RAMBØLL

Bright ideas.
Sustainable change.

CopenCloud Rapport

Intelligent data storage can optimize organizations' costs and lower their climate footprint.

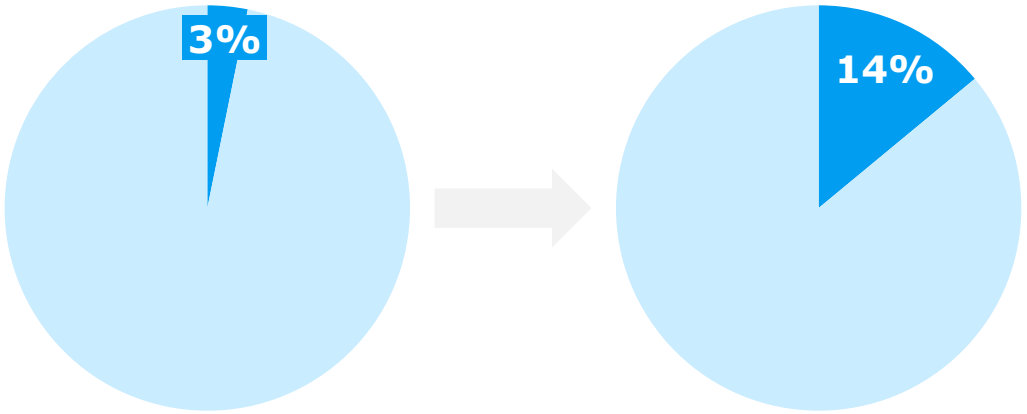
17 May 2023

This report has been prepared in connection with a competence development and counseling course, supported by the Beyond Beta program.

The purpose of the report is to demonstrate what effect Intelligent data storage has on optimization of organizations' costs and reduction of their climate footprint.

Greenhouse gas emissions from data storage are a growing global challenge, which creates a great demand for solutions

Global data traffic is expected to reach approximately **175 ZB** i 2025 of which **10%** or **~17.5 ZB** data is expected to be stored

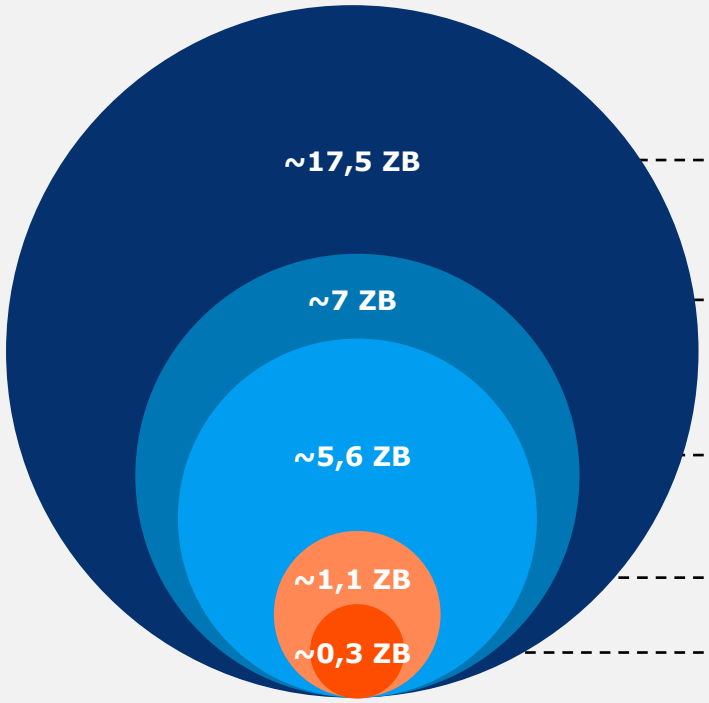


Data centers are expected to make up **3.2% of global greenhouse gas emissions in 2025**. And in **2040** it is expected this proportion to rise to **14%**, which is why there is a great need for effective data storage solutions[3]

[1]Tape to Play Critical Roles as the Zettabyte Era Takes Off, Tape Storage Council (2022); [2]Accelerating Green Datacenter Progress with Sustainable Storage Strategies, IDC (2021); [3]Copenhagen Centre on Energy Efficiency;
*1 ZB = 1 Zettabyte = 1bn terabyte

Cold/inactive data constitutes the fastest growing data segment and expected to account for ~5.6 ZB of commercial data by 2025[1,2]

Total amount of stored data globally in 2025



The total amount of stored data globally is expected to reach approximately **~17.5 ZB in 2025**

IDC estimates that around **40%** or **~7 ZB** of the global data traffic in 2025 will be commerciial[2]

