

# IP 스토리지 환경에서 네트워크 트래픽 패턴분석\*

류준길, 남영진, 박찬익, 안종석†  
포항공과대학교 컴퓨터공학과  
† 동국대학교 컴퓨터공학과

{lancer, yjnam, cipark}@postech.ac.kr, jahn@dgu.ac.kr

## Analysis of Network Traffic Patterns in IP Storage Environment

Junkil Ryu, Young Jin Nam, Chanik Park, Jong Suk Ahn†  
Dept. of Computer Science and Engineering, POSTECH  
† Dept. of Computer Engineering, Dongguk University

### 요 약

본 논문에서는 스토리지만을 위한 전용 네트워크인 파이버 채널대신, 최근 대중화되고 있는 IP 네트워크를 사용하는 IP 네트워크 스토리지에 QoS를 제공하기위한 연구의 일환으로 스토리지 입출력 트래픽에 의해서 발생하는 네트워크 트래픽의 변화를 분석하였다. iSCSI 프로토콜을 사용하여 구성된 IP 네트워크 스토리지 환경하에서, 스토리지 트래픽에 따른 네트워크 트래픽은 패턴에 있어 거의 동일한 양상을 보이지만 iSCSI 프로토콜의 영향으로 호스트 스토리지의 트래픽보다 네트워크 트래픽에 burstiness가 더해진다는 것을 알 수 있다.

### 1. 서론

기존에 스토리지는 버스를 이용하여 호스트와 스토리지를 직접 연결하던 방식과 스토리지를 파이버 채널을 사용하여 SAN과 같이 네트워크로 연결하는 방식이 존재하였다. 그러나 파이버 채널의 고비용과 기존의 갖추어진 인적, 물적 인프라를 활용하기 위한 방법으로 IP 네트워크를 활용하여 SAN을 구성하려는 노력이 진행되어왔고, 그 결과의 하나로 iSCSI 프로토콜(이하 iSCSI)이 만들어 졌다[1]. 한편, 스토리지 용량과 성능이 증가하면서 다수의 사용자들에 의해서 스토리지가 공유되는 형태가 일반화되었다. 더불어, 각 사용자들이 원하는 스토리지 서비스를 보장해주기위해서 스토리지에 QoS 기능추가에 대한 필요성이 대두되었고 이와 관련된 연구들이 진행되어 왔다[4,5].

본 논문에서는 iSCSI를 이용한 IP 네트워크 스토리지에서 QoS를 제공하는 연구의 기초로서 호스트 스토리지 트래픽에 따른 네트워크 트래픽의 변화를 분석하여 네트워

크 스토리지가 받는 트래픽이 호스트 스토리지 트래픽과 차이가 나는지 알아본다. iSCSI를 이용한 네트워크 스토리지와 관련된 기존 논문들은 블록 수준의 네트워크 스토리지(iSCSI)와 파일 수준의 네트워크 스토리지(NFS, CIFS)의 성능과 관련된 분석을 하는 것이 주였다[2,3]. 즉, 블록수준의 IP 네트워크 스토리지에서, 개입되는 IP 네트워크에 따른 트래픽의 변화를 분석한 연구는 아직 알려진 바 없다.

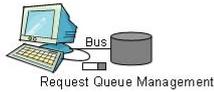
네트워크 스토리지 QoS와 관련하여 스토리지/네트워크 트래픽을 분석하기위해서 본 논문은 “2. 배경지식 및 연구 목표”에서 기존 연구들이 언급한 스토리지 QoS 요구사항과 네트워크 스토리지 관련 연구들을 살펴보고 “실험”을 통해서 호스트의 스토리지 트래픽과 네트워크 트래픽의 차이를 분석한다.

### 2. 배경지식 및 연구목표

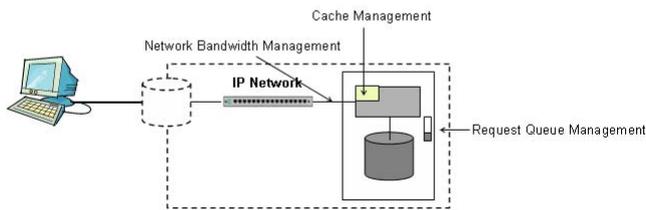
스토리지의 QoS 요구사항은 다음과 같이 기호로 정의할 수 있다[4,5]. 개별 사용자  $i$ 에 대해 QoS 요구사항  $Q_i = \{f_i, iops_i, sz_i, s_i, rt_i\}$ 로 표시되며, 여기서  $f_i$ 는 입출력 요구 중에서 읽기 요구 비율을 나타내며, 0과 1사이의

\* 본 연구는 한국과학재단 특정기초 연구과제 R01-2003-000-10739-0의 지원을 받아 수행되었습니다.

