

Dell PowerMax: 데이터 감소

인라인 압축 및 중복 제거

2022년 7월

H19254

백서

개요

PowerMax 스토리지 플랫폼은 인라인 압축 및 중복 제거와 같은 여러 데이터 감소 기술을 제공합니다. 또한 패턴 탐지와 효율적인 데이터 배치를 통해 성능과 효율성을 매우 균형 있게 조정합니다.

Dell Technologies

저작권

본 발행물의 정보는 "있는 그대로" 제공됩니다. Dell Inc.는 본 발행물의 정보와 관련하여 어떠한 진술이나 보증도 하지 않으며, 특히 상품성이나 특정 목적을 위한 적합성에 대하여 어떠한 묵시적인 보증도 부인합니다.

본 간행물에 기술된 일체의 소프트웨어를 사용, 복사, 배포하려면 해당 소프트웨어 라이선스가 필요합니다.

Copyright © 2022 Dell Inc. or its subsidiaries. All Rights Reserved. Dell Technologies, Dell, EMC, Dell EMC 및 기타 상표는 Dell Inc. 또는 해당 자회사의 상표입니다. 인텔, 인텔 로고, Intel Inside 로고 및 제온은 미국 및/또는 기타 국가에서 인텔의 상표입니다. 기타 모든 상표는 해당 소유주의 상표일 수 있습니다. 2022년 7월 미국에서 발행됨 H19254.

Dell Inc.는 본 문서의 정보가 해당 발행일 현재 정확한 것으로 간주합니다. 모든 정보는 예고 없이 변경될 수 있습니다.

목차

| | |
|--------------------|----|
| 핵심 요약 | 4 |
| 데이터 감소..... | 5 |
| 시스템 리소스 활용도 | 9 |
| 관리 및 모니터링 | 9 |
| 지원되는 데이터 서비스 | 17 |
| 결론..... | 18 |
| 참고 자료 | 19 |

핵심 요약

개요

DELL PowerMax를 통한 데이터 감소는 인라인 압축, 인라인 중복 제거 및 패턴 탐지를 결합하여 시스템 효율성을 향상합니다. 이러한 데이터 감소 기술을 사용하면 사용자가 용량을 크게 절감할 수 있습니다. 데이터 감소로 데이터를 압축하고 중복된 데이터 복제본을 제거합니다. 이 백서에서는 PowerMax 시스템의 데이터 감소 기능이 어떻게 작동하는지 설명하고 Unisphere for PowerMax, Solutions Enabler 및 Mainframe Enabler 소프트웨어와 같은 Dell 관리 애플리케이션을 사용하여 보고하는 방법을 설명합니다.

개정

| 날짜 | 설명 |
|----------|--------|
| 7월 2022년 | 최초 릴리스 |

Dell은 여러분의 의견을 소중히 생각합니다

Dell Technologies와 본 문서의 작성자들은 본 문서 자료에 대한 여러분의 피드백을 기다리고 있습니다. [이메일](#)로 Dell Technologies 팀에 문의하십시오.

작성자: Robert Tasker

참고: 이 주제에 대한 다른 설명서에 대한 링크는 [PowerMax 및 VMAX Info Hub](#)를 참조하시기 바랍니다.

데이터 감소

개요

데이터 감소는 인라인 압축, 인라인 중복 제거, 패턴 탐지, 효율적인 데이터 배치, ML(Machine Learning)을 결합합니다. 이 결합은 엔터프라이즈 스토리지 시스템에서 예상되는 성능을 달성하는 동시에 사용 가능한 총 물리적 용량보다 더 많은 호스트 데이터를 쓸 수 있는 시스템을 생성합니다. 이 기능은 기본적으로 켜져 있으며 스토리지 그룹 수준에서 활성화 또는 비활성화할 수 있습니다. 또한 PowerMax 2500 및 8500 시스템에서 사용할 수 있는 모든 데이터 서비스를 지원합니다. 이 지원 기능은 CKD 에뮬레이션에도 적용되지만 CKD에 대한 중복 제거는 포함하지 않습니다.

압축은 데이터 크기를 줄이고 중복 제거는 데이터를 단일 인스턴스로 저장합니다. 패턴 탐지에는 연속적인 0의 문자열이 압축된 데이터의 일부로 저장되지 않도록 하는 non-zero 할당 함수가 포함됩니다. 압축, 중복 제거 및 패턴 탐지는 시스템 내에 통합된 하드웨어 지원 기능을 사용하여 수행되며 이러한 기능 수행 중의 오버헤드를 줄입니다. 머신 러닝은 디스크에 저장된 가장 사용량이 많은 데이터를 식별하고 최적의 성능을 위해 감소되지 않도록 합니다. 효율적인 데이터 배치는 **밀집화(compaction)** 기능을 사용하여 데이터를 전략적으로 저장하여 낭비되는 공간을 최소화하므로 가비지 수집 또는 조각 모음 기능의 필요성을 줄입니다.

ABR(Activity Based Reduction)

ABR(Activity Based Reduction)은 자주 액세스하는 데이터를 압축 해제하여 발생하는 성능 비용을 줄입니다. 이 기능을 사용하면 가장 사용량이 많은 데이터의 최대 20%를 압축 해제된 시스템에 저장할 수 있습니다. 이 기능은 자주 액세스하는 데이터를 지속적으로 압축 해제하여 발생하는 성능 레이턴시를 최소화하므로 시스템에 이점을 제공합니다. 가장 사용량이 많은 데이터를 결정하기 위해 시스템은 IO 통계를 처리하는 ML 알고리즘을 사용합니다. 이 작업을 수행하면 데이터 감소 절감 효과와 성능 모두를 위해 균형 잡힌 최적의 환경이 유지됩니다.

Compression

압축은 수신되는 쓰기 워크로드를 최소 용량만을 소비하는 가장 작은 크기로 줄여줍니다. 데이터는 GZIP 압축 알고리즘을 사용하는 데이터 감소 하드웨어를 통과하면서 압축됩니다. 데이터 감소 하드웨어를 통과하면서 데이터는 효율성을 극대화하기 위해 병렬 압축된 4개의 섹션으로 나뉩니다. 4개 섹션의 합계는 디스크에 저장된 데이터가 최종적으로 축소된 크기입니다. 이 기능은 감소된 데이터에 대해 세분화된 액세스를 제공합니다. 각 섹션을 독립적으로 처리할 수 있으므로 부분적인 읽기 또는 쓰기 요청에 대해 요청된 데이터가 포함된 섹션만 처리됩니다.

중복 제거

중복 제거는 동일한 데이터 복제본을 식별하고 각 복제본의 단일 인스턴스를 저장하는 용량 절감 방법입니다. 중복 제거가 효율적인 용량 절감을 제공하기 위해 필요한 몇 가지 요소가 있습니다.

- **해시 ID:** 해시 ID는 중복 제거 관계가 필요한지 여부를 확인하는 데 사용되는 수신되는 데이터의 고유 식별자입니다. 시스템은 SHA-256 알고리즘을 사용하여 해시 ID를 생성합니다.
- **해시 ID 테이블:** 해시 테이블은 시스템 디렉터 간에 분배된 시스템 메모리의 할당입니다. 이 표는 중복 제거 프로세스에서 사용된 해시 ID를 카탈로그로 작성합니다. 표의 항목은 중복 제거 관계가 존재하는지 또는 새 항목이 필요하고 데이터를 디스크에 저장할 수 있는지 파악하는 데 사용됩니다.
- **DMO(Dedupe Management Object):** DMO는 중복 제거 관계가 존재할 때만 존재하는 시스템 메모리 내의 64바이트 개체입니다. 이러한 개체는 어레이에서 백엔드 용량을

사용하는 중복 제거된 데이터와 프런트엔드 디바이스 간의 포인터를 저장하고 관리합니다. DMO는 프런트엔드 디바이스 및 디스크에 저장된 데이터 간의 중복된 데이터에 대한 포인터를 관리합니다. 또한 중복 제거 관계가 존재할 때 해시 ID를 어떤 해시 테이블에 저장할지를 관리합니다.

