



# IBM Data Reduction Solutions

## 최소한의 인프라로 최대한의 데이터를 관리하는 방법

본 백서는 오늘날 많은 기업이 직면하고 있는 가장 시급한 문제 중의 하나인 급증하는 데이터의 관리 및 해결 방안을 다루고 있습니다. 자세한 내용은 다음과 같습니다.

### 요약사항

- 핵심 요약
- 스토리지 관리자에게 가중되는 압력
- 데이터 급증 문제의 해결: 효과적인 데이터 정리의 4 단계
  - 단계 1. 데이터 중복 방지
    - 백업 방법
    - Tivoli 스토리지 매니저 6의 점진적 중복 백업
  - 단계 2. 마이그레이션 및 삭제를 위한 데이터 분류
  - 단계 3. 마이그레이션, 아카이브 및 삭제 자동화
    - 편익 - 스토리지의 효율적인 이용
    - 정보 아카이브
    - IBM Optim 및 Tivoli 스토리지 매니저를 이용한 데이터베이스 아카이빙
  - 단계 4. 남은 데이터의 중복제거 및 압축
    - 단계 4a. Tivoli 스토리지 매니저 6를 이용한 중복 제거 및 압축
    - 단계 4b. IBM 시스템 스토리지 ProtecTIER® TS7650을 이용한 남은 데이터의 중복 제거
    - 단계 4c. 검토 - 데이터 중복제거를 위한 TSM 또는 ProtecTIER
- 왜 IBM인가?
- 다음 단계
- 추가 정보
- IBM Tivoli 소프트웨어

### 핵심 요약

그 어느 때보다 많은 데이터가 저장되며 사용되고 있습니다. 데이터 볼륨은 기하급수적으로 증가하고 있으며, 정부 규제 및 경쟁사들간의 경쟁이 강해짐에 따라, 더 많은 데이터를 더 오랜 기간 동안 보유할 수 밖에 없는 상황에 이르렀습니다. 예산은 그대로이거나 오히려 삭감되고 있습니다. 디지털 정보에 대한 의존도가 한층 높아지면서, 데이터 손실에 따른 비용적 손실은 더욱 심각해지고 있습니다.

데이터 시스템을 이용할 수 없게 되면, 수익 창출의 기회를 놓칠 수 밖에 없습니다. 지금과 같은 온디맨드 세상에서 이러한 문제에 신속하게 대응하지 못한다면 고객들은 다른 곳으로 가버릴 것입니다. 비즈니스 중단으로 인하여 기존 고객들과 파트너의 신뢰를 잃을 수 있습니다. 심각한 위반 사실은 항상 알려지기 마련이며 이는 시장에서 어려움을 겪는 상황을 초래합니다.

제품이나 서비스를 공급하지 못하면 계약상의 불이익이 초래될 수 있으며, 정부기관에 대한 보고를 제때 제출하지 못하면 과징금이 부과될 수 있습니다. 극단적인 경우에는 CEO나 CFO가 교도소에 수감될 수도 있습니다. 요는 데이터 자산 보호에 보다 만전을 기해야 한다는 것이며, 이러한 자산들이 증가하고 예산은 축소되면서, 최소의 투자로 최대의 성과를 거두어야 한다는 것입니다. 그러므로 더 똑똑한 솔루션이 필요합니다.



## 스토리지 관리자에게 가중되는 압력

스토리지 관리자들은 “데이터 급증”과의 전쟁에서 일선에 서 있습니다. 데이터가 급격하게 증가하면서 업무상의 요건을 충족해야 하는 이유로 백업 윈도우(Backup Window: 백업 가능 시간)는 더욱 줄고 있는 상황임에도 스토리지 관리자들은 할당된 시간(백업 윈도우) 내에 백업 작업을 완료해야 합니다. 그 결과, 일부 데이터는 전혀 백업이 이루어지지 않을 수도 있습니다.

백업 시스템의 데이터가 증가한다는 것은 문제 발생 시 복구에 더 오랜 시간이 소요된다는 것을 뜻합니다. 이로 인하여 가동 중지시간이 늘어나며, 더욱 큰 위험에 노출됩니다. 물론, 데이터를 제대로 백업하지 않으면 복구 역시 불가능합니다.

이 모든 데이터 증가에 대한 한가지 해결 방안은 보다 많은 스토리지를 구입하는 것이었습니다. 스토리지 자체 비용은 점차적으로 저렴해졌기 때문에, 그다지 심각하지 않았습니다. 그러나 이러한 장비들의 수용, 전력, 냉각 비용 및 관리 비용은 폭증하고 있습니다.

비즈니스가 항상 변화하는 것은 분명한 사실입니다. 스토리지 관리자들은 새로운 애플리케이션과 데이터 소스의 가동에서 인수 합병에 따른 데이터와 시스템의 통합, 그리고 새로운 기업 및 정부 데이터 관리 규정의 준수에 이르기까지 자사 환경 내의 수많은 변화에 적응해야 합니다.

데이터 증가로 인한 결과는 다음과 같습니다.

- 백업에 더 오랜 시간이 소요됩니다. 종종 허용 백업 윈도우를 맞추지 못합니다. 일부 데이터는 적절히 보호되지 않습니다.
- 복구에 더 오랜 시간이 소요됩니다. 가동 중지시간이 늘어나 수익 창출 기회의 손실로 이어집니다. 보호되지 않은 데이터를 복구할 수 없게 됩니다.
- 지속적인 스토리지 구입이 불가능해집니다. 수용 면적/전력 및 냉각용량이 한계에 도달합니다. 운영 및 관리 비용이 급증합니다.
- 새로운 데이터 소스는 문제를 악화시킵니다. 새로운 애플리케이션이 가동됩니다. 인수 합병으로 지원 대상 시스템의 수가 증가합니다.

IBM은 이러한 모든 문제들에 대응할 수 있는 다이나믹 스토리지 관리 인프라를 구축할 수 있게끔 지원합니다. IBM은 데이터 스토리지 풋프린트를 축소하는데 유용한 솔루션을 확보하고 있으며, 이러한 솔루션에 우리가 설정한 목표는 고객사의 자본 및 운영 비용 절감, 애플리케이션 가용성 및 서비스 수준 향상, 급변하는 여건 및 데이터 손실과 관련된 리스크 완화입니다.

