

# I/O-intensive 애플리케이션의 IP 네트워크 기반 스토리지의 최적 설계

신영호\*, 한성근\*, 김규석\*, 김주영\*, 박남규\*, 이혁로\*

\*한국과학기술정보연구원

e-mail:shinyh@kisti.re.kr

## A Optimized Design on IP-Networked Storage for I/O-intensive application

Young-Ho Shin\*, Sung-Geun Han\*, Kyu-Seok Kim\*, Joo-Yeong Kim\*,  
Nam-Kyu Park, Hyuk-ro Lee\*

\*Korea Institute of Science and Technology Information

### 요 약

웹을 통한 다수의 정보서비스와 전자자원을 공동으로 활용하는 경우에 있어서는, 구성자원에 대한 통합과 공동 활용이 상당히 중요한 요소이다. 일반적으로 웹 서비스를 위한 구성요소로서는 Web, WAS, 데이터베이스, 검색엔진 등의 다양한 소프트웨어를 통합, 운영하게 된다. 또한, 투입비용에 대한 효과적인 측면과 서비스 연속성 측면에서 서버의 공유 및 공유 스토리지에 대한 요구를 피할 수 없게 된다. 본 논문에서는 전체적인 서비스의 성능에 커다란 영향을 미치는 Back-end Tier의 검색서버를 통하여, IP 네트워크 기반의 스토리지 I/O에 대한 성능측정을 수행하였고, 그 성능실험의 결과에 따라 해당 애플리케이션에 최적의 스토리지 설계하고, 애플리케이션의 특성별로 최적의 스토리지를 할당하여 정보서비스의 전체적인 성능향상을 이루었다.

### 1. 서론

KISTI는 과학기술 정보 분야의 전문연구기관으로써, 국가 과학기술 R&D 효율성의 극대화를 목표로 국내 산학연 연구자들에게 첨단연구정보를 제공해 오고 있다[1]. 과학기술정보에 대하여 시멘틱, 가시화 등 다양한 정보유통 기술을 통하여 연구자들에게 정보서비스를 제공하고 있다.

KISTI의 정보시스템은 Web-WAS-DB의 전형적인 3-tier 구조이며, 정보시스템의 핵심 구성요소인 서버와 스토리지를 통합된 구조로 서비스하고 있다. 그리고, 무중단 서비스를 위하여 서버의 이중화 및 Load Balancing을 구축 운영 중이며, 단위 서버가 다수의 서비스를 수행하는 구조이고 데이터 공유 등의 요구에 따라 통합 스토리지인 Unified 형태의 스토리지를 채택하고 있다.

Web과 WAS Tier에서는 I/O보다는 공유의 개념과 특성이 강한 반면, Back-end Tier에서는 I/O-intensive 요구가 많아지게 된다. 특히, 검색서버의 경우, I/O 성능이 전체적인 서비스의 성능에 미치는 영향은 상당한 것이다.

본 논문에서는 다양한 정보서비스 유형에 따른 최적의 스토리지 구성안을 제시하고자 한다. 2장에서는 대표적인 스토리지 연결 형태인, DAS, NAS, SAN 및 네트워크 스토리지에 대해 각각의 프로토콜별로 정리하고, 검색엔진의 구조 및 특징에 대해 설명한다. 3장에서는 검색엔진을 통한 색인 생성에 소요된 시간을 기준으로 프로토콜별로 성능을 비교한다. 4장에서는 성능실험에 따른 비교 결과를

통하여 애플리케이션별로 최적의 스토리지를 할당하여 나타난 결과를 정리하고, 이를 토대로 한 향후 연구방향 등을 기술하였다.

### 2. 관련연구

#### 1) 연결방식에 따른 스토리지

일반적인 스토리지 연결방식은 크게 DAS, NAS, SAN으로 나눌 수 있다. DAS(Direct Attached Storage)는 컴퓨터에 직접 연결된 스토리지를 말하며, NAS(Network Attached Storage)는 TCP/IP 기반의 네트워크(LAN or WAN)으로 연결된 스토리지를 말한다. 그리고, 파일공유와 파일 액세스를 위하여 CIFS와 NFS로 특화된 I/O 프로토콜을 통하여 액세스된다. SAN(Storage Area Network)는 일반적으로 Fibre Channel로 프로세서와 스토리지간에 any-to-any 연결을 제공하는 전용 네트워크로 연결된 스토리지를 말한다[2].