중복 제거는 압축의 경우와 동일한 데이터 감소 하드웨어를 사용하여 수행되며, 하드웨어에서 데이터를 처리할 때 고유한 해시 ID가 생성됩니다. 그런 다음 해시 ID는 동일한 ID를 찾는 해시 ID 테이블과 비교됩니다. 일치하는 데이터가 발견되면 데이터가 디스크에 저장되지 않고 중복 제거 공유가 생성됩니다. 포인터는 해시 ID 테이블의 프런트엔드 볼륨과 고유 ID 간에 설정됩니다. 포인터는 디스크에 저장된 데이터의 단일 인스턴스를 볼륨에 연결하여 향후 데이터에 대한 액세스를 제공합니다. DMO는 데이터, 데이터에 액세스하는 프런트엔드 볼륨 그리고 해시 ID 테이블 간의 포인터를 관리합니다. 해시 ID 테이블에 일치하는 항목이 없으면 향후 해시 ID 비교를 위해 새로운 항목이 추가됩니다.

중복 제거 알고리즘

PowerMax 시스템은 데이터 감소 하드웨어에 구현된 SHA-256 해싱 알고리즘을 사용하여 중복 데이터를 찾습니다. 그러면 여러 소스에서 공유할 수 있는 단일 인스턴스로 데이터가 저장됩니다. 이 프로세스는 데이터 무결성을 유지하면서 향상된 데이터 효율성을 제공합니다.

SHA-256 알고리즘은 각 32KB 데이터 블록에 대해 32바이트 코드를 생성합니다. 하루에 5%씩 업데이트되는 쓰기 데이터가 1PB에 달하는 시스템인 경우 100만 년간 운영 시 해시 충돌이 발생할 가능성은 20%입니다. 각 128KB 트랙이 4개의 32KB 블록으로 처리되므로 동일한 128KB 트랙의 4개 블록 모두에서 해시 충돌이 있어야만 실제 해시 충돌이 발생할 수 있습니다. 4개 블록 모두에서 모두 충돌이 발생하는 일은 이론상으로는 가능합니다(1조 년간 운영해도 1% 미만). 또한 중복 제거의 비교 단계에서 일치하는 항목이 발견되면 바이트 대 바이트 비교가 수행됩니다. 이 비교는 테이블을 업데이트하고 데이터에 대한 액세스를 허용하도록 포인터를 설정하기 전에 일치하는 항목이 있는지 확인하기 위해 수행됩니다.

밀집화 (compaction)

데이터 배치는 밀집화(compaction) 프로세스를 이용하여 수행됩니다. 밀집화(compaction)는 디스크에서 감소되었거나 감소되지 않은 데이터를 사용 가능한 최상의 위치에 직관적으로 배치합니다. 디스크에 데이터를 저장하는 작업에는 쓰기 개체를 사용합니다. 각 개체는 시스템에 구성된 드라이브 전반에 걸쳐 인접해 있는 6MB의 백엔드 데이터 디바이스 용량입니다. 쓰기 개체는 1K 경계에 정렬되며 단일 용도로 순차적으로 사용됩니다. 쓰기 개체는 쓰기를 최적화하기 위해 지원 대상인 모든 RAID 유형에 대해 전체 스트라이프에 분산됩니다. 각 개체는 FBA 및 CKD 에뮬레이션 모두에 대해 감소 또는 감소되지 않은 데이터를 지원합니다.

- **FBA 쓰기 개체:** 감소되지 않은 쓰기 개체는 48개의 FBA 트랙으로 구성됩니다. 감소된 쓰기 개체는 1,000개의 감소된 트랙으로 구성됩니다. 쓰기 개체에 대한 항목 감소 범위는 1KB~96KB입니다.
- **CKD 쓰기 개체:** 감소되지 않은 쓰기 개체는 108개의 CKD 트랙으로 구성됩니다. 감소된 쓰기 개체는 1,000개의 감소된 트랙으로 구성됩니다. 쓰기 개체에 대한 항목 감소 범위는 1KB~52KB입니다.

EDC(Extended Data Compression)

PowerMax 2500 및 8500 시스템에는 이미 압축된 데이터를 압축하여 용량을 추가로 절감하는 EDC(extended data compression)라는 추가 기능이 포함되어 있습니다. 이 작업은 오랜 시간 동안 액세스되지 않은 데이터를 확인하여 실행됩니다. 데이터를 EDC 후보로 만드는 요소는 아래와 같습니다.

- 데이터는 데이터 감소를 지원하는 스토리지 그룹에 속합니다.
- 지난 30일 동안 데이터에 액세스하지 않았습니다.
- 데이터가 아직 EDC를 통해 압축되지 않았습니다.

EDC에 적합한 데이터는 데이터를 저장하는 데 사용되는 용량을 더욱 줄이기 위해 Def9_128_SW 알고리즘을 사용하여 압축됩니다. 이 프로세스는 시스템 내에서 자동화된 백그라운드 프로세스입니다. 추가 절약은 다음과 같습니다. 스토리지 그룹에 포함됩니다. 수준 달성 압축 비율. EDC는 PowerMax 스토리지 어레이에서만 사용할 수 있습니다.

CKD 압축

ABR(Activity Based Reduction)은 자주 액세스하는 데이터를 압축 해제하여 발생하는 성능 비용을 줄입니다. 이 기능을 사용하면 가장 사용량이 많은 데이터의 최대 20%를 압축 해제된 시스템에 저장할 수 있습니다. 이렇게 하면 자주 액세스하는 데이터를 지속적으로 압축 해제하여 발생하는 부정적인 성능의 영향을 제거하므로 시스템에 이점이 있습니다. 데이터 사용량 수준을 파악하기 위해 시스템은 수신되는 I/O에서 수집된 통계를 프론트엔드 디바이스로 처리하는 ML 알고리즘을 사용합니다. 이 작업으로 데이터 감소 효과와 성능을 동시에 얻을 수 있도록 시스템 리소스 간 균형을 유지하여 최적의 환경을 유지할 수 있습니다.

압축은 수신되는 쓰기 워크로드를 가능한 가장 작은 크기로 줄여 최소 용량을 사용합니다. 데이터는 GZIP 압축 알고리즘을 사용하는 시스템에 내장된 데이터 감소 하드웨어를 통과하면서 압축됩니다. 데이터 감소 하드웨어를 통과하면서 데이터는 하드웨어의 효율성을 극대화하기 위해 병렬 압축된 4개의 섹션으로 나뉩니다. 4개 섹션의 합계는 디스크에 저장된 데이터가 최종적으로 축소된 크기입니다. 이렇게 하면 부분 읽기 또는 쓰기 요청이 있을 때 감소된 데이터에 대한 세분화된 액세스가 제공됩니다. 각 섹션을 독립적으로 처리할 수 있으므로 요청된 데이터가 포함된 섹션만 처리됩니다.

