

Using Simulation to Design Scalable and Cost-Efficient Archival Storage Systems

James Byron¹ Darrell D. E. Long¹ Ethan L. Miller^{1,2}
Center for Research in Storage Systems, University of California, Santa Cruz¹
Pure Storage²
{jbyron,darrell,elm}@ucsc.edu

Abstract—The need for reliable and cost-effective data storage grows as digital information becomes increasingly ubiquitous. Archival systems must store valuable data for years while adapting to changing user needs, capacity, and performance requirements. Storage devices differ in terms of performance, capacity, reliability, acquisition cost, power consumption, and the rates at which their features change over time. As a result, choosing the best storage technology to use for an archive has become increasingly challenging with the proliferation of new technologies alongside existing ones.

We have designed a simulator that models the capacity, performance, acquisition cost, and power cost of an archival system using the characteristics of the drives and media that comprise it. We simulate and compare four storage technologies that exhibit different cost and performance characteristics: tape, optical disc, hard disk, and NAND flash SSD. We evaluate the total cost of ownership for each storage technology within an archival system, and we explore the effect that prospective technological advancements and growth rates over time may have on the relative cost and viability of each storage technology for archival systems. We show that the lifecycle and upgrade cost of drives are significant cost factors for removable media archives. We observe that increasing performance requires adding more drives to an archival system, and the cost of each drive dominates the cost to increase performance. We compare trends in storage technologies to suggest developments that could minimize the long-term total cost of ownership for archival systems. We show that hard disks and flash could become cost-competitive with tape-based archives by adopting new designs to minimize infrastructure and electricity costs.

Index Terms—archival storage, simulation, total cost of ownership, performance, power consumption

I. INTRODUCTION

Storage technologies vary widely in terms of their performance, cost, power consumption, reliability, and pace of development over time. With a growing need for low cost data storage in archival systems, the selection of a good archival technology can yield significant long-term cost savings and better performance. We present an archival storage simulator that utilizes trends in storage technology development to predict the long-term cost of archiving data using different storage technologies. We also analyze the best and worst case scenarios for each technology to understand how potential breakthrough improvements might impact archival system design over time.

This research was supported in part by the National Science Foundation under award IIP-1266400 and industrial members of the Center for Research in Storage Systems.

The information age has given rise to large digital storage systems that record vast amounts of valuable information. Demand predictions suggest a 30% *compound annual growth rate* (CAGR) through 2025 in the amount of digital information that must be stored, much of which may have significant financial or personal value [1]. In order to minimize the cost of reliably storing large amounts of data for long periods of time, archival systems must utilize cost-effective and power-efficient storage devices.

We compare the features and cost of tape, optical disc, hard disk drives (HDDs) and solid state drives (SSDs). We show that tape and optical disc become more expensive than hard disk as the performance requirements for an archival system increase. Tape and optical disc archives are most cost-effective in archives with minimal performance requirements. HDDs and SSDs offer high performance at somewhat higher total cost than tape or disc. We show that a simple and low-cost network attached storage adapter for HDDs and SSDs can reduce their cost in an archive. We demonstrate that increasing the longevity of HDDs will slightly reduce their cost in long term archival storage. Optical disc archives may be cost-effective for data that requires infrequent access and few changes. We describe the relationship between slow performance and high power consumption. Finally, we predict the future cost of archival storage as the rate for development slows for each storage technology.

II. APPROACH

The economics of long-term storage, including both the value of data and the cost to store it, is an important factor for long-term data preservation [2], [3]. Archival systems must achieve high capacity, performance, and reliability with a low *total cost of ownership* (TCO). The economic value of archival systems may be expressed as a function:

$$V_{\text{Archive}} = V_d - TCO(d, \text{time}). \quad (1)$$

V_d is the economic value of the data d for the remainder of all time. We assume V_d remains constant in different archival systems; however, different levels of performance can influence V_d . For example, in some situations, V_d may increase if the archive can access data within a certain number of seconds. We leave a study of this to future work. $TCO(d, \text{time})$ is a function that calculates the total cost to store and maintain the data over the period specified by *time*. A positive value for

TABLE I
ARCHIVE PARAMETERS

Parameter	Value	CAGR
Capacity [1]	1 PB	30%
Read Bandwidth	100 MB/s	30%
Write Bandwidth	100 MB/s	30%
Electricity Cost [4]	\$0.11/kwh	1.3%
Data Read Yearly [5]	3–75%	-
Data Modified Yearly	0%	-
Data Scrubbed Yearly	100–1200%	-
Simulation Length	25 years	-

V_{Archive} indicates that the data is worth storing with the archival system, and a negative value indicates that the archival storage system costs more than the data that it stores is worth. We have designed a simulator to compare different storage systems and predict whether they will provide positive value in archival systems.

We compare the suitability of storage technologies for an archive based on their total cost of ownership (TCO). The search for better archival storage technology may be expressed as a function of cost:

$$\max(V_{\text{Archive}}) = V_d - \min_{m \in M}(\text{TCO}_m(d, \text{time})). \quad (2)$$

Here, m is a single storage technology medium in the set M of all archival storage technology mediums at a single point of time. Observe that since V_d remains constant for all storage technologies based on our assumption, we can compare the storage technologies in M without knowing beforehand the value of V_d . We compare different technologies in archival systems by calculating the cost of constructing an archive that stores the data d using each storage technology.

