

NS2 를 이용한 iSCSI 성능분석

송점기, 최새봄, 임효택
동서대학교 컴퓨터공학과
{ d7003090, d7003154, htlim }@dongseo.ac.kr

iSCSI Performance Analysis Using NS2

Jeom-Ki Song, Sae-Bom Choi, Hyo-Taek Lim
Dept. of Computer Engineering, Dongseo University

요 약

iSCSI(Internet Small Computer Systems Interface)프로토콜은 TCP/IP 네트워크를 통해 블록 중심형 스토리지 데이터에 접근을 가능하게 하는 SCSI 전송프로토콜이다. iSCSI 는 서버에서 스토리지 장치로의 데이터 전송을 위해 표준 이더넷 스위치와 라우터를 사용하며, 또한 기존의 IP 네트워크를 사용하여 SAN 의 확장이 가능하다. 본 논문에서는 이러한 iSCSI 를 사용하여 여러가지 다른 환경에서 iSCSI 프로토콜의 성능을 분석한다. 첫째는 스토리지 장치에 따른 iSCSI 의 성능변화와 둘째는 현재의 네트워크 망에서의 변화이다. 이더넷에서 스토리지 데이터를 액세스하는데 있어서 시급했던 문제중에 하나는 충분한 대역폭과 낮은 지연율을 제공하는 것이었다. 최근의 이더넷 기술의 발전은 IP 네트워크를 통해 스토리지 장치를 액세스하는 것이 가능해졌으며 백본망을 비롯한 기가비트 이더넷이 활성화된 오늘날의 네트워크 망에서 iSCSI 를 이용한 스토리지 서비스가 활성화 될 것이다. 본 논문에서는 iSCSI 의 기반이 되는 IP 네트워크에서의 iSCSI 의 성능을 분석한 내용을 기술한다.

1. 서론

인터넷 SCSI 또는 iSCSI 는 TCP/IP 네트워크상에서 블록 중심의 스토리지 데이터의 전송을 가능하게 하는 SCSI 전송프로토콜로서 스토리지 네트워크 영역에서 빠르게 위치를 자리잡아가고 있다. iSCSI 는 기존 IP 네트워크의 인프라를 활용하여 스토리지 네트워크를 확장하고 있으며 이를 위해 호스트와 스토리지 네트워크 사이에서 TCP/IP 프로토콜을 사용하여 SCSI 명령과 데이터를 전송한다. 일반적으로 SAN 에서는 호스트와 스토리지 시스템을 상호 연결하는 별도의 전용 인프라가 필요했다. 이러한 상호연결을 위한 기본수단은 SCSI 전송을 제공하는 피이버채널 네트워크이기 때문에 결과적으로 IP 네트워크를 기반으로 하는 어플리케이션과 스토리지와의 직접적인 연결은 불가능했다. 하지만 IP 네트워크에서 iSCSI 를 사용하면 피이버채널 네트워크를 교체하지 않아도 피이버채널을 기반으로 하는 SAN 에 도달할 수 있다. 호스트에 호스트 어댑터를 추가할 필요가 없으며 별도의 WAN 인프라 역시 필요하지 않다. 호스트에서 기존의 IP/이더넷 네트워크 연결을 사용하여 스토리지에 직접 액세스함으로써 스토리지를 통합하고 활용도를 높이며 비용을

상당히 절약할 수 있다. 다시 말해 iSCSI 는 기존의 IP 네트워크를 이용함으로써 비용면에서의 효율성뿐만 아니라 보안과 확장성, 호환성, 네트워크 관리 및 스토리지 관리 기능등의 많은 이점을 갖는다. 그러나 iSCSI 를 통해 원격지의 스토리지를 OS 에 의해 마치 로컬의 디스크 장치처럼 사용해서 블록 데이터를 전송한다 하더라도 결국 그 중간에는 IP 네트워크가 존재하게 된다. 그렇기 때문에 스토리지 데이터를 전송하는데 있어서 네트워크 트래픽과 같은 문제가 발생할 수 있다. 사실상 iSCSI 가 IP 네트워크를 이용하는데 있어서 가장 큰 문제 중에 하나는 IP 네트워크를 통해 스토리지 데이터를 전송하는데 있어서 충분한 대역폭과 낮은 지연율을 제공해야 하는 점이다.

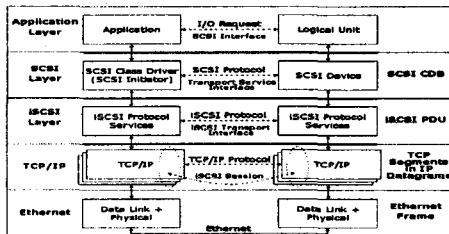
본 논문에서는 스토리지 네트워크 영역에서 빠르게 성장하고 있는 iSCSI 프로토콜과 관련해서 프로토콜에 대한 설명과 시뮬레이션 과정 그리고 그 결과를 분석한다.