"Cold data" i.e. expectation for access to data is more than 30 days are expected to amount to min. **80%** of organizations' data in 2025

Hot data" i.e. Expectation of access to data is within 30 days are expected to amount to approx. **15%** of organizations' data in 2025

"Very hot data" expectation for access to data is within 10 days are expected to amount to approx. **5%** of organizations' data in 2025

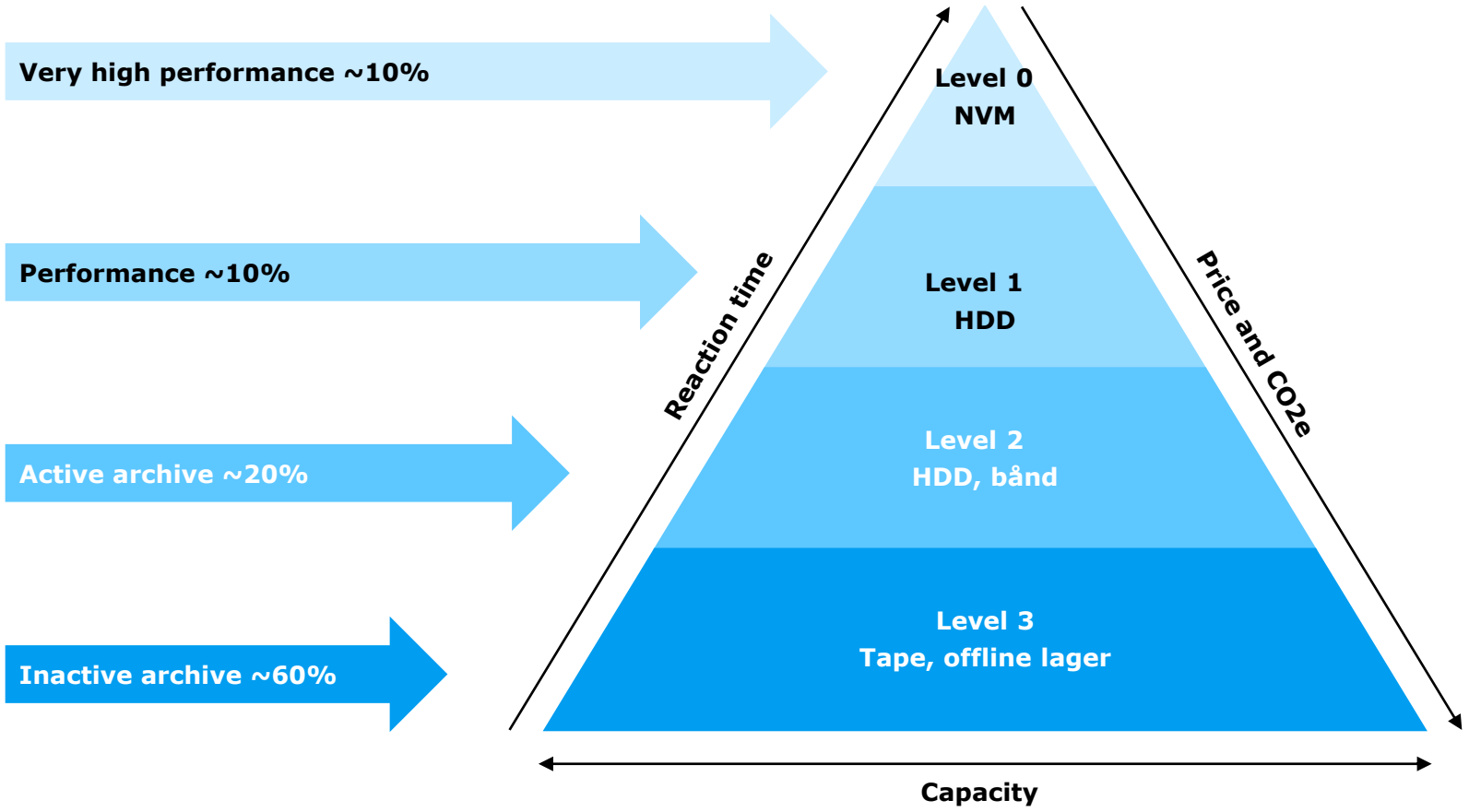
- IDC expects that **commercial data**, i.e. data belonging to organisations, **amounts to approx. 40% of the total amount of stored data globally**
- The total amount of stored commercial data can be further segmented as hot/ cold, depending on when the data is expected to be accessed by the user
- Horison Information Strategies estimates that **60% of data** today can be classified as cold/inactive data, i.e. data that a user rarely or never accesses or changes, and that this proportion may rise to **80%in 2025**[1]
- **Cold/inactive data** thus constitutes the **largest and fastest growing data segment**
- Thus, there is a **great need for** solutions that can be **stored cold/inactive data in an energy- and cost-efficient manner**