A. Archive Parameters

The *archive parameters* listed in Table I prescribe the minimum requirements that the archival system, which is composed of individual storage devices, must meet or exceed. The parameters also define the cost of electricity, the length of the simulation in years, and the workload for the archive in terms of read, write and data maintenance operations. Data maintenance or scrubbing requires reading all of the data in the archive in order to verify its correctness.

B. Storage Device Parameters

Storage device parameters describe the individual storage devices that are available to use in the archive. The parameters include values for cost, performance, endurance, and power consumption. Storage devices can include hard drives or solid state drives that combine the storage medium and the mechanism to read or write data. It can also include tape or optical disc that require separate storage media and drives to read or write the data. We also include libraries or networked attached storage systems as appropriate to measure and model the cost of automating the archival system and attaching it to a network.

C. Total Cost of Ownership (TCO)

We include the cost of storage drives, storage media such as tapes or optical discs, library systems, and power cost as we calculate the TCO for archival systems. We have not included networking, administrative, cooling, or other infrastructure costs in order to limit the complexity and variability of our simulation model, but we anticipate including more variables in future work. We define TCO with the formula:

$$\text{TCO}(d, \text{time}) = \sum_{t=1}^{\text{time}} (U(d, t) + M(d, t) + P(d, t)). \quad (3)$$

An archive's TCO is the sum of its annual upgrade costs $U(d, t)$, maintenance costs $M(d, t)$, and power costs $P(d, t)$ over the number of years specified with time. Upgrade cost is the capital expense of increasing the archive's capacity and performance by purchasing new drives, media, and libraries. Maintenance is the expense of replacing drives and media that have failed at random or because they have reached the limit of their endurance and worn out. Power cost is the cumulative expense of operating the archive at standby and during read or write operations. We describe how these three functions work in Chapter III.

D. Growth Rate

Some of the factors that contribute to the economic cost of archival storage grow over time. We consider the growth rates of data storage needs, storage device capacity, read and write throughput, the probability of device failure, and electricity costs to simulate changes in archival system and device behavior over time. We use the compound annual growth rate (CAGR) to describe the rate of change based on historical data. We calculate CAGR with the following formula:

$$\text{CAGR} = \left(\frac{\text{val}_i}{\text{val}_0} \right)^{\frac{1}{(\text{time}_i - \text{time}_0)}} - 1. \quad (4)$$

val_0 and time_0 are the parameter's initial value and the starting time, and val_i and time_i are the parameter's final value and time, respectively. We use the CAGR formula to determine growth rates for parameters in our simulations since it calculates the annual growth rate needed to reach val_i , starting with val_0 , after $\text{time}_i - \text{time}_0$ years. CAGR has been used previously to describe changes in storage technology. Kryder's Law, which describes the growth of hard disk areal density, predicts a CAGR of 58% so that HDD capacity would double every 18 months [6]. Hard disk capacity has increased more slowly in recent years [2], [7]. We use CAGR in our simulations to update parameters for each year using the formula:

$$\text{val}_i = \text{val}_0 \times (\text{CAGR} + 1)^{\text{time}_i - \text{time}_0}. \quad (5)$$

val_0 and time_0 are the starting value and time, and val_i is the value after a growth rate of $\text{CAGR} + 1$ for $\text{time}_i - \text{time}_0$ years.

Archival System Components

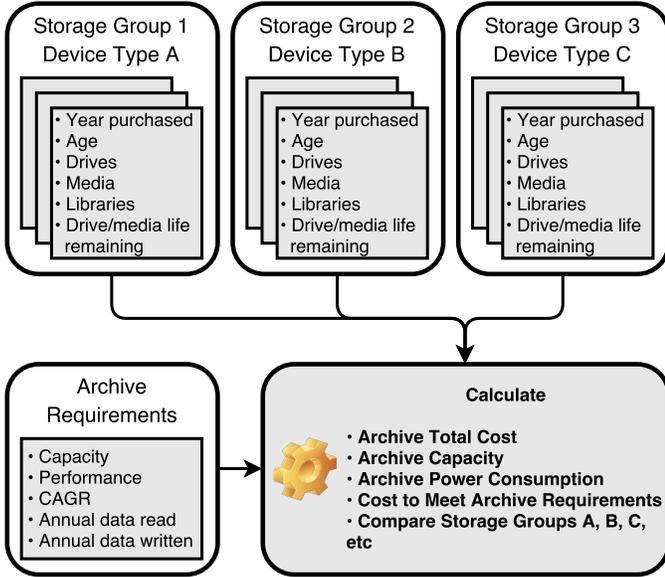


Fig. 1. An archive consists of one or more *Storage Groups*, each of which contains a list of storage devices such as drives, storage media, and libraries. *Storage Groups* also record the age and usage history of each device in order to calculate when each device has worn out and must be replaced.

III. IMPLEMENTATION

Our simulator calculates the capital and operating costs for an archival system over time. The simulated archival system fulfills a set of requirements called *archive parameters* for capacity and throughput. We include simulated storage devices for tape cartridges and tape drives, optical discs and optical drives, hard disks, and solid state drives. We compare the cost of acquiring and operating archival systems that use one of these storage technologies. We run our experiments to simulate 25 years of operation, which is long enough to show the divergence of long-term trends of different storage technologies.

The simulator consists primarily of two classes: the *Archive* class and the *Storage Group* class. The *Archive* class simulates the actions of a data center that may combine several storage technologies to store archival data. The *Storage Group* class simulates a single storage technology, including the drives, media, and libraries used to store data.

