

데이터 보호를 위한 최적화 방안

디스크 기반 백업/복구 기술 최신 동향과 분석

저자 - 김현규 Principal Sales Consultant, 한국오라클 시스템 사업부

Abstraction

모든 분야에서 많은 새로운 기술과 유행이 나타났다가 사라지기를 반복한다. 그 중에서 일부는 일시적으로 히트 상품이 되기도 하고, 일정 기간 동안 큰 흐름이 되기도 한다. IT 분야도 이러한 흐름에는 대체로 순응하고 있다. 특히, 데이터 보호의 가장 큰 분야인 백업/복구 분야에서도 새로운 기술이 등장하면서 변화가 일어나고, 기존 기술과 경쟁하면서 사라지기도 하고 새로운 자리를 차지하기도 한다.

백업/복구 기술의 변화

최근에는 백업/복구 기술을 데이터 보호 (Data Protection) 기술로도 많이 표현한다. 넓은 의미에서 데이터 보호 기술은 업무용 스토리지 장치에서 흔히 사용되는 RAID (Redundant Array of Inexpensive Disks) 기술부터 물리적 저장 매체의 외부 반출시 의도되지 않는 데이터 유출을 방지하기 위한 암호화 기술에 이르기까지 많은 분야를 포함하고 있다. 개인용 하드 디스크 드라이브에 비해 업무용 하드 디스크 드라이브는 많은 Read/Write 부하를 수행하여 개별 드라이브의 장애 확률이 높고, 이를 대비하여 교체 시 데이터를 재생할 수 있는 RAID 기술은 스토리지 안정성을 높이는 기본적인 구축 사항이다. 마찬가지로, 많은 사람들이 데이터 보호라고하면 백업/복구를 가장 먼저 떠올릴 정도로 기본적인 업무이며, 데이터 센터에 최적화된 백업/복구 시스템을 구축 하기 위해 많은 시간과 비용을 투자하고 있다

기술	테이프백업	디스크백업	VTL	De-Duplication	CDP
적용분야	<ul style="list-style-type: none"> 전동적 백업방식 안정적, 지속적 기술성 속도(매우 높음) 대용량백업 장기보관/소산 	<ul style="list-style-type: none"> 전동적 백업방식 안정적, 지양적 기술성 속도(높음) 대량/작은 파일 운영 서버의 부하 	<ul style="list-style-type: none"> 최근 백업 방식 안정적, 지속적 기술성 속도(높음) 대량/작은 파일 서버/디스크 성능 제약 	<ul style="list-style-type: none"> 최신 백업방식 확산단계 기술성 속도(별견중) 파일 특성 중요 VTL 부가 기능화 업체별 솔루션 다양 	<ul style="list-style-type: none"> 최신/영후백업방식 초기단계 기술성 속도(낮음) 범위 제한적(현재) 솔루션 업체 제한적
어키백처					

