

Flash SSD를 이용한 CDP(Continuous Data Protection)의 성능개선

On Performance Enhancement of CDP(Continuous Data Protection) System Using Flash SSD

고대식*

Dae-Sik Ko*

요 약

본 대부분의 기업들은 업무 프로세스가 전산화되어 있기 때문에 전산장애나 재해로 인해 시스템이 다운되면 기업의 신뢰성과 사업의 연속성에 치명적인 악영향을 미칠 수 있다. 이에 따라 기업에서는 데이터 손실이 없는 비즈니스 연속성에 대한 관심이 고도화되는 실정이다. 본 논문에서는 시스템 장애를 물리적 장애와 논리적 장애로 분리하여 정의하고 실시간 데이터백업에 필요한 IOPS 성능을 개선하기 위한 Flash SSD를 이용하는 CDP 시스템을 제안하였다. Flash SSD를 사용한 CDP의 IOPS 성능을 측정하기 위하여 실험 시스템을 구성하고, 블록크기의 변화에 따른 IOPS를 실험한 결과, Flash SSD를 이용한 CDP 시스템의 IOPS 성능이 S-ATA에 비해 약 50배 향상되는 것으로 확인하였다.

Abstract

If the System is downed by computer or disasters, it can have a bad influence on the reliability of the corporation and business continuity because all companies are computerized. Accordingly, interest on the business continuity without the loss of data in the corporate is increasing. In this paper, system faults have been defined as physical faults and logical faults and CDP solution using Flash SSD has been proposed for enhancing IOPS which is needed for realtime-backup. In order to measure IOPS performance of the CDP using Flash SSD, we constructed an experimental system. From the results we can see that IOPS performance of CDP using Flash SSD is about 50 times more effective than that of the S-ATA.

Key words : Continuous Data Protection, Business Continuity, Snapshot, Mirroring, Back Up

I. 서 론

2011년 4월 12일에 발생한 N은행의 전산 장애 사태는 악의적인 목적을 갖고 침입한 해커에 의해 데이터가 손실되고 열흘이 넘게 서비스 및 업무가 중단

되었다. 이처럼 대부분의 기업들은 전산 시스템의 프로세스 자체가 업무 프로세스화 되어있기 때문에 전산 장애나 재해로 전산시스템이 다운되면 업무수행이 불가능하다. N은행 사태처럼 데이터 손실은 기업의 업무 차질 및 금전적 손실을 야기할 뿐만 아니라,

* 목원대학교 전자공학과 교수
· 제1저자 (First Author) : 고대식
· 투고일자 : 2011년 8월 9일
· 심사(수정)일자 : 2011년 8월 10일 (수정일자 : 2011년 10월 11일)
· 게재일자 : 2011년 10월 30일

기업의 신뢰성과 사업의 연속성에 치명적인 악영향을 미칠 수 있기 때문에 최근 인터넷 서비스를 제공하는 기업을 중심으로 비즈니스 연속성에 대한 관심이 고조화 되고 있는 실정이다. 장애란 정보기술서비스관리의 통제 가능성 관점에서 협의의 장애 개념으로서, 통제 불가능한 재해, 즉 자연 재해와 인적 재해를 제외한 발생원인 관점에서 직접적으로 영향을 미치는 인적 장애, 시스템 장애, 기반구조 장애 즉 운영 장애, 설비 장애 포함 등과 같은 통제 가능한 요인들에 의한 정보시스템의 기능 저하, 오류, 고장으로 정의한다[1]. 물리적인 장애와 논리적인 장애의 분류기준은 장애가 발생했을 시 장비의 교체유무로 분류를 하였다. 물리적인 장애는 하드웨어의 기계적인 장애와 공유된 스토리지의 고장, 그리고 자연적인 재해가 있다. 논리적 장애에는 소프트웨어의 이상, 운용자 실수, 바이러스, 해커, 시스템 충돌, 데이터베이스의 손상 등이 있다. 시스템 장애가 발생했을 때의 주요 발생 요인을 분석한 것이다. 하드웨어 고장으로 인한 장애가 44%에 이를 정도로 발생 빈도가 높았으며, 그 외에 관리자의 실수, 소프트웨어 장애, 바이러스 피해 등의 순으로 발생 확률이 높게 측정되었다. [1],[2]

