

백업 및 복원 시스템의 설계 및 구현

이 문 구
김포대학 인터넷정보과

Design and Implementation of Backup and Recovery System

Moon Ku Lee
Dept. of Internet Information, Kimpo College
E-mail : yeon0330@kimpo.ac.kr

Abstract

The proposed high-speed backup and recovery system is a high performance backup software based on multi streaming I/O and high speed/high compression technology, and can realize an increase efficiency of backup volume without any increase of backup H/W equipments since it provides a high performance in data backup of a user and executes high compression very fast. Particularly, it minimizes consumption of network band width through a development of algorism for high compression in connection with backup equipments in a large-scaled system environment, is automated when it is recovered due to danger and troubles, and allows users to conveniently access in various ways since GUI and CLI for execution and management are all provided.

I. 서론

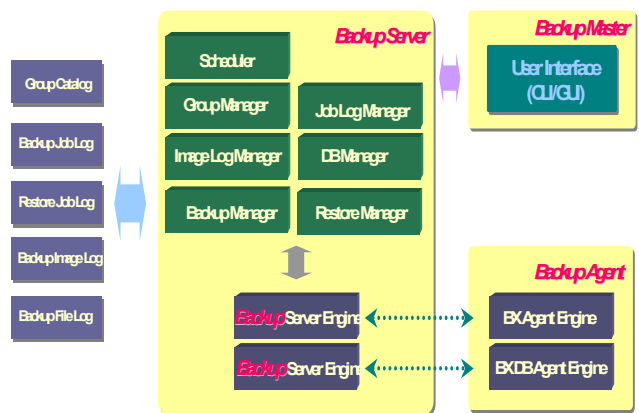
광대역 통신의 보급 확산을 배경으로 정보의 대용량화 및 다양화가 급진전되면서 스토리지 소프트웨어분야의 기술은 높은 시장 성장세를 보이고 있다. 스토리지 소프트웨어는 물리적 저장 매체에 저장된 정보의 접근성(accessibility), 가용성(availability), 성능(performance)을 관

리하고 보증하는 소프트웨어로 정의 되어진다.

기업이나 기관의 주요 데이터가 전산 처리되고 있는 현 상황에서 주전산센터의 시스템장애 및 정전, 천재지변 등 비상사태 발생으로 인한 데이터 손상은 해당 기업이나 기관의기관의 존립기반 자체를 위협하는 결과를 낳을 수도 있다. 따라서 백업센터 구축 및 활용의 필요성은 아무리 강조되어도 지나침이 없는 상황이다. 그렇기 때문에 본 연구는 정보시스템의 보안을 위한 고성능·고압축 백업 및 복구시스템을 설계·구현하였다.

II. 제안하는 백업 및 복원 시스템

2.1 시스템의 구성 요소



[그림 1] 백업 및 복구시스템의 구조

제안하는 백업 및 복원 시스템은 백업마스터, 백업 서버, 백업에이전트로 구성된다. 백업마스터는 GUI, CLI 를 통해, 백업 서버와 사용자의 인터페이스 제공하며, 백업 서버등록, 백업 정책(그룹, 클라이언트, 스케줄, 백업 데이터) 등록·관리 그리고 백업·복구 모니터링, 백업 로그(이미지로그, 파일로그, 에러정보) 검색, 수동 백업·복구 요청 기능을 수행한다.

백업서버는 백업 데이터와 백업 로그 관리, 마스터로부터의 백업 정책 등록과, 수동백업 복구 명령을 받아서 수행 한다. 그리고 등록된 스케줄 정보에 따라 자동 백업하며, 만기일이 지난 백업 이미지(데이터) 자동 삭제하도록 하였다.

백업에이전트는 서버로부터의 백업 요청시 조건에 맞는 데이터를 검색하여 백업서버로 전송하고, 서버로부터의 복구 요청시 데이터를 받아서 복구한다.

2.2 백업 및 복구를 위한 핵심기술

1) 고성능의 대용량 데이터 압축/복원 기술

우선 대용량의 데이터를 어떠한 가공도 하지 않고 백업 시스템에 보관하는 것은 데이터 백업에 사용되어지는 미디어의 소모량 등에 많은 부담을 주게 되며, 다른 시스템으로 백업 데이터를 전송하고자 할 때, 네트워크 대역폭에 부담을 주게 된다. 그러므로 제안하는 기술은 초고속 실행과 고압축률을 동시 구현하여 백업시스템 자원 및 네트워크 대역폭 소모를 크게 감소시키도록 설계되어 시스템 및 네트워크 리소스를 획기적으로 감소시킬 수 있다.