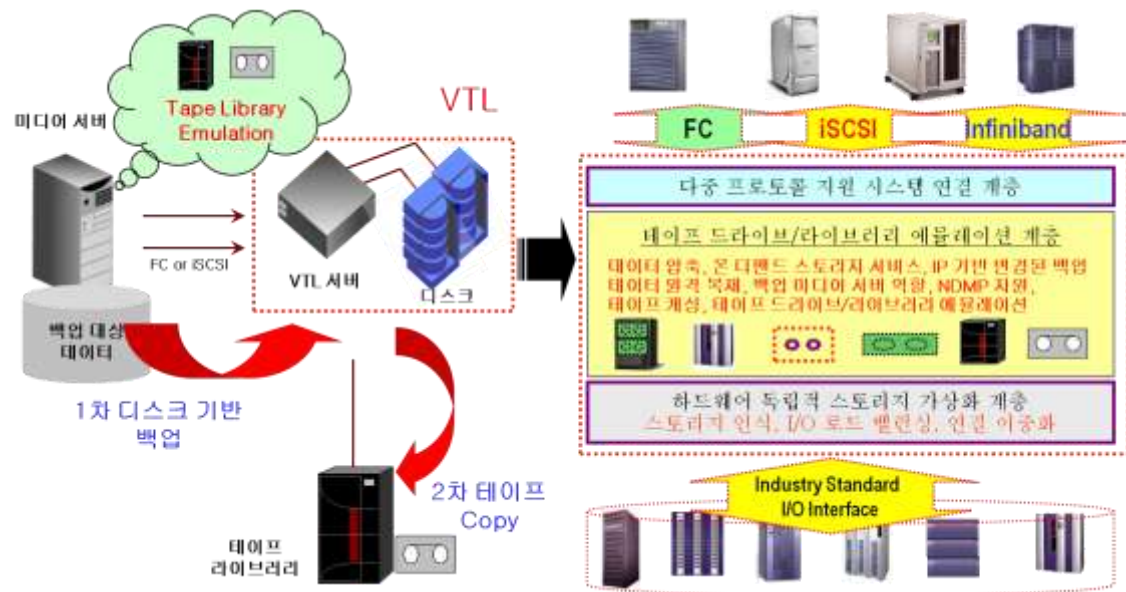
<그림 1> 백업/복구 기술의 변화

백업/복구 기술에도 전통적인 테이프 백업 방식과는 다른 새로운 기술들이 나타나서, 각각의 장단점을 가지고 데이터 센터 요구에 맞게 사용되고 있다. 초창기 테이프 백업 방식에서 낮은 백업/복구 성능, 잦은 미디어 에러, 백업 자원의 비효율성 등의 문제점이 디스크 백업을 고려하게 하였고, 초기 디스크 백업에서는 기존 테이프 백업에서 익숙해 있던 미디어 관리 등의 불편함이 해소되지 않자, 디스크를 테이프 장치처럼 사용할 수 있는 VTL (Virtual Tape Library) 을 출시하게 하였다. 또한, 늘어만 가는 데이터 량의 백업 저장 효율을 높이기 위한 중복 제거 (De-Duplication) 기술의 출현과 RPO (Recovery Point Objective) / RTO (Recovery Time Objective) 단축을 최우선 목표로 하는 지속적 데이터 보호 (CDP, Continuous Data Protection) 기술에 이르기 까지 다양한 백업/복구 기술이 존재한다. 그런데, 여기서 재미있는 것은 몇몇의 특정 업체들이 자기들의 새로운 기술을 출시하면서 예전 기술은 사라질 것이라고 주장 하였지만, 실상은 다른 분야의 기술처럼 백업/복구 기술에도 새로운 기술이 예전 기술의 장점을 완벽히 승계하고 단점을 보강하지 않는한, 보완 관계로 끊임없이 경쟁하면서 공존하는 모습으로 정착되고 있다는 것이다.

앞서 이야기 된 것처럼 최근의 백업 기술 경향은 디스크 기반 백업 기술을 필두로, 백업 매체 저장 용량의 증대, 다중 백업 기술의 필요, 재해 복구를 위한 백업 매체 소산 요구가 주요 관심 사항이다. 그리고, 최적화된 백업/복구 시스템을 구성할 때, 가장 우선시 되어야 하는 사항은 복구에 있다는 것이 강조되고 있다. 즉, 복구할 필요가 없는 데이터는 백업 받을 필요가 없듯이, 복구 목표 시점이나 복구 목표 시간에 맞추지 못하는 백업 시스템은 최적화된 시스템이 아니라는 것이다. 이에 최대의 화두가 되고 있는 디스크 기반 백업 기술의 장단점에 대해 알아보고, 어떻게 적용하는 것이 최적화된 백업/복구 시스템인지 고민해보고자 한다.

디스크 기반 백업 기술 최신 동향과 분석

첫째, 디스크 기반 백업 기술에서 가장 널리 사용되는 기술은 VTL (Virtual Tape Library) 이다. VTL (Virtual Tape Library) 은 <그림 2> 에서와 같이 백업 어플리케이션에게 디스크 스토리지를 마치 테이프 장치처럼 생각하게 하는 가상화 솔루션이다. 그럼으로써, 물리적 매체로서의 디스크 장치의 장점과 논리적인 측면에서 전통적인 테이프 백업 방식의 장점을 모두 얻을 수 있는 솔루션으로서 많은 고객들이 채택하고 있다.



<그림 2> VTL (Virtual Tape Library) 개념

사실 VTL (Virtual Tape Library) 이 처음 출시되었을 때 테이프 시대는 마감될 것이라고 주장하는 사람들이 많았다. 왜냐 하면, VTL (Virtual Tape Library) 이 테이프라는 매체가 가지는 단점을 충분히 제거하고, 성능 적인 측면에서도 디스크가 테이프에 비해 월등할 것이라고 생각하였기 때문이다. 또한, 가격적인 측면에서도 테이프를 충분히 따라 잡을 수 있다고 생각하였다. 그러나, 어느 정도의 시간이 지난 지금 VTL (Virtual Tape Library) 솔루션이 테이프 시장을 많이 잠식한 것은 사실이지만, 그렇다고 해서 일방적인 상황은 아니다. VTL (Virtual Tape Library) 만을 공급하는 업체 조차도 VTL (Virtual Tape Library) 과 테이프 백업을 함께 구성해야 한다고 권장하고 있는 것이 현실이다. 그럼, VTL (Virtual Tape Library) 솔루션을 도입할 때 어떤 요소들을 생각해야 하는지 알아보자.