실수 값을 갖는다.  $iops_i, sz_i, rt_i$ 는 초당 도착하는 입출력 요구의 수, 입출력 요구의 평균 크기, 그리고 보장하고자 하는 입출력 응답시간을 나타낸다.  $s_i$ 는 순차적 접근정도를 나타내며 0과 1사이의 실수 값을 갖는다. 1인 경우는 순차적인 접근, 0인 경우는 임의 접근을 나타낸다. 이러한 QoS 요구사항을, 스토리지가 호스트에 버스를 통해 직접 연결되어 있는 경우에는 각 스토리지 사용자에게 대해 스토리지로의 입출력 요구 큐(Queue) 관리를 통해서 만족시켜 줄 수 있지만[7], 호스트와 스토리지 사이에 IP 네트워크가 개입되는 상황에서는 스토리지로의 입출력 요구 큐 관리



가. 호스트와 버스를 통해서 직접 연결된 스토리지



나. 호스트와 IP 네트워크를 통해서 연결된 스토리지

그림 1 호스트와 스토리지간 연결

뿐만 아니라, 네트워크 대역폭에 대한 관리, 캐시에 대한 관리도 필요하다(그림1 참고). IP 네트워크 스토리지(본 논문에서는 iSCSI를 이용한 네트워크 스토리지를 말함)의 경우, IP 네트워크의 특성상 대역폭을 관리하지 않은 면 각 호스트들은 더 많은 대역폭을 가지기위해서 경쟁을 하게 되고, QoS 요구사항에 적합하지 않은 결과를 초래하게 된다(네트워크에서 병목현상을 보일 때, RTT가 작은 호스트가 QoS 요구사항에는 관계없이 더 많은 대역폭을 할당받음[8]). 따라서 QoS 요구사항을 충족시키기 위해서 일정 네트워크 대역폭이상을 유지해야하는 호스트들은 대역폭 유지에 실패하게 되고, 결국에는 QoS 요구사항을 만족시키지 못하게 된다. 또한 IP 네트워크 스토리지 경우, 개입되는 IP 네트워크로 인해서 delay가 발생하기 때문에 이러한 delay를 네트워크 스토리지에서 보완할 수 있는 방법이 제시되어야 한다. delay를 보완하는 방법들 중 하나로 다음과 같이 캐시를 사용하는 방법을 생각해 볼 수 있다. 적용되는 캐시 placement/replacement 알고리즘에 있어서는 호스트와 네트워크 스토리지에 있는 캐시에 동일한 정책을 취하는 것보다는 서로 차이를 두는 것이 스토리지에서 캐시 hit율을 높이면서 발생하는 delay를 최소화하는 방법이다. 일례로, 호스트의 캐시에서 eviction되는 것을 네트워크 스토리지에서 미리 캐시에 가져다 놓는 (placement) 방법[9]과 호스트에서 LRU replacement 알고리즘을 취하면 네트워크 스토리지에서는 LRU방식이 아닌 Multi Queue replacement 알고리즘을 취하는 방법[10]이

스토리지에서 캐시 hit율을 높인 것을 보면 알 수 있다. 또한 캐시를 각 호스트별로 개별화하고, 각 호스트에 대해서 할당되는 캐시사이즈를 차별화함으로써 해서 각 호스트들의 스토리지 QoS 요구사항을 충족시키는데 도움을 준다[11].

스토리지 트래픽에 따른 네트워크 트래픽의 변화 유무와 정도는 위에 제시한 방법들에 영향을 미치게 되는데, 우선 네트워크 대역폭관리에 있어서 할당되는 대역폭 계산과 할당 방법을 변경해야한다. 즉, 스토리지 트래픽에 따른 네트워크 트래픽이 일정하다면 스토리지 트래픽만을 기준으로 대역폭을 할당할 수 있겠지만, 그렇지 않다면 네트워크 트래픽의 변화를 고려해야한다. 또한 QoS 요구사항을 만족시키기 위해서 각 호스트에게 개별 버퍼/캐시를 할당하는 경우, 버퍼/캐시 크기를 결정할 때 중요한 요인으로 작용할 수 있다. 그러나 스토리지 트래픽은 네트워크 버퍼와 TCP/IP 스택을 거치기 때문에 스토리지 트래픽에 따른 네트워크 트래픽의 변화 유무와 정도는 쉽게 추측할 수 없다. 그래서 본 논문은 IP 네트워크 스토리지에 QoS를 보장해주기 위한 기초연구로서 스토리지 트래픽의 변화에 따른 네트워크 트래픽의 변화 정도를 분석하고자 한다.