A. Archive Class

The *Archive* class represents the entire archival system that contains one or more types of storage. Each simulation uses one instance of the *Archive* class to represent a data center. The *Archive* class can combine one or more storage technologies to meet the requirements for the archival system that are defined in the *archive parameters*. The *Archive* class also calculates the total cost, performance, and power consumption of the archive by querying each individual *Storage Group*.

B. Storage Group Class

The purpose of the *Storage Group* is to represent a single type of storage within the archive and to record data about each device. The *Archive* class creates one *Storage Group* for each storage technology in the simulation so that each *Storage Group* represents one storage technology within the archive. Storage technologies consist of storage drives, media, and library systems or network attached storage devices to manage the drives and media. *Storage Groups* for removable media like tape and optical disc record the drives and media separately. HDDs and SSDs, on the other hand, are represented only as drives since they do not have removable storage media. *Storage Groups* also contain information about the age and usage history of each device.

Storage Groups provide functions to write or read data from storage devices, calculate power consumption, and calculate the cost to purchase more drives or media to increase the *Storage Group's* capacity or total performance. When adding new drives or media, *Storage Groups* also calculate the number of library systems or network attached storage devices needed to manage the new drives and media. Tape cartridges and optical discs can only be written and read by a compatible drive. For this reason, *Storage Groups* may add new tape or optical drives in order to read and write new tape cartridges or optical discs. *Storage Groups* record the amount of data that has been read or written to each drive and storage medium. Devices that have reached their read or write endurance limits are said to have reached the end of their useful lives, and they are removed from the *Storage Group*. We also use each device's failure probability to select devices at random to fail within each *Storage Group*. The failure probability simulates the chance that storage media or drives will fail during a year, and the growth rate for failure probability simulates the increasing probability of device failure as time passes.

C. Calculate Cost to Meet Requirements

The *Archive* class calculates the required total capacity, read bandwidth, and write bandwidth for the archive using the *archive parameters* for capacity and bandwidth, their growth rates, and the current simulation time. The *Archive* class then calculates how much additional capacity, read bandwidth, or write bandwidth is needed to meet the requirements.

$$Inc_p = \min(p) - \sum_{s=1}^{SGroups} (GetValue(p, s)) \quad (6)$$

To determine the required increase for each archive parameter Inc_p , we calculate the difference between the current minimum for the parameter, $\min(p)$, and the total of the parameter in all *Storage Groups*, $SGroups$. The set of p parameters includes the archive's capacity, read bandwidth, and write bandwidth. The *Storage Groups* use the values Inc_p to calculate the cost to meet the archive's requirements. The *Storage Groups* use the functions $U(d, t)$, $M(d, t)$, and $P(d, t)$ to upgrade capacity and bandwidth, replace any devices that have failed, and calculate power consumption to determine their total upgrade cost.

1) *Upgrade Cost*: The upgrade function calculates the cost of purchasing new drives and media to store additional data or to meet minimum performance requirements for the archive.

2) *Maintenance Cost*: Maintenance replaces any failed or worn out devices with new devices of the same type.

3) *Power Cost*: The function to calculate power cost measures the total active, standby, and idle power consumption of the *Storage Group* and uses the current cost of electricity to determine the total power cost during a simulation year. Active power consumption is the power used while a *Storage Group* reads, writes, and scrubs the archive’s data. Idle and standby power consumption are the electricity used when a *Storage Group* is not active.

IV. ARCHIVE PARAMETERS

The *archive parameters* describe the minimum requirements, environment variables, and usage patterns for the simulator. Storage systems in general support a wide range of usage needs. Storage systems differ by the type of workload that they support. High intensity workloads are considered *hot*, while low intensity or archival workloads are considered *cold*. High performance workloads may require high throughput and low latency to support many simultaneous users. Cloud services frequently support high intensity workloads as thousands of remote clients may simultaneously access databases to retrieve and store information, small files to load websites, and large files to stream movies. Low intensity workloads may exhibit intermittent access patterns to data, and a few users may perform most of the storage system’s operations. Low intensity workloads may include daily backups of important databases, background data maintenance tasks like scrubbing or deduplication, or a peripheral online service that typically serves a few simultaneous clients at most. Users may not access cold data storage frequently, but the data stored in such systems is no less important or sensitive to the performance of the storage system. For example, database backups for a large organization are infrequently read after they are written to the backup server. If the organization’s database needs to be restored from the backups, the performance of the cold storage system that stores the backups becomes an important factor in how quickly the database can be restored. As the last example showed, performance requirements and intensity are orthogonal: some cold data still requires low latency and high throughput.

The definitions for cold storage and archival storage suggest a data set that may be accessed less frequently than hot or high intensity storage systems; however, there is no exact definition for an archival storage system. Previous studies have analyzed storage system behavior under different workloads, including both archival and high intensity workloads [5], [8], [9]. The challenge of rigorously characterizing storage system workloads does not outweigh the importance of using workload characteristics to understand how best to model and design storage systems [10].

Archival systems differ from other storage systems in terms of their typical workloads. Archival data may be written only

TABLE II
AVERAGE US ELECTRICITY COST TO COMMERCIAL USERS [4]

Year ^a	Price per kwh	CAGR to 2017 ^b
2017	\$0.1055	-
2016	\$0.1032	2.23%
2015	\$0.1030	1.21%
2014	\$0.1052	0.095%
2013	\$0.1012	1.05%
2012	\$0.0989	1.30%
2011	\$0.0998	0.93%
2010	\$0.0999	0.78%
2009	\$0.1026	0.35%
2008	\$0.1050	0.05%
2007	\$0.0979	0.75%

^aValues taken from the January report of the following year.