데이터 배치는 밀집화(compaction)라는 데이터 배치 프로세스를 사용하여 수행됩니다. 밀집화(compaction)로 디스크에서 감소되었거나 감소되지 않은 데이터가 사용 가능한 최상의 위치에 직관적으로 배치됩니다. 디스크에 데이터를 저장하는 작업에는 쓰기 개체를 사용합니다. 각 객체는 시스템에 구성된 드라이브 전반에 걸쳐 인접해 있는 6MB의 백엔드 데이터 디바이스 용량입니다. 쓰기 개체는 1K 경계에 정렬되며 순차적으로 단일 용도로 사용됩니다. 쓰기 개체는 쓰기 최적화를 위해 지원되는 모든 RAID 유형에 대해 전체 스트라이프에 분산됩니다. 각 개체는 감소된 또는 감소되지 않은 데이터를 지원합니다. 감소되지 않은 쓰기 개체는 108개의 CKD 트랙으로 구성됩니다. 감소된 쓰기 개체는 1,000개의 감소된 트랙으로 구성됩니다. 쓰기 개체에 대한 항목 감소 범위는 1KB~52KB입니다.

데이터 감소 I/O 흐름

모든 I/O는 캐시를 통과한 다음 시스템에서 처리됩니다. 데이터 감소 작업은 시스템에서 데이터를 받은 후 디스크에 배치되기 전에 이루어집니다. 인라인 프로세스를 사용하려면 데이터 감소가 적용되는 I/O 흐름 내에서 추가적인 검사가 필요합니다. 시스템은 이러한 검사를 사용하여 수신되는 데이터가 데이터 감소 하드웨어를 통과해야 하는지 확인합니다. 데이터 감소가 활성화된 스토리지 그룹의 수신되는 데이터는 데이터 감소 흐름을 따릅니다. 그러나 활동 기반 감소 ABR(Activity-Based Reduction)로 인해 데이터 감소가 활성화된 스토리지 그룹의 활성 데이터는 성능 최적화를

위해 데이터 감소 흐름을 건너뛰니다. ABR로 인해 압축되지 않은 데이터는 나중에 압축하여 압축 풀로 이동할 수 있습니다. 데이터 감소가 비활성화된 스토리지 그룹의 데이터는 데이터 감소 흐름을 무시하여 감소되지 않은 상태로 시스템에 기록됩니다.

읽기, 쓰기, 쓰기-업데이트 등 고려해야 할 몇 가지 다른 I/O 유형이 있습니다.

- **읽기:** 이미 어레이를 채우고 있는 데이터에 액세스하기 위한 요청입니다.
- **쓰기:** 디스크 공간을 사용하는 수신되는 I/O입니다.
- **쓰기-업데이트:** 어레이의 디스크 공간에 할당된 데이터를 변경할 수 있는 수신되는 I/O입니다.

아래 그림은 데이터 세트 또는 관련 스토리지 그룹의 특성에 따라 결정되는 I/O의 경로를 설명합니다.

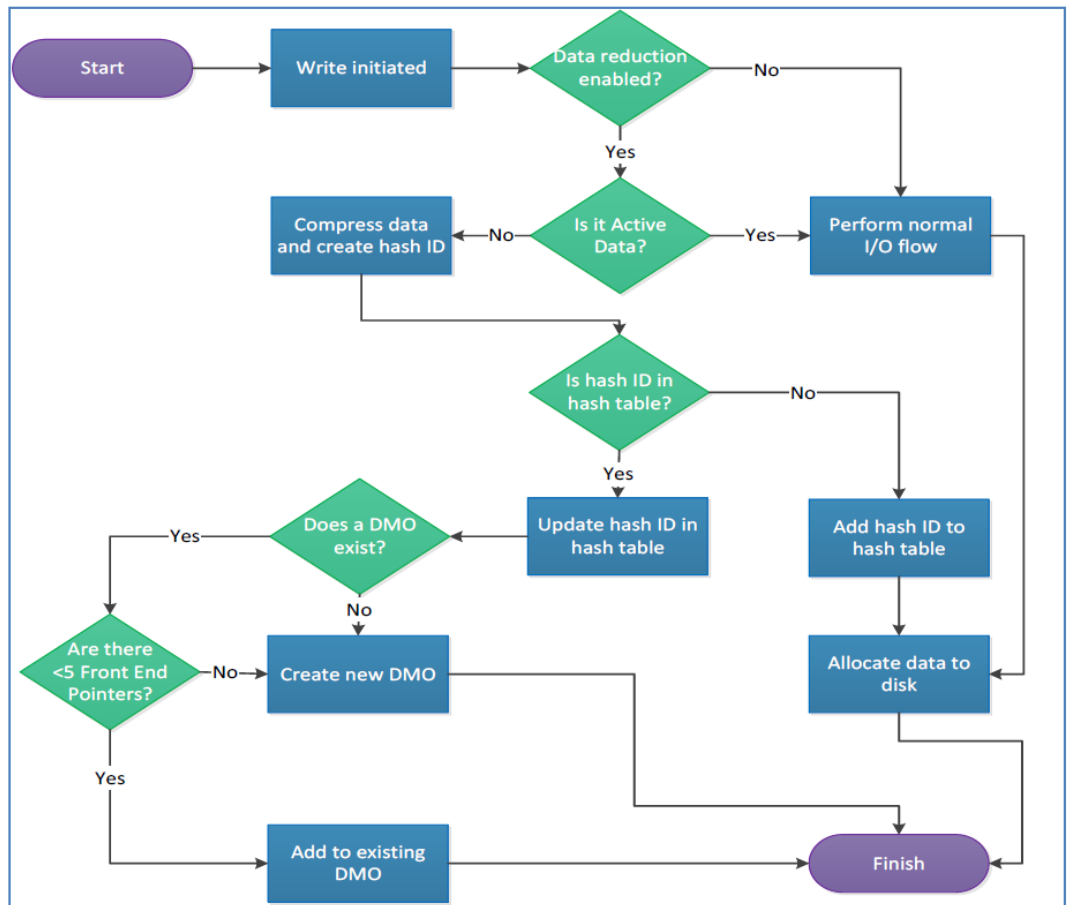


그림 1. PowerMax 엔터프라이즈 스토리지 시스템의 데이터 감소 I/O 흐름

시스템 리소스 활용도

개요

용량과 메모리는 각 PowerMax 시스템에 구성된 두 가지 기본 리소스입니다. 용량의 종류는 물리적 용량, 유효 용량, 프로비저닝된 용량으로 분류됩니다. 메모리는 시스템 리소스로 분류됩니다.

메모리 리소스는 프로비저닝된 용량뿐만 아니라 물리적 용량에 대한 메타데이터 구조도 지원합니다. 사용 가능한 유효 용량의 양은 물리적 용량의 양, 사용 가능한 시스템 리소스의 양 및 시스템에 기록된 데이터의 감소 가능성과 관련이 있습니다. 감소하기 매우 쉬운 기록된 데이터는 물리적 용량을 적게 사용하여 더 효과적으로 용량을 생성합니다. 그 반대로 마찬가지입니다. 감소하기 어려운 기록된 데이터로 인해 사용 가능한 유효 용량이 줄어들 수 있습니다. [데이터 감소](#) 섹션(용량, 시스템 리소스)에 설명된 정보는 PowerMax 2500 및 8500 시스템, Unisphere for PowerMax, Solutions Enabler 및 Mainframe Enabler 소프트웨어용 관리 애플리케이션에서 확인할 수 있습니다. Unisphere for PowerMax는 그래프, 차트 및 목록 형식으로 데이터를 제공하는 UI(user interface)입니다. Solutions Enabler는 동일한 데이터를 제공하는 표준 명령줄 인터페이스이지만 차트와 그래프의 형태는 제외됩니다. Mainframe Enabler는 메인프레임 환경에서 Dell 스토리지 시스템을 모니터링하고 관리하는 구성 요소 제품군입니다. 이 백서의 다음 섹션에 들어있는 이미지는 PowerMax 2500 또는 8500 시스템을 관리하는 Unisphere for PowerMax를 보여줍니다.