백업/복구 성능 적인 측면에서 볼 때, VTL (Virtual Tape Library) 이 테이프 장치보다 항상 빠를까? 이런 물음에 “예” 라고 대답하는 전문가는 이제 없다고 생각한다. 정답은 “데이터 성격에 따라 다르다” 이다. 즉, 디스크 장치는 Random I/O 에 더 적합하고, 테이프 장치는 Sequential I/O 에 더 적합하다. 우리가 실행하는 일상적인 I/O 는 디스크에 더 적합하기 때문에 디스크가 항상 빠를 것으로 생각하기 쉽다. 그러나, 백업/복구 업무에서 발생하는 I/O 의 유형은 다양해서, 어떤 장치에만 항상 적합하다고 할 수 없다. 예를 들면, 많은 수의 작은 파일을 백업/복구 하는 경우에는 VTL (Virtual Tape Library) 이 좋겠지만, 적은 수의 큰 파일을 백업/복구 하는 경우에는 테이프가 더 나을 수 있다. 일상적인 상황에서는 시스템 전체를 복구하는 경우가 거의 없기 때문에, 파일 단위의 복구만을 수행한다면 VTL (Virtual Tape Library) 이 더 적합할 수 있다. 그러나, 얼마 전 모 금융 기관의 사고에서 보듯이 시스템 전체를 복구해야 하는 상황에서는, 복구 성능 대역폭이 큰 테이프 복구가 훨씬

원 빠르다는 것을 알 수 있다.

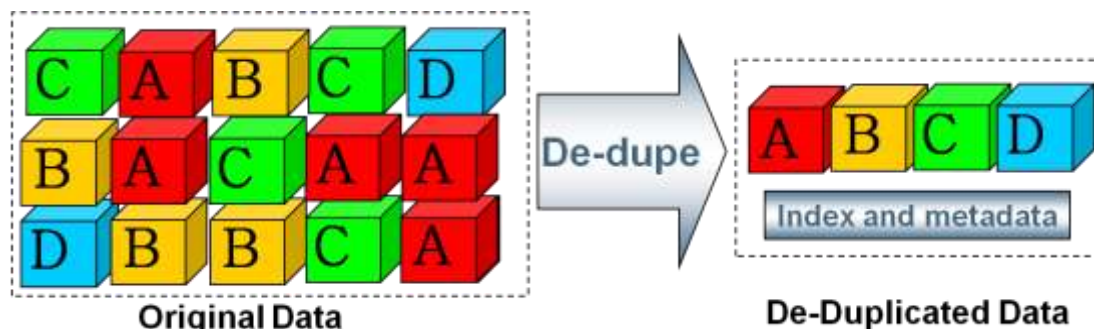
자원 사용의 효율적인 측면에서는 일반적으로 VTL (Virtual Tape Library) 이 테이프 장치를 많이 앞선다고 생각된다. 만약, 백업 성능이 잘 나오지 않는 많은 수의 작은 파일을 백업 받는 경우, 테이프 장치를 사용한다면 편도 1차선 언덕에서 저속의 차량이 앞서가는 것과 같이, 해당 작업이 끝날 때까지 할당된 테이프 장치를 다른 작업이 사용할 수 가 없다. 반면에, VTL (Virtual Tape Library) 장치는 가상 드라이브를 할당하는 개념이기 때문에, 다른 가상 드라이브를 사용하여 동시에 보다 많은 낮은 성능의 작업을 수행을 할 수 있다. 또한, 작은 용량의 백업을 대용량 테이프에 받을 경우, 백업 정책에 의해 해당 미디어 용량을 다 채우지 못하는 경우도 많기 때문에, 저장 공간의 비효율적인 사용이 된다. 반면에, VTL (Virtual Tape Library) 장치는 할당되는 단위가 테이프에 비해 무척 작기 때문에, 데이터를 쓴 만큼만 스토리지 영역이 사용되어, 보유한 저장 공간을 효율적으로 운영할 수 있다.

확장성 및 비용 측면에서 VTL (Virtual Tape Library) 의 경우, 성능과 용량을 고려할 때 VTL (Virtual Tape Library) 어플라이언스 서버 및 디스크 컨트롤러별로 한계가 존재하고, 이를 넘는 경우 또 다른 시스템을 구입해야 하는 계단식 비용 증가형이다. 반면에, 테이프의 경우는 성능 확장을 위해선 드라이브 추가를, 용량 확장을 위해선 슬롯을 추가하는 구조로 되어 있기 때문에, 상대적으로 직선형의 비용 증가 구조를 가지고 있다. 그리고, 일반적으로 VTL (Virtual Tape Library) 시스템 별로 제시되는 최고 성능을 발휘하기 위해선, 일정 수량 이상의 많은 하드 디스크 드라이브를 필요로 하기 때문에, 백업 용량은 많지 않지만 높은 백업/복구 성능이 요구될 경우, 필요 용량 이상의 하드 디스크 드라이브가 장착될 수도 있다.