[1]Tiered Storage – Storage Optimization for the Zettabyte Era, Horison Information Strategies (2021); [2]White Paper – Tape and Cloud: Solving Storage Problems in the Zettabyte Era of Data, Phil Goodwin (2019)

Today, many organizations store their inactive/cold data on energy-intensive solutions, which is inefficient[1]

- As organizations' volume of data and therefore storage needs grow, **correct segmentation and storage of data becomes more important**[1,2]
- Today, many organizations **store** their **cold/inactive data on energy-intensive solutions** such as HDD*, resulting in **high costs and a high climate footprint**[2]
- Organizations can thus **achieve savings in costs and climate footprint** by moving their cold/inactive data to storage solutions such as **tape storage**, which have a lower energy consumption
- Archive storage is suitable for cold/inactive data that an organization does not need to use every day
- The benefits of archival storage include **lower costs and climate footprint due to lower energy consumption** compared to SSD** and HDD[1]
- The **primary disadvantage associated with archive storage is the longer response time**, which means that data that needs to be accessed immediately or very quickly is not suitable for this type of data storage
- On the right, Horizon Information Strategies' proposal for how an organization's data is optimally stored in various storage solutions can be seen

Optimal % data allocation per tier for an average organization[1]

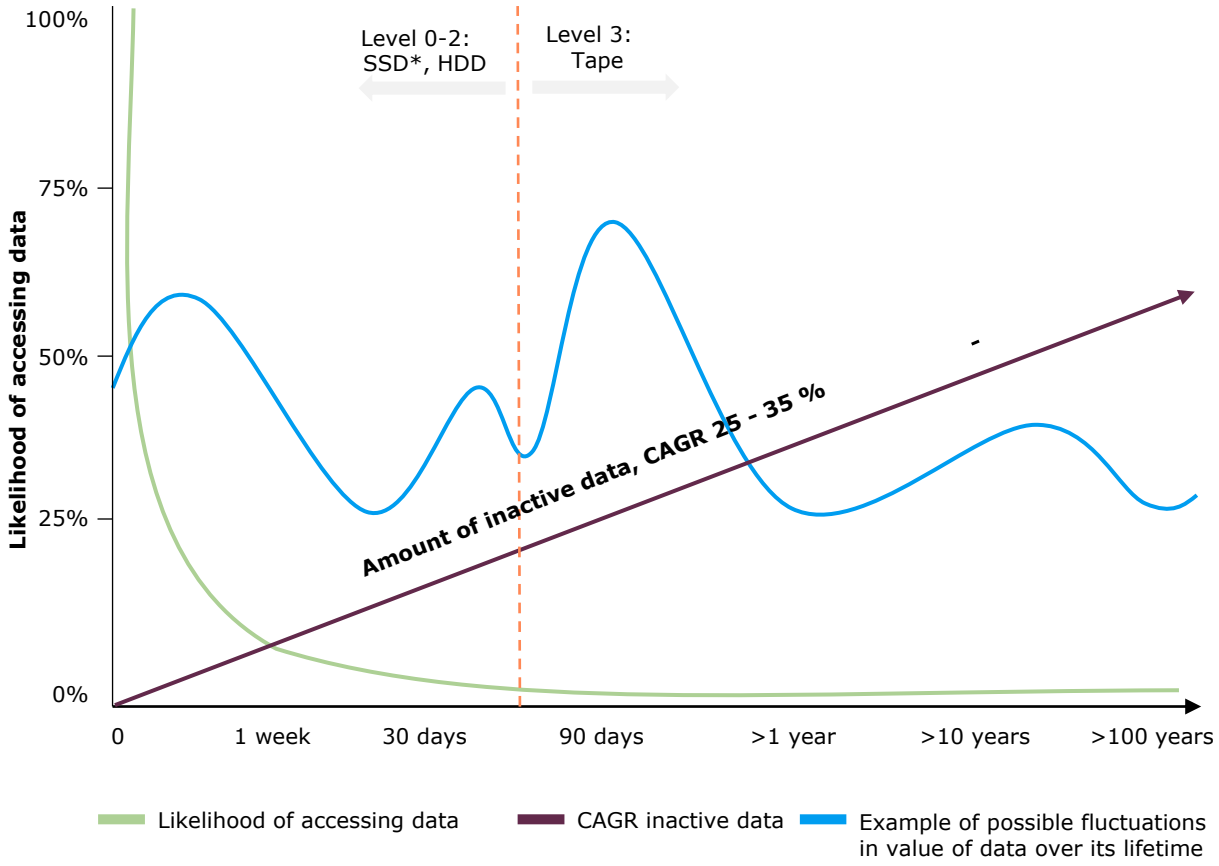


[1]Tiered Storage – Storage Optimization for the Zettabyte Era, Horison Information Strategies (2021); [2]White Paper – Tape and Cloud: Solving Storage Problems in the Zettabyte Era of Data, Phil Goodwin (2019); [3]Tape to Play Critical Roles as the Zettabyte Era Takes Off, Tape Storage Council (2022); *HDD = hard disk drive; **SSD = solid state drive

Data should be stored according to how often it is accessed and where valuable it is to the organization[1,2]

- **The challenge** for many organizations is to **understand which data is hot/ active and which data is cold/inactive**, and thus which can be moved to long-term storage[1,2]
- **The segmentation of data should take into account various factors** (see figure on the right):
  1. **The probability of accessing the vast majority of data decreases** with the age of the data. Some data already becomes inactive as it is generated, but for most types of data **the probability of user access drops one month after data was generated and typically below 1% after 90-120 days**
  2. **The value of specific data** to an organization may **vary over time**
  3. **The amount of stored data is steadily increasing** as more data is generated and data is generally stored for a longer period of time, i.e. a long life cycle is becoming more normal with storage periods exceeding 100 years. This also increases the amount of data that should be stored in level 3
- **Software that automatically and intelligently moves data** based on the frequency of access to data **therefore becomes attractive**, i.e.
  - ✓ **Data that is rarely accessed and becomes inactive is automatically moved** to archive storage
  - ✓ If the **frequency of data access increases**, data from archive storage is moved **to a higher storage level** (0, 1 or 2), while data whose **frequency of access decreases** over a period of time is **moved to a lower level**

Examples of factors that influence how data should be stored[1,2]

