

대용량 테이프 저장 장치의 효율적인 사용을 위한 사용량 제한시스템 구현

김성준*, 최윤근*, 장지훈*, 김종권*

*한국과학기술정보연구원

e-mail:sjkim@kisti.re.kr

Implementation of Quota system for efficiency using on HPSS

Sung-Jun Kim*, Youn-Keun, Choi*, Ji-Hoon Jang*, Joong-Kwon Kim*

*KISTI

요 약

HPSS는 수 백만 개의 파일들을 동시에 처리하며 수 petabyte의 용량과 초당 수 gigabyte의 전송률을 보여주는 고성능 소프트웨어이며 고성능 컴퓨팅 환경에서 사용되는 대용량 데이터 스토리지를 위하여 개발되었다. 본 고에서는 HPSS의 대용량 테이프 저장장치 서비스를 시작하면서 사용자별로 용량 제한을 두기 위해서 사용량 제한 시스템을 구현하였다.

1. 서론

현재 한국과학기술정보연구원(KISTI)에서는 기존의 슈퍼컴퓨터들을 이용하여 산출된 데이터를 보다 안전하고 효율적으로 관리하기 위해서 슈퍼컴퓨터 3호기(IBM p690, NEC SX-5)의 도입시에 대용량 스토리지 시스템(High Performance Storage System : HPSS)을 도입하여 설치, 운영 중에 있다.

HPSS에는 전 세계 유수의 슈퍼컴퓨터 센터에서 데이터를 안전하게 관리하기 위하여 널리 쓰이고 있는 시스템이며, 기존의 어떤 데이터 저장 시스템과도 비교할 수 없는 고성능 시스템이고 전 세계적 추세인 그리드 구축 부문에서 데이터 그리드 구축용 시스템으로 그 가능성을 높게 평가받고 있다.

KISTI에 설치된 HPSS 시스템은 최신 4.3 버전이며 1대의 HPSS & SFS 서버와 데이터를 실제적으로 처리하는 Mover 3대로 구성되어 있다. 약 3.6TB의 디스크 캐시와 60TB의 LTO 테이프로 구성되어 있는 HPSS 시스템은 컴퓨팅 작업을 통하여 얻은 어떠한 크기의 결과물도 효율적으로 처리하며 안전하게 보관하는 성능을 보장하며, 사용법은 일반 유닉스의 ftp 명령어와 유사하여 일반 유닉스 사용 경험이 있는 사용자라면 누구나 쉽게 이용할 수 있다.

HPSS시스템은 가상의 파일시스템을 구축하여 사용자들의 데이터를 보관해주도록 설계되어 있다. 따라서, 기존의 유닉스 시스템에서 사용자의 자원 사용을 제한하는 기법들은 적용할 수 없는 상태이다.

본 연구에서는 HPSS를 사용함에 있어서 일부 사용자의 과도한 사용을 막으면서, 많은 사용자들이 미리 정해진 정책에 맞게 사용할 수 있도록 사용량 제한 시스템(Quota System)을 추가 구현하였다.

2. 관련 연구

가. HPSS

HPSS는 highly-scalable, parallel, 고성능 계층적 스토리지 관리 소프트웨어이다. 이것은 IBM과 미국 에너지국 산하 4개의 연구소와 협동으로 개발되었으며 ASCII 프로그램에서 이 개발비용의 상당부분을 지원하였다. HPSS는 고성능 컴퓨팅 환경에서 사용되기 위한 고성능 대용량 스토리지 수요에 의해 개발되었다. HPSS는 병렬화와 네트워크 스토리지 디바이스를 통하여 초당 기가 바이트급(109)의 전송량과 펨타 바이트급(1015)의 데이터를 처리할 수 있도록 디자인되었다[1].

HPSS는 소켓 인터페이스를 가지는 TCP/IP 뿐만 아니라 Fibre Channel Standard(FCS)와 ATM와 같은 모든 네트워크 환경을 지원한다. 사용자는 사용자 인터페이스 유틸리티에 의해 선택된 서버노드로 데이터를 요청하면 HPSS 서버는 직접적으로 데이터를 읽고 쓰기 위한 network-attached 스토리지 디바이스를 지정하며, 이를 통해 high-speed 데이터 전송 네트워크를 통하여 병렬로 모든 작업을 수행하게 된다.