CDP(Continuous data protection)는 전산장비의 고장이나 다른 재해에 대비하여 파일 또는 데이터를 지속적인 캡처 또는 트랙 데이터 변경과 주 데이터의 변경을 저장하는 것으로 변경된 데이터를 실시간으로 백업 한다[3]. 임의의 과거 시점에서 발생하는 장애를 대비하여 데이터의 손실 없이 복구 하고 블록, 파일 또는 어플리케이션 기반으로 저장을 하며 다양한 복구시점을 제공하여 장애 시점 바로 전까지 복구가 가능하다. 실시간으로 변경되는 모든 데이터 및 커맨드를 지속적으로 복제하고 보관하여 장애 발생 시 원하는 시점으로 데이터 및 시스템을 복구하여 장애 시에도 서비스 중단 없이 실시간으로 복구 및 서비스를 재개 한다[4]. 하지만 기존 CDP는 IOPS(Input Output Operation Per Second) 처리속도의 제한점 때문에 CDP 서버로 전송된 백업 데이터는 보호대상 서버의 성능과 CDP서버의 성능차이로 발생하는 병목현상 때문에 동시에 많은 어플리케이션을 처리할 수 없고, Unix 서버와는 연동하는데도 제한점이 있었다.

[3],[4]

본 논문에서는 이러한 제한점을 해결하기 위하여 CDP의 캐시부분을 Flash SSD(Solid State Drive)를 이용하여 설계하고 제안된 시스템의 성능을 분석하였다. 이를 위하여 Flash SSD를 이용한 하드웨어와 소프트웨어 제안시스템을 구현하고 실험시스템을 구성하여 블록크기 변화에 따른 IOPS 성능을 실험 분석하였다.

II. CDP를 이용한 장애복구

2-1 백업시스템

기존 백업 시스템은 파일 시스템 기반으로 속도가 느리고 서로 다른 이기종 환경에 대한 제약이 있다. 또한 시스템 장애 시, 모든 데이터에 대한 보장이 어려우며 시스템 재구성 후 서비스를 재가동해야 하는 문제점이 있기 때문에 복구 및 재설치 비용과 사용자에게 서비스 하지 못하는 시간 동안의 경제적 손실이 발생한다. [5]-[8]. 비즈니스 연속성은 비즈니스 환경에서 산재하는 위협에 대비하기 위한 것으로 최대 요건은 데이터 장애 발생 시 얼마나 빨리, 그리고 정확히 복구할 수 있는가 이다. 백업 솔루션은 비용 효율적으로 다양한 잠재 위협을 최소화 하고, 비즈니스의 연속성을 보장할 수 있어야 한다.[5][8]

이러한 문제점을 해결하기 위해 블록 단위의 백업으로 속도를 향상시키고 환경에 대한 제약을 해결할 수 있다. 시스템 장애 시에는 가장 최신의 데이터에 대한 보장이 가능하며 원격 부팅 서비스를 이용한 신속한 서비스 재개가 가능하다. 또한, 비즈니스 연속성의 제공으로 사용자에게 미치는 유형적, 무형적인 경제적인 손실을 최소화 할 수 있었다. 추후 IP/FC 기반의 스토리지로 통합이 가능하며 시스템 및 DR(Disaster Recovery) 환경으로 확장이 가능하다.

