

병렬 가상 파일 시스템의 성능 분석을 통한 Cluster 파일 시스템 구축에 대한 연구

박 준 완*, 윤 천 균**

*호남대학교 대학원 인터넷 비즈니스 학과, **호남대학교 정보기술학부

*junid@itc.honam.ac.kr, **chqyoun@honam.ac.kr

Research about Cluster file system construction through performance analysis of Parallel file system

*Junwan Park, **Chunkyun Youn

*Dept. of Internet Business, Honam University

**Devison of Information Technology, Honam University

요 약

최근 리눅스의 보급이 활발하게 이루어지면서 저비용으로 슈퍼컴퓨터의 성능에 버금가는 클러스터 시스템이 활발히 구현되고 있다. 이에 맞추어 클러스터 전체 노드 디스크의 저장 공간을 효율적으로 사용할 수 있는 네트워크 파일 시스템이 필요로 하게 되었다. 클러스터 파일 시스템 구축시 가장 보편적으로 사용되는 파일 시스템인 Network File System(NFS)은 보안의 위협성과 데이터 access시 성능이 떨어지는 단점 때문에 NFS를 대체할 수 있는 파일시스템을 필요로 하게 되었다. 본 논문에서는 리눅스에서 사용하는 Ext3 파일 시스템의 성능을 기준으로 최근 활발한 연구가 진행중인 파일 시스템중 하나인 가상 병렬 파일 시스템과 NFS와의 성능 비교 및 시험을 통하여 최적화된 Cluster 파일 시스템에 대해서 연구하고자 한다.

1. 서론

Cluster란 PC 또는 워크스테이션을 고속 네트워크로 연결하여 고성능 또는 고가용성을 얻을 수 있도록 하는 시스템을 말한다. 클러스터 기술은 상용으로 시중에 판매되는 장비들을 이용하여 구축하기 때문에 적은 비용으로 높은 성능을 낼 수 있어 확장성이나 업그레이드 등의 장점들을 많이 가지고 있다. 또한 클러스터의 운영체제로 많이 사용되는 Linux는 코드의 이식비용을 현격히 감소시키고 공동작업의 가능성을 높인다. 바로 이러한 특성들이 클러스터를 계산 과학 분야에서 새로운 컴퓨팅 동향으로 자리잡아 가고 있는 이유들이다.[1]

많은 클러스터 시스템이 구축되면서 여러개의 CPU나 메모리를 하나로 묶어 사용하는 방법들이

많이 연구되고 있다. 이와 더불어 여러 노드에 나누어져 있는 디스크들을 하나의 디스크처럼 사용할 수 있는 Network를 이용한 파일 시스템이 활발하게 연구되고 있다.

현재 Cluster구현에 가장 많이 사용되는 운영체제인 Linux는 기본 파일시스템으로 Ext3 파일 시스템을 사용하고 있다. Ext3 파일 시스템은 리눅스에서 사용하던 Ext2 파일 시스템에 Journaling 기능을 추가한 파일 시스템으로서 안정적이고 신뢰성을 보장받고 있다[2]. 현재 Cluster구현을 위해서 Network를 이용한 파일 시스템으로 상용의 Intermezzo, Global 파일 시스템을 일부 사용하고 있지만 공개 파일 시스템으로 가장 보편적으로 사용하고 있는 파일 시스템은 썬 마이크로 시스템즈에서 개발한 네트워크상에서

파일시스템을 공유하도록 설계된 파일시스템의 한 종류인 NFS를 사용한다[3].

이외에도 Coda FS, AFS, PVFS등 리눅스 클러스터를 위한 여러 가지 파일 시스템이 연구중이다.

Coda 파일 시스템은 NFS보다 개선된 네트워크 분산파일 시스템으로 비연결 작업 지원, 읽고 쓰기 서버 복제, 인증, 암호화 등의 개선된 기능을 지원한다.