**이러한 솔루션을 갖추면, 필요한 스토리지 수의 저감, 관리해야 할 데이터 저감, 가동 중지시간 저감, 경쟁력 제고 등을 이룰 수 있습니다.**

## 데이터 급증 문제의 해결: 효과적인 데이터 정리의 4단계

경쟁사들이 정적 문제의 원인이 될 수도 있는 개별 제품만을 제공하는 것과 달리 IBM은 데이터 관리에 대하여 전체적인 시각의 접근법을 취하고 있습니다. 일례로, 데이터 증가의 주범은 전체 백업을 실시할 때마다 반복되는 대량의 데이터 중복입니다.

따라서 첫 번째 단계는 최종 백업 이후 변경된 데이터만을 백업함으로써 데이터 증가의 근본적인 원인을 제거하는 것입니다.

둘째, 보유하고 있는 데이터의 종류를 파악하고 분류하여 사용 빈도가 낮은 데이터를 저렴한 스토리지 공간으로 옮기고 더 이상 필요치 않은 데이터는 삭제함으로써 가장 효과적으로 데이터를 관리할 수 있게 만드는 것입니다. 프로덕션(실사용) 시스템의 스토리지를 정리하면, 백업 소요 시간이 단축되며, 애플리케이션 성능이 향상될 것입니다.

셋째, 비즈니스 요구 사항 및 서비스 수준 계약에 의거한 자동화된 프로세스를 구축하여 프로덕션 시스템을 가능한 한 건전하게 유지합니다.

그리고 마지막으로 중복 데이터를 제거하고 압축하여 데이터 보호 및 보존 시스템에 저장합니다. 데이터 중복제거는 본 백서의 후반부에 상세하게 다루어져 있습니다.

### 단계 1. 데이터 중복 방지

증상이 아닌, 문제의 원인을 처리해야 합니다. 매주 당 데이터 변화율이 5%라면, 95%의 데이터는 이번 주에 변경되지 않았습니다. 주말에 그 데이터에 대한 전체 백업을 실시한다면, 지난 주말에 백업했던 것과 거의 대부분이 중복될 것입니다. 이는 상당한 스토리지 용량을 차지할 뿐만 아니라 더 오랜 시간이 소요되며, 신규 데이터 생성이 늘어날수록 이러한 문제는 더욱 심각해지지만 할 것입니다. 따라서 이 모든 중복 데이터를 제거하기 위한 데이터 중복제거 제품들이 그렇게 일반화되는 것은 당연한 일입니다. 그 제품들이 백업 스토리지 풋프린트를 95% 이상까지 축소할 수 있다고 주장하는 것은, 바로 이 데이터 중복 제거를 말하는 것입니다.

최초 백업 이후로 단 한번도 전체 백업을 실시하지 않았다면 어떨까요? 항상 신규 데이터와 변경 데이터 만을 백업하였다면, 고가의 중복 제거 솔루션을 요하는 그 모든 중복 데이터가 생성되지 않았을 것입니다. 백업 윈도우 단축, 필요 스토리지 저감, 스토리지 취득 비용 절감 등이 모두 주간 전체 백업 배제에 따른 편익이 될 것입니다.

IBM은 전체 백업 필요성을 방지하는 데 도움이 될 수 있는 똑똑한 솔루션을 보유하고 있습니다. 데이터 보호 솔루션 IBM Tivoli 스토리지 매니저 제품들은 증분(Incremental-only) 백업 기능을 제공합니다.

- Tivoli 스토리지 매니저는 최종 백업 후에 변경된 파일만을 백업합니다. 데이터베이스 파일과 같이 아주 큰 파일의 경우에는, 파일의 변경된 부분만을 복사하는 서브파일 백업(Sub-file Backup)을 실시할 수 있습니다.
- Tivoli 스토리지 매니저 패스트백은 디스크에 기록되는 데이터의 변경된 개별 블록만을 백업함으로써, 관리 수준을 한 단계 끌어올립니다.
- Tivoli 컨티뉴어스 데이터 프로텍션(CDP) 포 파일은 마찬가지로 신규 또는 변경된 파일만을 복사하여, 데스크톱 및 랩톱 컴퓨터의 데이터를 보호합니다. 손실 위험이 있는 데이터의 양을 지속적으로 현저히 줄입니다.

#### 백업 방법:

##### 전체 + 증분 (Full + incremental):

- 백업 - 우선적으로 전체 백업이 필요하며 그 후 증분 백업을 실시합니다. 일반적으로 매 주말마다 한차례 전체 백업을 그리고 이후 6일 간은 증분 백업을 실시합니다. 전날에서 변경된 데이터만이 테이프에 전송됩니다. 이후 또 다시 전체 백업을 실시해야 합니다.
- 복구 - 전체 데이터를 복구한 다음, 각 일자의 증분 데이터를 그 위에 적용시켜야 합니다. 동일 파일에 대하여 한 차례의 전체 백업과 세 차례의 증분 백업이 이루어진 경우, 네 차례의 복구가 필요한 셈입니다. 이는 돈과 시간 낭비이며, 위험을 야기합니다.

##### 전체 + 차분 (Full + Differential):

- 백업 - 우선적으로 전체 백업이 필요하며 그 후 차분 백업을 실시합니다. 일반적으로 매 주말마다 한차례 전체 백업을 그리고 이후 6일 간은 차분 백업을 실시합니다. 이는 최종 전체 백업 이후에 변경된 모든 데이터가 백업됨을 의미합니다. 증가율이 10%라고 하면, 1일차에는 100%(전체)가, 2일차에는 10%, 3일차에는 20%, 4일차에는 30%, 5일차에는 40%, 6일차에는 50% 그리고 7일차에는 60%가 백업될 것입니다. 이는 필요치 않은 환경의 절반 이상을 백업하게 된다는 것을 의미합니다.
- 복구 - 전체 데이터를 복구한 다음에 복구하고자 하는 최종 차분 데이터를 그 위에 적용시켜야 합니다. 이는 환경의 150%를 복구하는 셈일 수도 있다는 것을 의미합니다.

#### 영구 증분(Incremental Forever)

##### (점진적 증분 : Progressive Incremental):

- 백업 - 이는 최초 백업 시에 전체 백업이 필요하며 그 후에는 증분 백업만을 실시합니다. 추가 데이터 전송이 없으므로, 네트워크 대역폭, 전송 시간이 절감되고, 백업 및 복구 시간이 단축되며, 수 천불의 테이프 비용을 절감할 수 있습니다.
- 복구 - 복구하고자 하는 시점을 선택한 다음, 필요한 파일을 복구합니다. 복구는 단 한차례입니다. 추가 트래픽 또는 시간이 필요하지 않습니다. 나머지 두 방법에 비하여 훨씬 빠릅니다.