## 2. 실험

IP 네트워크 스토리지 실험환경은 다음과 같다. 사용자 호스트 시스템과 스토리지의 하드웨어 사양은 CPU Pentium III dual, 메모리 512MB, Intel Gigabit NIC, 그리고 100Mbps이고, 각각 리눅스(커널버전 2.4.18) 운영체제를 탑재하였다. 사용자 호스트와 스토리지간에는 1Gbps 이더넷을 이용하여 연결하였고, 호스트와 스토리지를 모니터링하기 위해서 100Mbps 이더넷을 이용하여 외부 서버와 연결하였다. 사용자 호스트 시스템과 스토리지내에는 미국 뉴햄프셔 대학에서 개발한 iSCSI 장치 구동기와 iSCSI 타겟 장치 구동기[10]를 각각 탑재하였다. 스토리지에서는 iSCSI 타겟 장치 구동기를 통해 도착한 입출력 요구가 상위 SCSI 서브시스템과 SCSI 디스크를 통하여 실제적인 입출력이 발생한다. 사용자와 스토리지간의 네트워크 트래픽을 보기위해서 IP 네트워크 모니터링 소프트웨어인 IPTraf를 사용하였다.

실험은 작은 시간구간에서 Poisson distribution으로 스토리지 트래픽을 발생시켰을 때, 네트워크 트래픽과 스토리지 트래픽이 관계(패턴 변화)를 알아보고, 실제 시스템 환경에서 얻은 스토리지 입출력 요구 데이터를 이용하여 실험을 하였다. 이 실험에서는 패턴을 보는 것이 아니라 스토리지 트래픽의 burstiness 정도와 그에 따른 네트워크 트래픽의 burstiness 정도를 측정하여 변화 정도를 알아보

고, 그 원인을 분석해본다. Burstiness 정도는 Hurst parameter를 사용하여 측정하였다.

실험 1은 1분 동안 일정한 스토리지 트래픽을 발생시켜서 그에 따른 네트워크 트래픽 변화 유무를 자세히 확인하기 위한 것이다(스토리지 입출력 요구는 Poisson Distribution을 따름). 그림 2는 입출력 요구 블록의 크기가 1KB인 random write를 나타낸 것이다. 그림 2에서 Network Traffic(H->S)는 iSCSI에서 명령과 데이터에 의해서 측정된 값이고, Network Traffic(S->H)는 iSCSI 명령을 실행한 후에 발생하는 Response이다. Total은 이 두 값을 합친 것이다. 스토리지 트래픽보다 네트워크 트래픽(Total)이 위에 위치하는 것은 iSCSI와 TCP/IP에 의한 프로토콜 오버헤드 때문이다. 그림 2에서 보듯 네트워크와 스토리지 트래픽 패턴은 일치함을 알 수 있다(random/sequential, r/w에 관계없이). IP 네트워크로 인해서 약간의 차이가 발생할 가능성이 있다고 생각했지만 실험에서는 동일한 패턴을 보이는 것을 확인할 수가 있었다.

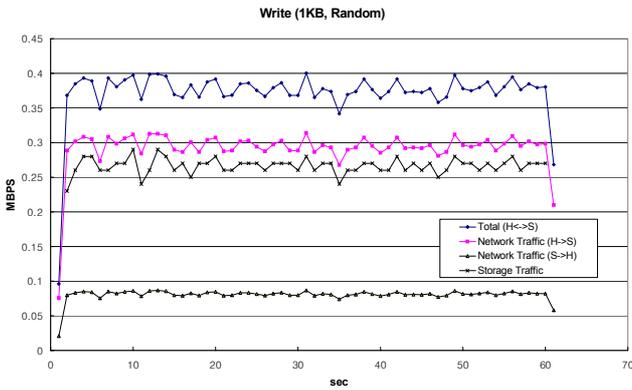


그림 2 일정한 스토리지 트래픽에 따른 네트워크 트래픽

실험 2는 실제 시스템 환경에서 얻어진 Trace를 사용하여(burstiness 성질을 지님, 네트워크 트래픽 변화 유무와 정도를 확인하는 것이다. 이 실험에서 사용한 스토리지 트래픽 트레이스는 cello trace[12]이다. 이러한 트래픽 트레이스를 설치된 IP 네트워크 스토리지에 재생을 하였다.