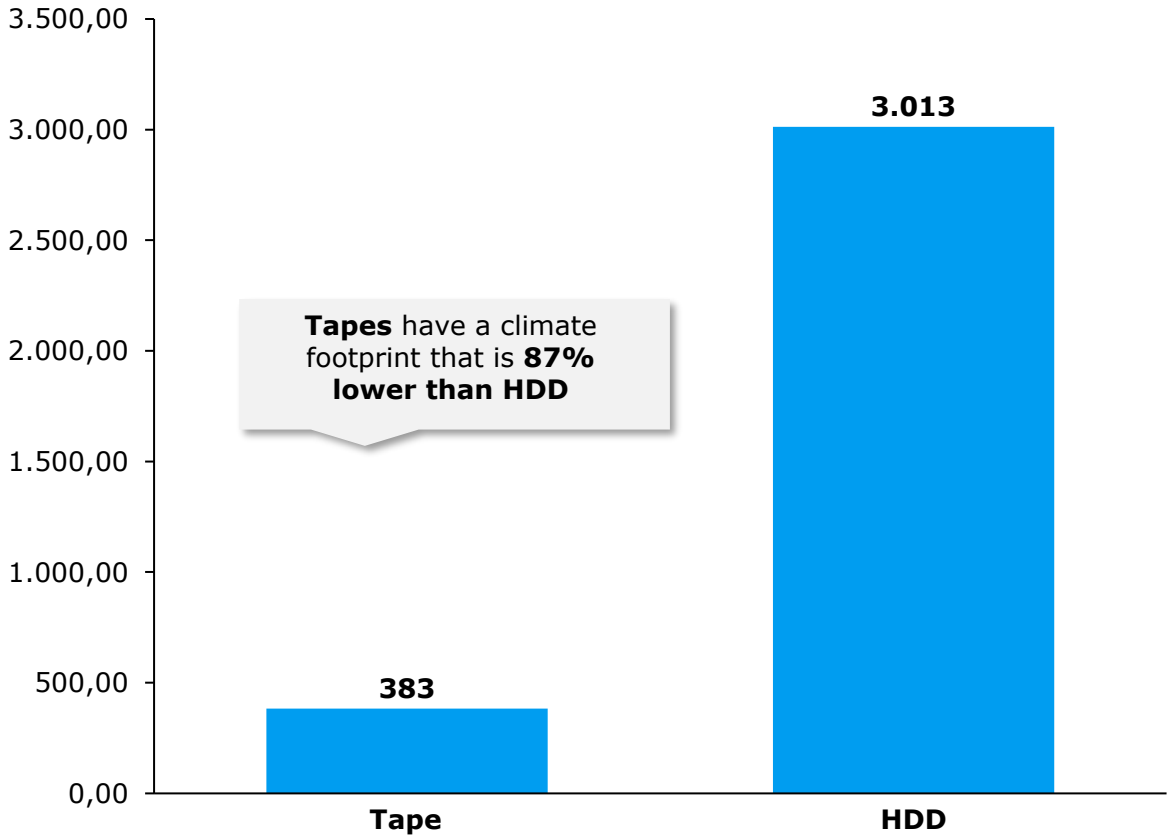


[1]Tiered Storage – Storage Optimization for the Zettabyte Era, Horison Information Strategies (2021); [2]Tiered Storage – Building the Optimal Storage Infrastructure, Horison Information