각 연결 스토리지 형태에 따른 장점과 단점을 보면, DAS는 컴퓨터와 스토리지를 전용 케이블로 직접 연결하는 방식으로서, 전용 케이블을 사용하기 때문에 주어진 성능과 안정성이 보장되나, 파일시스템을 공유하도록 하는 방법이 없기 때문에 파일 공유가 불가능하며, 또한 확장성과 유연성도 상대적으로 떨어진다.

SAN은 컴퓨터와 스토리지 간에 대용량 데이터를 고속으로 전송할 수 있는 SAN전용 광 네트워크를 구성해야 한

다. 따라서 뛰어난 성능을 보장받을 수 있으며, 확장성이 뛰어나다. 그러나 전용 네트워크 구축을 위해 컴퓨터에 HBA와 Fiber Channel Switch 및 광케이블 등에 의해 구축비용과 복잡성이 높아지게 된다. 또한 컴퓨터들 사이에 파일 공유가 불가능하다.

NAS는 LAN, WAN 등의 데이터 네트워크를 통한 스토리지 접속을 말하며, 파일 공유의 장점이 있으나, 성능이 다소 낮은 단점이 있다.

## 2) IP-Network Storage

기존에는 IP-Network 스토리지가 여타의 스토리지 연결 방식에 비해, 구축의 편의성은 만족스러웠으나 성능에 있어서 부족한 점이 많았다. 그러나 최근에는 NFS와 iSCSI와 같은 IP-networked 스토리지 프로토콜은 요즘의 LAN 환경에서 일반적으로 사용되고 있으며, 꾸준히 증가하고 있는 추세이다[3]. 또한, 최근의 정보 세계에서는 데이터 스토리지가 CPU와 RAM에 대해 2차적인 역할로부터 일차적인 중요한 역할로 바뀌고 있고, iSCSI는 미래의 스토리지 시스템을 구축하기 위한 훌륭한 솔루션으로 제안되고 있다[4].

본 장에서는 SAN의 뛰어난 안정성과 성능에도 불구하고 별도의 전용 네트워크 및 모든 클라이언트에서 HBA 어댑터를 필요로 하기 때문에 다수의 컴퓨팅장비를 운영하는 경우에 있어서는 비용과 불편함이 따르게 된다. 따라서 구축과 운용편리성측면에서, LAN을 위해 개발된 컴퓨터 네트워크 기술인 Ethernet기반의 IP-Network 스토리지 기술의 대표적인 NFS, iSCSI 그리고, FCoE 기술에 대해 알아본다.

NFS(Network File System) 프로토콜은 네트워크 파일 공유의 표준이 되었고 오랜 기간 동안 시장의 변화와 요구사항에 적응 발전해 오고 있다. 버전별로 주요특징을 살펴보면, NFS v2는 클라이언트와 서버 간에 UDP상에서 RPC(Remote Procedure Call)을 통하여 통신하며 설계의 주요 특징은 stateless 특성이다. NFS v3는 가변길이 파일 핸들들 32bit에서 64bit로 확장하고, 파일 전송 최대사이즈 제한인 8KB를 없앴으며, 성능향상을 위해 비동기 쓰기를 지원했다. 또한, 전송 프로토콜로서 UDP에 추가적으로 TCP를 지원하기 시작했다. 최근 버전인 NFS v4는, 정밀한 데이터 공유 애플리케이션을 위해 성능과 Locking을 개선하는 것이 목적이었다. 주요 특징은, 방화벽을 통한 접근을 위해 프로토콜을 통합하고, 이전버전의 stateless 속성과 달리, 클라이언트가 서버와 stateful로 상호작용한다. 또한, 클라이언트가 cache 파일 데이터를 위임하는 개념을 소개했으며, GSS API를 통해 강화된 보안성을 높였다.[3]

iSCSI(Internet Small Computer Systems Interface) 프로토콜은 SCSI command를 TCP/IP 패킷으로 캡슐화하는 블록레벨 프로토콜이다. 따라서 기존의 투자된 IP 네트워크를 활용하면 된다. FCP(Fiber Channel Protocol)를 사용

하지 않으며, iSCSI는 표준이 잘 되어있어서 이기종간 호환이 가능하며, 유연성이 높고 구축비용이 저렴하다.