#### Tivoli 스토리지 매니저 6의 점진적 증분 백업

- 점진적 백업은 신규 또는 변경된 파일만을 백업하기 때문에, 보다 정교하며 지능적인 백업 방식입니다. TSM 관계형 데이터베이스는 각 개별 파일을 추적하여 매일 컴퓨터의 상황을 정확히 파악할 수 있기 때문에 TSM이 이러한 종류의 백업을 할 수 있게 됩니다. 복구가 필요할 경우, 파일의 필요한 버전만이 복구됩니다.
- 다수의 경쟁업체들이 사용하고 있는 전체 + 증분 또는 전체 + 차분 방법은 낭비가 크고 많은 시간이 소요됩니다. 모든 데이터를 주간 단위로 재백업해야 합니다. 그런 다음 일주일 내내, 변경되는 파일을 증분 또는 차분 백업해야 합니다. 복구 시에는 먼저 전체 데이터를 복구한 다음에 증분 또는 차분 데이터를 복구해야 합니다. 이는 더 많은 테이프를 장착해야 하고, 복구할 데이터와 복구 시간이 늘어나며, 복구 정확도가 떨어진다는 것을 의미합니다. (이름이 변경되었거나 의도적으로 삭제한 파일도 복구)
- 점진적 백업은 더 적은 대역폭, 시간, 스토리지 용량, 테이프 드라이브가 소요됩니다. 훨씬 빠르며 이전처럼 장비를 복구합니다.

여타 파일 기반 백업 솔루션들은 적당한 복구 시간을 유지하기 위해서는 일반적으로 주당 1회의 주기적인 전체 백업을 실시해야 합니다. TSM의 독자적인 점진적 증분 방법은 베이스를 마련하기 위한 최초 전체 백업 이후 다시 전체 백업을 할 필요가 전혀 없습니다. 그 결과, 그림 1과 같이, 매 개월마다 수 테라바이트의 백업 용량이 절감될 수 있습니다. 또한 그 모든 데이터를 쓰는데 얼마나 많은 시간이 걸릴지도 고려해 보시기 바랍니다.

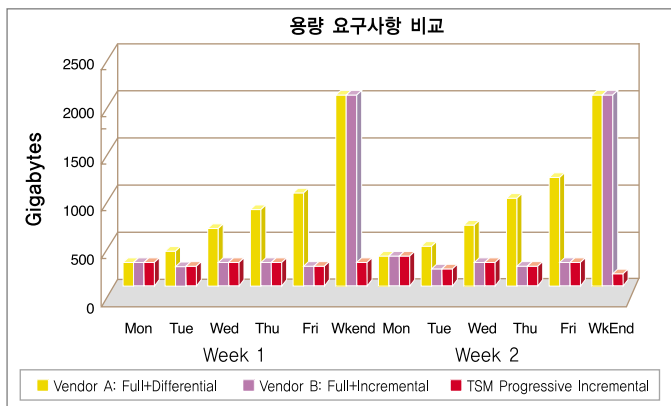


그림 1

차트에서 확인할 수 있듯이, 2 테라바이트의 데이터로 시작하여 매일 200 기가바이트를 추가하거나 변경합니다. 베이스를 마련하기 위하여 이미 전체 백업을 실시했다고 가정합니다.

황색으로 표시된 전체 + 차분은 매주 한 차례 전체 백업이 실시되고 전체 백업을 하는 동안 매일 최종 전체 백업 이후의 모든 신규 데이터를 복사한다는 것을 보여줍니다. 다음 전체 백업이 실시될 때까지 얼마나 많은 양의 백업 데이터가 증가하는지 알 수 있을 것입니다. 이 시나리오에서는, 1개월 치의 백업을 저장하는데 26TB가 필요할 것입니다.

청색으로 표시된 전체 + 증분은 매주 한 차례 전체 백업이 실시되고 전체 백업을 하는 동안 매일 최종 전체 백업이나 증분 백업 이후의 신규 데이터 만을 복사한다는 것을 보여줍니다. 이 시나리오에서는, 1개월 치의 백업을 저장하는데 14TB가 필요할 것입니다.

적색으로 표시된 TSM의 독자적인 점진적 증분 방식은 후속 전체 백업이 전혀 필요치 않으며, 이 시나리오에서 1개월 치의 백업을 저장하는 데에는 7TB만의 용량이 필요할 것입니다.

**1개월간 필요한 백업 용량:**

**벤더 A: 26TB      벤더 B: 14TB      IBM TSM: 7TB**

**단계 2. 마이그레이션 및 삭제를 위한 데이터 분류**

데이터 스토리지 풋프린트의 축소를 위한 다음 단계는 데이터의 유형과 그 수명 주기 상의 위치 측면에서 보유하고 있는 데이터를 평가하는 것입니다. 다수 기업들은 모든 데이터를 플랫폼 파일 시스템(Flat File System)에 보관하고 있습니다. 이런 경우라면, 파일 시스템은 거의 또는 전혀 액세스할 필요가 없는 데이터로 가득 차있을 것입니다. 더 심각한 것은, 법적 소송에서 밝혀질 경우 큰 손해가 따를 수 있는 오래된 이메일과 메모 등과 같이 누군가 봐서는 안될 데이터도 있을 수 있습니다.

목표는 어떤 데이터를 저가의 스토리지 티어로 이동시킬 수 있는지, 그리고 환경에서 완전히 삭제할 수 있는 지를 파악하는 것입니다. 그럴 수 있다면, 주 스토리지 용량의 구입 필요성을 줄일 수 있으며, 보유 데이터의 관리와 보호가 한층 간편해질 것입니다. 백업 및 복구 성능이 향상될 것이며, 데이터 보존 및 만료 정책의 준수를 입증하기가 수월해질 것입니다.

IBM은 이 중요한 단계를 위한 Tivoli 스토리지 프로덕티비티 센터 포 데이터를 제공하고 있습니다. 이 소프트웨어를 이용하면, 액세스 또는 저장 일자 별, 사용자 별, 생성한 애플리케이션 별, 그 외 여러 필터 별로 정렬된 데이터 위치 관련 보고서를 입수할 수 있습니다. Tivoli 스토리지 프로덕티비티 센터 포 데이터 보고서를 통해 입수하는 정보를 기반으로 데이터 관리 소프트웨어에 효과적인 정책을 설정하여, 주 시스템을 저해하지 않도록 데이터에 적절한 조치를 자동으로 취할 수 있을 것입니다.

Tivoli 스토리지 프로덕티비티 센터 포 데이터는 중복 데이터, 고아 데이터, 임시 데이터, 비업무 데이터를 식별하여 제거하는 데에도 유용합니다.