2. 관련연구

본 절에서는 원격스토리지 시스템과 관련해서 iSCSI 프로토콜에 간단히 설명한다.

2.1 iSCSI

iSCSI 는 IP 를 통해 스토리지 장치에 액세스를 가능하게 한다. 스토리지의 연결에 있어서 지금까지 사용되어왔던 물리적 SCSI 연결 방식 대신 물리적으로는 이더넷을 사용하며 내부적으로는 SCSI 프로토콜을 사용하려는 기술 표준이다. 이것은 TCP/IP 상에서 캡슐화된 SCSI 정보를 전송하는 것을 기반으로 한다. iSCSI 는 로컬 혹은 원격지에서 서버와 스토리지를 연결하며 스토리지가 여러 지역에 분산돼 있고 IP 네트워크의 관리가 쉬운 고성능 SAN 에서 사용된다. iSCSI 이전에 스토리지 영역을 구축하기 위한 유일한 방법은 비용이 비싼 파이버 채널뿐이었다. 10M, 100Mbps, 이더넷에 비해 빠른 기가비트의 속도를 제공하는 것은 파이버채널뿐이었기 때문이다. 하지만 파이버 채널은 표준이 설정돼 있지 않았고 벤더들 간에 호환성 문제나 높은 비용문제로 인해 스토리지를 구축하는데 많은 어려움이 있었다. iSCSI 에 많은 관심과 연구가 이루어지는 이유는 정보의 디지털화가 각 분야에서 이루어지면서 정보를 저장하는 스토리지 시장의 규모가 지속적인 성장을 거듭하기 때문이다. 즉 스토리지 영역을 구축하는데 있어서 기존 IP 네트워크를 통해 스토리지들의 상호 연결이 가능하게 되면 사용자에게 더 나은 서비스를 제공할 수 있기 때문이다. 사용자가 IP 네트워크를 이용해서 스토리지에 접근, 자원의 이용을 가능하게 할 수 있다. iSCSI 는 IP 스토리지라는 기술로 분류된다. 다시 말해 IP 와 기가비트 이더넷을 이용해 스토리지 네트워크를 구축하는 것을 말한다. 기가비트 이더넷과 IP 기술이 보편화되면서 IP 스토리지는 사용하기 쉬운 네트워크 프로토콜로 자리잡았다.



[그림 1] iSCSI 프로토콜 스택

기본적인 iSCSI 시스템 모델은 위의 그림에서도 보는 것과 같이 Initiator SCSI 장치와 Target SCSI 장치로 이루어진다. Initiator 와 Target 의 로그인 과정이 수행되면 iSCSI Initiator 는 TCP/IP 연결이 성립되며, 가상의 케이블을 통해 SCSI 명령과 데이터 정보를 캡슐화한 다음 Target 으로 보내게 되고 Target 은 받은 SCSI 명령을 수행한다. iSCSI 는 현재 IETF 에 의해 표준이 되었으며 스토리지 분야에서는 혁신적인 기술로 자리잡아 가고 있다.

3. 시뮬레이션 환경

본 논문에서는 iSCSI 프로토콜의 성능을 두 가지

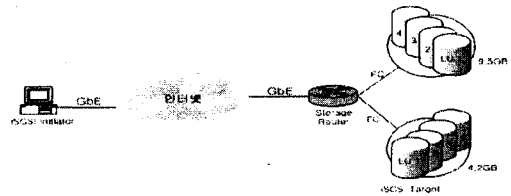
환경에서 분석한다. 첫째는 스토리지 장치에 따른 iSCSI 의 성능 변화와 둘째는 현재의 네트워크 상에서 제공하는 대역폭을 이용해 실제 네트워크상에서의 성능을 분석하는 것이다. 테스트 환경으로는 단일의 iSCSI 클라이언트, 스토리지 라우터, SCSI 하드디스크로 구성하며 OS 로는 Linux kernel 2.4.18 을 사용하고 iSCSI initiator driver 는 iSCSI 클라이언트에 설치된다. iSCSI 클라이언트는 스토리지 라우터와 연결되며 스토리지 라우터는 FC 인터페이스(Disk Controller)를 통해 네 개의 SCSI 하드 디스크와 연결된다. 스토리지 라우터는 블록 수준 데이터 액세스를 위한 전송 수단으로 사용한다. Initiator 는 스토리지 라우터에 접속하여 iSCSI 로그인 과정을 수행하게 되며 로그인 후 Full Feature Phase 단계로 넘어가면서 Target 의 정보를 전송 받게 된다. Full Feature Phase 단계가 끝나면 initiator 는 물리적으로 스토리지를 로컬의 디스크처럼 사용하게 되는 것이다. 이러한 테스트 환경을 NS2 를 이용하여 iSCSI 클라이언트, 스토리지 라우터, SCSI 장치를 구현할 것이다.