나. 특징

HPSS은 고속 병렬 처리를 위한 기존의 슈퍼컴퓨터 및 클러스터 컴퓨터에게 높은 가용성과 범용성을 제공하는 스토리지 서비스이다. HPSS는 연구소, 정부, 의료기관 등에서 많은 양의 대용량 데이터로 인하여 발생될 수 있는 운영상의 문제를 고속 네트워크와 계층적 스토리지 시스템을 이용하여 해결함으로써 애플리케이션들의 높은 효율성을 제공한다. HPSS의 주된 특징은 <표 1>과 같다[2].

<표 1> HPSS 특징

특징	내용
Scalable 스토리지 시스템	단일 내입 스페이스 내에서 수 GB/sec 이상의 다중 파일 전송 속도
	2 ^m 이상의 파일 size 지원
	수 PB 이상의 단일 내입 스페이스
	다중 sub-system 지원 병렬 I/O를 위한 소프트웨어 striping 지원
분산 스토리지 시스템	단일 내입 스페이스 안에 여러 서버가 존재
	네트워크와 연결된 다른 스토리지 시스템 지원

HPSS의 핵심적인 목표는 오늘날의 어떤 상업적인 저장 시스템들 보다 훨씬 빠르게 병렬 또는 클러스터링 된 시스템과 저장장치 사이에 대용량 파일들을 신속히 이동시키는 것이다. 이를 위해, HPSS는 다음의 개념들에 의해서 디자인되고 구현되었다[3].

· Network-centered 구조

고성능 병렬컴퓨터뿐만 아니라 전통적인 슈퍼컴퓨터나 클러스터, 워크스테이션을 위한 scalable parallel 스토리지 시스템을 제공한다. 즉, 이기종의 시스템들이 단일 이름 공간에 단 하나인 저장 서비스처럼 사용자들에게 보여지게 되며, 이를 통하여 하나의 통합된 스토리지 시스템과 같이 활용할 수 있다.

· Third Party Data Transfer

데이터가 디스크나 테이프 컨트롤러에서 직접 사용자에게 전해지도록 디자인된 구조는 network-

centered 구조와 밀접한 관계가 있다. 일단 전달 세션이 구성되면, 실제 데이터 전송은 사용자와 저장장치 컨트롤러 사이에 직접적으로 이루어진다. 보통 이런 작업은 IPI-3 서드 파티 프로토콜을 이용하여 이루어지지만 TCP/IP도 사용된다. 기존의 시스템과 다른 점은, 전송된 데이터는 메모리 버퍼를 거치지 않고 HPSS에 의해 분산 처리함으로써 단일 서버에 집중되는 병목현상을 방지할 수 있다는 것이다.

· Parallel Operation

병렬파일 처리를 위해 디스크나 테이프를 스트라이핑 그러나 파일의 병렬처리 과정은 사용자들에게는 감추어져 있다. 이런 병렬 접근을 통하여 일반적인 단일 디스크나 테이프 또는 RAID array에서 볼 수 있는 전송 속도 보다 훨씬 빠른 데이터 전송률을 보여 준다. HPSS Application Program Interface(API)는 애플리케이션을 실행하는 사용자들에게 저장장치로 병렬 또는 순차적으로 접근할 수 있도록 지원하며 소프트웨어 스트라이프 기법을 이용함으로써 월등한 데이터 전송률을 나타낸다. 하나의 파일을 여러 대의 드라이브에 동시에 고속으로 전송할 수 있으므로 데이터 처리 속도의 향상을 가져온다.

· Component 기반 설계

커널의 수정 없이 유닉스 기반에서 작동되며 ANSI C 구현되었다. HPSS는 분산 트랜잭션 처리를 위해서 DCE(Distributed Computing Environment)와 Encina를 사용한다.

· 다중 계층 및 Class of Services

파일들은 서비스 클래스 안에 존재하며, 사용자는 파일 크기나 성능들을 고려하여 클래스들을 선택할 수 있다. 서비스 클래스들은 다양한 스토리지 클래스로 구성된 스토리지 계층에 의해 정해진다. 스토리지 클래스는 논리적인 저장 매체의 그룹이다. 한 계층은 하나의 테이프로 구성되거나 또는 2계층 이상의 디스크, 디스크여레이, 테이프로 구성될 수 있다. 데이터가 HPSS로 전송되면, 상위 레벨 계층인 디스크 캐시에 보관되며, 일정 시간이 경과된 후에 다음 단계(보통 테이프)의 계층으로 이동하게 된다. 디스크 있는 파일들은 일정 시간동안은 보관되며, 더 이상 사용되지 않을 경우에는 디스크 캐시에서 삭제되게 된다.