일반적으로 iSCSI가 최근의 버전인 NFS v4보다 더 좋은 결과를 보여준다. 그것은 Metadata-intensive 특성 때문이다. NFS는 서버와 클라이언트 사이에 교환되는 메시지를 분석해 보면, 이 메시지의 약 65%가 Metadata관련 메시지로 나타난다[3].

NFS와 iSCSI의 두 프로토콜의 가장 큰 차이는 네트워크 파일 액세스 프로토콜에 있다. 각각 파일 액세스에 있어서, 파일레벨로 운영되는 기본적으로 stateless프로토콜인 NFS는 데이터 통신이 일어나기 전에 몇 가지의 command를 교환할 것이 요구된다. 반면에, iSCSI에서는 Client-side에 파일의 상태를 보관한다. 또한 iSCSI는 TCP protocol을 사용하지만, NFS는 TCP protocol위해 RPC(Remote Procedure Call)을 사용한다.[5].

FCoE 프로토콜은 일반적인 기존의 네트워크 인프라상에서는 구현되지 않으며, CEE(Converged Enhanced Ethernet) 또는 DCE(Data Centre Ethernet)이라 불리우는 새로운 이더넷 환경에서만 동작한다. 또한 10Gb 이더넷 기반의 CEE/DCE는 해당 기능이 지원되는 스위치와 CNA는(Converged Network Adapter)라는 NIC와 HBA 기능을 수행하는 별도의 네트워크 어댑터가 필요하다.

CNA는 파이버 채널 HBA와 이더넷 NIC 둘 다 동일한 하나의 어댑터 카드에 포함하고 있다. CNA는 하나 이상의 물리적인 이더넷 포트를 가진다. 그리고 WWN(World Wide Name)은 SCSI의 64비트 고유 식별자로서 NIC의 MAC address와 비슷하다. 스토리지에 1개의 포트가 있고, 볼륨이 여러 개라면 포트에 하나의 WWN이 있고, 그 아래 볼륨들에 여러 개의 LUN이 할당되어 사용될 것이다.

FCoE 프로토콜은 iSCSI 프로토콜에 비해 최근에 발표된 기술로서, 지금도 계속 발전하고 있는 프로토콜이지만, 현재는 장비 간 호환성, 10Gb기반 및 CNA 카드와 같은 추가적인 장치가 필요하다는 단점이 있다.

## 3) 검색엔진

검색엔진이란 인터넷, 각종 파일, 데이터베이스 등의 데이터로부터, 찾고자하는 주제의 핵심어를 입력하면 그와 유사한 문서를 결과로 제공해주기 위해 사용되는 엔진으로써, 검색알고리즘의 주요 내용은 색인어에 대한 역색인(어떠한 키워드가 주어졌을 때, 어떠한 문서에서 나타났는지를 알려주는 자료구조) 구축과 질의어에서 색인어를 추출하여 색인어를 포함한 문서를 추출하고 가중치를 계산하는 절차이다. 여기에서 색인어란 빠른 검색을 위해서 문서에서 검색의 조건이 될 만한 것들(색인어 : 최소의 의미를 가진 작은 단위인 형태소분석을 통하여 추출된 용어)들의 존재여부와 위치(Posting File)을 미리 추출하여 빠른 탐색 자료구조를 통하여 구축하여 놓은 것이다.[6]