물리적 용량

물리적 용량은 설치된 디스크와 적용된 RAID 보호에 따라 시스템에 구성된 디스크 공간의 양입니다. 데이터 감소가 사용되지 않는 구성에서 물리적 용량은 호스트 데이터에 사용할 수 있는 총 용량입니다. 예를 들어 물리적 용량이 100TB인 시스템은 데이터 감소를 사용하지 않는 100TB의 호스트 데이터를 수용할 수 있습니다.

유효 용량

유효 용량은 데이터 감소를 사용 중일 때 사용할 수 있는 공간의 양입니다. 초기 설치 시의 총 용량은 시스템에 구성된 메모리 양에 따라 다르며 4:1의 기본 데이터 감소 절감 효과(CKD 에뮬레이션의 경우 3:1)를 기준으로 합니다.

예를 들어 물리적 용량이 100TB인 동일한 시스템에는 유효 용량이 400Tb으로 표시됩니다. 이 400Tb라는 값은 유효 용량의 시작 지점이며 데이터가 시스템에 기록되고 데이터 감소가 적용되면 변경됩니다.

프로비저닝된 용량

프로비저닝된 용량은 시스템에서 물리적 또는 유효 용량을 사용하려는 호스트 및 애플리케이션에 생성되고 제공되는 디바이스의 형태로 사용 가능한 용량을 나타내는 것입니다.

관리 및 모니터링

개요

Unisphere for PowerMax는 시스템의 용량 및 리소스 사용량을 관리하고 모니터링하는 데 사용되는 사용자 인터페이스입니다. 시스템 수준에서 용량 사용량, 데이터 감소, 시스템 리소스에 대한 정보가 용량 대시보드에 표시됩니다. 용량 대시보드에서 사용자는 유효 용량, 프로비저닝된 용량, 스냅샷 용량 및 데이터 감소와 시스템 리소스에 대한 정보를 표시하는 화면으로 이동이 가능합니다.

용량 대시보드

Unisphere for PowerMax에는 용량 사용과 관련된 정보를 표시하는 여러 디스플레이가 있습니다.

기본 대시보드에는 시간 경과에 따른 유효 용량 사용량 및 데이터 감소를 보여주는 대화형 그래프가 표시됩니다. 이 디스플레이는 유효 용량 사용 내역 및 데이터 감소 비율이 유효 용량과 어떻게 관련되어 있는지를 보여줍니다. 이 정보를 사용하면 표시되는 데이터 감소 비율과 관련된 유효 용량 사용 추세를 모니터링하고 추적할 수 있습니다. PowerMax 2500 또는 8500 시스템은 동일한 스토리지 리소스 풀 내에서 FBA 및 CKD 에뮬레이션으로 구성될 수 있지만 기간별 그래프는 선택한 에뮬레이션 보기에 따라 다르게 표시됩니다.

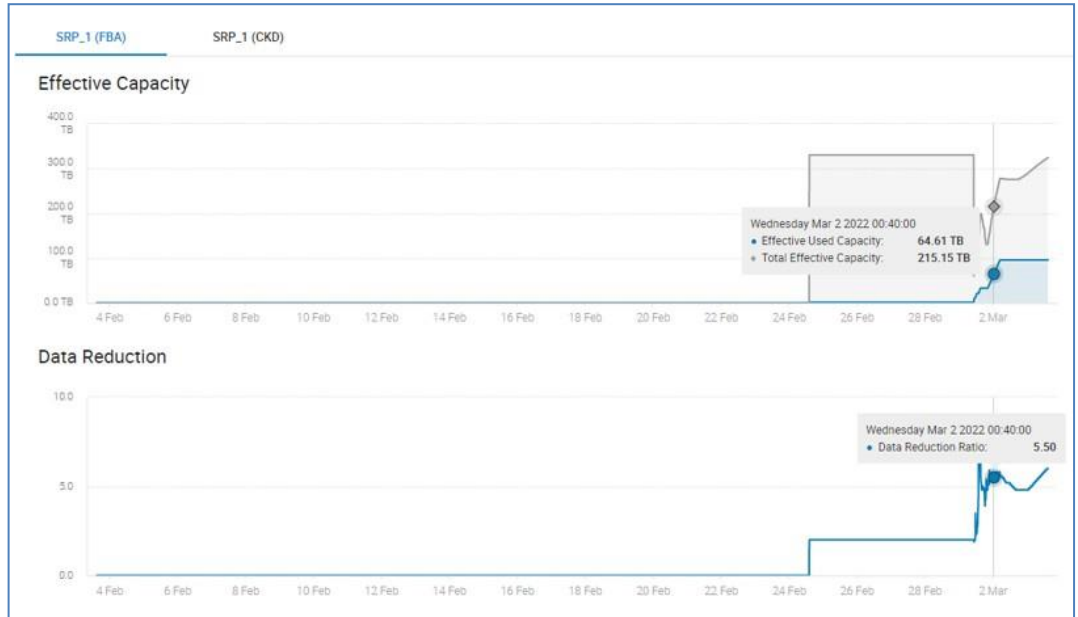


그림 2. FBA 에뮬레이션에 대한 유효 용량 및 데이터 감소를 보여주는 용량 대시보드 기간별 그래프

기본 대시보드는 프로비저닝된 용량, 유효 용량, 스냅샷 사용량, 및 데이터 감소에 대한 막대 그래프 형태의 데이터도 제공합니다. 각 섹션을 각 항목에 대해 보다 세분화된 데이터를 표시하는 더 자세한 디스플레이로 확장할 수 있습니다.