Strategies (2020); *SSD = Solid State Drive; **CAGR = Compound Annual Growth Rate

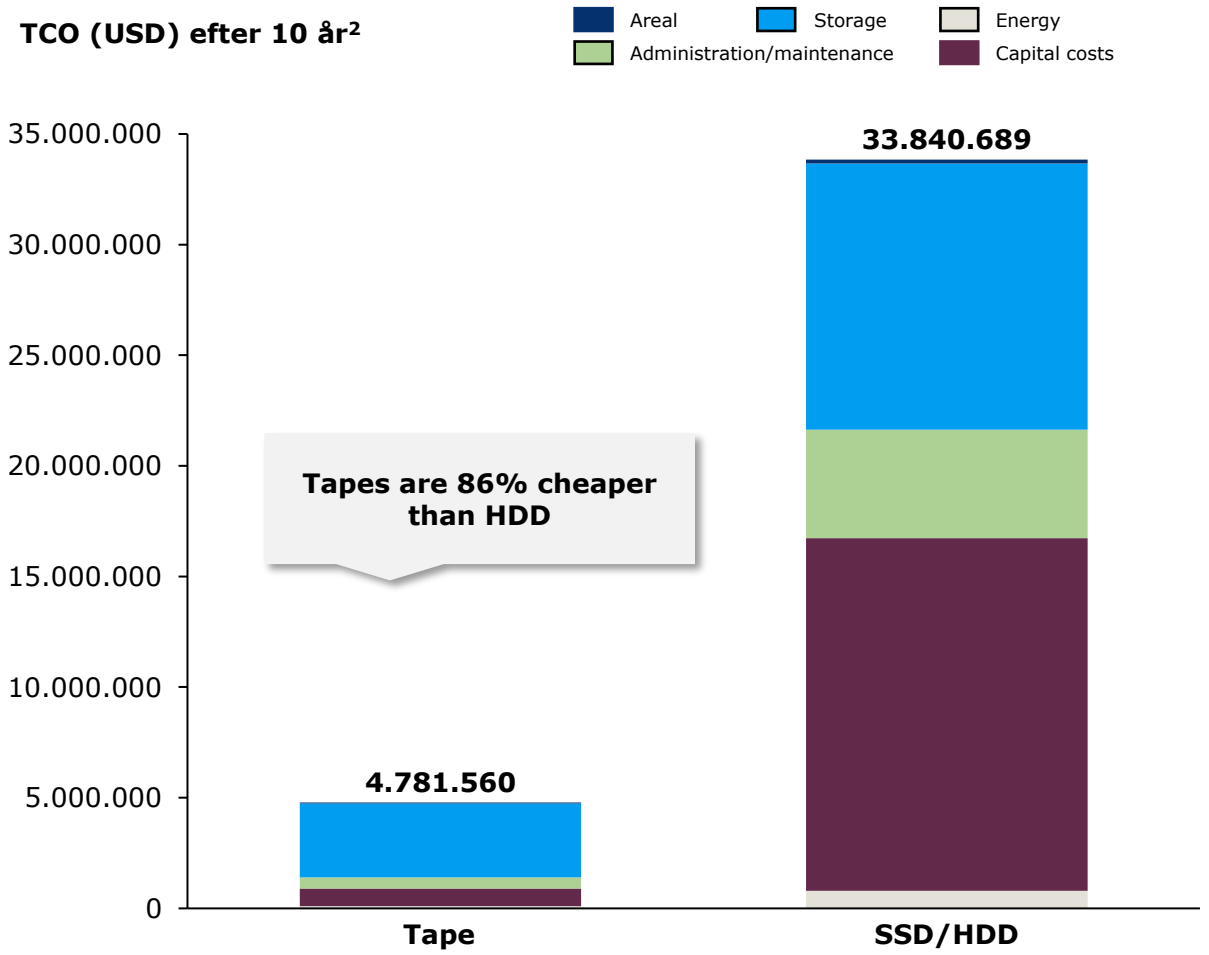
Tape storage appears to be the most cost-effective energy-efficient solution compared to SSD and HDD[1]