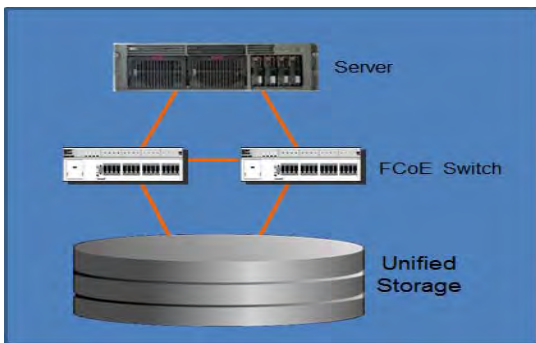
일반적으로 파일이나 데이터베이스에서 레코드를 빨리

검색하기 위해 별도의 inverted file이라고 하는 색인파일을 만들게 된다. 대부분의 검색엔진의 경우 다방향 탐색트리로써, 대용량의 파일을 효율적으로 검색하고 갱신하기 위해 고안된 자료구조인 B-tree 등과 같은 자료구조를 기반으로, 형태소 분류를 통한 Term을 색인 하고, 해당 Term에 어떤 문서가 연결되어 있는지 찾게 된다. 이렇게 색인을 생성하는 과정은 정보검색에 있어서 핵심적인 과정이고, 정보검색 시스템의 성능에 크게 영향을 미치는 요소이다.

### 3. 성능실험

본 연구에서 성능테스트 대상을 검색엔진의 색인생성으로 채택하게 된 배경은 검색대상문서인 원시데이터로부터 검색엔진의 색인기가 색인을 생성하는 과정의 시간이 빠른 정보서비스 개시에 있어서 상당히 중요한 시간이며, 특히, 실시간 색인과 같은 색인 동기화의 필요성이 있는 경우, 색인생성 시간이 더욱 중요하기 때문이다.

색인생성을 위한 성능 테스트를 위한 시스템의 구성은 (그림 1)과 같으며, 사용된 서버 등의 장비 구성 사양은 <표 2>과 같다.



(그림 1) 시스템 구성도

<표 2> 구성장비 간략 사양

서버 (VM)	H/W	Intel Xeon CPU E5-2670 2.60GHz
	Core	32
	Type	x86_64
	OS	RHEL 6 (2.6.32-279) x86_64
	NIC	10Gbps Ethernet / MTU 9000
Switch	FCoE (C사의 Switch) - 10Gbps 기반	
스토리지	Unified Storage (N사)	

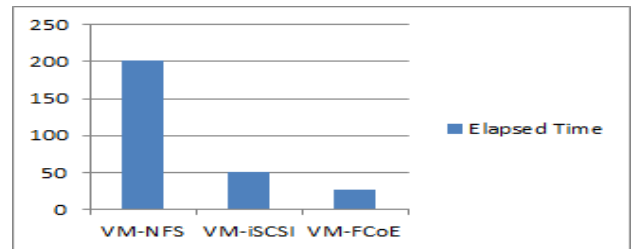
성능테스트를 위한 서버는 스토리지 유형별로 VM-NFS, VM-iSCSI, VM-FCoE로써, 동일사양의 2노드 클러스터로 구성하였으며, 스토리지에서는 연결방식에 따라 NFS, iSCSI, FCoE에 해당하는 스토리지의 볼륨을 각각 생성하였다.

먼저, 성능비교를 위하여 Postmark Benchmark 틀을 통하여 스토리지 유형별로 성능 테스트를 수행하였으며, 성능테스트 조건은 블록사이즈 4KB, 버퍼링제거, 10MB의 파일사이즈, 1000개의 파일, 트랜잭션 수는 1000으로 하였

다. 성능테스트 수행결과는 <표 3>와 같으며, 해당하는 그래프는 (그림 2)과 같다.

<표 3> Postmark 성능비교

구분	수행시간 (초)
VM-NFS	201
VM-iSCSI	50
VM-FCoE	27



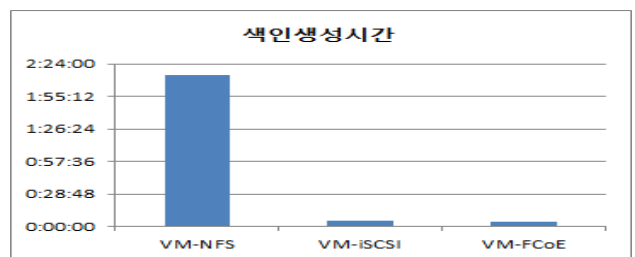
(그림 2) Postmark 성능비교

색인생성 테스트에 사용된 데이터 셋은 2,966,711건의 사용자 데이터로 구성된 정보로서, 데이터 셋 사이즈는 약 1GB, 색인생성 후 색인 사이즈는 약 1.5GB이다.