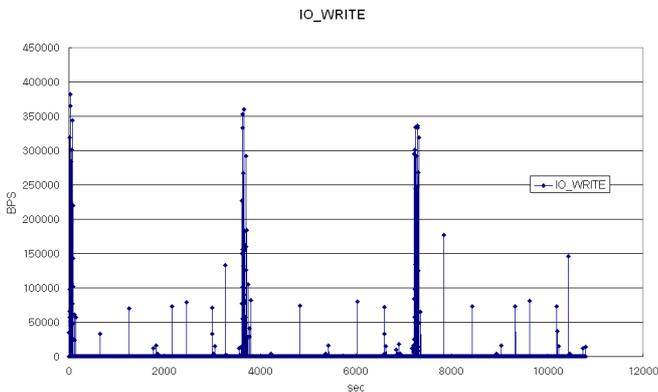


그림 3 스토리지 쓰기 트래픽

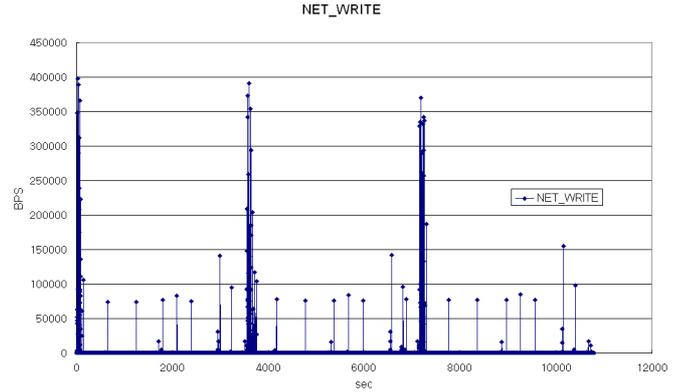


그림 4 네트워크 트래픽 (호스트->스토리지)

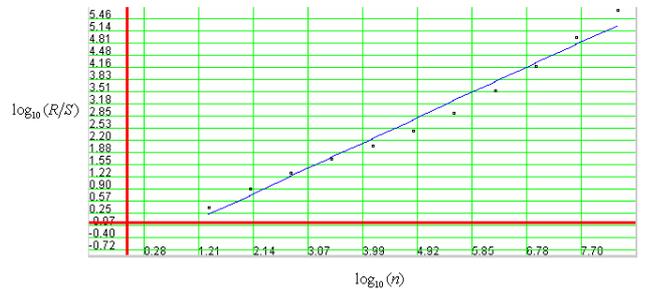


그림 5 스토리지 쓰기 트래픽의 R/S 그래프

그림 3과 그림 4는 스토리지 쓰기 트래픽과 그에 따른 네트워크 트래픽을 나타낸 것이다. 각 트래픽들에 대해서 위해서 그림 5와 같이 R/S 그래프를 사용하여 Hurst parameter를 구하였다[13]. 실험 전 예상으로는 스토리지 트래픽이 네트워크로 가기 전에 버퍼를 거치기 때문에 burstiness의 정도가 더 작아질 것이라고 예상을 하였지만, 실험에서 구해진 데이터는 예상을 뒤엎는 결과를 내놓았다. 표 1에서 보듯이 스토리지 트래픽의 Hurst parameter가 네트워크 트래픽의 Hurst parameter보다 작음을 알 수 있다. 이것은 네트워크 트래픽이 스토리지 트래픽보다 burstiness 정도가 크다는 것을 보여주는 것이라고 할 수 있다.

	스토리지 트래픽		네트워크 트래픽	
	스토리지->호스트	호스트->스토리지	스토리지->호스트	호스트->스토리지
Hurst Parameter	0.801	0.732	0.83	0.754

표 1 스토리지/네트워크 트래픽의 Hurst Parameter

이것의 원인으로 생각해 볼 수 있는 것은 우선 정반대의 결과를 예상했을 때와 마찬가지로 버퍼이다. 즉 버퍼가 트래픽의 burstiness 정도를 작게하는 것이 아니라 크게 할 가능성이 있다는 것이다. 이것 외에 생각해볼 수 있는 것은 iSCSI 프로토콜이다. iSCSI 쓰기/읽기 명령의 경우, 호스트에서 스토리지, 스토리지에서 호스트 방향으로만 데이터 흐름을 발생시키는 것이 아니라 반대 방향으로도 적지

만 데이터 흐름을 발생시키기 때문에 그 영향으로 네트워크 트래픽에 burstiness 정도를 크게 나타나게 할 수 있을 것이다. 먼저 버퍼의 영향으로 인한 것인지에 검증해보기 위해서 실험을 하였다.

