

SAN환경에서 NDMP를 이용한 백업소프트웨어

황홍연, 송석일, 복경수, 강태호, 남상수, 허정필, 유재수*

*충북대학교 정보통신공학과

e-mail:{hwang1021, prince, ksbok, thkang, ssnam, hjungpil}@netdb.chungbuk.ac.kr
yjs@cbucc.chungbuk.ac.kr

The Backup Software for SAN with NDMP

Hong-Youn Hwang, Seok-II Song, Kyoung-Soo Bok,

Tae-Ho Kang, Sang-Su Nam, Jung-Pil Heo, Jae-Soo Yoo*

*Dept. of Computer & Communication Engineering, Chungbuk National University

요 약

최근 SAN과 NAS와 같은 기술의 발전과 함께 대용량의 데이터를 효과적으로 백업하기 위한 요구들이 증가되고 있다. 본 논문에서는 백업을 위해 제안된 개방형 표준 프로토콜 NDMP를 이용하여 SAN 환경에 적합한 백업 소프트웨어를 설계하고 구현한다. 본 백업 소프트웨어는 일반 파일 시스템의 백업은 물론 데이터베이스, 원시 데이터의 백업이 가능하다. 또한 SAN 환경에 적합하도록 Lan-free 백업과 온라인 백업을 지원한다.

1. 서론

최근 데이터 통신 기술의 발전과 더불어 대용량의 데이터를 가공 처리하기 위한 많은 연구들이 진행되고 있다. 이러한 상황에서 사용자의 실수 또는 천재지변으로 인한 개인 혹은 기업의 정보 손실은 경제적으로 막대한 손해를 미치게 된다. 사용자의 실수, 디스크나 하드웨어의 결함, 소프트웨어의 에러로부터 각종 데이터 및 파일 시스템을 보호하기 위해서는 데이터를 주기적 또는 비주기적으로 백업하는 것이 중요하다.

NAS(Network Attached Storage), SAN(Storage Area Network) 등의 보급으로 저장 매체의 대용량화는 가속화되고 있다. 이러한 상황에서 백업할 데이터도 계속해서 증가하고 있다. 따라서 백업 데이터의 신뢰성을 보장하고 통합 관리하기 위한 백업 전략과 백업 소프트웨어가 필요하다[1]. Network Appliance와 IntelliGuard Software에 의해 NAS 기반에서 이기종 간의 백업을 위한 개방형 표준 프로토콜로써 NDMP(Network Data Management Protocol)가 제안되었다. NDMP는 현재까지

IETF(Internet Engineering Task Force)에 버전 5가 제안되었다[2].

메리랜드 대학에서는 LAN(Local Area Network) 기반의 백업을 위한 AMANDA(Advanced Maryland Automatic Network Disk Archiver)를 제안하였다. AMANDA에서는 자체적인 프로토콜을 설계하여 백업을 수행한다. 그러나 SAN 환경에서 대용량 데이터의 백업을 지원하지 않는다.

본 논문에서는 백업을 위한 프로토콜인 NDMP를 이용하여, NAS 뿐만 아니라 SAN에 적합한 백업 소프트웨어를 시스템을 설계하고 구현한다. 제안하는 백업 소프트웨어는 SAN 상의 방대한 데이터를 보호할 뿐 아니라 향후 SAN 장비를 백업 센터로 사용할 수 있도록 하는데 그 목적이 있다[3].

본 논문의 구성은 다음과 같다. 2장에서는 기존에 이루어져 왔던 연구를 살펴보고, 3장에서는 제안하는 백업 시스템의 구조에 대해서 설명한다. 4장에서는 제안하는 백업 소프트웨어의 특징을 언급하고, 마지막으로 5장에서는 결론 및 향후 연구 계획에 대해서 기술한다.

2. 관련 연구

기존의 AMANDA는 전형적인 LAN 기반의 중앙 집중식 백업 형태를 지니고 있다[4][5]. 테이프 장치 이외에도 테이프 체인저(tape changer)와 같은 백업 장치에 대한 제어 및 관리를 한다. AMANDA는 자체적인 백업 프로토콜을 가지고 있으며 NDMP를 지원하지 않는다.