백업을 수행하기 위해서 백업된 데이터를 저장할 저장 공간을 확보해야 한다. 백업 소프트웨어에서는 백업된 데이터를 저장할 저장 공간으로 디스크와 테이프 장치를 사용한다. 디스크 장치의 경우 다수의 디스크 장치를 풀이라는 논리적인 단위를 사용하여

각 사용자에게 할당한다. 백업 사용자는 자신에게 할당된 풀 리스트에서 백업 데이터를 저장할 풀을 선택할 수 있으며 사용자에게 신속한 서비스를 재개하기 위해서 적절한 시점에 필요한 유형에 따라 데이터를 백업할 수 있다.

서버의 장애는 하드웨어의 여러 부분에 즉각적인 피해를 가져오기 때문에 백업은 서버에서 분리해 낼 수 있는 매체를 선택하여 보관한다. 데이터의 보존 가치가 높은 기업들을 중심으로 하여 비즈니스 요구에 맞게 B2T(Backup To Tape), B2D(Backup To Disk), VTL(Virtual Tape Library)가 있으며 각각의 특징은 표 2와 같다. [2][5]

표 1. 백업 저장매체의 특징

Table 1. Back up Media characteristic

매체	장점
B2T	가장 널리 사용되고 있는 백업방식 폭넓은 소프트웨어 지원 미디어 제거한 후 별도 보관 가능 저렴한 비용으로 용량 확장 편리함
B2D	백업보관 용량 큼 여러 주기의 보관이 용이 소산 보관이 용이 디스크 RAID를 이용한 백업 및 복구 속도와 안정성 보장 Random Access방식으로 성능 좋음 IT 효율 높음 미디어 오류 적음 원격 복제 본 생성 가능
VTL	빠른 데이터 백업 및 복구 분산된 파일을 복구/검색 하는데 최적 테이프 스토리지용 관리 및 백업 복구 소프트웨어 바로 사용 가능 테이프와 디스크의 장점을 결합한 방식

2-2 CDP를 이용한 비즈니스 연속성

CDP는 전산장비의 고장이나 다른 재해에 대비하여 파일 또는 데이터를 지속적인 캡처 또는 트랙 데이터 변경과 주 데이터의 변경을 저장하는 것으로 변경된 데이터를 실시간으로 백업 한다. [2] 임의의 과거 시점에서 발생하는 장애를 대비하여 데이터 의 손실 없이 복구 하고 블록, 파일 또는 어플리케이션 기반으로 저장을 하며 다양한 복구시점을 제공하여 장애 시점 바로 전까지 복구가 가능하다. 실시간으로 변경되는 모든 데이터 및 커맨드를 지속적으로 복제

하고 보관하여 장애 발생 시 원하는 시점으로 데이터 및 시스템을 복구하여 장애 시에도 서비스 중단 없이 실시간으로 복구 및 서비스를 재개할 수 있으며 그림 1은 이와 같은 CDP의 동작구성도를 보여준다. [3]

