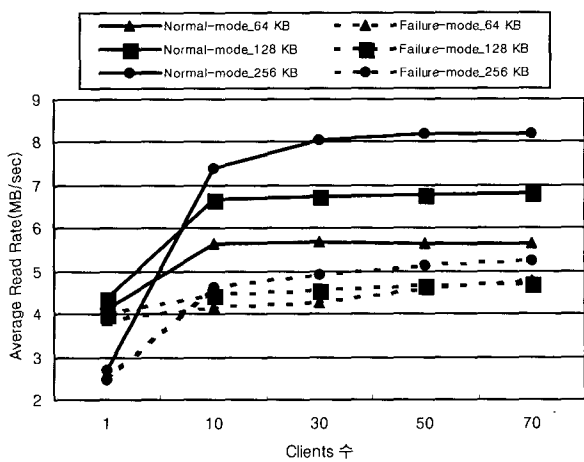
(그림 7) 채널 수에 따른 평균 데이터 전송률

4.6 단일 디스크 고장시 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 8)과 같이 매번 재건(rebuild)시켜 요구에 응답하여야 하므로 그에 따른 부하로 인해 성능은 급격하게 저하된다[17]. 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[18]. 본 절에서는 단일 디스크 고장시 다양한 디스크 배열 특성변수에 따른 성능을 분석해 보았다. (그림 9)에서 보듯이 실험은 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 8) 실패모드에서의 읽기 요구



(그림 9) 단일 디스크 고장시의 성능 분석

5. 결 론

본 논문에서는 실제 멀티미디어 서버 시스템과 같은 환경을 구축하고, 디스크 배열을 저장 시스템으로 하는 서버의 여러 가지 특성변수들을 실측을 통해서 분석하여 보았다. 최근 가장 많이 사용되는 동영상 매체로서 MPEG-1 데이터를 실시간으로 재생하기 위해서는 1초동안 약 185KB의 데이터간 전송되어야 한다. 주어진 여러 가지 시스템 환경에서 최대한의 데이터 전송률을 보장하기 위해서 최적의 특성변수 값들을 실측을 통해서 구해보았다. 먼저, 데이터 요구 블록 크기에 있어 블록 크기에 비례하여 전송률이 증가하였으나, 메모리 관리상의 문제를 고려하여 256KB일 때 최고의 전송률을 보임을 확인하였다. 또한 이 데이터 요구 블록 크기에 있어 스트라이핑 단위가 64KB에서 256KB 까지 증가할 경우 전송을 역시 비례적으로 증가하나, 작업 부하가 커질수록 256KB의 큰 스트라이핑 블록 크기가 가장 우수한 전송률을 보임을 확인하였다. 그 이유는 접근되어질 블록의 수가 감소하여 탐색시간과 회전지연시간에 대한 오버헤드가 줄어들기 때문이다.

그리고, 실시간 응용을 위한 저장시스템으로써 RAID 구

조의 결정에 있어 레벨 0와 레벨 5가 전송률은 거의 비슷함을 보였다. 이러한 이유는 대부분 멀티미디어 서버의 요구가 읽기 요구, 즉 재생서비스 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. 디스크 블록 크기에 따른 평균 데이터 전달률은 디스크 블록 크기가 2KB일 때, 가장 우수한 전송률은 보임을 알 수 있었다.

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. 또한, 단일 디스크 고장시 평균데이터 전송률의 비교에서 고객의 수가 증가할 수록 작업부하가 커짐에 따라 성능의 저하가 두드러짐을 알 수 있었으며, 스트라이핑 단위가 증가할 수록 데이터 전송률의 증가에 따른 상대적인 성능 저하가 발생하였다.

본 연구에서 실제 멀티미디어 서버 시스템상의 실측을 통해 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 블록, 채널 수 등의 결정에 중요한 자료로 폭넓게 활용될 것이다.

참 고 문 헌

- [1] M. H. Vin, P. Goyal, A. Goyal and Anshuman Goyal, "A Statistical Admission Control Algorithm for Multimedia Servers," *Proc. of the ACM Multimedia 94*, pp.33-40, Oct. 1994.
- [2] Jenwei Hsieh, Mengjou 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.
- [3] Ahn L. Chervenak, D. A. Patterson, Randy H. Katz, "Choosing the Best Storage System for Video Service," *Proc. ACM Multimedia '95, San Francisco, CA*, pp.109-119, Nov. 1995.
- [4] Banu Ozden, Rajeev Rastogi, Avi illberschatz, "Demand Paging for Video-On-Demand Servers," *Technical Report AT&T Bell Laboratories*, 1995.
- [5] Gerhard Weikum, Peter Zabback, Peter Scheuermann, "Dynamic File Allocation in Disk Arrays," *Proc. ACM SIGMOD Conference, Denver, May*, 1991.
- [6] Gerhard Weikum, Peter Zabback, "Tuning of Striping Units in Disk Array-Based File Systems," *Proc. 2nd IEEE International Workshop on Research Issues in Data Engi-*

neering, Feb. 1991.

[7] Shenze Chen, Don Towsley, "A Queuing Analysis of RAID Architectures," *Department of Computer and Information Science, University of Massachusetts*, 1992.

[8] 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.

[9] 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.

[10] 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.

[11] 박정연, 안병철, 김정두, "VOD서버를 위한 디스크 배열 파라미터에 관한 연구", *정보처리학회논문지, 제4권 제11호*, pp. 2662-2670, 1997.

[12] Michelle Y. Kim, "Synchronized Disk Interleaving," *IEEE Transaction on Computers, Vol.C-35, No.11*, Nov. 1986.

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

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

[15] 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.

[16] Prashant J. Shenoy, Harrick M. Vin, "Efficient Striping

Techniques for Multimedia File Servers," *Proc. of the 7th IEEE NOSSDAV'97*, pp.25-36, May, 1997.

[17] J. Menon and R. Mattson. Comparison of sparing alternatives for disk arrays. *Proceeding of International Symposium on Computer Architecture*, May, 1992.

[18] R. R. Muntz and J. Lui, "Performance Analysis of Disk Arrays Under Failure," *Proceedings of 16th VLDB Conference*, pp.162-173. 1990.



전 상 훈

e-mail : shjeon@kdtc.ac.kr

1992년 영남대학교 전산공학과 졸업(학사)

1994년 영남대학교 대학원 전산공학과
(공학석사)

2000년 영남대학교 대학원 컴퓨터공학과
(공학박사)

1999년~현재 경동정보대학 인터넷정보개발 전임강사
관심분야 : 멀티미디어시스템, 디지털영상편집, 컴퓨터게임



안 병 철

e-mail : ahn@yu.ac.kr

1976년 영남대학교 전자공학과 졸업(학사)

1986년 오레건 주립대 대학원 전기 및 컴퓨터공학과(공학석사)

1989년 오레건 주립대 대학원 전기 및 컴퓨터공학과(공학박사)

1978년~1984년 국방과학연구소 연구원

1989년~1992년 삼성전자 컴퓨터부문 수석연구원

1992년~현재 영남대학교 공과대학 전자정보공학부 부교수

관심분야 : 컴퓨터 구조, 그래픽스, 멀티미디어 및 실시간 운영체제