이상의 비교에서 VTL (Virtual Tape Library) 이 테이프를 완전히 대체하지 못하고, 경쟁과 보완을 함께 하는 관계로 갈 수 밖에 없는 이유를 알 수 있다. 실제로, 많은 고객 사이트에서 VTL (Virtual Tape Library) 은 이제 디스크 버퍼처럼 최근 데이터를 단기간 보관하고 순환 사용하는 영역으로, 테이프는 장기 보관 또는 소산 영역으로 업무 분장을 하고 있다. 이는 최근의 고객 요구 사항이 큰 용량의 VTL (Virtual Tape Library) 보다는 비교적 작은 용량의 VTL (Virtual Tape Library) 을 더 필요로 한다는 것에서도 확인된다.

둘째, 디스크 기반 백업 기술 중에서 많이 언급되고 있는 솔루션 중의 하나는 중복 제거 (De-Duplication) 솔루션이다. 중복 제거 솔루션의 핵심은 백업 대상 데이터에서 기존에 백업된 내용은 다시 저장하지 않고 새로운 또는 변경된 내용 만을 저장 해서, 스토리지 저장 효율을 향상 시킨다는 것이다. <그림 3> 에서 보는 것과 같이 백업 대상 데이터에서 중복된 데이터는 한 개만 저장하고, 나머지는 일종의 인덱스 형태로 메타 데이터에 저장하는 솔루션이다. 이 때, 중복을 확인하는 단위는 대부분의 솔루션에서 가변적 또는 고정적 블럭 단

위 형태가 된다. 사실 우리가 운영하는 정보 시스템에서 중복 제거 효과가 가장 큰 업무는 백업 업무라고 생각한다. 일반적으로 하루에 변경되는 데이터량을 10% 정도로 생각할 때, 매일 전체 백업을 받는다면 매일 90%의 데이터는 중복적으로 받고 있는 것이다. 만약, 이 부분을 절약할 수 있다면 엄청난 스토리지 절감 효과가 나타날 것이다. 즉, 중복 제거 (De-Duplication) 기술은 백업 저장량을 획기적으로 다이어트 할 수 있는 솔루션인 것은 분명하다. 또한, 원격지에 백업을 분산 배치하고자 할 때도 백업 대상 전체 용량에 비해 매우 작은 데이터 량을 전달함으로써, 손쉽게 원격지 백업 구축을 할 수 있는 장점도 아울러 존재한다. 그리고, 현재의 중복 제거 (De-Duplication) 기술은 VTL (Virtual Tape Library) 장치의 부가 기능으로써 주로 보급되고 있는 관계로 VTL (Virtual Tape Library) 도입 시 중복 제거 (De-Duplication) 기술이 있는 VTL (Virtual Tape Library) 을 사용할 것인지, 중복 제거 (De-Duplication) 기술이 없는 일반적인 VTL (Virtual Tape Library) 을 사용할 것인지를 결정하게 된다.



<그림 3> 중복 제거 (De-Duplication) 기술 개념

다음은 중복 제거 (De-Duplication) 솔루션 도입시 고려되어야 할 중요한 요소들에 대한 이야기이다. 일반적으로, 중복 제거 (De-Duplication) 솔루션은 안정성 측면과 데이터 정합성 측면에서 고려되어야 할 사항이 많다. 예를 들면, 중복 제거 데이터 확인 작업 시 대부분의 솔루션이 SHA-1, SHA-256 등과 같은 해쉬 함수 (Hash Function) 를 사용하는 데, 해쉬 함수 특성상 같지 않은 데이터를 같다고 생각할 확률이 “제로 (Zero)” 가 아니라는 것이다. 물론, 상기 해쉬 함수를 사용하는 경우, 대부분의 경우에 있어서 Undetected/Uncorrected 하드웨어 에러 위험보다도 더 낮은 데이터 손실 확률인 페타바이트당 $13 \times 10^{-49}\%$ 이지만, “제로 (Zero)” 는 아니다. 이는 백업 업무가 데이터 보호의 최후의 보루라는 입장에서 볼 때, 비즈니스 크리티컬한 업무에 적용하기엔 위험 부담이 크다. 여기서 백업 업무의 아이러니한 모습이 보이는데, 백업 업무는 완벽한 복구를 목표로 여러 겹으로 다중 백업을 받는 것이 신뢰성과 안정성을 높인다고 생각하는 관습이 있다. 그런데, 중복 제거 (De-Duplication) 기술은 가장 많은 중복을 가지는 업무에 중복을 제거하여 효율성을 높이자는 전략이기 때문에, 트레이드 오프 (Trade-Off) 가 발생할 수 밖에 없다. 또, 중복 제거에 대한 성능을 높이기 위해 약한 해쉬 함수를 사용한다면 위험성이 증가하고, 위

협성을 줄이기 위해 강한 해쉬 함수를 사용한다면 성능 문제가 발생하는 트레이드 오프(Trade-Off)도 존재한다. 그리고, 중복된 데이터 정보를 메타 데이터화 해서 관리하는데, 메타 데이터 장애에 대한 대책도 필수적으로 세워야 한다.