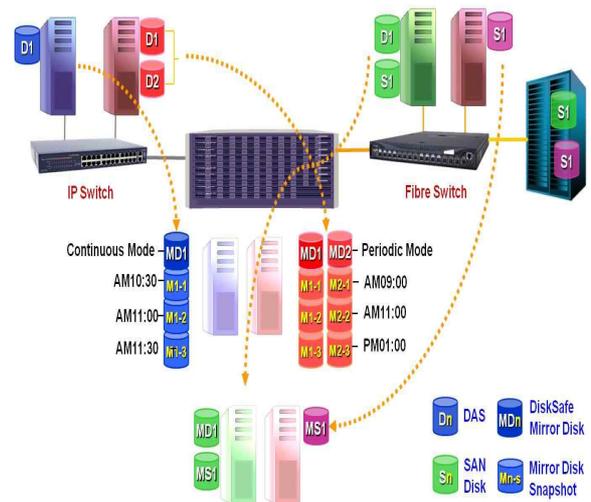


그림 1. CDP 동작구성도

Fig. 1. Configuration of the CDP system

2-3 장애복구 기법

(1) 물리적 장애 복구

재해나 하드웨어의 고장, 스토리지의 고장으로 인한 물리적인 장애는 실시간으로 데이터를 동기식 미러링 및 원격지 복제를 통하여 이중화함으로써 데이터를 보호할 수 있다. 이중에서 동기식 미러링은 여러 개의 스토리지를 하나의 논리적인 스토리지 풀로 구성하고 구성된 스토리지 풀에서 SAN, NAS를 통하여 서버에 부하를 주지 않고 스토리지간에 미러링을 한다. 원격지 복제는 논리적인 스토리지 풀의 데이터를 원격지의 백업서버의 스토리지 풀로 IP기반 네트워크 환경을 통해 비동기 방식의 블록단위로 데이터 복제를 한다.

(2) 논리적 장애 복구

소프트웨어 오류, 운용자 실수, 바이러스, 해커침입, 시스템 충돌로 인한 데이터베이스 및 데이터의 손상 등은 스냅샷 기능을 이용하여 데이터와 Command를 실시간으로 백업 및 보관하여 어떤 시점

으로도 Roll Back 가능하게 함으로써 장애를 복구 한다. 이 중에서 스냅샷(Snapshot)은 사용자가 원하는 특정 시점에서의 데이터 상태를 저장, 유지 시켜주는 기법이다[3][9]. 스냅샷 기법은 그림 2와 같이 데이터 전체가 아닌 데이터에 대한 이미지만을 복사하여 스냅샷 생성 시점의 데이터를 그대로 유지하게 된다. 스냅샷은 기본적으로 미러링 기술을 통해 백업 작업이 수행되는 동안 변경된 부분을 캡처하고 이미지를 복사하는 기술이다[2][10]. Snapshot은 이미지 복제된 데이터를 시/분 단위로 스케줄링 하여 복제한다.

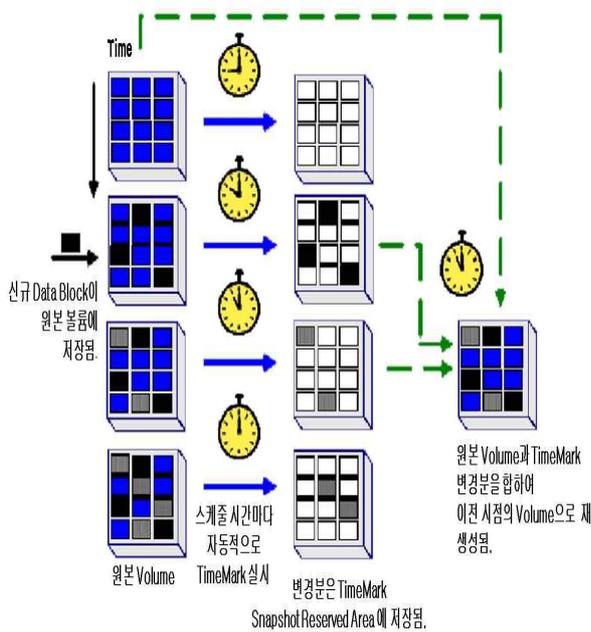


그림 2. 스냅샷 이미지 복제 및 복구
Fig. 2. Replication of the snapshot image and recover

이미지 복제 시 데이터 이미지 외에도 데이터 블록에 대한 인덱스를 매 스케줄에 의해 복제되는 원본 이미지 인덱스를 비교 변경분만을 복제한다. 그림 2와 같이 스케줄화된 이미지 복제 용량을 효과적으로 관리하고 복제된 이미지를 복구할 경우에도 빠른 복구 시간으로 데이터를 복제 한다.

스냅샷 에이전트(Snapshot Agent)는 비 정합성으로 인한 DB/FS 검사 없이 신속하게 DB 복구를 보장한다. DB 및 어플리케이션 시스템의 100% 데이터 정합성 및 무결성을 보장한다.