^bWe calculate the CAGR from the given year to the price in 2017.

once, and much of the archive’s data may never be read other than to verify its correctness. The process of verifying the data on storage devices is called *scrubbing*. Workloads for scrubbing vary depending on the physical conditions of the storage system, the probability that the storage devices will gradually and detectably decay, and the importance of the data. We select baselines for data read, modified, and scrubbed based on previous research that characterized the workloads of archival systems [5].

The set of requirements include the archive’s minimum capacity, total read bandwidth, and total write bandwidth. It also includes each parameter’s compound annual growth rate (CAGR), which we use to upgrade the archive as time passes. In general, we anticipate that archival systems will need to increase their capacity and throughput over time, and we choose CAGR parameters that match historical trends. Environment variables include the cost of electricity, its CAGR, and the number of years to simulate the archive’s operation. Finally, usage patterns control the workload that the archival system will support. The workload values are the portion of the archive’s data that must be read, modified, or verified during each each year of the simulation. We use the workload values to ensure that the read and write throughput values will be sufficient to complete all data operations during each simulation year.

A. Archive Power Cost

The cost of power is an important factor to the total cost of ownership for virtually any storage system, including archival storage systems. Electricity cost varies widely by region, country, and locality. In our simulations we assume that electricity is available and reliable. We determine the cost of electricity using data from the US Energy Information Administration’s monthly survey of electricity prices by state [4], and we calculate the CAGR for electricity by determining the price increase over a ten year period.

The price for commercial electricity has grown slowly over the last decade as shown in Table II. The future price will vary depending on the cost of generating electricity. The price

of electricity has increased at an annual rate of 0.75% since 2007, but we use the more pessimistic CAGR of 1.3% for electricity cost from 2012 to accommodate for the possibility that future increases in energy costs will outpace the CAGR from the last decade. We run additional experiments varying electricity costs in Section VI-G to explore the effect of energy cost in archival storage systems.

B. Archive Capacity

We arbitrarily set the initial capacity requirement for the archival simulator to 1 petabyte or 1,000 terabytes. The parameter for capacity minimum enforces the constraint that the archival system must store 1PB of data as the simulation begins. The capacity parameter increases as the simulation progresses. The arbitrary initial capacity of 1PB demonstrates the cost of each storage technology as it scales to accommodate added capacity over time. Estimates vary of the annual growth rate for datacenter capacity. An annual increase of 50% would lead to a doubling of archival data every 21 months; however a recent estimate provided for Seagate predicts a growth rate of 30% for data storage needs during the years 2015 to 2025 [1]. We use this value for capacity growth in our simulations.

The minimum values for read and write throughput are also arbitrary; however, the simulator automatically increases the minimum throughput requirements if they are not sufficient to perform all required read and write operations during a simulation year.

V. STORAGE DEVICES

A. Tape Drives and Media

Tape is a widely adopted archival storage technology. Tape storage has been used in the music and audio recording industry to reliably store analog data for decades. In audio applications, the sequential nature of audio media lends itself to the lengthy tape medium that winds a tape between two reels to read and write information. Archival tape uses the same underlying technology as audio tape; however the need for vast amounts of archival storage has led to significant advancements in tape storage technology. We base the configuration parameters of tape drives on the StorageTek LTO-8 drive [16]–[21], and we use the HPE LTO Ultrium tape cartridge [22]–[24] to set parameters for tape media. We consider archival tape within a scalable data center environment where power consumption, capital cost, and reliability affect the total cost of ownership for the archival system.

Tape-based archival systems require at least one drive to read and write to tape media, and the drives and tapes are stored within a library system that robotically loads and unloads tapes from the drives as users read or write data. We measure the capital cost of the archival tape system by calculating the purchase cost for the required number of tapes, drives, and library systems. Next, we use the specifications of each device to determine its parameters for read and write throughput, latency, power consumption, and durability. The LTO Consortium, which creates and licenses specifications for tape media and drives, expects that tape capacity will continue

TABLE IV
LTO TAPE CAPACITY AND SPEED, 2000 TO 2017

Year	Capacity	Speed	CAGR Since 2000	
			Capacity	Speed
2000 [26]	100 GB	20 MB/s	-	-
2002 [26]	200 GB	40 MB/s	41.4%	41.4%
2005 [26]	400 GB	80 MB/s	32.0%	32.0%
2007 [26]	800 GB	120 MB/s	34.6%	29.2%
2010 [26]	1500 GB	140 MB/s	31.1%	21.5%
2012 [27]	2500 GB	160 MB/s	30.8%	18.9%
2015 [28]	6000 GB	315 MB/s	31.4%	20.2%
2017 [17]	12000 GB	360 MB/s	32.5%	18.5%

TABLE V
STORAGE MEDIA PARAMETERS

Item	Tape [22]		Optical Disc [12]	
	Value	CAGR	Value	CAGR
Cost	\$175	0%	\$20.83	0%
Capacity	12 TB	32.5%	300 GB	17.1%
Read Endurance	200	-	∞ (est)	-
Write Endurance	200	-	1	-
Fail Probability	0.011%	-	-	-

to increase. Currently, the roadmap from LTO Ultrium forecasts that tape capacities will increase to 192 TB [25].

Tape drives may read or write tapes that were purchased for an older tape drive model; however, tape drives may deprecate their support for older tapes after two or three generations. We use tape and drive compatibility to simulate the cost of upgrading to a new tape drive system when a new technology replaces an older one.