검색엔진에 의한 색인 성능테스트 방법은, 위의 사용자 데이터 셋에, 필요한 색인 조건에 따라 색인을 생성하는 시간을 측정하는 것으로, 색인생성 회수는 서버 타입 별로 각각 5회씩 수행하였으며, 총 수행결과의 합에 대한 평균값으로 결과는 <표 4>과 같으며, 해당하는 그래프는 (그림 3)와 같다.

<표 4> 색인생성 성능비교

구분	색인생성시간 (시:분:초)
VM-NFS	02:14:13
VM-iSCSI	00:04:49
VM-FCoE	00:04:36



(그림 3) 색인생성시간 비교

### 4. 결론

위의 성능실험결과를 분석해보면, 파일시스템 벤치마킹을 위해 일반적으로 많이 사용되고 있는 Postmark Benchmark 틀을 통한 I/O 성능은 iSCSI를 기준으로 NFS는 약 4배정도 느렸으며, FCoE는 2배 정도 빠른 성능을 보여주고 있다. 그러나 I/O-intensive Workload인 검색엔진에 의한 색인생성 workload의 경우에 있어서는, iSCSI를 기준으로 NFS는 30배 이상으로 성능이 낮았으며,

FCoE의 경우는 근소한 차이로 성능이 우수함으로 나타났다.

본 연구를 통하여, IP기반의 Network Storage를 사용할 경우에 있어서, 일반적인 I/O환경에서와 달리, Workload의 형태에 따라 상당한 성능차이가 나타날 수 있다는 것을 실증하였으며, iSCSI와 FCoE에 있어서는 색인생성과 같은 Workload에 있어서, 현재 기준으로 iSCSI가 FCoE에 비해 색인성능은 다소 FCoE에 비해 떨어지나, FCoE를 사용할 수 없는 조건에 있어서는 적당한 대안이 될 수 있다는 것을 확인하였다. 그러나 FCoE는 현재에도 발전 중인, 부상하는 기술로서 눈여겨 보아야 할 것이다.

본 연구의 결과를 Unified Storage환경의 실제 업무에 적용하여 검색엔진에 의한 색인생성시간을 대폭적으로 단축하였으며, 상당한 정보서비스의 성능향상을 이루게 되었다.

### 참고문헌

- [1] [http://www.kisti.re.kr/sub/sub00/sub00\\_01.jsp?htxt\\_code=001&htxt\\_code=12463259246251054943356398449951&htxt\\_nm=%C0%CE%BB%E7%B8%BB](http://www.kisti.re.kr/sub/sub00/sub00_01.jsp?htxt_code=001&htxt_code=12463259246251054943356398449951&htxt_nm=%C0%CE%BB%E7%B8%BB)
- [2] David Sacks, "Demystifying Storage Networking DAS, SAN, NAS, NAS Gateways, Fibre Channel, and iSCSI", IBM Storage Networking June 2001
- [3] Peter Radkov, Li Yin, Pawan Goyal, Prasenjit Sarkar, Prashant Sheony, "A Performance Comparison of NFS and iSCSI for IP-Networked Storage"
- [4] MS. Sulbha V. Hemke, Dr. A. D. Gawande and Prof. L. K. Gautum, "iSCSI-The Future of the Storage Network", IJAIEEM, Volume 2, Issue4, April 2013
- [5] Ju Dapeng, Liu Chuanyi, Wang Dongsheng, Liu Hong, Tang Zhizhong, "Performance Comparison of IP-Networked Storage", Volumn 14, Number 1, February 2009
- [6] W. Bruce Croft, Donald Metzler, Trevor Strohman, "Search engines: information retrieval in practice"