AFS(Andrew File System)는 분산 컴퓨팅 환경의 작업부담을 줄이면서 성능을 증대시키기 위해 로컬캐시를 사용하는, 장소 독립적인 파일시스템이다. 동일한 하나의 파일 공유 공간을 제공하며, 캐시에 의한 파일 사용 성능이 향상되고 안정성, 확장성이 뛰어나다.[4]

PVFS(Parallel Virtual File System)은 다중 디스크들을 통한 데이터의 물리적 분산이 가능한 파일시스템이다. 다중 디스크를 통해 파일 데이터를 stripe 하여 네트워크의 대역폭을 늘이고 병목현상을 최소화 한다.[5]

본 논문에서는 리눅스 파일 시스템인 Ext3 파일 시스템의 성능을 기준으로 공개용 파일 시스템 중 가장 보편화된 NFS와 뛰어난 성능이 알려진 PVFS와의 성능 비교 및 차이를 benchmark하여 Test하고 결과를 비교 분석한다. Ext3 파일 시스템은 Network를 거치지 않은 Local에서의 성능을 test하여 가장 빠른 경우의 속도를 측정 한 것이며, PVFS와 NFS는 두개의 노드를 직접 연결하여 가장 최적의 성능을 내도록 구성했다. Test 방법은 File I/O Test와 File Create Test 2가지를 통하여 파일 시스템의 성능을 test 하였다.

본 논문의 구성은 다음과 같다. 제 2장에서는 실험 환경 및 조건에 대해 기술하고, 제 3장은 Benchmark 결과를 분석하며, 제 4장에서 결론을 맺는다.

2. 실험 환경 및 조건

2.1 실험 환경

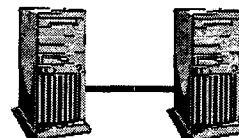
본 논문에서는 3가지 파일 시스템을 가지고 Benchmark를 하였다. 먼저 Ext3 파일 시스템은 그림 2.1과 같이 네트워크를 거치지 않고 local에서 데이터를 Access 하여 파일 시스템의 가장 최적의 성능을 낼 수 있도록 했다. 이 데이터 수치는 다음에 비교할

파일 시스템들의 성능 비교 기준으로 쓰인다. 다음 NFS와 PVFS는 그림 2.2와 같이 최적의 성능을 분석하기 위해 2대의 시스템을 직접 연결하여 오차 범위를 최소화 했다.



Ext3 File System

그림 2-1
Ext3 File System
Benchmark 구성도



NFS & PVFS File System

그림 2-2
NFS & PVFS File System
Benchmark 구성도

2.2 실험 조건

정확한 Benchmark 실험을 위해서는 같은 OS 및 Application을 사용하여야 한다. 본 실험에서는 OS는 Red Hat Linux 8.0을 사용하였고 파일 시스템 Benchmark tool로는 Bonnie++를 사용하였다. Bonnie++ tool은 파일 시스템 benchmark tool로서 크게 File I/O test와 File Create test 2가지를 측정한다. 파일시스템에 대한 CPU점유율까지 측정해 주기 때문에 보편적으로 많이 사용하는 benchmark tool이다.

파일 시스템의 성능을 비교하기 위해서는 물리적 시스템을 동일한 시스템에서 Test 하여야 한다. 본 실험에서는 Pentium3-900M, 256SDR RAM 및 Seagate Barracuda ATA IV 5400rpm 20G HDD를 이용하여 NFS와 PVFS 파일 시스템을 구축하였다. 또 NFS, PVFS 파일 시스템은 Network을 이용한 Test 이기 때문에 외부의 Traffic에 영향을 받지 않아야 한다. 이를 위하여 외부 네트워크가 연결되지 않은 그림 2.2와 같은 독립된 시스템을 구성하였다.

3. Benchmark 결과 분석

3.1 File I/O Test

아래 표는 Bonnie++ Tool의 결과중 File I/O Test로서 File의 크기를 이용하여 Test한 결과이다. 표 3-1은 순차 출력을 character와 block 그리고 rewrite로 나누어 test하였고, 표 3-2는 순차 입력과 무작위 검색성능을 test하여 결과를 나타내었다. 결과는 초당 Kbyte(K/sec)의 Access 속도와 CPU의 점유율로 표시했다. 초당 Kbyte의 Access 속도는 높은 수치가 CPU의 점유율은 낮은 수치가 더 좋은 결과를 나타낸다.[6]