4. Disk 측정

이 절에서는 특정 스토리지 장치를 사용해서 iSCSI 프로토콜의 성능을 분석하고 논의할 것이다.

4.1 환경설정

[그림 2]는 디스크의 성능에 따른 iSCSI 의 성능을 분석하기 위한 환경을 그림으로 나타낸 것이다.



[그림 2] Disk 측정 모델

시뮬레이션 구성은 위의 그림에서 보는 것과 같이 두 가지 형태로, 9.5GB, 4.2GB 의 디스크 네 개를 Target 으로 설정한다. 다음은 기가비트 이더넷을 통해 iSCSI initiator 에 의한 SCSI 명령어를 블록 길이에 따른 전송 속도와 디스크의 액세스 타임, 그리고 응답시간의 변화등을 반복적으로 테스트 한다. 특정 제품의 디스크 정보는 Disk Simulator 인 DiskSim 2.0 으로 시뮬레이션을 위해 사용된 제품의 데이터베이스 파일의 데이터를 이용하여 Segate ST39102FC 와 ST 34502LC 의 HDD 를 구현하였다. 다음 표는 시뮬레이션에 사용된 제품에 대한 Parameter 값과 제품 정보를 보여준다.

<표 1> SEGATE_ST39102FC

Disk brand name	SEGATE_ST39102FC
Drive capacity	9.090bytes(formatted)

Disc rotational latency	10.025 rpm
Avg rotational latency	2.99 msec
Average seek(read/write)	5.4 msec / 6.2 msec

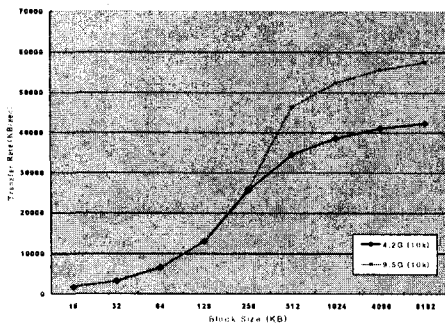
<표 2> SEEGATE_ST34502LC

Disk brand name	SEEGATE_ST34502LC
Drive capacity	4.293bytes(formatted)
Disc rotational latency	10.025 rpm
Avg rotational latency	2.99 msec
Average seek(read/write)	7.7 msec / 8.7 msec

이번 실험은 디스크 성능에 의한 iSCSI 의 성능측정으로 다음과 같이 가정한다. 첫째, 케이블은 매체 자체의 특성으로 인한 지연율이 발생한다. 이러한 지연율은 매체 자체의 특성으로 간주하고 그 값은 배제한다. 둘째, initiator 와 target 의 연결은 단일 Session 상에서 하나의 연결만으로 설정한다. iSCSI 의 로그인 프로세스를 통해 initiator 와 target 간의 인증 및 프로토콜의 수행을 위한 협상시 한 세션 내에 한 개 이상의 TCP 연결이 가능하다. 셋째, 성능 측정 시 우리는 가장 많이 사용하는 두 개의 SCSI 명령어인 Read 와 Write 명령어를 사용하며 명령어 비율은 50:50 로 한다. 같은 명령 내에서 데이터를 순서대로 전달하기 위해 데이터 시퀀싱을 사용하게 되는데 읽기 작업의 경우 DataSN 은 0 부터 시작하고 이후의 각 데이터는 PDU 에 대해 1 씩 증가시킨다. 쓰기 작업의 경우에도 R2T 에 대한 응답인 첫 번째 데이터 PDU 는 DataSN 을 0 부터 시작하고 이후의 PDU 에 대해서는 1 씩 증가시킨다. 넷째, iSCSI PDU 전송간에 오류는 발생하지 않는 걸로 한다. iSCSI 에서는 오류에 대한 처리를 재시도 와 재할당 두개의 메커니즘을 제공하는데 전송 도중 누락된 CmdSN 을 다시 보내는 경우나 TCP 연결이 끊어진 경우는 발생하지 않는 것으로 한다.