2) 고성능 스트리밍 기술

대용량의 데이터는 복수 개의 디스크에 나누어져 저장되는 것이 일반적이며 대부분의 대용량 시스템에서는 입출력 채널을 이용하여 복수개의 디스크들을 병렬적으로 동시에 접근할 수가 있다. 따라서, 다중 프로세스기법, 다중 쓰레드 기법 등의 병렬 기법을 이용하여 여러 개의 디스크를 동시에 접근해서 데이터를 고속으로 백업 서버로 전달하는 기술을 이용함으로써 대용량 데이터를 백업 받는데 효율적이다.

3) 기존 백업 시스템과의 연동 기술

고가의 백업 시스템을 이미 사용하고 있는 사용자가 새로운 다른 백업 시스템으로 교체하기란 일반적으로 쉽지 않은 것은 자명한 사실이다. 하드웨어의 의존도가 거

의 없고, 기존 시스템과 연동이 용이하게 설계함으로써, 기존에 운용하고 있는 백업 시스템의 성능을 향상시킬 수 있으며 비용 절감 및 운용성을 극대화시킨다.

4) 3-Tier방식 구성

백업 시스템 구성의 융통성과 확장성을 부여하기 위하여 3-Tier방식으로 구성되어 중앙집중 관리 및 백업대상 시스템의 그룹화 관리 등이 용이하다.

Master Layer :

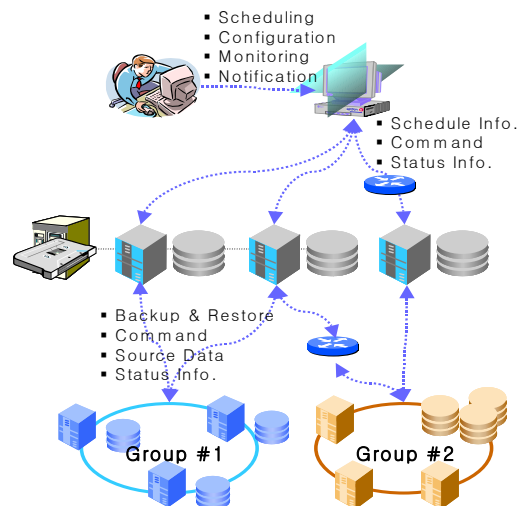
User Interface를 제공, 사용자와 Communication Backup Server에 구성정보/수행명령어 전송 Backup Status등을 사용자에게 전달 단일 마스터로 SPI 및 중앙 집중 관리 체계 제공

Server(Engine) Layer :

Scheduler나 Master의 명령에 따라 백업수행 Agent에 백업/복구 start 명령 전달 Agent로부터 데이터수신, 압축, 저장 백업 대상 증가에 따른 유연한 확장성 제공

Agent Layer :

백업데이터를 보유한 n개의 운영서버로 구성 대상데이터를 고속 read하여 Server로 전송 Grouping 관리로 백업정책의 적용/관리 용이 원거리에 있는 서버의 백업도 가능



[그림 2] 3-Tier방식의 구성도

5) 편리한 User Interface

백업/복구작업 자동화구성, 수행, 관리를 위한 다양한 기능을 UI전문 개발팀이 제작한 GUI 및 CLI로 모두 제공되어 편리하고 다양한 사용이 가능하다.

III. 백업 및 복원 시스템의 알고리즘

3.1 데이터베이스 백업

데이터베이스 복구절차는 발생 가능한 데이터 장애 상황에 대하여 복구 절차를 체계화함으로써 신속하고 완벽한 복구가 가능하도록 관리체계를 유지한다. 장애가 발생하여 적용했던 복구 절차 또는 발생 가능한 장애에 대비한 복구 절차등에 대해 장애 시 복구 활동에 대한 이력관리 등을 통해 최적의 복구 절차를 수립하여 관리한다. 복구 완료 후에 DB정상 작동 여부, 복구된 데이터의 건수 등을 확인하여 이력을 관리한다. 모든 복구절차는 데이터베이스 및 시스템 관리자의 책임 하에 실시한다. 협력 직원은 즉각적인 복구가 될 수 있도록 신속히 장애에 대응하여야 한다.