표 3-1, 3-2를 살펴보면 character당 access 속도는 Ext3 > PVFS > NFS 순으로 차이가 나고 약간의 차이를 보이고 있다. 하지만 block당 access 속도를 살펴보면 Ext3 파일 시스템과 PVFS가 약간의 속도 차이를 보이고 있는 반면에 NFS는 상당히 큰 차이를 보이고 있다. CPU점유율도 NFS가 Ext3보다는 낮지만 더 나은 성능을 보이고 있는 PVFS보다 더 높은 것을 알 수 있다. 하지만 무작위 검색 속도는 NFS와 Ext3가 거의 비슷한 속도를 내고 있는 반면에 PVFS는 검색속도에서 많은 차이를 보이고 있다.

	size (MB)	Sequential Output					
		Per Char		Block		Rewrite	
		K/sec	% CPU	K/sec	% CPU	K/sec	% CPU
Ext3	496	12744	98	25804	20	5100	4
NFS	496	7747	72	7745	4	4042	3
PVFS	496	9102	67	23761	9	7093	4

표 3-1 File I/O Test(순차 출력)

	size (MB)	Sequential Input				Random Seeks	
		Per Char		Block			
		K/sec	% CPU	K/sec	% CPU	/sec	% CPU
Ext3	496	9883	73	30239	8	131.2	0
NFS	496	5639	83	14264	5	138.6	1
PVFS	496	9135	69	24344	6	78.6	0

표 3-2 File I/O Test(순차 입력 및 무작위 검색)

3.2 File Create Test

아래 표는 Bonnie++ Tool의 결과중 File Create Test로서 File의 개수를 이용하여 Test한 결과이다. 파일개수의 기준은 50개로 benchmark tool을 사용할 때 지정했다. 표 3-3은 순차 생성을 생성과 삭제로 나누어 test하였고, 표 3-4는 무작위 생성을 생성과 삭제로 나누어 test하여 결과를 나타내었다. 결과는 초당 파일의 개수(/sec)와 CPU의 점유율(%CPU)로 표시했다. File I/O Test와 같이 초당 파일의 개수는 많을수록, CPU의 점유율은 낮을수록 더 좋은 성능을 나타낸다.

표 3-3, 3-4에서는 기준점인 Ext3 파일 시스템이 성능은 가장 높지만 CPU 점유율이 다른 파일 시스템에 비해 너무 높아 기준점의 역할을 할 수가 없는 상태이다. 이에 반해 NFS와 PVFS 두 파일 시스템 모두 Network을 이용한 파일시스템이기 때문에 Client System CPU에는 영향을 거의 미치지 않았다. NFS와 PVFS 두개의 파일 시스템을 비교하면 먼저 파일 생성에서는 NFS가 PVFS보다 약 5.5배가량 높은 성능을 나타내고 있다. 반면에 파일 삭제에서는 PVFS가 NFS보다 더 높은 성능을 나타내고 있다.

	Num Files	Sequential Create			
		Create		Delete	
		/sec	%CPU	/sec	%CPU
Ext3	50	173	99	2884	97
NFS	50	116	1	579	4
PVFS	50	21	0	779	2

표 3-3 File Create Test(순차 생성)

	Num Files	Random Create			
		Create		Delete	
		/sec	%CPU	/sec	%CPU
Ext3	50	176	99	538	74
NFS	50	122	1	72	0
PVFS	50	21	0	185	0

표 3-4 File Create Test(무작위 생성)

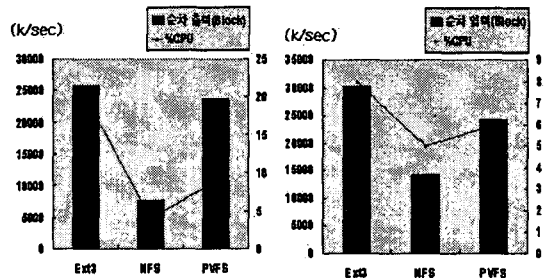


그림 3-1 Sequential I/O 결과값(Block)

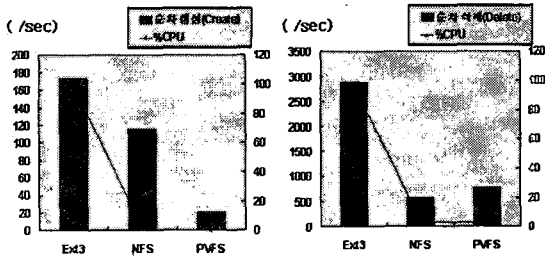


그림 3-2 Sequential Create (Create or Delete)