Illustrative empirical example

Tons of CO2e after 10 years[2]



TCO (USD) efter 10 år[2]



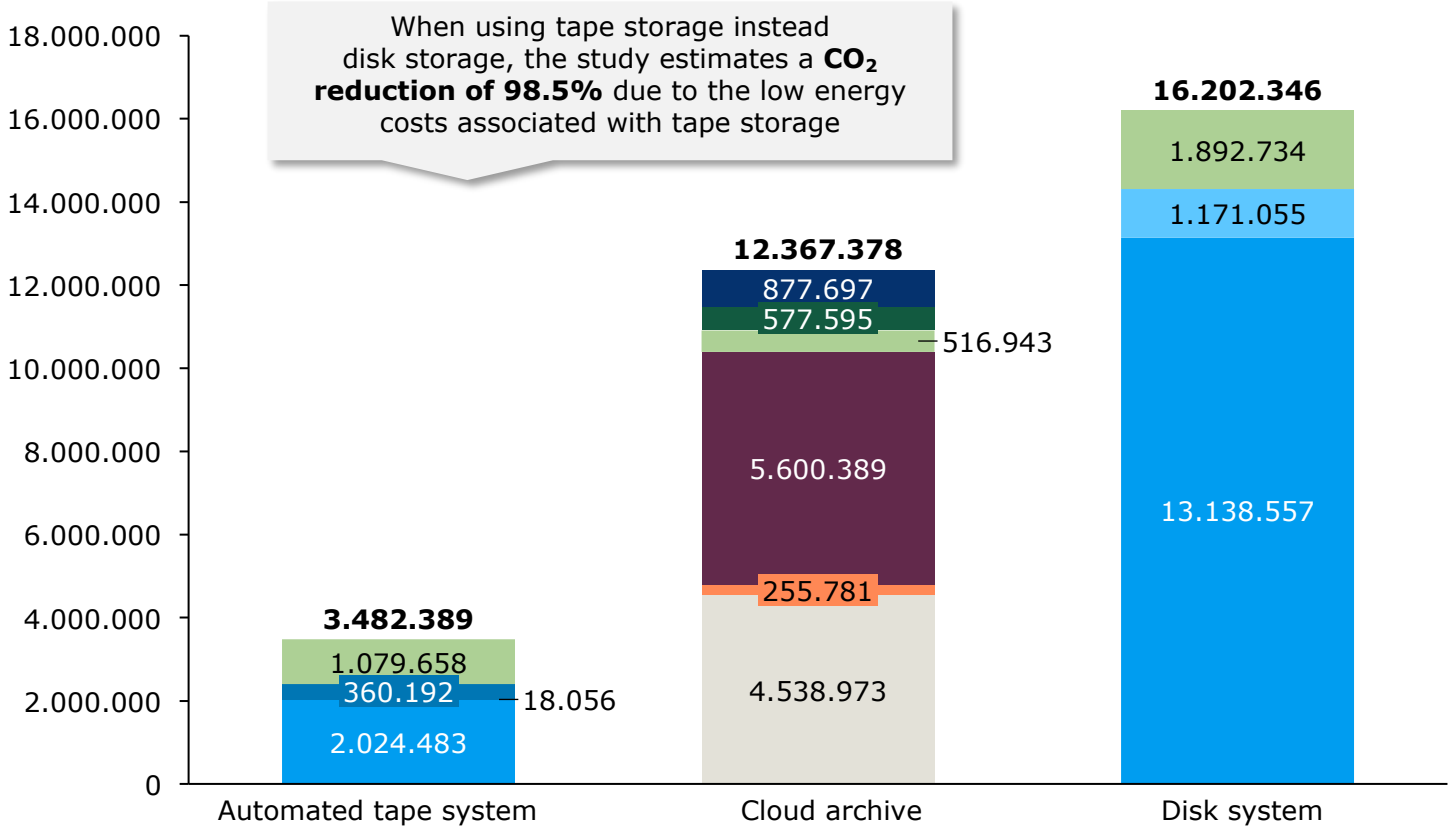
Ramboll

[1]Reducing Data Center Energy Consumption and Carbon Emissions with Modern Tape Storage, Brad Johns Consulting LLC (2020); [2]10PB with an annual growth of 35% over a 10-year period. The study assumes that the electricity to data storage comes from gas.