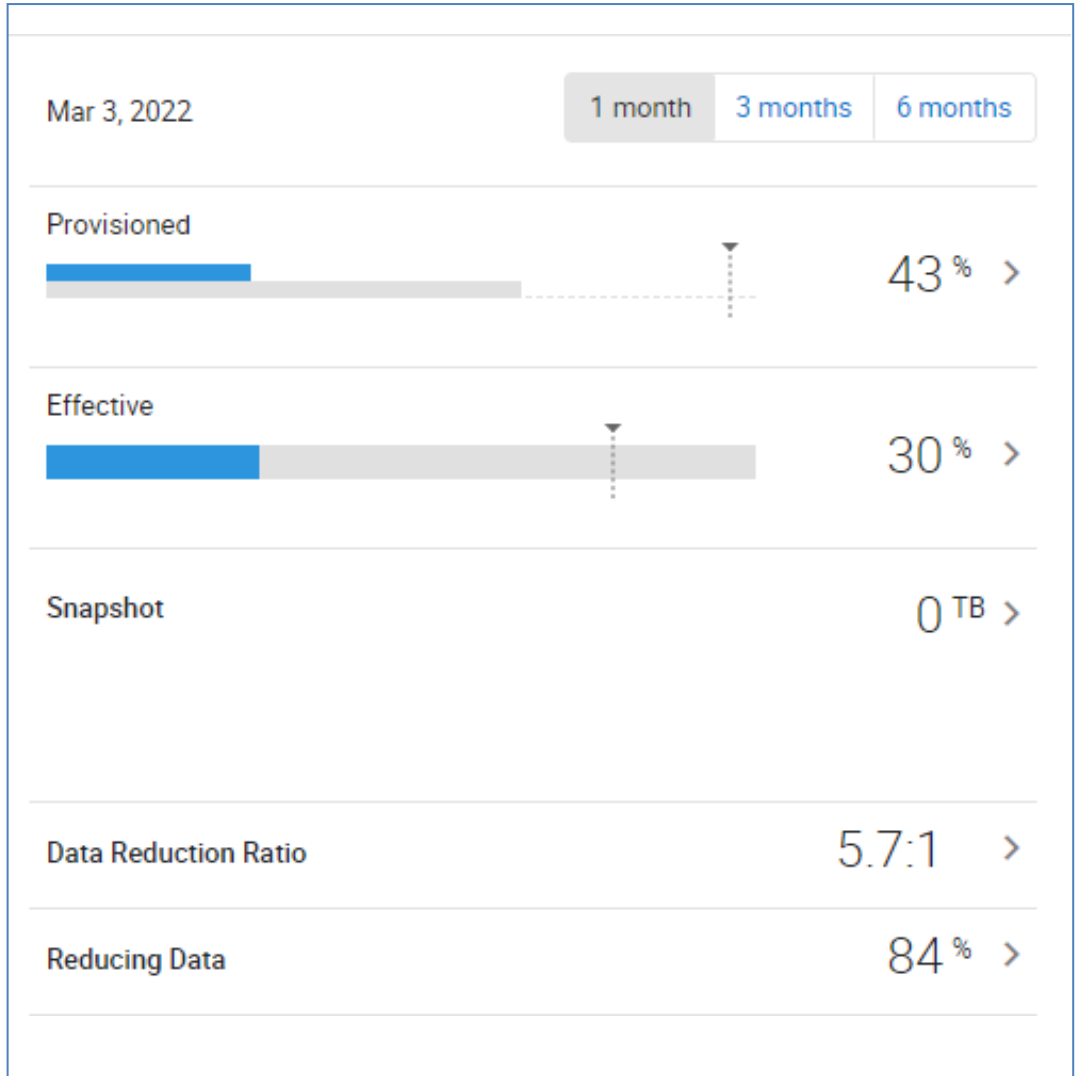


그림 3. 프로비저닝된 용량, 유효 용량, 스냅샷 사용량 및 데이터 감소에 대한 용량 대시보드 막대 그래프

프로비저닝됨

프로비저닝된 용량은 호스트 및 애플리케이션에 사용 가능한 용량으로 표시되는 디바이스 형태로 프로비저닝되는 용량입니다. 프로비저닝된 용량은 SRP 용량 및 시스템 리소스라는 두 가지 메트릭을 사용하여 추적됩니다.

- SRP 용량**은 프로비저닝된 TB 값으로 프로비저닝된 용량을 보여주고 사용 가능한 유효 용량을 보여줍니다. 유효 SRP 용량의 양은 처음에는 4:1(CKD 에멀레이션의 경우 3:1)의 기본 데이터 감소 절감 효과를 기준으로 합니다. 이 용량은 시스템에 구성된 물리적 용량을 사용하여 계산됩니다. 사용자가 디바이스를 생성하면 프로비저닝된 용량은 증가합니다. 표시된 백분율 값은 구독 백분율이며 프로비저닝된 분량과 유효 용량을 사용하여 계산됩니다.

- 시스템 리소스**는 메타데이터 형태의 프로비저닝된 용량을 지원하는데 사용할 수 있는 메모리 양을 나타냅니다. 표시되는 총 값은 캐시가 시스템에 추가된 경우에만 변경됩니다. 사용된 용량은 그 값을 나타냅니다. 사용 가능한 용량은 시스템이 지원할 수 있는 추가 프로비저닝된 용량이 어느 정도인지를 나타냅니다. 사용자가 디바이스를 생성하면 사용된 양이 증가합니다. 두 값 사이에 차이가 있는 경우 데이터 감소 또는 기존 장치의 스냅샷 생성과 같은 메모리를 사용하는 기타 기능을 지원하기 위해 메모리 리소스가 사용됩니다.

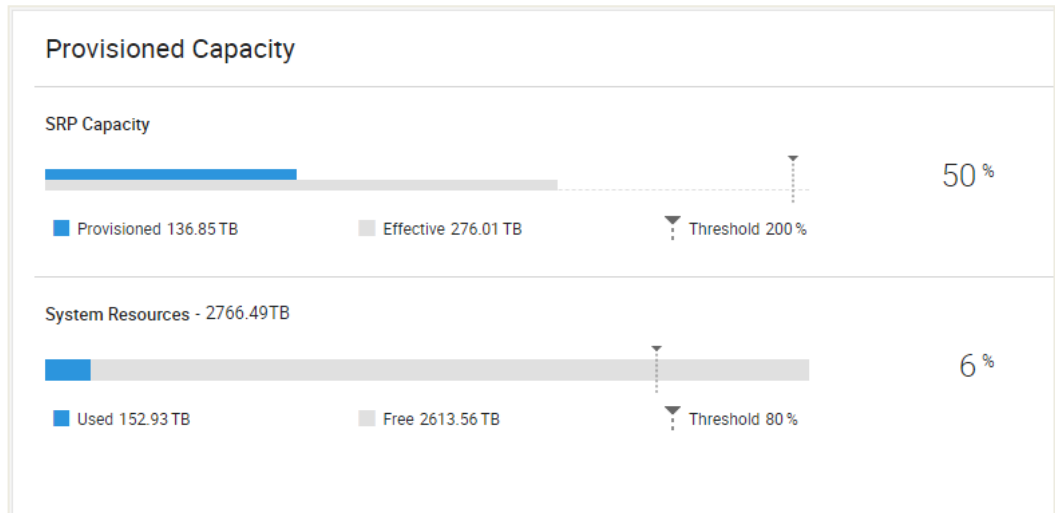


그림 4. 프로비저닝된 용량 디스플레이 내의 시스템 리소스 사용량

유효

유효 용량은 데이터 감소 사용으로 인한 절감 효과에 대한 기대에 따라 사용자가 사용할 수 있는 용량을 나타냅니다. 유효 용량 디스플레이에서 사용 가능한 물리적 리소스와 효과적인 리소스를 자세히 볼 수 있습니다. 이 섹션은 물리적 용량, 유효 용량 리소스, 유효 용량 사용량의 세 가지 섹션으로 표시됩니다.

- **물리적 용량**은 시스템에 구성된 하드 드라이브에서 사용할 수 있는 물리적 용량을 보여줍니다. 표시된 용량은 포맷 및 RAID 보호가 적용된 이후의 값입니다. 표시된 값은 데이터 감소가 사용되지 않을 때 시스템이 호스트 데이터에 대해 지원할 수 있는 용량입니다.
- **유효 용량 리소스**는 현재 시스템 리소스 사용량을 기준으로 달성 가능한 값을 나타냅니다. 표시된 유효 용량 리소스 값은 현재 데이터 감소 절감 효과와 물리적 용량 및 유효 용량 사용량에 따라 조정됩니다.
- **유효 용량 사용량**은 시스템 리소스 사용량과 현재 데이터 감소 절감 효과를 기준으로 사용할 수 있는 유효 용량의 현재 양을 표시합니다. 원형 차트에 표시되는 값은 현재 사용 가능한 유효 용량입니다. 오른쪽에 제시된 값은 사용량을 스냅샷 사용량, 사용자가 사용한 사용량, 이용 가능한 사용량의 세 가지 범주로 분류합니다.