Tape library systems store a specified number of tapes and drives. Library systems in general require a base unit, which includes a robot, drive bays, and slots for the media. The library may also support several expansion packs, each of which can support additional drives and media. The expansion units cost less and require less electricity than the base unit. We derive library parameters from the StorageTek SL-150 [11], [16], [18].

When constructing a simulated archive using tape, our simulator first determines how many tapes are needed to meet the capacity requirements and how many drives are required to meet the read and write bandwidth requirements for the archival system. Next, the simulator selects the minimum number of libraries and library expansion units to store the tapes and drives. The simulator totals the cost for each of these devices. Finally, the simulator determines the expected annual power consumption for the library system and adds the power cost to the capital cost for the tape archival system.

Tape archives read and write data to serve user requests, and the archive verifies or scrubs every tape in the archive to ensure it remains readable. The simulator records the reading, writing, and scrubbing operations in terms of power consumption, active time, and wear on the media and drives. Tape drives and media that reach their maximum read or write endurance are considered failed, and they are removed from the pool of

TABLE III
STORAGE DRIVE PARAMETERS

Item	Tape Drive [11]		Optical Drive [12]		Enterprise HDD [13]		Archive HDD [14]		SSD [15]	
	Value	CAGR	Value	CAGR	Value	CAGR	Value	CAGR	Value	CAGR
Drives and Media	separate	-	sep.	-	combined	-	comb.	-	comb.	-
Drive Cost	\$8300	0%	\$12000	0%	\$500	0%	\$200	0%	\$500	0%
Capacity	0	-	0	-	14 TB	32%	8 TB	32%	2 TB	44%
MB/sec Write	300	18.5%	60	6.2%	233	15%	150	15%	500	44%
MB/sec Read	300	18.5%	60	6.2%	233	15%	150	15%	530	44%
Active Watts	29.5	-	20.75	-	6.4	-	7.5	-	0.15	-
Idle Watts	9	-	7 (est)	-	5.2	-	5	-	0.11	-
Standby Watts	9	-	7 (est)	-	1.2	-	1	-	0.03	-
Drive Lifetime (est)	30	-	30	-	5	-	5	-	10	-
Fail Probability (est)	1%	10%	1%	10%	2%	10%	2%	10%	2.1%	10%
Write Endurance (est)	1000000	-	∞	-	∞	-	∞	-	200	-
Read Endurance (est)	1000000	-	∞	-	∞	-	∞	-	∞	-
Sec to Load Media	59	-	59 (est)	-	0	-	0	-	0	-
Sec to Get Data	15	-	5 (est)	-	5	-	5	-	1 ms	-
Generation (years)	2	-	5-10	-	1	-	1	-	1	-
Reverse Compatibility	1	-	∞	-	-	-	-	-	-	-

TABLE VI
OPTICAL DISC CAPACITY AND SPEED, 1982 TO 2017

Year	Capacity	Speed	CAGR Since 1982	
			Capacity	Speed
1982 [31]	0.7 GB	7.8 MB/s	-	-
1996 [32]	8.5 GB	30.2 MB/s	10.1%	16.8%
2006 [33]	50.0 GB	60.1 MB/s	8.9%	15.0%
2016 [12], [30]	150.0 GB	60.0 MB/s	17.1%	6.2%

TABLE VII
HARD DISK CAPACITY, 2005 TO 2017

Year	Capacity	CAGR since 2005
2005 [36]	500 GB	-
2007 [37]	1000 GB	41.4%
2009 [38]	2000 GB	41.4%
2011 [39]	4000 GB	41.4%
2017 [40]	14000 GB	32.0%

active storage devices.

B. Optical Disc

Optical disc was released in 1982. The physical format of discs has scarcely changed in the decades since its initial release. Optical discs feature greater reverse compatibility than tape drives. Similar to tape, optical disc has been used in large scale archival systems for cold storage [29]. Since 1982, optical disc has increased in capacity at a CAGR of 17.1%. Modern discs are also double sided, which increases the current maximum capacity to 300 GB. The performance of optical discs has increased more slowly than other storage devices. Panasonic plans to increase optical disc capacities to 500 GB and 1 TB [30]. Archival optical discs feature a long lifespan of at least 50 years, and optical discs promise more durability and require less environmental control than other devices.

Panasonic's Freeze-ray archival system was developed in conjunction with Facebook to minimize migration, storage media replacement, and power costs. We use the Panasonic Freeze-ray to model optical drives, discs, and libraries [12], [16], [29], [34], [35].

C. Hard Disk Drive

Hard disks differ from tape and optical disc in several ways. First, HDDs do not separate the storage medium from the read and write mechanism. Hard disks do not need to physically load a storage cartridge into the drive before reading data.

When used in a large storage system, hard disks support much greater parallelism than either tape or optical disc because each drive can read and write data to itself independently of any other drive. Storage systems that use tape and optical disc, on the other hand, must queue user requests until the drive is available to read or write data to the storage medium.

While hard disks feature higher throughput and lower latency than tape or optical disc per byte of storage, they also consume more power. Tapes and discs consume virtually no energy when they are not being used; but each hard disk consumes a small amount of power during standby as well as when it is actively reading or writing data. We use baseline parameters from two hard drive models. The HGST HS14 [13], [41] features a mean time between failures (MTBF) of 2.5 million hours, and the Seagate Archive HDD v2 [14] features an MTBF of 800,000 hours.