중복 제거(De-Duplication) 기술에 대한 소개시 가장 많이 받는 질문 중의 하나가 중복 제거율이다. 이러한 질문에도 “몇 %” 또는 “몇 배”로 단정적으로 대답하는 전문가는 더 이상 없다. 정답은 “데이터 성격에 따라 매우 다르다”이다. 중복 제거(De-Duplication) 기술 역시 데이터 압축 기술의 하나로 생각할 수 있다. 데이터 압축율이 데이터 성격에 따라 차이가 많이 나듯이, DB 파일, 일반 파일, 문서 파일, 이미지 파일, 멀티 미디어 파일 등 각각의 특성에 따라 중복 제거율도 편차가 심하게 발생한다. 일반적으로 압축이 잘되는 데이터일수록 중복 제거율도 높게 나오는 것으로 간주한다. 그런데, 작은 파일이 많을 수록 중복 제거(De-Duplication) 효과가 크고, 큰 파일이 많을 수록 데이터 압축 효과가 큰 경향도 있기 때문에, 데이터 성격에 따라 중복 제거보다는 압축 기술을 더 선호할 수도 있다. 극단적인 예로, 스냅샷 파일 및 암호화(Encryption) 파일 등은 중복 제거(De-Duplication)의 결과가 원본 파일보다 더 커질 수도 있다. 그렇기 때문에, 중복 제거율을 예측하는 가장 좋은 방법은 실 데이터를 가지고 테스트 하는 방법이다.

성능적인 측면에서도 중복 제거(De-Duplication) 기술은 고려되어야 할 사항이 존재한다. 중복 제거(De-Duplication) 기술 초창기에 몇몇 업체들은 백업 성능도 수십 배가 빨라진다고 강조한 적이 있다. 그러나, 오늘날 이렇게 단정적으로 이야기 하는 전문가는 없다고 본다. 중복 제거(De-Duplication) 방식에 따라 그리고 백업 대상 데이터의 성격에 따라 빨라질 수도 있고 오히려 늦어질 수도 있다. <그림 4>에서 보듯이 중복 제거를 어디서 하느냐에 따라 “소스 기반 방식”과 “타겟 기반 방식”으로 나뉜다. 일반적으로 작은 데이터 규모에서는 소스 기반 방식이 효과가 크고, 데이터 규모가 커질 수록 소스 기반 방식에서 발생하는 중복 제거에 따른 오버헤드가 운영 업무 서버에 영향을 주기 때문에 타겟 기반 방식을 선호한다. 또한, 중복 제거를 언제 하느냐에 따라 “인라인 방식”과 “포스트 프로세싱 방식”으로 나뉜다. 인라인 방식은 데이터를 백업하는 것과 동시에 중복 제거(De-Duplication) 기술을 적용하는 방식이고, 포스트 프로세싱 방식은 데이터 백업이 끝난 후에 별도로 중복 제거를 백그라운드 방식으로 수행하는 방식을 말한다. 보다 큰 대용량 백업의 경우에 포스트 프로세싱 방식보다도 인라인 방식에서 중복 제거 장치의 병목 현상이 심화될 가능성이 더 높다.