실험 3에서는 호스트 시스템의 네트워크 버퍼 크기에 변화를 주어 시스템 네트워크 버퍼가 스토리지와 네트워크 트래픽 사이에 어떠한 영향을 주는 지 확인하였다.

	스토리지 트래픽	네트워크 트래픽
Send/Receive Buffer	호스트->스토리지	호스트->스토리지
32KB	0.732	0.75
64KB		0.754
128KB		0.751

표 2 네트워크 버퍼크기에 따른 Hurst Parameter

표 2는 호스트의 네트워크 버퍼크기에 따른 네트워크 트래픽의 burstiness parameter의 값이 변화하지 않는 것을 보여준다. 이것으로 보아 스토리지와 네트워크 사이에 있는 버퍼는 트래픽 변화에 그다지 영향을 주지 않음을 알 수 있다. 그러면 네트워크 트래픽을 조금 더 burstiness하게 하는 원인으로 생각해 볼 수 있는 것은 iSCSI이다. 스토리지 요구 명령에 따른 iSCSI 요구 명령은 response/status PDU를 추가적으로 발생시키기 때문에 그것이 burstiness를 더 크게 만드는 것이다(그림 2에서 스토리지에서 호스트로의 트래픽).

## 2. 결론 및 향후 연구내용

스토리지 트래픽에 따른 네트워크 트래픽은 패턴에 있어 거의 동일한 양상을 보이지만 iSCSI 프로토콜의 영향으로 호스트 스토리지의 트래픽보다 네트워크 트래픽이 burstiness가 어느 정도 더해진다는 것을 알 수 있다. 이것을 통해, 네트워크 대역폭을 각 호스트에게 할당해 주기 위해서, 스토리지 트래픽을 기준으로 필요한 최소의 대역폭 계산할 때는 네트워크 트래픽이 좀 더 burstiness 하다는 것을 고려해야 한다.

향후 연구 내용으로는 스토리지 트래픽에 따른 네트워크 트래픽이 더 burstiness하다는 점을 고려한 각 호스트의 QoS 설정에 맞는 네트워크 대역폭 계산방법을 고안하는 것이다.

## 참고문헌

- [1] Julian Satran, et al. iSCSI. <http://www.ietf.org/internet-drafts/draft-ietf-ips-iscsi-20.txt>
- [2] Stephen Aiken et al., "A Performance Analysis of the iSCSI Protocol," Proceedings of 20th IEEE / 11th NASA Godard Conference on Mass Storage Systems and Technologies, 2003 (MASS'03)
- [3] Peter Radkov et al., "An Experimental Comparison of Block- and File-Access Protocols for IP-Networked Storage", 3rd USENIX Conference on File And Storage Technologies, 2004 (FAST'04)
- [4] J. Wilkies, "Traveling to Rome: QoS specifications for automated storage system management," Proceedings of International Workshop on Quality of Service, June 2001.
- [5] G. Alvarez et al., "Minerva: An automated resource provisioning tool for large-scale storage systems," ACM Trans. on Computer Systems, Nov. 2001.
- [6] Thomas R. Henderson et. al., "On Improving the Fairness of TCP Congestion Avoidance," Proceedings of Glovecom'98
- [7] Y. Nam, "Dynamic Storage QoS Control for Storage Cluster and RAID Performance Enhancement Techniques," Ph.D Thesis, POSTECH, 2004
- [8] Zhifeng Chen et. al., "Eviction Based Cache Placement for Storage Caches," Proceedings of the 2003 USENIX Annual Technical Conference
- [9] Y. Zhou et. al., "The multi-queue replacement algorithm for second level buffer caches," Proceedings of the 2001 USENIX Annual Technical Conference
- [10] Pawan Goyal et. al., "CacheCOW: QoS for Storage System Caches," Proceedings of 7th International Workshop on Quality of Service, 2003.
- [11] UNH. iSCSI reference implementation. <http://www.iol.unh.edu/consortiums/iscsi/>
- [12] Chris Rummmler and John Wilkes, "UNIX disk access patterns," Proceedings of the 1993 USENIX Winter Technical Conference
- [13] W. Hsu and A. Smith, "Characteristics of I/O traffic in personal computer and server workloads," IBM System Journal, vol.42, no.2, 2003