4.2 결과분석

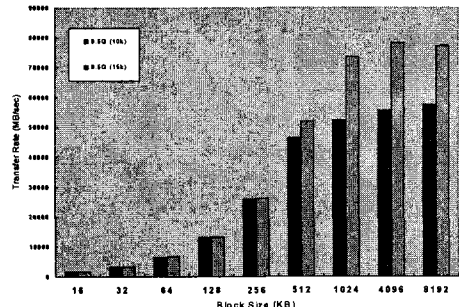
이전장에서의 시뮬레이션 구성을 통해 데이터의 전송길이에 따른 전송속도 1GB 의 대역폭에서 반복적으로 측정했다. 다음 그림은 전송길이에 따른 처리량을 나타낸다



[그림 3] 전송길이에 따른 처리량

위 그림에서 블록 길이가 적을 경우에는 성능에 있어 큰 차이가 보이지 않지만 512KB 이상이 되면 Disk 의 성능에 따른 처리량의 폭이 커짐을 알 수 있다. 4MB 이상의 데이터 전송부터는 처리량에 있어 디스크 성능에 한계를 보이기 때문에 완만한 곡선을 보이게 된다.

RPM 은 분당회전속도로 HDD 의 속도를 평가하는 지표로 HDD 의 속도를 결정짓는 요소 중 하나이다. 다음은 디스크의 특성 중 RPM 에 따른 iSCSI 프로토콜을 사용했을 때 성능이 어떻게 변화되는지를 나타낸다.



[그림 4] RPM 의 변화에 따른 처리량

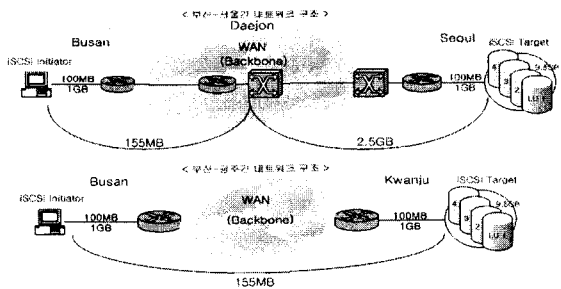
HDD 의 액세스 타임, 전송률 등의 모든 수치는 바로 RPM 에 의해서 많은 영향을 받게 되는데 시뮬레이션을 통해 동일한 용량의 HDD 에서 RPM 을 다르게 설정했을 때 처리량이 어떻게 변화되는지 알 수 있다.

5. RealNetwork 측정

다음은 기존의 IP 네트워크 상에서 iSCSI 프로토콜의 성능을 분석을 통해 네트워크에서의 전송속도 및 응답시간을 알아본다.

5.1 환경설정

[그림 5]는 iSCSI 의 성능 분석을 위해 사용된 현재의 IP 네트워크를 그림으로 보여주는 것이다.



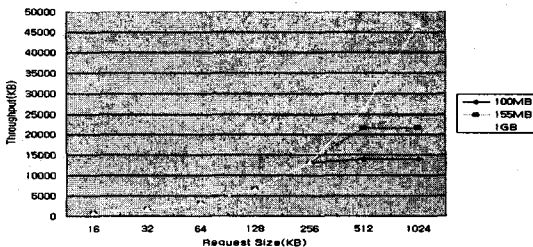
[그림 5]네트워크 구조

위 [그림 5]은 부산지역을 Initiator 로 Target 을 서울지역으로 해서 현재 제공되는 네트워크의 대역폭을 활용, 그 성능을 측정하기 위해 3 가지 형태로 시뮬레이션을 한다. 먼저 첫째는 기존의 이더넷의 고속 버전인 고속 이더넷 상에서 성능을 분석하게 되는데 고속 이더넷은 100Mbps 의 전송속도를 지원하고 있다. 이더넷과 관련해서 한국통신 인터넷은(KORENET)은 주요대도시간은 10Gbps 를 중,소도시간 622Mbps-2.5Gbps 의 대용량 전송로를 제공하고 있다. [그림 5]에서 보는 것과 같이 부산에서 KORENET 백본망까지는 100Mb 의 전송율을 가지며 부산의 노드로부터 서울노드까지의 백본망은 10Gb 의 대역폭을 제공하고 있다.