그림 2: Tivoli Storage Productivity Center

### 단계 3. 마이그레이션, 아카이브 및 삭제 자동화

Tivoli 스토리지 매니저 제품군에는 여러 스토리지 티어 간의 데이터 마이그레이션을 자동화하기 위한 두 가지 솔루션이 들어 있습니다. TSM 6 포 스페이스 매니지먼트는 AIX 데이터 용이며, TSM HSM 포 윈도우즈는 윈도우즈 서버 용입니다. 이 솔루션들은 백그라운드에서 작동하며, 설정된 관리 정책에 의거하여 자동으로 파일을 선별하여 이동시킵니다. 이 솔루션들은 데이터가 원래 저장되었던 곳에 포인터나 스텝(Stub) 파일을 남겨, 사용자나 애플리케이션이 그 데이터가 어디로 이동되었는지를 신경 쓰지 않아도 됩니다. TSM이 이동된 파일에 대한 요구를 그대로 재할당(Re-Route) 합니다.

#### 특징

- Tivoli 스토리지 매니저 6는 자동 공간 관리 기능을 제공하여 고객들이 데이터 증가 및 그와 관련된 스토리지 비용을 통제하고 효율적으로 관리할 수 있도록 도움을 줍니다.
- 스토리지 풀 “가상화”는 TSM 스토리지 자원의 이용률을 극대화하는데 유용합니다.
- 계층 내의 데이터 위치에 입각한 최적화된 복구 관리
- 사용자 및 애플리케이션에 명료합니다.
  - 간단한 포인터(스텝 파일)가 원래 위치에서 데이터를 대체합니다
  - 디스크에서 클라이언트로의 신속한 직접 복구
- 피크 시간 중의 네트워크 트래픽을 최소화하는 지정 마이그레이션 (Scheduled Migration)
- 백업 윈도우 외의 시간에 자동 마이그레이션이 이루어집니다.

TSM의 데이터 마이그레이션 솔루션은 주 스토리지 시스템이 효율적으로 작동하도록 정리를 도와줍니다. 이 솔루션들을 이용하여 새로운 스토리지 기술을 배치할 때 데이터를 간편하게 이동시킬 수 있습니다. 그리고 이 모든 것이 백그라운드에서 이루어져, 어느 프로덕션 또는 백업 프로세스에도 영향을 주지 않습니다.

이 공간 관리 솔루션은 파일 사이즈, 파일을 열지 않은 기간 등에 대하여 설정한 정책에 의거하여 TSM 클라이언트에서 TSM 서버로 파일을 자동 이동시킵니다. 실제 TSM 클라이언트 상에 작은 스텝 파일을 남겨 두어, 그 파일을 액세스할 경우 사용자나 관리자가 개입할 필요 없이 자동으로 TSM 서버에서 그 파일을 호출하여 재사용할 수 있도록 TSM 클라이언트로 원위치 시킵니다.

더 이상 “디스크 공간 부족” 메시지가 나타나지 않게 됩니다. TSM은 설정된 정책에 의거하여 적절한 미디어로 데이터를 이동시켜, 활성 파일을 위해 소중한 디스크 공간을 늘리고 필요시 이 파일들에 대한 자동 액세스를 제공합니다.

파일들을 TSM으로부터 이동시키면 재난 발생시 이동된 파일들을 복구할 필요가 없으므로, 복구 시간을 단축할 수 있습니다.

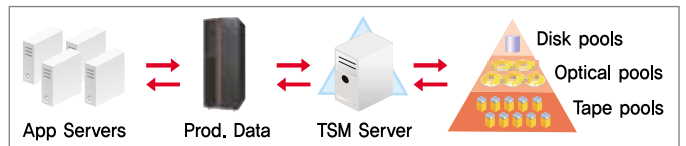


그림 3: TSM 스토리지 계층

#### 편익 - 스토리지의 효율적인 이용

- 비활성 데이터를 오프로드시켜 파일 서버의 응답 시간을 개선
- 저활성 또는 비활성 파일을 저가의 스토리지 계층으로 이동
- 기존의 스토리지 자산을 보다 효율적으로 이용
- 활성 파일에만 집중함으로써 백업 시간과 자원 이용을 저감
- 수작업 파일 시스템 정리 작업을 배제

## 정보 아카이브

아카이빙은 특정 유형의 데이터에 대한 또 다른 중요한 데이터 정리 기법입니다. 그 일례로는 재무 보고 데이터(주간, 월간, 분기, 연간 데이터)를 들 수 있습니다. 이는 향후 추이/요구사항/감사를 위해 보존되어야 하는 반면, 활성 데이터가 존재해야 하는 귀중한 디스크 공간을 소비할 필요는 없습니다. 진로 이력 기록과 고객 명세서도 흔히 이 범주에 속할 것입니다.

아카이빙은 장기적인 기록 보존용입니다. 한정된 기간 동안 파일을 보존하다가 (반면에 백업은 특정 버전의 파일을 보존) 주 프로덕션 스토리지 시스템에서 완전히 삭제한다는 점에서 백업과는 차이가 있습니다.

### 특징

- 비용 효과적인 미디어 상에서의 장기 스토리지
- 실시간 카피(Point in Time Copy); 수정 이력 및 감사 가능성
- 보존 기간 및 '보존 유지(Retention Hold)' 강화
- 신속한 만기 프로세스 (Expiration Processing)

### 장점 - 기록 보존

- 저가의 스토리지 계층으로 파일 아카이브를 이동시켜 파일-서버 복구 시간 단축-활성 데이터만을 복구
- 활성 파일에만 집중하여 백업 시간 및 자원 이용 저감
- 아카이브된 파일을 저가 스토리지 계층으로 이동
- 이력 정보 소재 파악에 도움이 되도록, 아카이브된 파일들을 기술 메타데이터(Descriptive Metadata)로 색인화

IBM은 고객 선호도 및 관련 애플리케이션에 따라 아카이빙을 위한 여러 가지 솔루션을 제공하고 있습니다.