- (1) 장애 발생 : 데이터 유실 및 서비스 파일 유실 시에 DB관리자의 책임 하에 장애 현상과 원인분석 및 특정 시간 내 복구 가능여부 판단
- (2) 복구 계획 수립 : 장애의 현상과 원인에 부합하는 복구 계획 수립 및 전략 수립
- (3) 백업 : 복구 작업에 실패할 경우를 대비하여 복구 실시 전에 전체 백업 실시, 즉시 복구해야 할 상황이거나 백업 시간의 장기화로 원하는 시간 내에 복구가 힘들다고 판단되면 복구 전 백업은 생략
- (4) 복구 실시 : 필요한 경우 백업 장치 전문가와 데이터베이스 전문가의 도움을 받아 복구 작업을 실시
- (5) 복구 결과 확인
- (6) 백업
- (7) 담당자에게 보고 및 이력관리

백업 및 복구를 위한 단계별 대응 방안을 마련하여 시스템의 설계에 반영하였으며, 적용된 기술요소를 아래의 표로 정리하였다.

[표 1] 백업 및 복구 단계

구분	1단계	2단계	3단계
서버(오픈시스템 기반)	주요부품의 이중화	운영체제 디스크 이중화 구성	업체별 지원체제 유지 즉각적인 복구/기술지원
저장장치(메인프레임)	물리적 장애를 대비한 Raid 구성	즉각적인 복구/기술지원	
데이터(메인프레임 DB2 데이터)	일부데이터 오류 시 백업 본 이용	일정데이터 오류 시 백업 본 및 로그를 이용하여 복구	즉각적인 복구/기술지원

시스템의 신뢰성을 확보하기 위해 각 구성요소별 이중을 통한 안정적인 환경을 구축하고 데이터 장애에 대비한 RAID 구성과 운영체제 영역복제, 백업장치를 이용하여 데이터베이스 및 시스템백업을 주기적으로 실시한다.

```
// process Backup Server Message
protected String sendMessage( String pMessage )
{
    master config file에서 port 정보를 얻어온다.
    연결할 server의 hostname을 설정한다.
    (if aServerHN is null, call setDefaultServerHostname)
    // constructs a backup server connected address (
    hostname + portnum )
    // connects backup server
    // sends Message
    // receive the Result and confirm error state
    // if error, constructs errorMessage
    // return Result or errorMessage
}
// sets Default Server's Hostname
protected void setDefaultServerHostname()
{
    config 파일에서 default server를 읽어서 aServerHN을
    setting한다.
}
// sets the Backup Server Hostname
protected void setServerHostname( String pServerLabel )
{
    주어진 server label의 hostname을 aServerHN으로
    setting 한다.
}
```

[그림 3] 백업명령의 상위 클래스 설정 알고리즘

```
/ main process
public void process()
{
    // parses arguments
    // sets a backup server hostname ( selected or
    default backup server )
    // sets attributes
    // calls related method
    // prints the result ( aOutStr.toString() )
}
public void runScheduledBackup ( )
{
    String sMsg = "" // message to Backup
    Server
    String sResult = "" // result from Backup
    Server
    // constructs a message for sending to Backup
    Server "sMsg"
    sResult = sendMessage( sMsg );
    // append result to aOutStr
}
public void runClientFilesBackup ( )
{
    String sMsg = "" // message to Backup
    Server
    String sResult = "" // result from Backup
    Server
    // constructs a message for sending to Backup
    Server "sMsg"
    sResult = sendMessage( sMsg );
    // append result to aOutStr
}
```

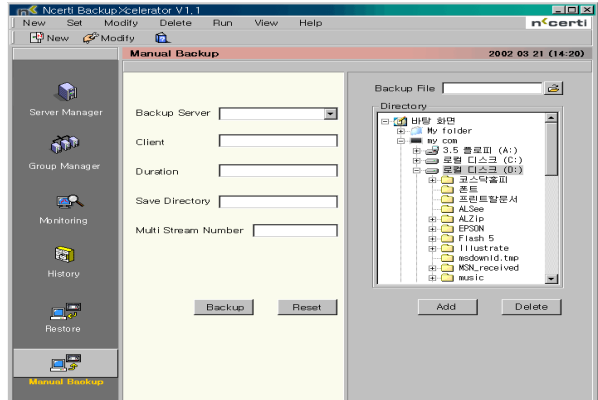
[그림 4] 백업 주요 알고리즘

3.2 고압축 기능제공

기존의 압축기술보다 1/3 ~ 1/5 의 압축률[표 2]과 복구검증 틀을 제공하여, 압축데이터의 안정성 보장한다.