다음은 기가 비트 이더넷으로 1000Mbps 의 대역폭을 제공하며 기가 비트 이더넷이 보편화되면서 기가비트 이더넷의 활용이 점차 증가될 것이다. 마지막으로 초고속 선도망(KOREN)에서 테스트를 한다. 초고속 선도망은 현재 선도망을 이용하는 많은 기관에게 155Mb ~ 1Gb 의 네트워크 서비스를 제공하고 있으며 국내 차세대 인터넷의 백본망으로의 역할이 강화되고 있는 망이다. [그림 5]에서 보는 것과 같이 부산-대전간은 155Mb 를 제공하며 대전-서울간은 2.5Gb 의 전송로를 제공하고 있다. [그림 5] 역시 부산-광주를 Initiator 와 Target 으로 설정하여 iSCSI 의 성능을 그림에서 보는 것과 같은 대역폭으로 설정한다. 다음은 iSCSI 의 성능 분석을 위해 세운 가정으로 첫째는 이더넷의 전파지연시간이나 네트워크 상황에 따라 발생할 수 있는 에러는 없는 걸로 간주하며 송신 측에서 보낸 데이터는 반드시 수신 측에 전달되는 것으로 한다. 둘째는 디스크 측정에서 세운 가정을 그대로 사용한다.

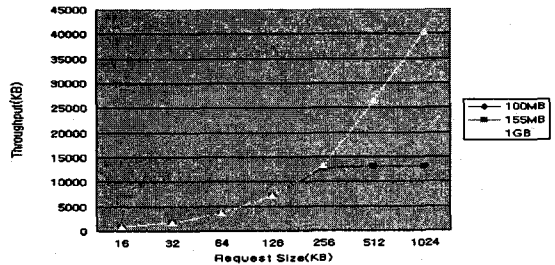
5.2 결과 및 분석

이전장에서의 네트워크 구성을 통해 데이터의 전송 길이에 따른 처리량을 측정한다. 디스크는 Segate ST39102FC HDD(9.5Gb)를 사용한다. 다음 그림은 전송 길이에 따른 처리량을 나타낸다.



[그림 6] 부산-서울간 iSCSI 처리량

[그림 6]에서 보는 것과 볼 수 있듯이 256KB 이상의 데이터를 전송했을 때 100MB 와 155MB 는 대역폭의 한계로 인해 더 이상의 처리 속도를 갖지 못하며 1GB 의 대역폭에서는 그 이상의 속도를 낸다. 1GB 상에서의 대역폭에 의해 100MB 와 155MB 에 비해 평균 전송률은 44.5%의 향상을 보이고 있다.



[그림 7] 부산-광주간 iSCSI 처리량

[그림 7]에서 보는 것과 같이 256KB 이상의 데이터를 전송했을 때 제공되는 대역폭에 따라 처리속도가 달라짐을 알 수 있다. 155MB 상에서의 대역폭에 의해 100MB 비해 전송률은 15.7%의 향상을 보이며 1GB 는 42.8%의 향상을 보인다.

6. 결론 및 연구방향

본 연구에서는 네트워크 시뮬레이터로 널리 사용되고 있는 NS2 를 사용하여 IP 를 통해 스토리지 장치에 액세스가 가능한 iSCSI 프로토콜을 구현하여 스토리지 장치에 따른 그리고 현재 네트워크상에서 데이터의 전송길이에 따른 성능을 측정, 비교했다. 디스크 장치 및 현재의 네트워크에서, 유선환경으로 iSCSI 프로토콜의 성능을 분석함으로써 비용 및 성능 면에서 효율적인 스토리지 설계를 위한 사항들을 제안했다는데 의미가 있다고 할 수 있다. 향후에는 기존의 이더넷상에서 iSCSI 를 통해 높은 대역폭을 요구하는 비디오 및 멀티미디어 정보를 지속적으로 전송하기 위해 스토리지 서비스의 품질을 보장할 수 있는 방안이나 iSCSI 가 현재 개방된 IP 네트워크를 통해 전송되기 때문에 IP 네트워크에서의 보안이나 인증과 관련된 여러 가지 기술적 사항에 대한 연구가 가능할 것이다.

참고문헌

- [1] IETF, "http://www.ietf.org"
- [2] Julian Satran, iSCSI Draft 20, "http://www.ietf.org/internet-drafts/draft-ietf-ips-iscsi-20.txt"
- [3] NS-Manual "http://www.isi.edu/nsnam/ns/ns-documentaion.html"
- [4] NS-Tutorial "http://www.isi.edu/nanam/ns/tutorial/"
- [5] KOREN "http://www.korean21.net"
- [6] KRNIC "http://isis.nic.or.kr"
- [7] DiskSim2.0 "http://www.seagate.com"
- [8] Segate "http://www.seagate.com"
- [9] KORNET "http://www.kornet.net"
- [10] Cisco System, Inc. White Paper, "Introduction to iSCSI"
- [11] Stephen Aiken, Dirk Grunwald, Andrew R.Pleszkun "A Performance Analysis of the iSCSI Protocol"
- [12] Kalman Z.Meth, Julian Satran "Design of the iSCSIProtocol"