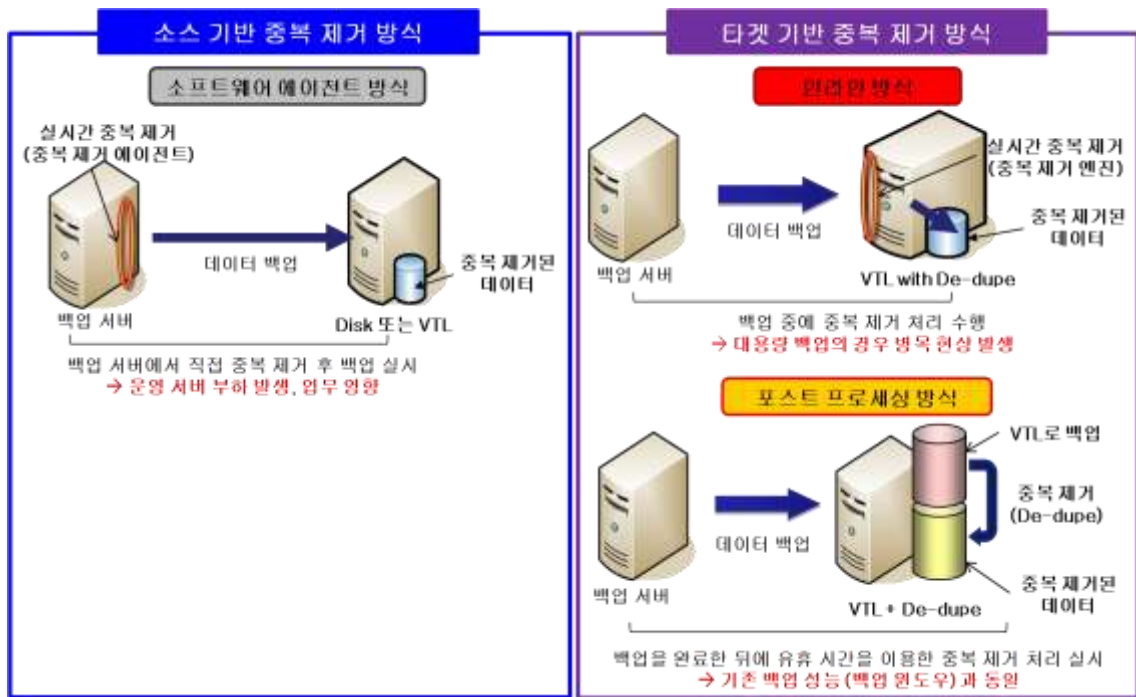
그럼, 백업 성능이 빨라질 가능성이 있는 방식은 어떤 것이 있는 지 알아보자. 일반적으로 백업 실행 시간은 크게 백업 대상 데이터 확인 및 목록 작성, 데이터 전송, 데이터 저장 등의 시간으로 이루어진다. 그리고, 중복 제거(De-Duplication) 기술을 사용하는 경우에 중복 제거 작업 시간이 더해지는 것이다. 그런데, 백업을 실행하는 절차가 중복 제거(De-

Duplication) 기술을 적용할 경우 더 많아짐에도, 백업 실행 시간이 더 빨라질 수도 있는 것은 어떤 경우 일까? 이는 중복 제거 효율이 매우 좋은 경우, 소스 기반 방식의 경우에 데이터 전송과 저장 시간 모두를 많이 절약할 수 있기 때문이고, 인라인 방식의 경우에는 데이터 전송 시간에서는 절약할 수 없지만 데이터 저장 시간을 많이 절약할 수 있기 때문에 가능한 것이다. 그래서, 동일 조건인 경우 소스 기반 방식이 인라인 방식에 비해 백업 성능이 더 높을 수 있다. 반면에, 포스트 프로세싱 방식에서는 백업을 수행한 후 중복 제거를 적용하는 방식임으로, 일반적인 백업에 비해 중복 제거에 의한 백업 성능 개선 효과는 전혀 없다.

그러면, 중복 제거 효율이 안 좋은 경우나 극단적으로 중복 제거가 전혀 되지 않는 상황이라면 백업 성능에 대한 개선 효과가 있을까? 이에 대한 대답은 분명하다. 중복 제거 (De-Duplication) 기술 적용에 드는 오버헤드 시간이 중복 제거 효율에 따른 절약 시간보다 더 크기 때문에, 오히려 백업 시간이 더 늘어날 수도 있다는 것이다. 그렇기 때문에 중복 제거 (De-Duplication) 기술 도입 시 백업 성능 개선 효과를 명시한다는 것은 다소 무리가 있다. 중복 제거 (De-Duplication) 기술 도입에 따른 백업 성능 개선 효과를 예측하기 위한 가장 좋은 방법도 실 테스트이다.

또한, 백업 성능 보다 더 중요한 것은 복구 성능이라고 생각한다. 중복 제거 (De-Duplication) 기술을 적용한 경우 복구 성능도 빨라질 수 있을까? 일반 백업의 경우보다도 중복 제거가 된 데이터를 원상 조합 하는 오버헤드가 더 실행되어야 하기 때문에 복구 성능은 일반적으로 더 느려지는 것으로 생각한다. 거의 모든 기술이 안고 있는 것처럼 중복 제거 (De-Duplication) 기술도 순방향 기능이 있으면 역방향 기능도 있기 때문에, 순방향 기능의 기대 효과가 더 크게 나타날 수 있도록 적용하는 것이 최적화된 사용이다.