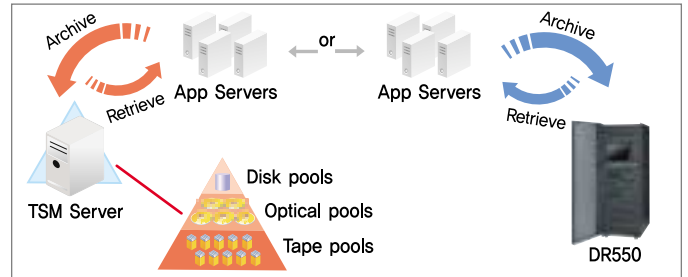


그림 4: IBM 아카이빙 솔루션

Tivoli 스토리지 매니저 6에는 클라이언트 백업 소프트웨어에 직접 통합되어 있는 아카이빙 기능이 포함되어 있습니다. 이는 정책 기반으로, 운영자가 보존 기간을 설정할 수 있습니다. 변경을 위한 보존 기간에 대한 요구사항에 있어 운영자가 해야 할 일은 정책을 업데이트하는 것이 전부이며, 따라서 TSM은 이미 아카이브된 파일들을 소급하여 업데이트 합니다. 일부 경쟁사의 제품처럼 복구하여 다시 아카이브할 필요가 없습니다. Tivoli 스토리지 매니저를 통해 여러 종류의 애플리케이션의 데이터를 아카이브 리포지토리로 통합시킬 수 있으며, 그 아카이브 리포지토리를 이기종 스토리지 시스템의 가상화 풀로 만들 수 있습니다.

시스템 스토리지 아카이브 매니저라고 하는 전문화된 버전의 TSM이 포함되어 있는 IBM 시스템 스토리지 DR550은 독립형 아카이브 어플라이언스입니다. 이는 메시징, 헬스케어 및 메디컬 이미징, 설계 및 엔지니어링, 문서 관리 등의 광범위한 산업을 지원하는 40가지 이상의 애플리케이션에서 직접 데이터를 가져옵니다.

## IBM Optim 및 Tivoli 스토리지 매니저를 이용한 데이터베이스 아카이빙

IBM Optim 데이터 그로우(Growth) 솔루션은 데이터베이스 테이블에서 불필요한 기록들을 이동시켜 보조 스토리지에 저장하는 독특한 데이터베이스 아카이빙 솔루션입니다. TSM의 공간 관리 및 아카이브 솔루션과 마찬가지로, Optim은 데이터베이스 및 스토리지 관리자에게 광범위한 비용 및 성능 편익을 제공합니다.

Optim 데이터 그로우 솔루션과 연계하여 TSM을 이용한다는 장점도 있습니다. Optim은 TSM의 API와 매끄럽게 연동하여 아카이브된 데이터베이스 기록들을 직접 TSM의 스토리지 계층으로 이동시킵니다. 여타 벤더들의 백업 솔루션과 마찬가지로, 아카이브된 기록들을 파일 시스템에 저장하여 이를 백업하는 2단계 프로세스입니다. 필요시, 백업으로부터 데이터를 복구하는 경우에도 이와 같은 과정을 거칩니다.

- IBM Optim은 데이터베이스 기록들을 자동으로 아카이브합니다. Optim은 TSM API를 이용하여 데이터를 TSM 스토리지 풀 계층에 저장/아카이브합니다.
- 여타 파일 기반 백업/복구 제품과 같이, Optim은 표준 파일 조작을 이용하여 디스크 기반 파일 시스템에 저장/아카이브 하며, 백업 제품은 지원되는 백업 미디어로 파일을 백업할 수 있습니다.
- Optim과 TSM API 통합으로 인한 장점은 무엇일까요? Optim은 테이프에 직접 데이터를 아카이브하거나 TSM이 자동으로 아카이브된 데이터를 테이프로 이동하도록 할 수 있습니다. Optim 아카이브 데이터는 국부 가용성, DR 또는 원격지 불팅을 위하여 TSM이 관리할 수 있는 TSM 스토리지 풀로 점증 백업됩니다.

## 단계 4. 남은 데이터의 중복제거 및 압축

앞서 언급하였듯이, 데이터 중복제거는 데이터 스토리지 용량 요구사항을 축소하기 위한 신기술입니다. 일부 중복 데이터는 여러 소스 시스템 전반에 존재하지만, 대부분의 중복 데이터는 반복적인 전체 백업 세트의 저장으로 인하여 유발됩니다. 주기적인 전체 백업을 요하는 백업 솔루션을 사용하는 경우, 중복제거는 관리하는 데이터 양을 줄이는 데에 매우 효과적인 것입니다. 그렇지 않다면, 중복제거로부터 기대할 수 있는 축소율은 사실상 환경에 달려 있을 것입니다.

데이터 중복제거의 기본 취지는 어떤 데이터 객체든 단 하나의 사본만을 보관하고, 중복 데이터가 삭제되는 단일 사본에 포인터를 두는 것입니다. 일부 솔루션들은 이를 파일 레벨에서 수행하는데 예를 들어, 파일들은 삭제될 것과 정확히 동일시되어야 합니다. 이를 흔히 싱글 인스턴스 스토리지 또는 SIS라고 합니다.

여타 솔루션들은 고정 또는 가변 블록 길이로 중복 데이터를 제거합니다. IBM의 솔루션들은 데이터 사이즈를 기반으로 한 혼합 방식을 이용합니다. 작은 파일의 경우 파일 기반으로, 큰 파일의 경우 가변 블록(또는 서브파일)을 기반으로 합니다.

어떠한 상황에서도, 중복제거 프로그램은 선택된 데이터에 대하여 체크섬 알고리즘을 실행하여 시그니처(Signature)를 생성한 다음, 그 시그니처가 이전에 나타났었는지 확인합니다. 나타났을 경우, 그 데이터를 폐기하고 이미 저장된 데이터에 대한 포인터를 그 자리에 둡니다.

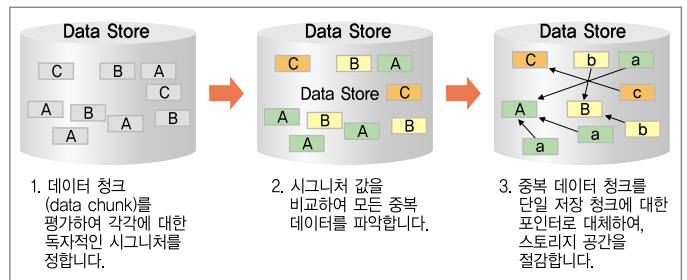


그림 5: 데이터 중복제거



데이터 중복제거는 데이터 생성 및 관리 수명주기의 여러 시점에서 이루어질 수 있습니다. 일반적으로, 이러한 중복제거 시점들을 데이터가 생성되는 '소스 측' 과 데이터가 저장되고 관리되는 '타겟 측'으로 세분화할 수 있습니다. 일례로, Microsoft Exchange와 같은 이메일 애플리케이션은 일정 기간 동안 소스 측 중복제거를 실시하여, Exchange 데이터베이스에 단 하나의 첨부물 사본만을 저장합니다. 백업 애플리케이션도 LAN이나 WAN을 통하여 이전에 백업된 데이터를 전송하지 않음으로써 소스 측 중복제거를 실시하여, 대역폭을 절감할 수 있습니다.