NDMP는 서론에서 언급한 바와 같이, NAS 기반에서 이기종 간의 백업을 위한 개방형 표준 프로토콜이다. NDMP는 DMA(Data Management Application)와 DSP(Data Service Provider)로 구성되며 이들 사이의 연결을 NDMP 세션이라 한다. DMA는 NDMP의 정책, 사용자 인터페이스, 그리고 DSP를 관리, 제어한다. DSP 타입은 Tape Service, Data Service, SCSI Pass-through Service, Translate Service가 있으며 실제적인 백업을 수행하기 위한 각각의 서비스를 제공하고 있다. Data Service는 주 저장장치에 대한 NDMP 인터페이스를, Tape Service는 보조 저장장치에 대한 NDMP 인터페이스를 제공한다. SCSI Pass-through Service는 DMA가 SCSI장치에 SCSI 명령어를 사용할 수 있는 기능을 제공한다. 버전 5에서 새로 등장한 Translate Service는 데이터 스트림의 멀티플렉싱 및 압축 기능을 제공한다.

NDMP에서는 각 개발자들에 의해 백업 소프트웨어를 쉽게 개발할 수 있도록 NDMP SDK를 제공하고 있다. 이 개발 도구는 NDMP에 명시된 인터페이스를 제공하고 있다. 다양한 백업 장치에 대한 지원, 백업 이미지의 생성, 백업 및 복구를 위한 사용자 인터페이스 부분은 많은 부분이 테스트 수준이다. 따라서, 스케줄링, 압축, 온라인 백업 등의 다양한 백업 형태를 지원하기 위해서는 많은 부분이 추가적으로 고려되어야 한다. 현재 NDMP SDK는 버전 3까지의 NDMP를 지원한다.

3. 백업 소프트웨어의 설계

제안하는 백업 시스템은 크게 두 부분으로 구성된다. 백업과 복구 명령을 처리하는 백업/복구 서버와 클라이언트 데이터에 대한 백업 및 복구를 수행하는 NDMP 연결로 구성된다. [그림 1]은 설계된 백업 소프트웨어의 시스템 구성도를 나타낸다.

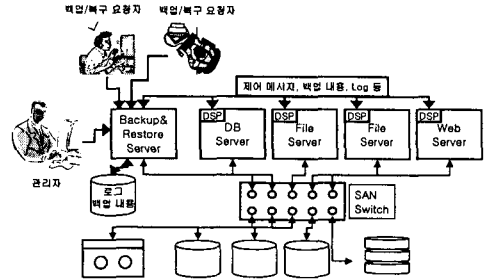


그림 1 시스템 구성도

백업/복구 요청자는 지역 혹은 원격에서 백업/복구 서버에 접속하여 백업/복구 요청을 한다. 백업/복구 서버는 사용자의 요청을 받아서 백업/복구와 관련된 모든 작업을 수행한다. 실제적인 백업/복구에 대한 수행은 SAN 클러스터 내에 존재하는 데이터를 보호하기 위하여 디스크 또는 테이프 장치에 백업을 수행하는 구조를 가지고 있다.

또한 중앙 집중식 구조를 통해서 백업 및 복구의 원활한 관리가 이루어질 수 있다. 일반 파일 시스템, 데이터베이스, 웹서버 등의 데이터를 백업/복구할 수 있다. 내부적으로 백업/복구 데이터의 이동은 NDMP와 호환이 되도록 NDMP SDK의 모든 기능을 수용하였다.

3.1 백업

SAN 환경에서는 크게 두 가지의 백업을 고려해야 한다. 첫째는 SAN 클러스터 내의 모든 데이터를 테이프 등의 보조저장장치로 백업을 하는 경우이다. 여기서 백업되는 데이터는 SAN 환경 내의 일반 파일 및 파일시스템, 데이터베이스 등의 방대한 데이터이다. 둘째는 원격으로 떨어져 있는 클라이언트의 주기억장치에 있는 데이터를 SAN의 데이터영역에 백업을 하는 것이다. 이 경우는 향후 백업 소프트웨어를 백업센터로 확장하여 이용하는 것에 해당하며, 현재는 첫 번째 백업 방법만을 고려한 것이다.