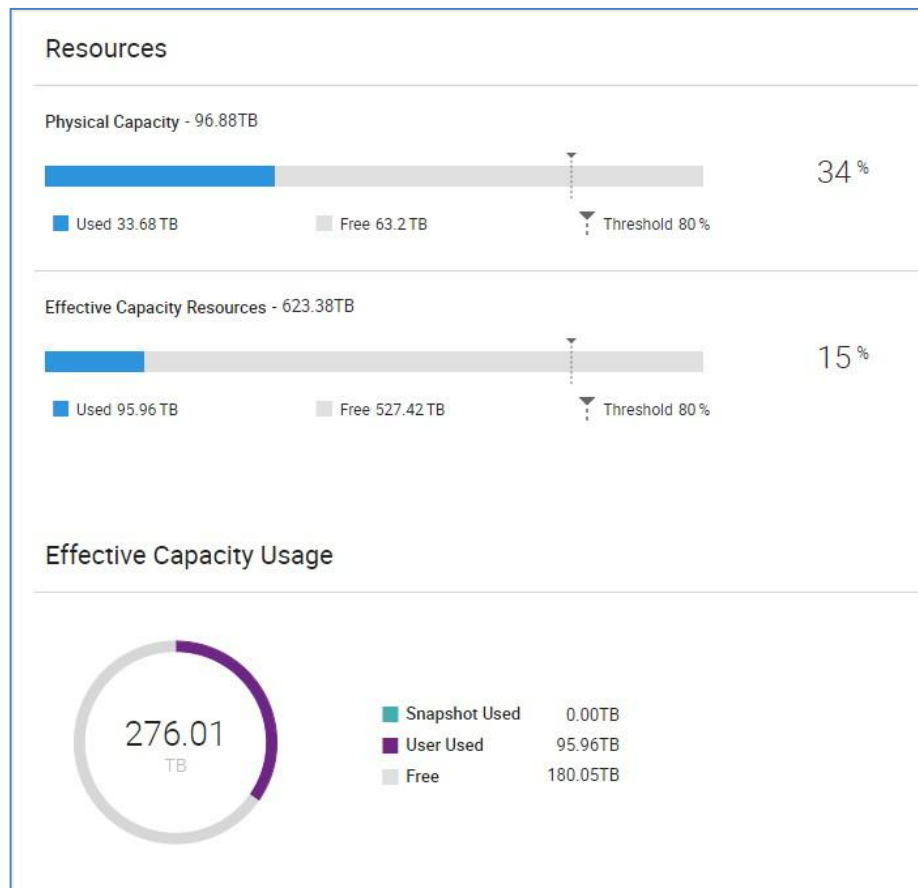


그림 5. 유효 용량 디스플레이를 보여주는 리소스 사용량

스냅샷

백엔드 스냅샷 용량은 공유 할당 및 데이터 감소와 같은 기능의 효율성으로 인해 스냅샷 당 델타보다 훨씬 적을 수 있습니다.

용량 대시보드의 스냅샷 막대 그래프 위에 마우스를 올려 놓아 자세한 내용을 확인합니다. 스냅샷 값은 다음과 같이 정의됩니다.

- **사용된 용량:** 스냅샷 변경 데이터에 사용된 유효 용량
- **이용 가능한 용량:** 사용된 양과 나머지 메타데이터에 따라 남은 스냅샷 변경 데이터 용량
- **합계:** 사용 + 무료
- **임계값:** 스냅샷 변경 데이터에 대한 알림 임계값

유효 용량 대시보드 및 스냅샷 용량 대시보드로 이동하려면 막대 그래프를 클릭합니다.

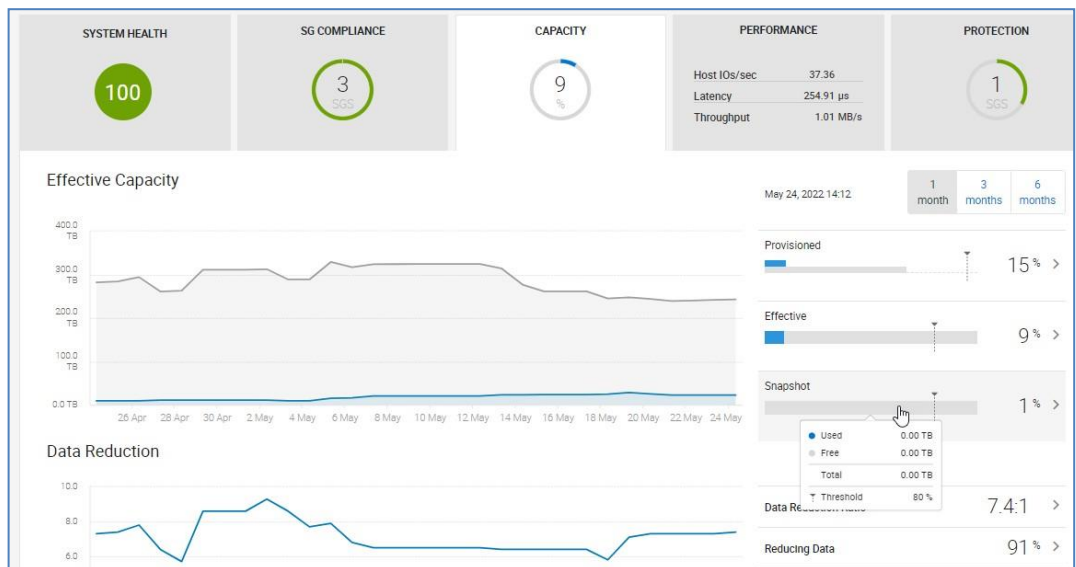


그림 6. 용량 대시보드

유효 용량 대시보드에서 다음을 수행합니다.

- **사용된 스냅샷:** 스냅샷 변경 데이터에서 사용되는 유효 용량

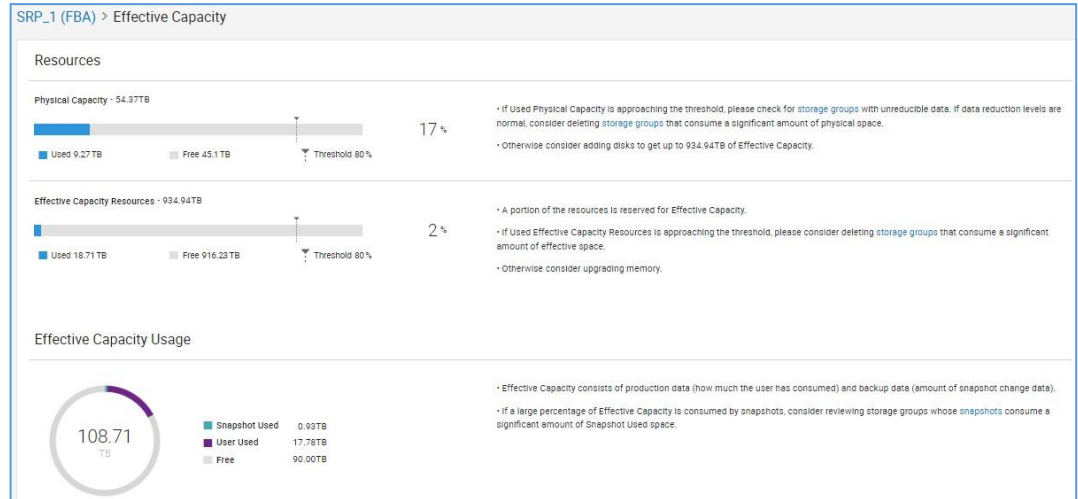


그림 7. 유효 용량 대시보드

스냅샷 용량 페이지의 값은 다음과 같이 정의됩니다.