The cost of long-term storage varies depending on the underlying storage technology. Storage density and cost per byte of storage are inversely proportional. As storage devices increase in bit density and capacity over time, the cost per byte of storage decreases concurrently. The availability of solid state drives as a faster and more energy-efficient alternative to hard disks may encourage hard drive manufacturers to focus on increasing capacity over performance since capacity, not latency or throughput, is the key advantage of hard disks over SSDs. Google has proposed designing hard disks in new form factors or with relaxed reliability requirements to maximize

their utility in data centers [42].

Previous studies have observed that hard disk capacity has recently increased more slowly than Kryder’s Law predicted [2], [7]. Similarly, the cost per gigabyte of storage on hard disk is not decreasing as quickly as it once did since the cost per gigabyte of storage is inversely proportional to the areal density of hard disks. HDD capacity has increased at a CAGR of 32% between 2005 and 2017; however a study of hard drives from 1981 to 2017 showed that the cost per gigabyte of storage dropped from \$500,000 in 1981 to less than \$0.03 in 2017, an annual decrease of 37% [41]. Both trends suggest a capacity growth rate that is considerably below the 58% annual growth rate anticipated with Kryder’s Law. Increasing the capacity of hard disks requires significant capital investment to develop and deploy new technologies with greater bit densities [7], [43], and therefore future increases in hard disk capacity will require significant capital investment by HDD manufacturers.

Hard drive reliability varies depending on each drive’s make, model, capacity, operating environment, and workload. In general, we expect that a drive’s probability of failure will increase as it ages and as it is used more. Backblaze, a cloud backup service provider, reports data about the longevity of hard drives in their data center [44]. The most recent report shows an annualized failure rate of 2.0% across all hard drive models. Backblaze replaces drives when they fail or when the cost of purchasing a new drive is less than the cost of operating two to four older drives that have lower capacities.

D. Solid State Drive

Solid state drives use less power per drive than tape drives, optical drives, or HDDs; however, they also cost more per byte of data storage. As with other storage devices, the cost of storing data on SSDs decreases as the density of data storage increases. Recent reports suggest that SSD cost-per-gigabyte decreases at 44% each year [45], [46]. The lifespan of SSDs depends on the environment in which they are used. NAND flash SSDs have no moving parts like other types of storage. Instead, SSDs write data by electrically changing cells that are arranged into large erase blocks. The lifespan of the drive is not limited by the physical wear-out of moving components. Instead, the program-erase cycle exerts wear on the NAND cells, and they can endure a limited number of program-erase cycles before they can no longer reliably store data [47], [48]. Triple layer cell (TLC) NAND flash, which stores three bits of data per flash cell, has a P/E cycle limit on the order of 1000 to 5000. In an archival system, the number of write operations is small compared with the number of scrubbing or user read operations. We conservatively set the write endurance of SSDs to 1000. We set the life expectancy of SSDs to 15 years since, having no moving parts, SSDs can outlast hard drives in archival systems [7]. We derive SSD parameters from the Micron 1100 [15], [45], [46], and Micron claims a MTBF of 1.5 million hours for this drive.

Numerous storage technologies have been optimized for archival storage systems. Tape, optical disc, and some HDDs

TABLE VIII
ARCHIVAL LIBRARIES AND NAS

Item	Tape ^a	Optical Disc ^b	HDD/SSD	
			Type A ^c	Type B ^d
Cost, base	\$7000	\$17250	\$10000	\$100
Cost, expansion	\$1000	8750	\$2500	-
Active watts, base	38	25	350	10
Active watts, exp.	1	12.5	10 (est)	-
Idle watts, base	35	12.5	10 (est)	1 (est)
Idle watts, exp.	1	1 (est)	10 (est)	-
Max drives, base	2	3	12/24	1
Max drives, exp.	2	3	12/24	-
Max media, base	30	912	-	-
Max media, exp.	30	912	-	-
Max exp. modules	14	6	7	0
Typical lifetime	30 (est)	50 (est)	10 (est)	10 (est)

^aBased on the StorageTek SL-150 [11], [16], [18].

^bBased on the Panasonic Freeze-ray [12].

^cA high performance NAS based on the HPE MSA 2050 SAN [49], [50].

^dA prospective low-cost NAS device based on the Banana Pi [51].

have been designed for intermittent use in archival storage systems. SSDs have been more expensive on a cost-per-gigabyte basis than other technologies, but continuously declining prices may make SSDs useful for some archival purposes. We propose a design change for SSDs that could decrease the cost of using them in archival storage systems. The Micron 1100, which is the baseline for our consumer SSD, features a capacity of up to 2TB. We propose an archival SSD that features an initial capacity of 8TB on a single drive but with the same cost-per-gigabyte as the 2TB drive. The high-capacity SSD would require more NAND flash chips than the consumer SSD, but it would likely require no additional circuitry for controllers compared with lower-capacity drives. Hereafter, we refer to the baseline SSD that is based on the Micron 1100 as the consumer SSD, and we refer to the proposed 8TB SSD as the high-capacity SSD.

E. Network Attached Storage

Hard disks and solid state drives can be connected to an archival storage system using networked attached storage (NAS) devices. Drives like the Micron 1100 and the HGST HS14 support a SATA or SAS interface. The NAS device connects many drives to a single network through which users access the individual storage devices. We base our high-performance NAS devices on the HPE MSA 2050 SAN [49], [50], which can support up to 12 HDDs or 24 SSDs per module.

NAS devices that are optimized for performance consume comparatively more power and cost more than slower devices. Today, scalable NAS devices target high-performance workloads instead of low-cost, power efficient archival workloads; however, simpler devices like the Banana Pi M2 Berry [51] can be used as a NAS device for a single HDD or SSD with a SATA interface. Integrating such a device into a single archival storage system relies on software to manage the redundancy and distribution of data across the storage devices.