백업 과정은 [그림 2]와 같다. 백업 방법은 그 형태에 따라 자동 백업과 수동 백업으로 나뉜다. 자동 백업은 사용자가 입력한 백업요구 중에서 정해진 시간에 의해 주기적으로 백업을 수행한다. 수동 백업은 사용자가 백업을 요청한 즉시 수행한다. 사용자가 자동백업과 수동백업을 요구하면 각각의 정보는 백업 요구 파일로 저장된다. 자동 백업의 경우, 백업 요구 시간에 백업이 수행되도록 클럭 데몬에서

4. 설계된 백업 소프트웨어의 특징

제안하는 백업 소프트웨어는 NDMP 프로토콜을 이용해서 SAN 환경에서 적합하도록 설계되었다. 여러 데이터 서버 및 테이프 장치, 그 외의 많은 백업 장치가 한 SAN 구성 내에 존재할 수 있다. 이러한 경우 소스 데이터의 위치, 백업 데이터의 위치, 백업 시간, 백업 종류 등 많은 정보를 유지해야 한다. NDMP 프로토콜을 이용해서 이러한 정보를 모두 유지할 수 있다. 필요한 경우에 백업 히스토리 정보를 이용해서 백업 내용의 일부분이나 전체를 복구할 수 있다.

현재 TAR, CPIO를 이용한 백업 및 복구가 가능하며, 또한 DBMS의 백업 및 복구가 가능하다. DBMS의 백업 및 복구에서는 현재 Oracle 9i만을 지원하고 있다. 물리적 백업에서는 데이터, Redo 로그, 제어파일을 tar, cpio, dd, restore를 이용해서 백업 및 복구를 수행할 수 있다[6]. On-line 및 Off-line 백업이 가능하다. 논리적 백업을 지원하기 위해서 Oraback 스크립트의 exp, imp를 이용하였다. 이로써 전체 데이터베이스 혹은 특정 사용자의 데이터베이스, 테이블을 백업할 수 있다[7].

본 백업 소프트웨어는 C 언어와 셸 스크립트 언어로 구현중이며, 현재 SAN 장비와 테이프 장치, 리눅스 서버에서 백업 및 복구 과정을 수행하였다.

[표 1]은 AMANDA와 제안하는 백업 소프트웨어 BUMGR를 비교 분석한 것이다. 제안하는 백업 소프트웨어는 AMANDA에 비해 SAN 환경에 적합할 뿐만 아니라 다양한 백업 타입을 지원하는 특징을 가진다.

표 1 백업 소프트웨어의 특징 비교

기능	AMANDA	BUMGR
중앙집중식 관리	○	○
NDMP 호환	×	○
파일시스템 백업	○	○
DBMS 백업	×	○
Raw 데이터 백업	×	○
압축	○	○
On-line 백업	×	○
LAN-free 백업	×	○
LAN 백업	○	○

5. 결론 및 향후 연구 계획

본 논문에서는 SAN 상의 방대한 데이터를 보호 뿐만 아니라, SAN 장비를 백업 센터로 사용할 수 있는 백업 소프트웨어를 제안하였다. 제안하는 백업 소프트웨어는 다양한 백업 타입을 제공하며 SAN 환경에 적합하도록 LAN-free 백업과 온라인 백업을 수행할 수 있다. 또한 백업을 위한 개방형 표준 프로토콜인 NDMP를 지원한다.

향후 연구 방향으로 다양한 테스트를 통해서 백업 소프트웨어에 대한 안정화를 수행하고 다양한 사용자의 요구를 폭넓게 수용할 수 있도록 인터페이스를 보완할 계획이다. 또한 원격으로 떨어져 있는 클라이언트의 주 저장장치에 있는 데이터를 백업하기 위하여 SAN을 이용할 수 있도록 할 것이다.

참고문헌

[1] A. L. Chervenak, V. Vellanki, Z. Kurmas, and V. Gupta. Protecting File System, "A Survey of Backup Techniques", Pro. Joint NASA and IEEE Mass Storage Conference, 1998

[2] <http://www.ndmp.org/>

[3] 조성훈, 김성주, 이준호, 이주영, 박석천, "SAN의 구조와 기술 요소", 한국정보처리 학회지, 제 8권, 제 4호, pp.19-28, 2001

[4] J. da Silva and O. Guomundsson, "The Amanda Network Backup Manager", Pro. USENIX Systems Administration (LISA VII), pp.171-182, 1993

[5] J. da Silva and O. Guomundsson, and D. Mosse, "Performance of a Parallel Network Backup Manager", Pro. USENIX, pp.17-26, 1992

[6] TAR & CPIO, "Red Hat Linux release 7.3 (Valhalla) man page"

[7] The oraback.sh Script, "<http://www.geocities.com/orabackup/>"