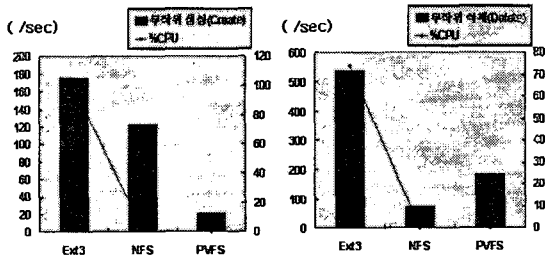


그림 3-3 Random Create (Create or Delete)

위 세개의 그림은 전체 결과 값중 일부를 막대그래프와 꺾은 선으로 표시하였다. 막대그래프는 파일 I/O 또는 파일 Create시 측정값이고 꺾은 선은 각각의 CPU사용량을 나타낸다. 그림 3-1은 순차 입출력 값을 그림 3-2는 순차 생성,삭제시 결과값을 그림 3-3은 무작위 생성, 삭제시 결과값을 막대그래프로 나타내었다. 그림에서 보이는 것처럼 일정 크기의 데이터에 대한 입출력의 성능은 PVFS가 NFS보다 약 2배에서 3배까지 차이가 남을 알 수 있다. 또 순차 또는 무작위로 파일을 삭제하는 경우에도 PVFS가 NFS보다 2.5배 이상 성능이 뛰어난 것을 확인 할 수 있다. 이에 반해서 순차 또는 무작위로 파일을 생성하는 경우는 NFS가 PVFS보다 약 6배가량 빠름을 알 수 있다.

4. 결론

Cluster System에서 파일 시스템이 차지하는 비중은 개별적인 성능 면에서는 CPU나 Memory에 비해서 크지 않다. 하지만 Hardware가 더 이상 고속으로 발전하지 않는 상황에서 Cluster 시스템 최적의 성능을 내기 위해서는 파일 시스템의 Data Access

Time이 상당히 중요한 요소이다. 본 논문에서는 현재 가장 보편적으로 사용되고 있는 NFS와 점차 사용자가 증가하고 있는 PVFS의 성능을 비교 분석하여 Cluster 시스템을 최적의 환경으로 구축 할 수 있는 좀 더 나은 파일 시스템을 찾는 것을 목적으로 하고 있다.

그림에서도 알 수 있듯이 파일을 생성하는 경우를 제외하고는 거의 모든 부분에 있어서 NFS보다 PVFS가 좋은 성능을 나타내는 것을 알 수 있다. NFS가 다른 파일 시스템에 비해 설치, 사용하기가 쉽다는 장점을 가지고 있지만 PVFS도 다른 파일 시스템과는 달리 상당히 간단한 설치구조를 가지고 있다.

PVFS는 파일 시스템의 클러스터링이라고도 불리우고 있을 만큼 여러 노드의 디스크를 병렬로 연결하여 데이터를 공유하고 또 여러 노드의 디스크로 데이터를 분할하여 저장 할 수 있다. 파일을 생성하는 시간이 오래걸린다는 단점이 있지만 효율적인 Cache 관리 방법으로 PVFS를 구현한다면 시간이 많이 단축될 수 있다.[7] Cluster를 구축하는데 있어 보편적으로 사용하는 NFS 보다는 PVFS를 사용하는 것이 시스템 최상의 자원을 활용 할 수 있는 하나의 방법이 될 것이다.

참고문헌

- [1] "클러스터 기술센터" KISTI/슈퍼컴퓨팅센터 슈퍼컴퓨팅연구실
- [2] Theodore Y. Ts'o. "Planned Extensions to the Linux Ext2/Ext3 Filesystem." USenix Annual Technical Conference, Freenix Track, June 2002
- [3] 이현, 김용대, 김인환, "OSI 참조 모델을 바탕으로 한 Linux 시스템의 NFS 성능 측정", 한국정보처리학회 1999
- [4] www.helloec.net/network/afs.htm
- [5] www.parl.clemson.edu/pvfs/introduction.html
- [6] <http://www.coker.com.au/bonnie++/readme.html>
- [7] 김호중, 황인철, "PVFS에서 캐쉬 관리자의 설계 및 구현", 한국과학기술원, 2002.8