

Implementing Data Deduplication

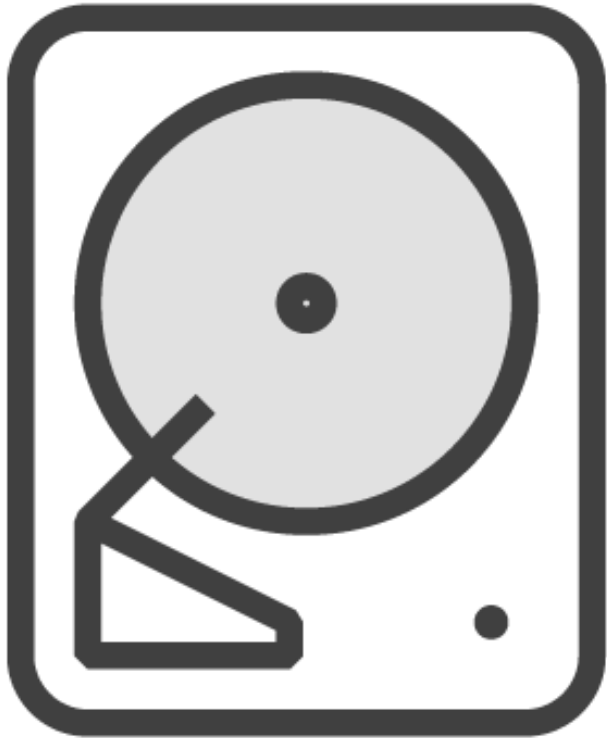


Mike Pfeiffer

@mike_pfeiffer [linkedin.com/in/mpfeiffer](https://www.linkedin.com/in/mpfeiffer)



Overview of Data Deduplication



Finds and removes duplications

Segments data into small chunks

Replaces redundant chunks of data with a reference

Compresses chunks of data

Overview of Data Deduplication

Supports volume sizes up to
64 TB

Support file sizes up to 1 TB

Removes duplications without
compromising data integrity

Support for optimized
volume backups

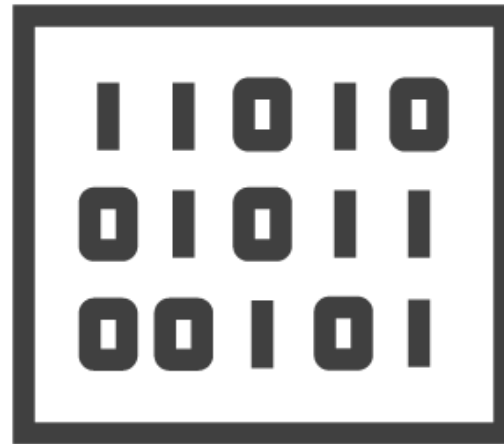


Data Deduplication Jobs



Optimization

Job type that performs both deduplication and compression



Scrubbing

Job type that analyzes corruption logs and makes repairs as needed



Garbage Collection

Job type that processes deleted or modified data on the volume



Use cases for deduplication

General file shares

User documents

Virtual hard disks

Backup volumes

Software deployment shares



Managing Backups with Data Deduplication



Block-based backups

Should maintain optimization on the backup media



File-based backups

Backup target must be large enough to hold the entire dataset

Summary



Common Data Deduplication Scenarios

Deploying Data Deduplication

Monitoring Data Deduplication

**Managing Backups with Data
Deduplication**