중복 제거 (De-Duplication) 기술에 대해 결론적으로 이야기 하고자 하는 것은, 문제가 발생할 수 있으니 사용하지 말아야 한다는 것이 아니라, 이러한 문제가 발생할 수도 있으니 적절하게 사용해야 한다는 것이다. 모든 데이터가 같은 가치를 가지고 있지 않은데, 무엇이든지 최고의 솔루션 만을 적용하려고 하는 것은 아닌지 생각해 볼 필요가 있다. 중복 제거 (De-Duplication) 기술도 약점을 가지고 있지만, 확장성 문제를 해결하고 중복 제거 효율을 더 높이는 글로벌 중복 제거 (De-Duplication) 기술이 계속해서 연구되고 있고, 시장 동향도 확장 추세에 있는 만큼 관심이 필요한 기술인 것은 분명하다. 즉, 백업 대상 데이터의 중요도와 중복 제거시 효과를 면밀하게 판단하여 중복 제거 (De-Duplication) 기술 적용 범위를 결정하는 슬기로움이 필요하다.



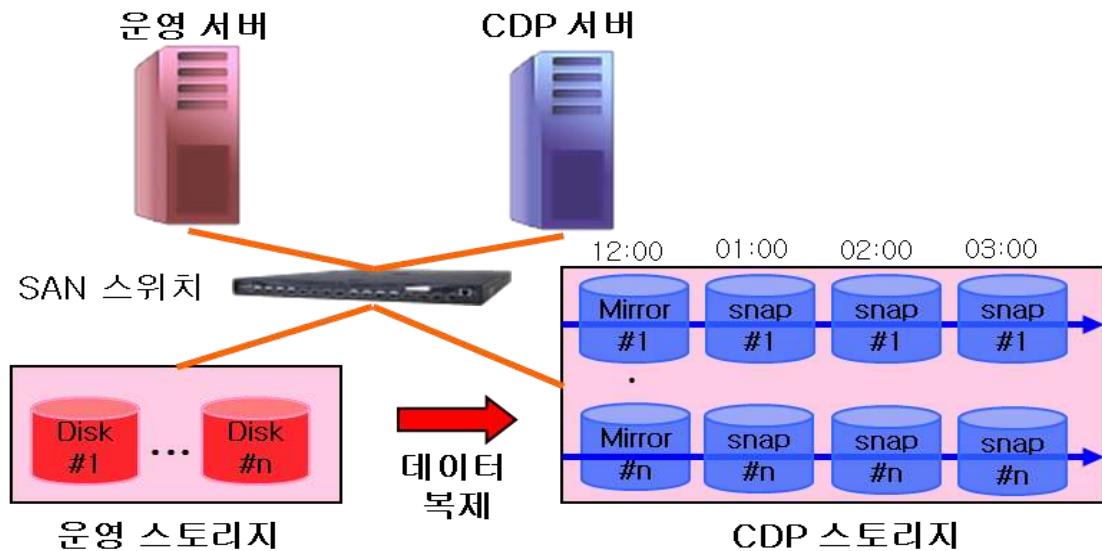
<그림 4> 중복 제거 방식 종류

셋째, 디스크 기반 백업 기술 중 또 다른 기술로 최근 언급되는 것이 지속적 데이터 보호 (CDP, Continuous Data Protection) 기술이다. 지속적 데이터 보호 (CDP) 기술은 지속적 백업 (Continuous Backup) 또는 실시간 백업 (Real-Time Backup) 으로서 불린다. 백업 대상 데이터가 변경되면 자동적으로 다른 저장 장소에 저장하는 기능을 가지고 있고, 기존의 백업 방식과는 다르게 백업 스케줄링 작업이 필요 없을 뿐만 아니라, 원하는 시점 어디로든지 복구할 수 있는 특징이 있다. 즉, 대용량 백업에서 전통적인 백업 방식이 안고 있는 RPO (Recovery Point Objective) / RTO (Recovery Time Objective) 단축을 최우선 목표로 하는 데이터 보호 기술이다. 지속적 데이터 보호 (CDP) 기술도 구현 방식에 따라 많은 차이가 있지만 여기서는 디스크를 기반으로 하는 어플라이언스 형태의 지속적 데이터 보호 (CDP) 기술에 대해서만 언급하기로 한다.

디스크 기반 지속적 데이터 보호 (CDP) 솔루션은 <그림 5> 에서와 같이, 기존의 전통적 스토리지 솔루션인 데이터 복제 (Data Replication), 데이터 미러링 (Data Mirroring), 스냅샷과 같은 유사 기술을 사용하여 원본 데이터의 변화를 복제본에 지속적으로 업데이트 하는 방식으로 구현되는 개념이다. 기술적 특성상 디스크 벤더들이 중심이 되어 솔루션을 출시하였고, 스토리지 가상화 솔루션을 가지고 있는 일부 벤더들도 스토리지 가상화 기술을 사용하여 이기종 스토리지를 지원하는 솔루션을 출시 하였다.