· 리모트 무버(Remote Mover)

데이터 무버는 HPSS 리소스에 접근하기 위하여

주 시스템으로부터 떨어져 있는 리모트 사이트에 설치할 수 있다. 이런 능력은 HPSS 서비스를 리모트 사용자들에게 확대시킬 수 있는 장점을 부여한다.

3. 시스템 설계 및 구현

가. 개요

HPSS 시스템은 사용자가 클라이언트에서 실행하는 명령어인 pftp_client 명령어와 사용자의 접속을 받아들여서 ftp서버의 역할을 하는 서버 부분으로 나누어진다[4].

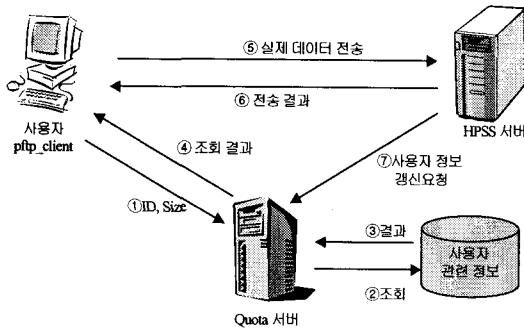
<표 2> 일반 유닉스와 HPSS 파일 전송명령어 비교

	일반 UNIX	HPSS 시스템
접속명령어	ftp	pftp_client
접속 대상	ftp서버	HPSS 서버(core)
file system	일반 UNIX 파일시스템	HPSS 가상 파일시스템
사용명령어	put, get, delete 등	put, get, delete 등

위에 <표 2>에서 보인바와 같이 hpss 시스템은 기존의 ftp와 사용상에 크게 다른 점은 없다. 다만, ftp는 실제 파일시스템을 대상으로 접속하는 반면, hpss는 가상의 파일시스템을 접속한다는 것이 다른 점이다. 이러한 이유로, 일반적인 유닉스에서 사용하는 quota 체계를 이용하여 사용자의 사용량을 제한할 수 없다.

나. 설계

아래 <그림 1>은 사용량 제한 시스템의 개략적인 흐름도이다.



<그림 1> 사용량 제한 시스템 구성도

사용자의 사용량 정보의 변경은 추가와 삭제로 구분된다. 사용량에 변화를 주는 명령어는 다음 표와 같다.

<표 3> 적용 대상 명령어

명령어	내용	명령어	내용
put	파일 업로드	mput	다중 파일 업로드
pput	병렬 파일 업로드	mpput	병렬 다중 파일 업로드
append	파일 덧붙임	delete	파일 삭제

사용자가 pftp_client를 통해서 파일을 전송하기 위해서 "put", "append" 등과 같은 명령어를 입력하면, 일단 사용자의 id와 전송하고자 하는 파일의 크기를 Quota 서버로 전송하여 사용자의 전체 사용량을 계산하여 전송 허용 여부를 결정하게 된다.

만약, 기존 사용량과 전송하고자 하는 데이터의 크기의 합이 제한량을 넘게되면 전송을 실패하게 되며, 가능한 경우에만 HPSS 서버로 데이터를 전송할 수 있다. 또한, 데이터의 전송이 성공적으로 끝난 후에는 사용량 관련 정보에서 사용자의 현재 사용량을 갱신하게 된다. 파일 삭제의 경우에는 삭제하고자 하는 파일의 정보를 HPSS 가상 파일시스템에서 읽어와서 해당 파일의 크기만큼 사용량 정보에서 빼게 된다.

다. 구현

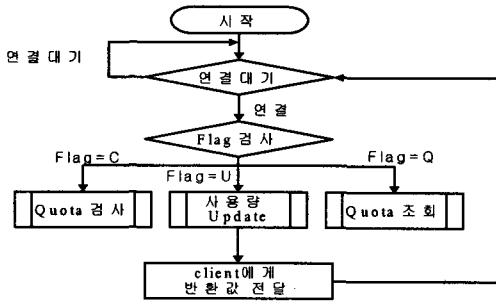
테스트 서버	IBM p690
테스트 클라이언트	HP SMP, IBM p690
프로그래밍 언어	ANSI c
데이터베이스	mysql

본 고에서 수정/추가한 명령어들은 다음과 같다.