타겟 측의 경우, 가장 널리 쓰이고 있는 중복제거는 가상 테이프 라이브러리 (Virtual Tape Libraries) 또는 VTL에서입니다. 이 디스크 기반 상자는 테이프 라이브러리 및 드라이브를 에뮬레이션 합니다. 또한 중복제거를 실시하여 비용 효과적으로 디스크 상에 동등한 양의 데이터를 저장함과 동시에 테이프보다 우월한 성능을 제공합니다. 테이프 기반 시스템 상에서 중복제거를 실시하는 것은 테이프의 휴대형 특성과 시간이 지남에 따라 재생해야 한다는 점을 감안할 때, 그리 좋은 아이디어는 아닙니다. 모든 포인터에 대하여 원래의 데이터를 유지하는 것 또한 보장하기 어렵습니다. Tivoli 스토리지 매니저 6과 같은 백업 애플리케이션은 그 디스크 기반 리포지토리에 대해서도 데이터 중복제거를 가합니다. 그리고 이제는 주 스토리지 시스템 상에서도 중복 데이터를 제거하는 솔루션을 찾아볼 수 있게 되었습니다.

현재, IBM은 두 가지의 강력한 데이터 중복제거 솔루션을 제공하고 있습니다. TSM 6 익스텐디드 에디션에서는, 여러 프로덕션 시스템으로부터 백업된 데이터 중복을 효과적으로 제거할 수 있습니다. 그러나 이 솔루션은 TSM 서버 상에서만 작동되므로, TSM 서버 전반에 걸친 중복제거는 이루어지지 않습니다.

IBM의 또 다른 솔루션은 IBM 시스템 스토리지 ProtecTIER 제품군의 중복제거 시스템입니다. ProtecTIER는 여러 TSM 서버, 여타 백업 시스템으로부터의 전체 백업, 또는 아카이브 소프트웨어 솔루션 등과 같은 광범위한 소스로부터의 데이터 정리에 유용한 동종 최고의 엔터프라이즈급 솔루션입니다.

- 다음 경우에 IBM Tivoli 스토리지 매니저 6 익스텐디드 에디션에 탑재되어 있는 중복 제거 기능을 사용합니다. TSM 서버가 한 대일 경우, 디스크 상에 백업 데이터를 더 많이 저장하여 TSM 복구 시간을 단축하고자 할 경우, 또는 여러 TSM 서버가 보호하는 시스템 전반에 대하여 중복 데이터의 양이 많지 않을 경우.
- 다음 경우에 IBM 시스템 스토리지 ProtecTIER TS7650 중복제거 솔루션을 사용합니다. TSM 서버가 여러 대일 경우, 다른 소스의 백업이 아카이브 데이터가 있을 경우, 또는 주기적인 전체 백업을 실시하는 여타 백업 제품을 사용하고 있는 경우.

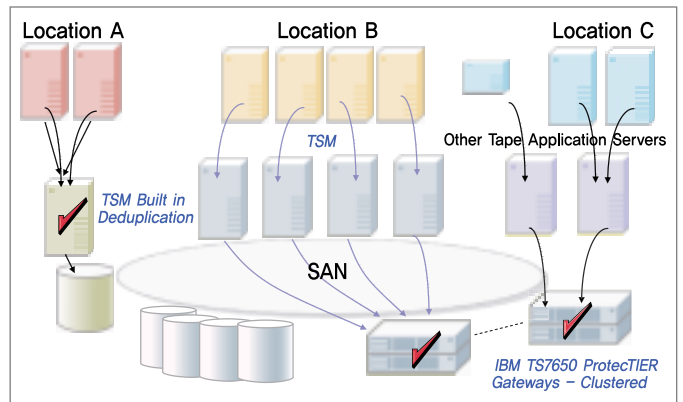


그림 6: 중복제거는 여러 위치에서 이루어질 수 있습니다.



#### 단계 4a. Tivoli 스토리지 매니저 6를 이용한 남은 데이터의 중복제거 및 압축

Tivoli 스토리지 매니저 6의 중복제거 기능은 다음과 같습니다. 익스텐디드 에디션 고객에게는 이 기능에 대한 추가 라이선스 비용이 부과되지 않습니다. 이 솔루션은 테이프 대신에 디스크에 더 많은 백업 데이터를 저장하여 복구 시간을 단축하는데 유용합니다. 일반적인 백업, TSM API를 통하여 임포트된 데이터와 아카이브 및 HSM 데이터 등, 모든 소스로부터의 데이터를 처리할 수 있습니다. TSM은 사후 프로세스로서 디스크 기반 데이터 풀의 중복을 제거하므로, 백업 성능에 전혀 영향을 주지 않습니다. 실행 후에는, 여유 공간이 늘어난 스토리지를 자동으로 회수합니다.

TSM은 중복 데이터의 가장 일반적인 원인 전체 백업을 이미 배제시켰기 때문에 TSM의 중복 제거 솔루션에서 기대할 수 있는 압축률은 평균적으로 그다지 높지 않은 약 40% 수준입니다. 그러나 점진적 증분 백업 방식과 내장된 데이터 압축 기능을 결합하면, TSM의 유효 데이터 압축률은 다음 사이트의 ESG의 보고서에 상술되어 있듯이, 출시되어 있는 여느 솔루션보다 극히 우수합니다.  
<http://www.servicemanagementcenter.com/main/pages/IBMRBMS/OMSA/ShowCollateral.aspx?oid=55497>

테이프 회수(Tape Reclamation)는 지능적인 테이프 이용을 통하여 TSM이 ROI를 증대시키는 또 다른 방법입니다. 시간이 지나면 테이프가 만료됩니다. 미디어 첨부만이 가능하기 때문에, 재사용할 수 없는 상당 부분의 “쓸모 없는” 공간을 남깁니다. 테이프 상의 모든 데이터가 만료되기까지 기다리는 대신, TSM 내에 정책을 설정할 수 있습니다. 이는 쓸모 없는 공간의 비율을 기준으로 나머지 유효 데이터를 다른 테이프로 이동시키고 그 유효 데이터를 통합하여 다른 테이프를 재사용할 수 있도록 여유공간을 확보하는 것입니다. TSM은 오프사이트 볼륨을 자동으로 회수한다는 점에서도 독특합니다. 오프사이트 테이프의 새로운 세트를 생성하고, 이전 세트를 회수하여 재사용합니다.