몇 년 전 디스크 기반의 지속적 데이터 보호 (CDP) 솔루션 출시 당시만해도 백업/복구 시

장의 판도를 변화시킬 기술로 주목을 받으면서, 이제 더 이상 백업 어플리케이션도 필요하지 않고, 지속적 데이터 보호 (CDP) 기술로 모든 백업이 구축될 것이라는 주장을 공공연히 들은 바 있다. 그런데, 여전히 진행중인 솔루션이며, 범용적이고 엔터프라이즈급 백업/복구 업무에 적용되기에는 미성숙된 점이 아직 많다. 디스크 기반 솔루션은 기본적으로 FC 인터페이스 기반의 SAN 스위치 상에 연결 되어야 하고, 솔루션 성능 확보를 위한 일정 수량의 포트가 확보 되어야 하는데, 주요 대상 서버의 경우 SAN 스위치 포트 수량 제한에 의해 구축이 불가능 할 수도 있다. 또한, 기존 서버의 SAN 연결을 재구성 해야하기 때문에 적용에 어려운 점이 많은 것도 사실이다. 변경이 많이 발생하는 업무에 적용할 경우, CDP 서버와 백업용으로 사용되는 저가의 디스크에 병목 현상이 발생되어, 운영 업무에 직접적인 영향을 주는 경우도 적지 않다. 그러나, 중복 제거 (De-Duplication) 기술처럼 발전 가능성이 많이 존재하는 기술이고, 전통적인 백업 방식이 안고 있는 RPO (Recovery Point Objective) / RTO (Recovery Time Objective) 감축 문제를 해결할 수 있는 솔루션인 만큼 관심이 필요하다.



<그림 5> CDP (Continuous Data Protection) 기술 개념

최적화된 백업/복구 방안

사람이 살아가는 사회에서도 모든 생활이 자연스럽게 최선의 상태로 유지되는 이상향 (理想郷) 이라면 법이 필요 없듯이, 정보화 시스템 내의 수 많은 시스템이 문제 없이 완벽하게 유지된다면 데이터 보호 솔루션은 필요 없을 것이다. 그러나, 최근 모 포털 업체의 고객 이메일 데이터 유실 사건과 모 금융사의 전산망 마비 사태에서 보듯이, 완벽한 정보 시스템은 고가용성 S/W 및 H/W 인프라 만으로 구축되는 것이 아니라, 높은 IT 보안 의식도 필요로 하고, 불의의 사고에 대비하여 철저한 데이터 백업도 함께 이루어 져야 한다는 것을 알수 있다. 철저하게 안정적으로 정보 시스템을 유지한다고 해도, 사용자 및 운영자의 실수, 프로

그럼의 오류, 시스템 구성 변경 작업 오류 등 데이터 손실이 발생할 확률은 늘 존재하며, 이를 대비해야 하는 것이 상식인 것이다.

최적화된 백업/복구 구성안은 데이터 백업의 목적은 복구에 있으므로, 업무 데이터의 가치에 맞게끔 복구 목표 시점 및 복구 목표 시간을 설정하고, 거기에 맞게끔 백업 정책 및 백업 시스템을 구축해야만 한다. 백업 장비 선정에 있어서도, 디스크 백업과 테이프 백업 중 무엇이 더 우수하다는 우열을 가리는 어리석은 논쟁보다는, 복구 목표를 최우선 고려하고 어떤 경우에도 백업된 데이터가 없어서 복구를 필요로 하는 업무가 복구되지 못하는 경우가 없도록 꾸미는 것이, 최적화된 백업/복구 시스템을 구축하는 것이다.

일반적인 작은 규모의 데이터 손실의 경우는 메인 센터와 보조 센터의 이중 삼중의 디스크 기반 솔루션에서 복구가 될 수 있지만, 앞서 언급된 최근 사례에서도 보듯이 불의의 사고에 의해 모든 디스크 데이터가 손실되어 디스크에서 복구가 될 수 없는 경우, 테이프 백업이 없었다면 상상하기도 싫은 상황이 발생하였을 것이다. 즉, 디스크와 테이프라는 서로 다른 물리적 매체에 백업을 동시에 해두었기 때문에, 디스크로 인한 복구가 어려워지자 테이프 데이터로 복구에 성공하여 시스템을 정상화 할 수 있었던 것이다. 일반적인 경우 흔히 신규 업무 또는 변경된 업무에 대한 백업이 이루어지지 않아 복구를 할 수 없는 경우가 종종 일어난다. 데이터 보호를 위한 백업 업무는 보험과 같은 것이고, 최후의 보루인 것이다. 그러므로, "이런 일이 일어날까?" 하는 생각보다는 "이런 일이 일어날 수도 있다"는 것에 대비하여, 철저한 복구 준비를 하는 것이 최적화된 백업/복구 방안이라고 생각한다. - 끝 -