TABLE IX
PROJECTED LIMITS OF STORAGE DEVICE CAPACITY

Technology	Roadmap	Years To Reach
Tape Cartridge [25]	192 TB	10
Optical Disc [30], [55]	1 TB	8
Archive HDD [56]–[58]	100 TB	10 to 18
High-Cap. SSD [57], [59]–[61]	100 TB	7

DAWN [52], Pergamum [53], and Pelican [54] propose tightly integrating hardware with software to create inexpensive and power-efficient arrays for archival storage. We compare the high-performance NAS with an imagined device similar to the Banana Pi and costing \$100 for acquisition, networking, and mounting hardware. While no device currently exists that can satisfy all the requirements of an archival system, we hope that our results will motivate hardware developments to accommodate cost-effective archival storage using hard disks and solid state drives.

Hereafter we refer to the high-performance HPE MSA 2050 SAN as the high-performance NAS device, and we refer to the imagined device that is optimized for archival storage as the low-cost NAS device.

F. Predictions of Slowing Development

Each of the storage technologies that we study has increased in capacity and performance in recent years as a result of development. Manufacturers invent new technologies and incorporate advancements into successive generations of storage devices to increase capacity and performance. Tape drives and media utilize higher-capacity tape with greater bit density and more tracks on each tape [25]. Optical discs now utilize multiple layers of data on each side of an optical disc to increase capacity, and greater areal bit density helps to improve read and write throughput [30], [55]. Hard disks provide higher capacity and throughput by increasing areal bit density on the hard disk platters through the use of technologies such as perpendicular magnetic recording (PMR) or heat-assisted magnetic recording (HAMR) [56]–[58]. Solid state disks rely on several techniques to increase capacity [57], [59]–[61]. NAND flash cells can be made smaller through the use of new manufacturing processes. Each flash cell can also store two, three, or four bits by adopting MLC, TLC, or QLC flash technology. NAND flash can be constructed in three dimensions to create multiple layers of flash on a single silicon wafer, and finally, individual wafers can be stacked to combine multiple wafers in a single NAND flash chip. SSD controllers also improve over time to provide better performance. Manufacturers of each storage technology continue to search for new ways to increase storage device capacity and performance or to reduce costs.

Storage device manufacturers introduce new devices with greater capacity or performance after years of research and development. Each advance in storage density or increase in performance requires significant investment from the manufacturer, and the future of developments for each technology

depends on its physical limits. Tapes may be limited in terms of capacity by the reliability of the magnetic material on the tape itself. The capacity of optical discs may be limited by the light frequency that can be used to read or write the disc’s data. Hard disks are limited by the size and number of the magnetic particles on the hard disk platters, and solid state disks are limited in how small the flash cells can be made while still reliably storing an electrical charge. As any storage technology approaches the physical limits of the storage medium, incremental improvements to capacity or performance may become increasingly difficult for manufacturers to achieve.

The limits of capacity and performance for storage technologies have been widely discussed [25], [30], [55]–[61], and while the history of each storage technology has yielded consistent growth of both capacity and performance, the future of continued advancements is not certain. Still, device manufacturers have provided guidance on their roadmap for prospective advancements in technology. Table IX lists the current roadmap for each storage technology’s maximum capacity. We also list the number of years that it will take to reach the roadmap capacity at the current rate of capacity growth for each technology. Seagate’s roadmap for future increases in hard disk areal density [58] includes a range between 15% and 30% CAGR. We use 15% as our pessimistic estimate and 30% as our optimistic estimate of hard disk capacity growth. The roadmap for each technology may change and extend over time since further advancements beyond the current roadmap projections are possible.

In Section VI-I, we use the current roadmap for each storage technology in order to explore the effect that physical limits may have on the relative value of each storage technology. We use the current or expected annual growth rate of each device’s capacity to determine the year in which the technology is expected to reach its roadmap capacity. In the years after the technology reaches its roadmap capacity, the growth rate for capacity decays exponentially. We also use the same decay of growth rate for read and write throughput because throughput is limited by the density of each storage device.

VI. EXPERIMENTAL RESULTS

A. Experimental Setup

We measure the cost of an archival system as the cumulative total of the capital costs for drives, media, and libraries, plus electricity cost. All our results are within a 2% margin of error. Each experiment is a single execution of the simulator, and the parameters remain constant during the experiment except for those parameters that grow over time. We vary only one parameter or device at a time to compare the effects of each change. Each simulation is executed with only one device, and we compare devices with separate runs of the simulator.

During the simulation, individual storage devices may fail at random based on the failure probability set in the archive parameters. Failure probabilities grow linearly over time as each device ages. Storage devices may also wear out when they have reached their maximum age or when they have endured their maximum number of read or write operations.