##### 테이프 회수의 특징과 장점

- 테이프 이용률을 높여, 비용을 절감합니다.
- 테이프 이용률을 지속적으로 모니터링 합니다.
- 사용자 정의 회수 쓰레시홀드
- 특정 시간이나 여유 공간이 쓰레시홀드 값에 달했을 때 회수가 이루어지도록 지정할 수 있습니다.
- 오리지널 테이프를 스크래치 풀(Scratch Pool)로 반환합니다.

#### 단계 4b. IBM 시스템 스토리지 ProtecTIER® TS7650을 이용한 남은 데이터의 중복제거

IBM 시스템 스토리지 ProtecTIER는 성능, 확장성, 데이터 무결성 및 신뢰성에서 가장 앞선 기술입니다. 비교 대상 간의 동등한 비교에서, 이 솔루션은 실제 고객 환경에서 출시된 제품 중 가장 빠릅니다. 단일 ProtecTIER 시스템은 성능(1000MB/sec)과 용량(1PB) 모두에 있어 간편한 확장이 가능합니다. ProtecTIER는 해쉬 알고리즘에 의존하지 않고 바이트 레벨 차등 백업을 실시하여 데이터 중복을 엔터프라이즈급 데이터 무결성으로 유지하는, 소수의 솔루션 중 하나입니다. ProtecTIER는 경쟁사 제품에서 흔히 볼 수 있는 OEM 부품이 아닌, 모든 IBM의 동급 최고 컴포넌트들을 갖추고 있습니다.

ProtecTIER는 대규모의 프로덕션 환경에서 성능이 입증되었으며 전세계적으로 IBM 서비스 조직의 지원을 받고 있습니다.

그 외에도, TS7650은 전세계 여느 벤더들에 비해 더 많은 용량(25PB)이 ProtecTIER 프로덕션 서버 뒤에 배치되어, 프로덕션에 있어 입증되었습니다.

TS7650 ProtecTIER 중복제거 제품군은 소규모(7TB)에서 중규모(18TB) 및 대규모(36TB) 어플라이언스에 이르기까지 다양합니다. TS7650G Gateway 제품은 선택에 따라 1PB까지 스토리지를 추가할 수 있도록 합니다. 또한, 활성-활성(Active-Active) 클러스터 구성은 고가용성 기능을 제공합니다.

#### 단계 4c. 검토 - 데이터 중복제거를 위한 TSM 또는 ProtecTIER

TSM이 ProtecTIER 환경에서 제 기능을 발휘하지만, TSM 중복제거와 ProtecTIER 중복제거 모두를 사용하지 않을 것입니다. 추가적인 편익 없이 두 배의 수고만 들기 때문입니다. 그렇다면, 둘 중 하나를 선택해야 한다면?

두 솔루션 모두 타겟 측 중복제거의 편익을 제공합니다. 스토리지 용량 요구사항이 크게 축소됩니다(특히 TSM의 점진적 증분 백업을 사용할 경우). 운영 비용, 에너지 사용량 및 총소유비용이 저감될 것입니다. 또한 디스크 상에 데이터가 더 많아도 복구 시간이 단축됩니다.

다음 경우에 TSM 6 내장 중복제거를 사용하십시오: 중복제거 작업을 TSM 내에 완전히 통합시키고자 하는 경우, 별도의 하드웨어 및 소프트웨어 비용을 들이지 않고 중복제거의 편익을 원하는 경우, 이는 TSM 6 Extended Edition에 무상으로 탑재되어 있습니다. 또는 최소한의 데이터 보관 요구사항으로 단대단(End to End) 데이터 수명주기 관리를 원하는 경우

다음 경우에 ProtecTIER를 사용하십시오.

- 1000 MB/초 이상의 고성능을 요할 경우
- 대량의 데이터가 있어 확장형 용량과 성능을 요할 경우
- 사후 프로세싱으로 인한 영향을 피하기 위하여 인라인(inline) 중복제거가 필요할 경우
- 여러 TSM(또는 여타 백업) 서버에 대한 중복제거가 필요한 경우
- TSM이 없고 주간 단위로 전체 백업을 실시하는 경우

#### 왜 IBM인가?

- IBM은 포괄적인 데이터 정리 기술을 갖춘 유일한 벤더입니다.
- IBM 데이터 정리 솔루션의 광범위한 포트폴리오는 가장 효과적인 기술로 고객의 문제를 해결할 수 있는 자유를 제공합니다.
- IBM의 고품질 글로벌 지원 서비스는 데이터 정리에 대한 투자가 고객의 니즈에 부응할 수 있도록 할 것입니다.
- IBM은 지속적으로 연구개발에 투자하여 고객들이 요구하는 보다 선도적인 고급 기능을 제공하고 있습니다.

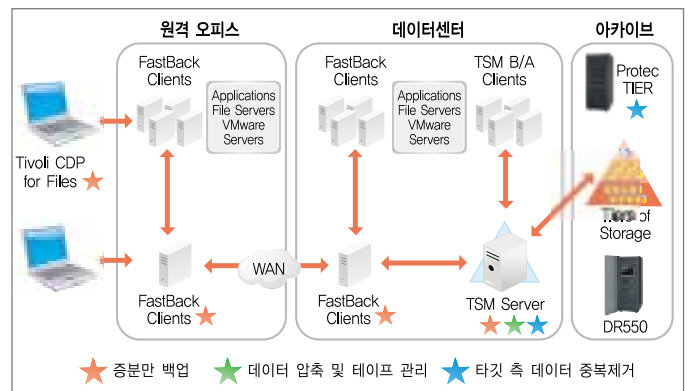


그림 7: IBM의 포괄적인 데이터 정리 솔루션 포트폴리오

현재 IBM은 Tivoli 스토리지 매니저, TSM 패스트백, Tivoli CDP 포 파일즈 등으로 구성되어 있는, 자사의 통합 복구 관리 포트폴리오에 대하여 증분 백업만을 이용하는 데이터 정리 솔루션을 제공하고 있습니다. 확장 기업 내에서 전체 백업을 다시 실시할 필요가 없을 것입니다.

Tivoli Storage Manager는 또한 동급 최고의 테이프 관리와 데이터 압축 기능을 제공하여 백업 및 아카이브 스토리지 요구사항을 한층 축소합니다.

TS7650 ProtecTIER 제품군의 어플라이언스 및 게이트웨이와 TSM 6에 타겟 측 중복제거가 내장되어 있습니다.