[표 2] 고압축 기능

Solution	압축율	속도	테스트 환경
BackupXcelerator™	68%	2.0 MB/sec	-CPU : Intel P4 1.3GHz -Mem : 256 MB -HDD : 40G UDMA/100 -O/S : WindowsXP -Data : 948MB Volume 776MB Used 73,000개 파일
WinRAR (Window기반)	55%	0.47 MB/sec	
Gzip (Linux기반)	59%	1.7 MB/sec	
A사 백업 S/W	53%	0.12 MB/sec	



[그림 8] 구현결과 인터페이스

3.3 구현결과

1) 제안하는 시스템의 지원 플랫폼

① Scheduler 설정 (Auto Backup)

백업 Configuration을 설정/갱신/삭제
 백업시간, 주기등을 지정, 스케줄링 기능 제공
 서버별 그룹관리 기능 제공
 대상별 백업담당자정보 (Email/Phone등) 등록

② Manual Backup/Recovery

One Time 백업 등을 위한 매뉴얼 백업 기능
 백업된 데이터를 특정 디스크로 복구하는 기능

③ Monitoring

백업 진행 상태 및 디스크 여분을 보여주는 기능
 장애발생 시 담당자에게 Notification

④ Backup History 관리

과거의 모든 백업 History정보를 관리하는 기능
 특정 기간에 발생한 모든 백업 결과를 검색

Agent	Server	BX Server	BX Agent	Compression
Solaris 2.6		○	○	S/C
Solaris 2.7		○	○	S/C
Solaris 2.8		○	○	S/C
Linux		○	○	S
Window NT			○	S
AIX 4.2			○	S
AIX 4.3		○	○	S
HP-UX 10.2			○	S
HP-UX 11.0		○	○	S
IRIX 6.3			○	S
IRIX 6.4			○	S
SUPER-UX			○	S

[그림 6] 제안하는 시스템의 지원 플랫폼\

Agent	Server	Solaris 2.6	Solaris2.7	Solaris2.8	Linux7.3	HP-UX11.0	AIX 4.3
Solaris 2.6		FN	FN	FN	FN	FN	
Solaris2.7		FN	FN	FN	FN	FN	
Solaris2.8		FN	FN	FN	FN	FN	
Linux 7.3		FN	FN	FN	FN	F	
Windows NT		FN	FN	FN	FN	F	
HP-UX10.20		FN	FN	FN	FN	FN	
HP-UX 11.0		FN	FN	FN	FN	FN	
AIX 4.2							
AIX 4.3							F
SUPER-UX (WEC)							F
IRIX 6.3							
RIX 6.4							

- AIX 4.3, SUPER-UX : Solaris, Linux, NT 와 연결 불가
 - IRIX 6.3, 6.4 AIX 4.2 : 테스트 장비 없음

[그림 7] 멀티 플랫폼 포팅 및 테스트

구분	제안하는 시스템	국외			국내	
		V사	L사	C사	B사	R사
특성	중대형 고용량 저렴한 가격	안정성위주 고가	소형 - 중형 저렴한 가격	초기 단계 Reference 미약	소형시장	Disk 백업 중소형 시장
백업 데이터 압축율	60-80%	50-60%	50-60%	50-60%	50%	50%
압축 백업 속도	빠름	보통	보통	보통	보통	보통
다중 스트리밍	○	○	○	○	x	x
타 솔루션과 연동	○	x	x	x	x	x
약점	브랜딩인지도 낮음	대용량 또는 파일의 개수가 많은 경우, 백업/복구 시간 및 데이터 압축률 성능 저하				

[그림 9] 성능비교

IV. 결론

본 연구는 기존의 백업 솔루션들이 갖는 대용량과 대역폭의 문제를 해결하기 위하여 3 Tier 구조로 구성 하였으며, 데이터의 고압축을 이용한 백업 알고리즘으로 초고속 백업 및 복구 시스템을 설계·구현하였다. 본 연구의 기대효과는 고속처리에 의한 백업 및 복구 시간 단축, 데이터 압축기술에 의한 투자 비용의 최소화, 기존 백업 시스템과의 연계를 통한 비용절감, 중앙 집중 제어/관리로 운영 업무 효율화를 들 수 있다.

참고문헌

- [1] 스토리지 네트워킹, 네트워킹타임즈, 2002.06.
- [2] Asia/Pacific: Server and Storage Market, 4Q01, Gartner, January 2002.
- [3] Storage Network Infrastructure 2002: A Guide to Market Definitions and Forecast Methodology, Gartner, June 2002.
- [4] Worldwide Fibre Channel, Host Bus Adapter Forecast and Analysis, 2001-2006, IDC, June 2002.