마지막으로 CDP 저널링(Journaling)은 실시간으로

모든 데이터와 I/O Command의 변경 분을 추적 및 저장하여 원하는 시점에 언제든지 원하는 시점 데이터에 접근할 수 있는 기법이다. 따라서 일반적인 백업은 24시간 동안 한 번의 백업을 하고 백업되는 시점의 데이터가 테이프와 디스크로 백업을 하지만 CDP 저널링은 지속적인 데이터 보호를 통하여 가장 최근의 시점까지의 데이터를 복구할 수 있다.

III. Flash SSD를 이용한 CDP 시스템 설계

본 논문에서는 기존 CDP 성능개선을 위하여 Flash SSD 를 사용하였으며 사용된 Flash SSD의 구조는 그림 3과 같다.

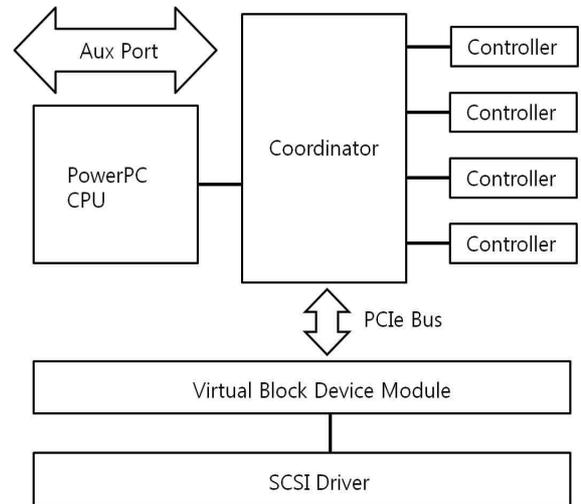


그림 3. Flash SSD 구조
Fig. 3. Architecture of the Flash SSD

그림 3에서 Flash메모리 기반의 SSD를 SCSI장치로 사용하기 위하여 블록장치를 SCSI장치로 변환시켜주는 SCSI 드라이버를 사용한다. 이와 같은 방법으로 CDP 시스템은 Flash SSD를 SCSI장치로 인식하여 동작시킨다. SSD는 HDD에 비하여 대용량의 경우 속도차이가 크게 나타나며 플래시메모리는 소음이나 충격에 강하고 임의접근이 가능한 장점이 있고 쓰기 속도가 상대적으로 느리므로 최근에는 DRAM기반의 SSD에 관한 연구도 진행 중이다. [12]-[14]

그림 4는 CDP 시스템 동작을 위한 전체 시스템 구성도이다.

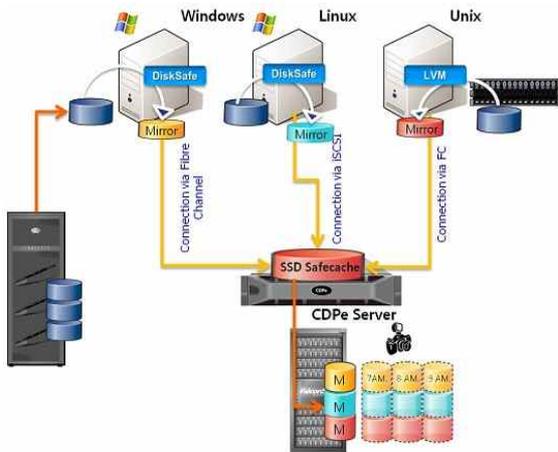


그림 4. SSD 를 이용한 CDP 시스템의 동작구성도
Fig. 4. Configuration of the CDP using SSD

그림 4에서, Windows, Linux, Unix와 같은 보호 대상서버에서 실시간 백업 데이터를 CDPe서버(CDP express server) iSCSI, FC프로토콜을 이용하여 백업 데이터를 전송한다. CDPe 서버로 전송된 백업 데이터는 보호대상 서버의 성능과 CDPe서버의 성능차이로 발생하는 병목현상을 해결하기 위하여 본 논문에서 제안하는 Flash SSD를 쓰기캐시로 사용한다. [11]

CDPe서버 내의 고성능 Flash SSD에 백업 데이터를 저장하고 보호 대상 서버의 실시간 백업을 완료 후에 CDPe서버의 로컬디스크에 백업 데이터를 저장한다. 그림 5는 CDPe 시스템의 백업데이터 처리절차이다. 보호대상서버에의 실시간 백업데이터를 CDPe 서버의 쓰기캐시영역으로 사용하는 Flash SSD에 저장한다.