Illustrative empirical example

Tape storage appears to be the most cost-effective solution compared to disk and cloud archive[1]

Comparison of TCO* for different storage solutions (total after 10 years in USD)[1]



- **Tapes are respectively 79% and 72% cheaper than disk and cloud storage after a 10-year period**
- **The savings of using tape storage** instead of storage on disks and Cloud is **more pronounced after a 10-year than a 5-year period**
- This is because **tape has a longer lifespan** (>30 years) **than disk systems**, which on average need to be replaced every 5 years
- At the same time, t**ape storage is cheaper than Cloud storage due to the latter's ongoing annual cost associated with the storage**, movement and retrieval of data

	Automated tape system	Cloud archive	Disk system
TCO after 5 years (USD)	1.508.715	3.594.966	5.285.564
TCO after 10 years (USD)	3.482.389	12.367.378	16.202.346

- Acquisition, incl. product warranty
- Bandwidth
- Energy costs
- Supplier fee for storage services Supplier
- fee for moving data Supplier fee for
- retrieving data
- Supplier support
- Management of storage
- Maintenance

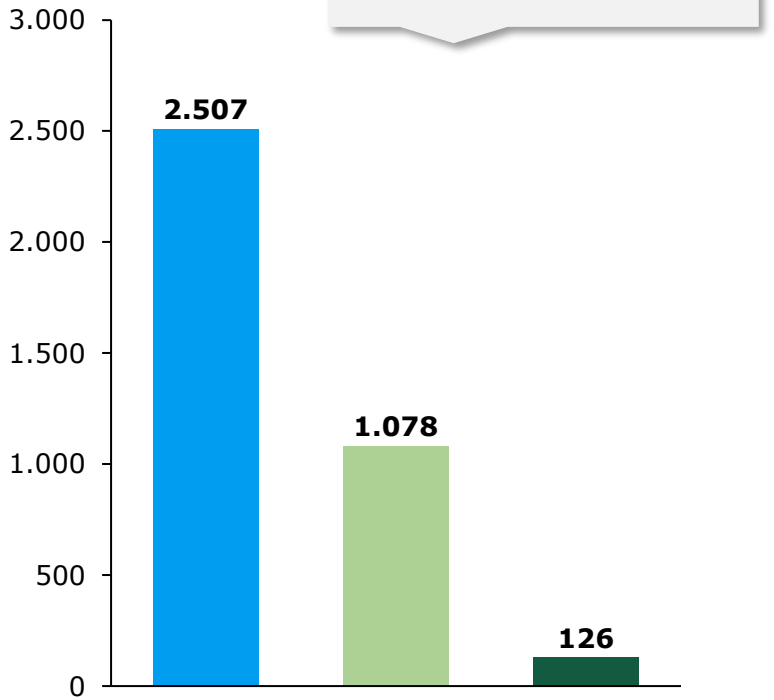
[1]TCO Calculator for Data Storage - How much can you save by using tape for enterprise backup/archive storage? Fujifilm (2023). *The model assumes 20PB of data stored in year 1 with an annual growth of 30% over 10 years and that 12% of the stored data must be downloaded per year. Note that TCO = total cost of ownership in the model includes the initial costs associated with purchasing hardware, software and warranties, operational costs associated with maintenance and support, electricity and cooling, etc., and replacement of technology

Illustrative empirical example

Tape storage has both a lower climate footprint and price as well generates less electronic waste compared to HDD

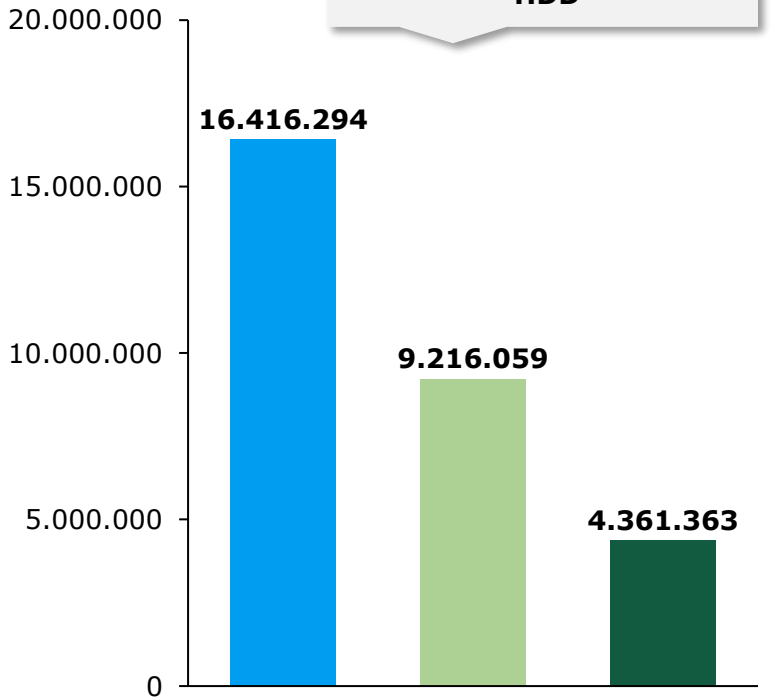
100 PB 10-year CO2e (tons)[1]