수정 : put, pput, mput, mpput, append, delete
 추가 : quota (현재 사용량 조회)

아래 <그림 2>은 사용량 제한 시스템의 서버측 데몬의 간단한 순서도이다. 클라이언트는 데몬으로 크게 3가지 동작만을 수행한다.

- Quota 검사(Check) : 파일의 전송 전에 미리 파일의 크기를 비교하여 전송 후에 최대 사용량을 넘는지 검사한다.
- 사용량 갱신(Update) : 사용자의 파일을 성공적으로 전송한 후에 파일 크기 정보를 데이터 베이스에 갱신하여 저장한다.
- Quota 조회(Quota) : 현재 사용자의 사용량을 조회하여 결과를 반환한다.



<그림 2> Quota 서버 순서도

명령어의 수정/추가는 클라이언트측 프로그램들을 대상으로 하였으며, 사용량 제한 시스템은 데몬형태로 서버측에서 동작하도록 구현하였다. 다음은 구현한 프로그램의 실행 예시를 보여준다.

```

ftp> quota - ①
Disk quotas for user hpss : Quota(0/10000)[KB]

ftp> put 9M.files - ②
200 Command Complete (9000000, 9M.files, 0, 1, 1048576).
200 Command Complete.
150 Transfer starting.
226 Transfer Complete.(moved = 9000000).
[Successful Update] Updated File Size : 9000[KB]
9000000 bytes sent in 0.80 seconds (10.73 Mbytes/s)
200 Command Complete.

ftp> quota - ③
Disk quotas for user hpss : Quota(9000/10000)[KB]

ftp> put 2M.files - ④
state : 0 Max: 10000 Used: 9000 Over:1000 [KB]
[Error] Exceed User Quota. Please delete before transfer!!

ftp> delete 9M.files - ⑤
200 PORT command successful.
150 Opening ASCII mode data connection for directory list.
226 Transfer complete.
64 bytes received in 0.01 seconds (5.40 Kbytes/s)
[Successful Update] Removed File Size : -9000[KB]
250 DELE command successful.

ftp> quota - ⑥
Disk quotas for user hpss : Quota(0/10000)[KB]

ftp> put 2M.files - ⑦
local: 2M.files remote: 2M.files
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for 2M.files.
226 Transfer complete.
2000000 bytes sent in 0.03 seconds (62.71 Mbytes/s)
[Successful Update] Updated File Size : 2000[KB]

ftp> quota - ⑧
Disk quotas for user hpss : Quota(2000/10000)[KB]
  
```

위의 실행 예시는 사용량 제한을 10M를 설정 후에 작업한 내용이다.

- ① quota 명령어를 사용하여 최대 사용량 10M를 확인.
- ② put명령어를 이용하여 파일 용량 9메가인 9M.files라는 파일을 전송함.
- ③ quota 명령어를 사용하여 전송 후 사용량 변화를 표시함.
- ④ put명령어를 이용하여 파일 용량 2메가인 2M.files라는 파일의 전송을 시도하였으나, 기존 사용량(9M) + 전송 파일(2M)의 결과가 최대 사용량은 10M를 초과하므로 전송 실패.
- ⑤ 전송을 위한 가용 공간을 확보하기 위해서, 기존의 9M.files를 삭제함.
- ⑥ quota 명령어를 사용하여 파일 삭제 후 후 사용량 변화를 표시함.
- ⑦ put명령어를 이용하여 파일 용량 2메가인 2M.files라는 파일의 전송 성공
- ⑧ quota 명령어를 사용하여 전송 후 사용량 변화를 표시함.

4. 결론

본 고에서는 한국과학기술정보연구원 슈퍼컴퓨팅센터의 HPSS 시스템에 적용하기 위해서 대용량 저장 장치의 사용량 제한 시스템을 설계·구현하였다. 현재 실제 시스템에 적용하여 사용자의 테이프 사용량 용량을 제한하고 있으며, 추후에는 사용자 데이터 패턴 분석을 위해서 통계 기능을 추가할 예정이다.

참고문헌

- [1] Per Lysne, Gary Lee, Lynn Jones, Mark Roschke, "HPSS at Los Alamis: Experience and Analysis", 6th IEEE Mass Storage Systems Symposium, p150-157, 1999
- [2] IBM, "Basics of the High Performance Storage System"
- [3] IBM, "HPSS System Administrator Guide"
- [4] IBM, "HPSS System User Guide"