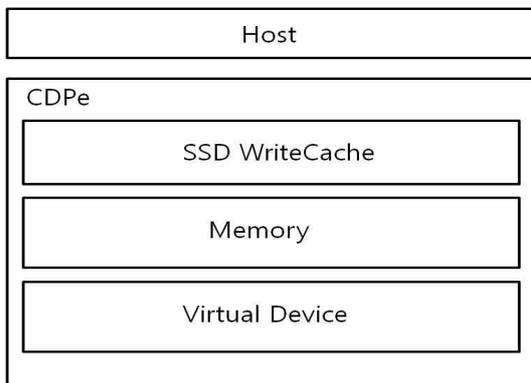


그림 5. 실시간 백업데이터 처리절차
Fig. 5. Process procedure for the real-time backup data

실시간 백업데이터를 SSD 캐시 영역에 저장할 때는 처리속도를 향상시키기 위해 순차쓰기 방법으로 백업 데이터를 저장한다. 저장된 백업데이터는 Flash SSD가 쓰기정책에 따라 CDPe서버의 로컬디스크로 랜덤쓰기방식으로 데이터를 저장한다.

IV. 실험 및 고찰

4-1 실험시스템 구성도

본 논문에서 제안한 CDPe의 성능을 평가하기 위하여 사용된 실험 환경은 (주)비엠스토리지에서 제공하는 그림 6과 같다. 그림 6에서 SATA 디스크와 CDP2의 성능을 비교하기 위하여, 보호대상 서버(Windows 2008 R2)에 대하여 보호대상 서버의 실시간 백업 데이터를 저장하는 오른쪽의 CDPe서버와 왼쪽의 MD 1000 스토리지로써 구성하고, 이때 CDPe를 위한 Flash SSD는 쓰기캐시로만 사용하도록 구성하였다.

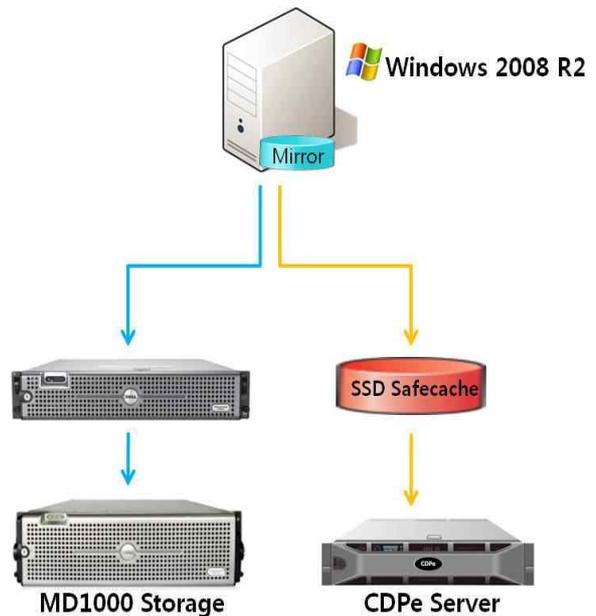


그림 6 실험 구성도
Fig. 6. Test Configuration

실험방법은 성능측정도구인 ipstor workload 툴을 사용하였고, 데이터는 4KB, 8KB, 16KB, 64KB 크기

의 랜덤데이터를 사용하여 6회에 실험을 하였다. 비교대상 시스템은 SATA디스크 스토리지 MD1000을 사용하였고 Flash SSD는 8G PCIe 타입의 SSD를 사용하였다. 평가기준은 가장 일반적인 방법으로 디스크 성능 측정 기준인 IOPS를 사용하였다. CDPe의 실험 환경은 문헌 [15]의 스토리지 성능평가와 유사하며 표 2와 실험 시스템별 환경의 세부사항은 표 3에 설명하였다.