- **유효 스냅샷 사용:** SRP에 대한 스냅샷 데이터에서 사용되는 유효 용량의 비율
- **물리적 스냅샷 사용:** SRP에 대한 스냅샷 데이터에서 사용된 가용 용량의 비율
- **스냅샷 리소스:** 전체 시스템에 사용된 스냅샷 메타데이터의 비율
- **스냅샷 리소스 사용:** 스냅샷 메타데이터 용량의 일부로 스냅샷에서 사용되는 용량

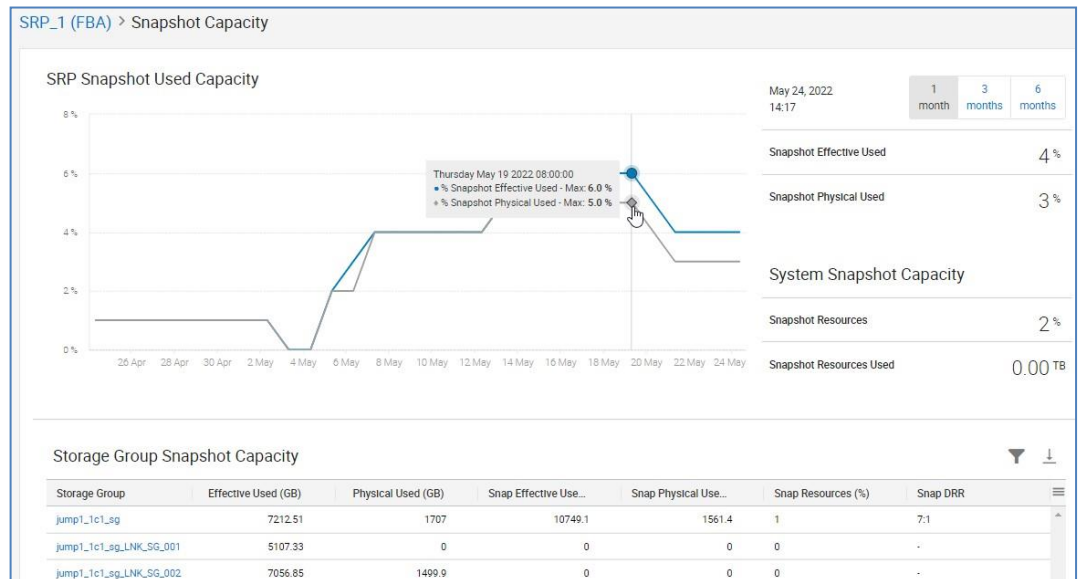


그림 8. 스냅샷 용량 사용량

Data Reduction Ratio

데이터 감소 디스플레이는 사용자가 데이터 감소 효율성을 볼 수 있는 단일 위치를 제공합니다. 데이터 감소 비율, 기간별 대화형 그래프, 모든 스토리지 그룹 표의 세 가지 섹션이 있습니다. 표시된 데이터 감소 비율은 시스템에 기록된 데이터인 경우만 활성화 및 감소를 고려합니다.

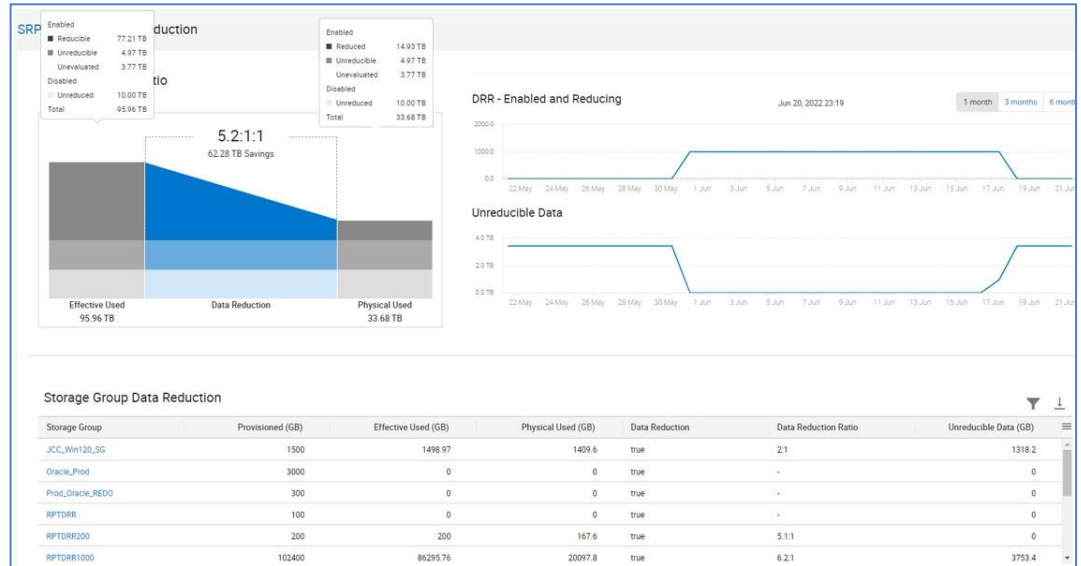


그림 9. 데이터 감소 비율

데이터 감소 비율은 유효 사용 용량, 데이터 감소 비율, 물리적 사용 용량을 나타내는 그래프로 표시됩니다. 물리적 사용은 사용 중인 실제 물리적 용량을 나타냅니다. 데이터 감소는 비용 절감을 비율로 나타냅니다.

유효한 사용은 데이터 감소가 적용될 때 절감을 달성하기 전에 시스템에 기록된 데이터를 나타냅니다. 표시된 모든 값은 호스트 또는 애플리케이션에서 작성한 전체 크기를 나타냅니다. 기록된 데이터는 Enabled과 Disabled이라는 두 가지 범주로 분류됩니다.

- 활성화**는 처리 중인 데이터가 데이터 감소 기능이 활성화되어 있고 데이터 감소 프로세스와 활동 기반 감소 기능의 적용을 받는다는 것을 나타냅니다. 데이터 감소가 활성화된 경우, 데이터는 감소 가능, 감소 불가능, 평가 불가능의 3가지 추가 범주에 속할 수 있습니다.
 - 감소 가능** 데이터는 데이터 감소 프로세스가 시스템에 기록된 것보다 적은 물리적 용량을 사용하도록 줄일 수 있는 데이터로 식별된 데이터의 양입니다.
 - 감소 불가능** 데이터는 줄일 수 없는 데이터입니다.
 - 데이터 감소 프로세스에서는 아직 **평가 불가능** 데이터를 평가하지 않습니다. 데이터가 감소 가능한지, 아니면 감소가 불가능한지 아직 파악되지 않았습니다.
- 비활성화**는 시스템에 기록된 데이터가 데이터 감소 절감의 대상이 아님을 나타냅니다. 비활성화로 식별된 모든 데이터는 감소되지 않음으로 표시됩니다.

물리적 사용은 디스크에 저장된 후 시스템에 기록된 데이터를 나타냅니다. 이는 활성화 및 비활성화된 모든 데이터와 감소된 데이터 및 감소되지 않은 모든 데이터를 고려합니다. 디스크에 저장된 데이터를 나타내는 두 가지 범주가 있습니다. 활성화 및 비활성화입니다.