Tapes have a climate footprint that is **95% lower than HDD**



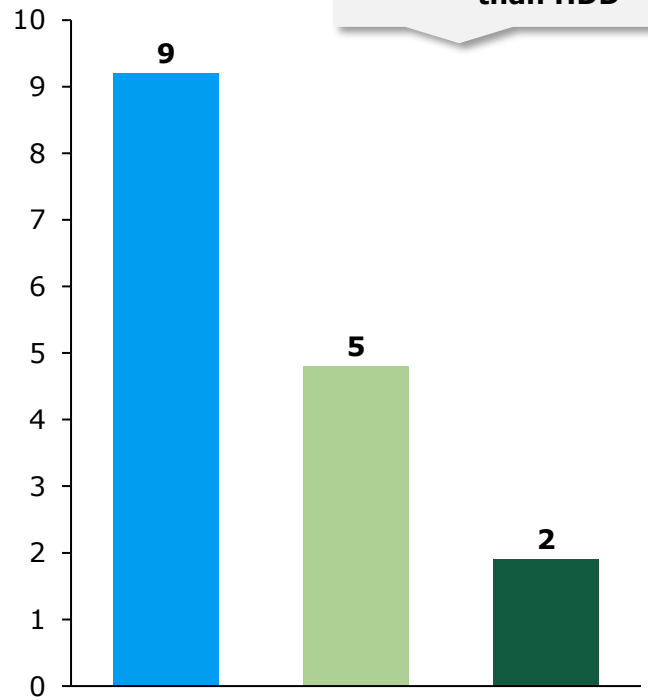
100 PB 10 year TCO (USD)[1]

Tapes have total costs which is **73% lower than HDD**



100 PB 10 years of electronic waste (tons)[1]

Tapes generate **80% less electronic waste than HDD**



100% HDD 40% HDD / 60% Tape 100% Tape

Ramboll

[1]Improving Information Technology Sustainability with Modern Tape Storage, Brad Johns Consulting L.L.C. (2021). This study looks at the life cycle phases of production, use and disposal.

Knowledge Resources | Intelligent data storage can optimize organizations' costs and lower their climate footprint

Title	Source	Year
Tape to Play Critical Roles as the Zettabyte Era Takes Off	Tape Storage Council	2022
Accelerating Green Datacenter Progress with Sustainable Storage Strategies	International Data Corporation (IDC)	2021
Tiered Storage – Storage Optimization for the Zettabyte Era	Horison Information Strategies	2021
White Paper – Tape and Cloud: Solving Storage Problems in the Zettabyte Era of Data	Phil Goodwin	2019
Tiered Storage – Building the Optimal Storage Infrastructure	Horison Information Strategies	2020
Reducing Data Center Energy Consumption and Carbon Emissions with Modern Tape Storage	Brad Johns Consulting L.L.C.	2020
TCO Calculator for Data Storage – How much can you save by using tape for enterprise backup/archive storage?	Fujifilm	2023
Improving Information Technology Sustainability with Tape Storage	Brad Johns Consulting L.L.C.	2021
The future of tape – white paper	Fujitsu	2022
Magnetic tape is alive and well: Now you can store 580 terabytes on one tape	Ingeniøren	2021

Knowledge Resources | Intelligent data storage can optimize organizations' costs and lower their climate footprint

Title	Source	Year
Effective Data Management Through Active Archives	The Active Archive Alliance	2023
Preservation or Deletion: Archiving and accessing the dataverse	John Monroe, Further Market Research	2023
Tape Becomes a Key Enabler for the Zettabyte Era	Horison Information Strategies	2021
How Tape Technology Delivers Value in Modern Data-driven Businesses	IBM and Fujifilm	2021
A more comprehensive TCO study	Brad Johns Consulting L.L.C.	2019

Bright
ideas.
Sustainable
change.

RAMBØLL