표 2. 실험 조건

Table 2. Experimental condition

데이터 크기	4, 8, 16, 64(KB)
데이터 입출력 방식	Random
실험 횟수	6회

표 3. 실험환경

Table 3. Experimental environment

시스템	시스템 환경
보호대상 시스템 (Dell 2950 III)	- OS : Windows 2008 Enterprise R2 64bit - CPU : Intel Xeon 2.00Hz - M/M : 4G
CDPe 시스템 (Dell 2950 III)	- OS : Oracle Enterprise Linux 5.3 64bit - CPU : Intel Xeon 2.00Hz - M/M : 4G

4-2 성능측정 결과

위에서 언급한 실험조건과 환경에서 Flash SSD기반 CDPe시스템의 초당 데이터 입출력횟수(IOPS)를 실험하였다. 이때, 전송 데이터 블록크기는 4K, 8K, 16K, 64K까지 변화시켜가면서 IOPS 측정하였으며 제안된 CDPe와 기존 SATA 디스크의 성능을 평가한 결과 값은 표 4와 같다.

표 4. 블록크기별 IOPS 실험결과

Table 4. Experimental IOPS results by variation of the block size

block	IOPS (SATA:A)	IOPS (SSD:B)	IOPS 비율 (B/A)
4K	391.81	20550.21	52.45
	402.46	21199.91	52.68
	382.2	20511.56	53.67
	368.23	20544.30	55.79
	366.47	20210.26	55.15
	339.7	20247.40	59.60
8K	383.43	18901.07	49.29
	389.85	18796.61	48.21
	374.85	18448.56	49.22
	363.42	18620.39	51.24
	361.48	18522.88	51.24
	339.38	18529.14	54.60
16K	367.04	15344.65	41.81
	383.74	15163.46	39.51
	351.93	15105.79	42.92
	358.66	15153.81	42.25
	346.88	15137.49	43.64
	330.8	15281.95	46.20
64K	301.69	7258.12	24.06
	327.44	7279.14	22.23
	317.8	7298.72	22.97
	310.03	7291.54	23.52
	302.65	7295.04	23.10
	279.83	7293.63	26.06

표 4에서, 전송 블록크기가 커질수록 입출력 처리 횟수가 줄어드는데 이것은 전송 블록크기가 커지면서 버퍼에 의한 시간 감소 효과가 없어지는 것으로 분석된다. 이분석 결과로부터 전송 블록크기는 크기가 작을수록 입출력 처리횟수의 성능 향상되는 것을 확인하였다. 블록크기에 따라 차이는 있지만 Flash SSD로 CDPe시스템 구성시 SATA에 비해 약 26배에서 50배까지 IOPS 성능이 향상되는 것으로 확인되었다.

V. 결 론

다양한 장애 발생에 따라 비즈니스 연속성을 확보하기 위하여 IT자원을 보호하고 장애발생을 대비하여 비즈니스 연속성을 보장할 수 있는 방안을 살펴보

았다. 비즈니스 연속성은 미러링을 통한 원격지 시스템이나 스냅샷을 이용하여 RTO, RPO를 모두 만족하는 장애 시스템을 복구하여 물리적인 장애나 논리적인 장애가 발생하였을 때 업무를 지속하는 것으로 24시간 정보시스템을 가동하는 기업에게는 필수적인 것이다. 본 논문에서는 실험시스템을 구성하고 CDPe와 SATA 디스크의 성능을 평가하였다.