- **활성화**는 데이터 감소 프로세스를 거친 데이터를 나타냅니다. 이 데이터의 세 가지 하위 범주가 있습니다. 감소, 감소 불가, 평가 불가입니다.
 - **감소** 데이터는 데이터 감소 프로세스를 통해 전송됩니다. 이 프로세스에는 데이터 감소 하드웨어를 통과하고 디스크에 저장되는 작업이 모두 포함됩니다. 디스크에 저장된 감소 데이터는 호스트 또는 애플리케이션에서 기록하는 경우 보다 적은 디스크 공간을 사용합니다.
 - **감소 불가**는 데이터가 데이터 감소 하드웨어를 포함하여 데이터 감소 프로세스를 통해 데이터가 전송되었지만 감소되지 못했음을 나타냅니다. 물리적 사용 섹션 내에서 처리되는 일부 감소 불가한 데이터는 중복 제거로 공유되는 데이터로 데이터 감소 절감에 기여할 수 있습니다.
 - **평가 불가** 데이터는 데이터 감소 프로세스에서 아직 평가되지 않는 데이터입니다. 따라서 데이터가 아직 감소 가능한지 불가능한지 결정되지 않았습니다.
- **비활성화**는 시스템에 기록된 데이터가 데이터 감소 절감의 대상이 아님을 나타냅니다. 비활성화로 식별된 모든 데이터는 감소되지 않음으로 표시됩니다.

DRR 활성화 및 감소, 그리고 감소 불가 데이터를 차트로 나타내는 대화형 그래프 차트는 기간별 데이터를 제공합니다. 이는 감소 불가한 데이터가 데이터 감소 비율에 미치는 영향을 보여줍니다. 이 그래프를 사용하면 사용자가 감소 불가한 데이터로 인해 발생할 수 있는 데이터 감소 비율의 변경 사항을 추적하고 모니터링할 수 있습니다.

스토리지 그룹 목록은 시스템의 각 스토리지 그룹에 해당하는 용량 사용량 및 데이터 감소 정보를 제공합니다. 대화형 그래프를 사용하여 데이터 감소 비율에 대한 변경 사항을 추적하는 경우, 스토리지 그룹 목록을 사용하여 데이터 감소 비율에 영향을 미치는 많은 양의 감소 불가한 데이터가 있는 스토리지 그룹을 파악할 수 있습니다.

효율성 비율 계산: 데이터 감소 비율을 계산하는 데 필요한 데이터는 데이터 감소 그래프의 팝업 창에서 사용할 수 있습니다.

- **데이터 감소율:** 데이터 감소 비율은 유효 사용에서 활성화 및 감소 가능을 사용하여 계산되며 물리적 사용에서 활성화 및 감소를 사용하여 계산됩니다.

활성화된 감소 가능 ÷ 활성화된 감소

- **전체 데이터 감소 비율:** 전체 시스템 데이터 감소 비율은 유효 사용 및 물리적 사용의 총 값을 사용하여 계산됩니다.

유효 사용 합계 ÷ 물리적 사용 합계

지원되는 데이터 서비스

개요

데이터 감소는 FBA 스토리지에서 지원됩니다. 혼합 FBA/CKD 시스템은 동일한 스토리지 리소스 풀 내에서 지원되지만 CKD 에뮬레이션을 위한 데이터 감소는 압축 및 활동 기반 감소만 활용합니다. PowerMax 및 VMAX All Flash 시스템에서 제공되는 기타 모든 데이터 서비스가 지원됩니다. 여기에는 로컬 복제(SnapVX), 원격 복제(SRDF), D@RE, VMware vSphere vVols(Virtual Volumes)가 포함됩니다.

로컬 복제(SnapVX)

SnapVX 스냅샷은 델타라고 하는 변경 데이터를 캡처하기 위해 타겟 볼륨을 사용하지 않고 애플리케이션을 보호합니다. 스냅샷 델타는 스토리지 백엔드에서 관련 시점 이미지에 대한 포인터를 사용하여 자동으로 유지됩니다. 리소스 공유 및 데이터 중복 제거는 이 설계를 자동으로 활용하여 캐시, 용량 및 성능 이점을 제공합니다.

결론

압축된 소스 데이터는 스냅샷 델타가 될 때 압축된 상태로 유지됩니다. 압축되지 않은 소스 데이터는 스냅샷 델타가 되거나 덜 활성화될 때 압축될 수 있습니다. 연결된 타겟을 통한 읽기 작업으로 인해 압축되지 않은 델타가 압축되거나 압축된 스냅샷 델타의 압축이 해제될 수 있습니다. 스냅샷 델타에는 중복 제거를 사용할 수 있습니다.

연결된 타겟에서 데이터 감소를 활성화하면 연결된 타겟 소유의 데이터에만 영향이 미칩니다. 연결된 타겟에 있는 데이터와 클론에는 중복 제거를 사용할 수 있습니다.

원격 복제(SRDF)

SRDF에 대한 압축은 지원되고 있으며 SRDF 압축이라고 합니다. SRDF 압축은 원격 복제를 사용하여 연결된 시스템과 데이터를 주고받을 때 대역폭 소비를 줄이기 위한 기능입니다. SRDF 압축과 데이터 감소는 모두 동일한 하드웨어를 사용합니다. 하지만 서로 다른 용도로 사용됩니다. 데이터 감소를 사용하여 압축된 데이터는 SRDF 링크를 통해 전송되기 전에 압축 해제됩니다. SRDF 압축 및 인라인 압축이 적용되는 경우에는 데이터가 압축 해제되고 SRDF 압축 기능을 사용하여 압축된 후에 원격 사이트로 전송됩니다.

D@RE(Data at Rest Encryption)

D@RE는 하드웨어 기반의 어레이 내 백엔드 암호화를 제공하며, 데이터 감소는 데이터의 인라인 프로세스로 수행됩니다. 데이터는 암호화 하드웨어를 통해 전송되기 전에 데이터 감소 하드웨어를 통과합니다. 따라서 데이터가 암호화되기 전에 압축, 중복 제거 또는 둘 모두가 적용됩니다. 디스크에서 암호화되고 D@RE가 활성화된 시스템 데이터는 이미 압축, 중복 제거 또는 둘 모두가 적용된 상태입니다.

Virtual Volumes

데이터 감소는 vVols에 대한 데이터 할당을 지원하며 다른 모든 데이터와 동일한 I/O 경로를 따릅니다. IO 경로는 [그림 1](#)에서 확인할 수 있습니다. vVol에 대한 스토리지 그룹이 없으므로 vVol 스토리지 컨테이너의 스토리지 리소스 수준에서 데이터 감소가 활성화됩니다.

결론

요약

물리적 스토리지 용량의 사용은 스토리지 업계에 있는 스토리지 관리자의 공통적인 우려 사항입니다. 지속적으로 증가하는 데이터 양으로 인해 물리적 용량의 사용 효율성을 높일 필요가 대두되었습니다. Dell PowerMax 2500 및 8500 데이터 스토리지 시스템은 이러한 효율성을 한 차원 끌어올립니다. 데이터 감소는 최적의 성능을 제공하면서 탁월한 용량 절감 효과를 제공합니다. 이를 통해 데이터 센터 설치 공간이 줄어들고 TCO가 전반적으로 줄어듭니다. 데이터 감소는 비용 절감 외에도 한 번의 클릭으로 간단하게 활성화하거나 비활성화할 수 있다는 이점도 있습니다. 이 모든 작업을 시스템이 알아서 처리합니다.

참고 자료

Dell Technologies 문서

다음 Dell Technologies 문서 자료에는 이 문서와 관련된 기타 정보가 나와 있습니다. 로그인 자격 증명에 따라 문서에 대한 액세스가 제한될 수 있습니다. 문서에 대한 액세스 권한이 없으면 Dell Technologies 담당자에게 문의하시기 바랍니다.

- [PowerMax and VMAX Info Hub](#)