## 다음 단계:

- IBM Global Technology Services와 IBM Business Partner는 고객사의 현황 분석을 도우며, 다음 단계를 권고해드립니다.
- IBM은 운용에 있어 가장 비용 효과적인 데이터 정리 기술을 식별하는데 도움을 드립니다.
- IBM의 포괄적인 BVA(Business Value Analyst) 도구를 이용하는 포괄적인 ROI 분석을 요청하십시오.

## 추가 정보:

IBM은 혼합 벤더 환경을 위한 광범위한 스토리지 관리 솔루션을 제공하는 유일한 벤더입니다. 우리는 고객의 변화하는 비즈니스 요구사항에 부응하여 확장할 수 있는 다이나믹 스토리지 인프라스트럭처를 구현하는데 도움을 드릴 수 있습니다. IBM Tivoli 스토리지 관리 솔루션이 자사 환경의 데이터 관리 문제 해결에 미치는 효과에 대해 자세히 알아보시려면, 해당 IBM 영업 대표 또는 IBM Business Partner에 문의하시거나, 다음 사이트를 방문하십시오.

[www.ibm.com/software/tivoli/solutions/storage](http://www.ibm.com/software/tivoli/solutions/storage).

다음 사이트에서 추가 정보를 입수하실 수 있습니다.

- <http://www.ibm.com/software/tivoli/solutions/storage/>
- <http://www.ibm.com/software/tivoli/products/storage-mgr-extended/>
- <http://www.ibm.com/systems/storage/tape/ts7530/index.html>

## IBM Tivoli 소프트웨어

Tivoli 소프트웨어는 고객이 자사 비즈니스의 운용을 파악하고 이해하는 가시성, 자사 비즈니스를 효과적으로 관리하여 리스크를 줄이고 자사 브랜드를 보호하는 통제, 자사 비즈니스를 최적화하여 운영 비용을 절감하고 신규 서비스를 보다 신속하게 제공하는 자동화를 통해 보다 높은 품질의 서비스를 제공할 수 있는 서비스 관리 플랫폼을 제공합니다. IT 중심의 서비스 관리와 달리, Tivoli 소프트웨어는 비즈니스 및 기술 요구사항 모두를 관리, 통합하고, 연계시키기 위한 공통 기반을 제공합니다. Tivoli 소프트웨어는 기업의 우선적인 서비스 관리 니즈에 신속히 대처하고 변화하는 비즈니스 요구에 사전 대응할 수 있도록 고안되어 있습니다. Tivoli 포트폴리오는 세계적 수준의 IBM 서비스, IBM 지원 조직 및 활발한 IBM 협력사들간의 생태계가 뒷받침하고 있습니다. Tivoli 고객들과 협력사 임직원들은 전세계에서 독립적으로 운영되고 있는 IBM Tivoli User Group([www.tivoli-ug.org](http://www.tivoli-ug.org), 국내의 경우 [cafe.naver.com/itug](http://cafe.naver.com/itug))에 참가하여 서로 간의 베스트 프랙티스를 활용할 수 있습니다.

또한, IBM 글로벌 파이낸싱은 고객사의 특정 IT 니즈에 적합한 맞춤형 파이낸싱 솔루션을 제공해드릴 수 있습니다. 낮은 금리와 유연한 상환 방법 및 대출, 그리고 자산 재매입과 처분에 대한 자세한 사항은 다음 사이트를 참조하십시오.

[www.ibm.com/financing](http://www.ibm.com/financing)



© Copyright IBM Corporation 2010

(135-270) 서울시 강남구 도곡동 467-12  
군인공제회관빌딩

한국아이비엠주식회사  
고객만족센터

TEL: (02)3781-7114  
[www.ibm.com/kr](http://www.ibm.com/kr)

2010년 6월

All Rights Reserved

IBM, IBM 로고, [ibm.com](http://ibm.com) 및 Tivoli는 미국 및/또는 다른 국가에서 사용되는 International Business Machines Corporation의 상표 또는 등록상표입니다. 상기 및 기타 IBM 상표로 등록된 용어가 본 서에 처음 나올 때 상표 기호(® or ™)와 함께 표시되어 있을 경우, 이러한 기호는 본 서가 출판된 시점에 IBM이 소유한 미국 등록 상표이거나 관습법에 의해 인정되는 상표임을 나타냅니다. 이런 상표는 다른 국가에서도 등록되어 있거나 관습법적인 상표일 수 있습니다. IBM의 최신 상표 목록은 [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml) 웹 페이지의 "저작권 및 상표 정보" 부분에서 확인할 수 있습니다.

Microsoft와 Windows는 미국 및/또는 다른 국가에서 사용되는 Microsoft Corporation의 상표입니다.

기타 다른 회사, 제품 및 서비스 이름은 다른 기업의 상표 또는 서비스 마크일 수 있습니다.

IBM 제품 및 서비스에 대하여 이 자료에서 참조한 내용은 IBM이 비즈니스를 전개하는 모든 국가에 해당되지 않습니다.

IBM Corporation의 서면 동의 없이는, 어떠한 형태로든 본 서의 일부를 전체 또는 전달할 수 없습니다.

제품 데이터는 본 서의 최초 출간 시점을 기준으로 정확성에 대한 검토를 마쳤습니다. 제품 데이터는 사전 고지 없이 변경될 수 있습니다. IBM의 향후 방향 및 의도와 관련하여 모든 언급은 사전고지 없이 변경되거나 취소될 수 있으며, 목적과 목표만을 나타냅니다.

여기 포함된 정보는 품질 보증과 관련된 어떠한 암시도 하고 있지 않으며, "있는 그대로" 배포되고 있습니다. IBM은 시장성, 특정 목적에 대한 적합성 또는 비침해에 대한 모든 종류의 보장을 명시적으로 거부합니다. IBM 제품은 IBM 고객 협약, 제한적 품질보증서, 국제 프로그램 라이선스 협약 등 관련 계약 조건에 따라 품질 보증을 받게 되어 있습니다.

고객은 법률 요구사항을 반드시 준수해야 합니다. 고객의 비즈니스에 영향을 미칠 수 있는 관련 법규 및 규제 사항 그리고 해당 법규나 규제 사항을 준수하기 위해서 고객이 취해야 하는 행위에 대한 식별과 해석에 대해서 능력 있는 법조인의 조언을 구하는 것은 전적으로 고객의 책임입니다. IBM은 법적 자문을 제공하지 않으며, IBM의 서비스 또는 제품을 통해 고객이 해당 법률 및 규제를 준수할 수 있다는 보증이나 표현을 일체 하지 않습니다