Flash SSD기반 CDPe시스템에서 초당 데이터 입출력횟수를 실험하였다. 전송 데이터 블록크기는 4K, 8K, 16K, 64K까지 변화시켜 블록크기와 초당 데이터 입출력의 관계를 분석하였다. 실험결과, 전송 블록크기가 커질수록 입출력 처리 횟수가 줄어드는데 이것은 전송 블록크기가 커지면서 버퍼에 의한 시간 감소 효과가 없어지는 것으로 분석된다. 이분석 결과로부터 전송 블록크기는 크기가 작을수록 입출력 처리횟수의 성능 향상되는 것을 확인하였다. 또한, Flash SSD로 CDPe시스템 구성시 SATA에 비해 최대 50배 성능이 향상되는 것으로 확인되었다. 향후, 워크로드의 변화에 따른 CDPe의 성능개선과 실제 어플리케이션의 적용에 대한 성능분석에 대한 연구가 추진되어야 할 것이다.

참 고 문 헌

- [1] 이현중 외 6인, “정보시스템 재해복구 지침”, 국무조정실, *정보통신부*, 2005.12
- [2] 김영호외 4명, “대용량 공유 스토리지 시스템을 위한 효율적인 스냅샷 기법”, *한국정보과학회 논문지 : 데이터베이스*, 제31권, 제2호, 2004
- [3] http://www.snia.org/forums/dmf/programs/data_protect_init/cdp/cdp_definition/
- [4] www.falconstor.com/dmdocuments/CDPDS.pdf
- [5] 조정현, 김윤희, “데이터집중 워크플로우 실행비용최소화를 위한 그리드 스케줄링 기법”, *한국정보과학회논문지*, 7권 2호, pp. 282-289, 2009.4
- [6] N. Zhu and T. Chiueh, “Portable and Efficient Continuous Data Protection for Network File Servers”, *Proceeding of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and*

Networks, Page 687-697, 2007

- [7] G. Laden, P. Ta-shma, E. Yaffe, and M.Factor, “Architectures for Controller Based CDP”, *Proc. Fifth Usenix Conf. File and Storage Technologies* 2007.
- [8] 박경희, 최진규, 이영훈, 정승국, “고속 데이터백업 성능분석”, *한국정보기술학회논문지 제7권1호*, pp. 32-37, 2009.2
- [9] M. Lu, S. Lin, and T. Chiueh, “Efficient Logging and Replication Technique for Comprehensive Data Protection”, *Proc. 24th IEEE Conf. Mass Storage Systems and Technologies*, 2007.9
- [10] K. Daudjee and K. Salem, “Lazy database replication with snapshot isolation”, *In SRDS. IEEE-CS Press*, 2005
- [11] 강윤희, 유진호, 정승국, “DRAM 기반 SSD 저장 시스템 IOPS의 성능평가”, *한국정보기술학회논문지 제7권1호*, pp. 265-272, 2009. 1
- [12] 이영훈, “SSD스토리지를 활용한 Real Machine 통합전략연구”, *한국정보기술학회논문지 제8권 4호*, pp. 169-178, 2010.4
- [13] 고대식, 정승국, “256GB 용량 DRAM기반 SSD의 설계”, *한국향행학회논문지*, 제13권 4호, pp 509-514, 2009.8
- [14] 주용완, “임베디드 시스템을 위한 Block 기반의 NAND 플래시 메모리 파일시스템 설계”, *한국정보기술학회 논문지*, 제8권 8호, pp 47-56, 2010.8
- [15] 고대식, “스토리지 성능평가지침 연구”, *한국향행학회논문지*, 제14권 2호, pp. 266-271, 2010.4

고 대 식 (高大植)



1982, 1985, 1991.2: 경희대학교 공학사, 공학석사, 공학박사

1989~현재: 목원대학 전자공학과 교수

2001~2003 : 목원대학 학술정보처장

2008~현재: 한국정보기술학회 회장

2009~현재 : KSN 포럼 부의장

2011~현재 : 국회산하 NewIT 분과위원장

관심분야 : 멀티미디어 통신, Healthcare, 클라우드컴퓨팅