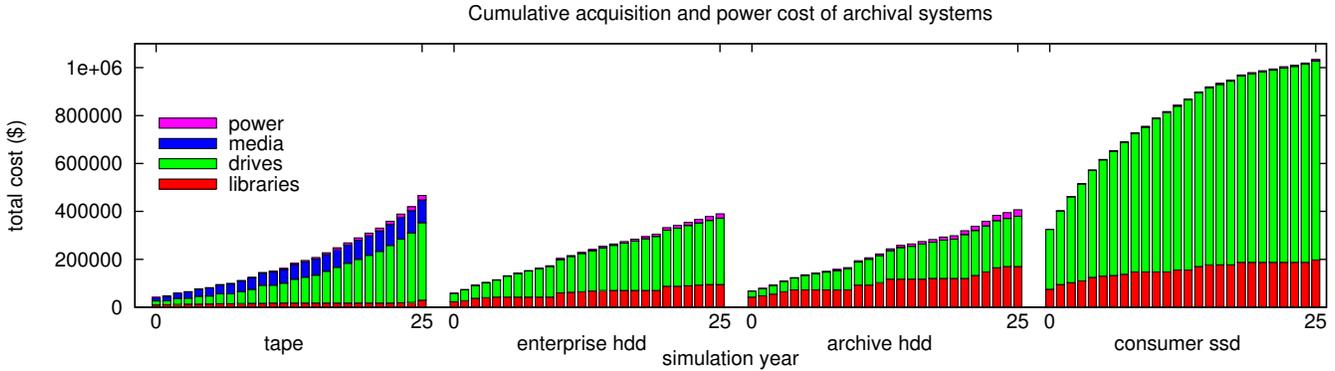


Fig. 2. The cost of archival storage depends on the cost, throughput, and CAGR for capacity and performance of the storage devices. The cost of electricity and power consumption of the archive also impact the archive’s total cost of ownership. Here we use the high-performance NAS for HDD and SSD devices.

The capacity and performance of new drives and media grow as time passes in the simulation. Each parameter grows linearly based on its CAGR; however, in some experiments, we use non-linear growth rates to understand the effects of a declining rate of technology development.

B. Growth of TCO

Figure 2 shows the cost of archives using tape, HDD, or SSD storage devices. The size of the archive grows by 30% each year. Simulation year 0 shows the initial cost of purchasing the drives, media, and libraries needed to store 1 PB of data, and each subsequent column is the archive’s cumulative cost after operating for a year and growing by 30%. The capacity of new tapes, HDDs, and SSDs grows by more than 30% each year in these simulations, and therefore the required number of new storage devices each year decreases as the rate of growth for capacity outpaces the growth of the data in the archive. Old tape drives do not read or write tape media from a newer generation, and new drives are reverse compatible with two generations of older tape. Archives using tape must add new drives in order to use new generations of tape media. Tape archives become more expensive than archives with hard disks because new tape drives are more expensive than hard disks.

The cost of libraries stabilizes within a few years for tape archives because the library systems continue to accommodate new drives and media throughout the simulation. New tape media feature higher capacity than old tape media, and the simulator replaces old media and drives with new media and drives rather than adding more libraries because the cost of replacing the old devices with new ones is less than the cost of acquiring more libraries to accommodate new tape drives and media. The cost of network attached storage for HDDs and SSDs grows as the simulation progresses because, based on our assumptions, the high-performance NAS devices last 10 years.

C. Archival Network Attached Storage

One of the main contributors to the acquisition cost of an archival system is the library for tape and optical disc or net-

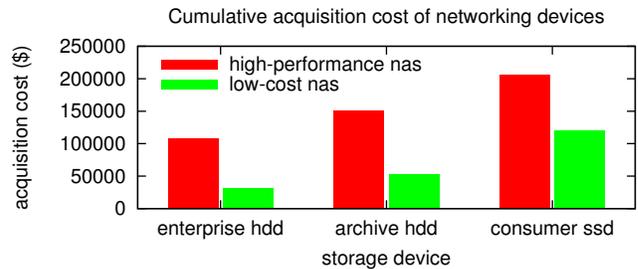


Fig. 3. The acquisition cost of the network attached storage devices depends on the capacity of the drives. The cost of the high performance NAS is also higher than that of the proposed low-cost NAS device.

worked attached storage for HDDs and SSDs. Many network attached storage systems are designed for high-performance instead of low-cost or low power consumption. In Figure 3, we show the cost of high-performance NAS devices with the low-cost devices described in Section V-E to demonstrate the potential cost advantage to archival systems of designing a network attached storage system that is optimized for archival storage.

D. Archival HDDs

Hard disks may be used in a variety of storage systems, including both hot storage for high-demand workloads and cold storage for archival systems. In archival systems, a failed hard disk results in a loss of capacity that must be replaced. Also, hard drives fail in increasing numbers as they age. Hard drive manufacturers may consider increasing the reliability of their hard drives for certain applications. We varied the failure rates and lifetime of HDDs in an archival system to show the effect of better drive reliability on the total cost of ownership for an archival system with HDD storage. Figure 4 shows the effect of varying the annual failure rate of hard disks from 2% to 1% and of increasing the lifetime of hard disks in the archival system. Our experiments showed that increasing HDD reliability does decrease the TCO for archival storage, but the variance was less than 3% in each case. We conclude that

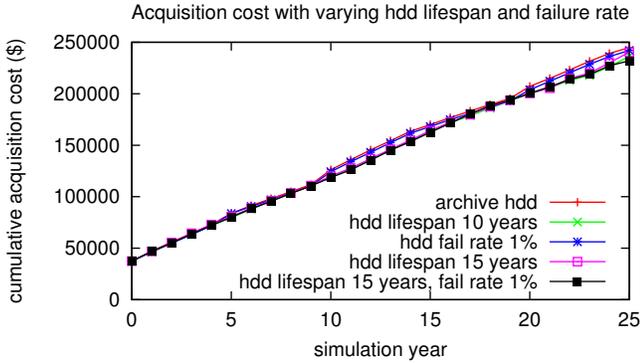


Fig. 4. The total acquisition cost of an archive using archival hard drives with different assumptions about the lifespan and the annual failure rate of the hard drives.

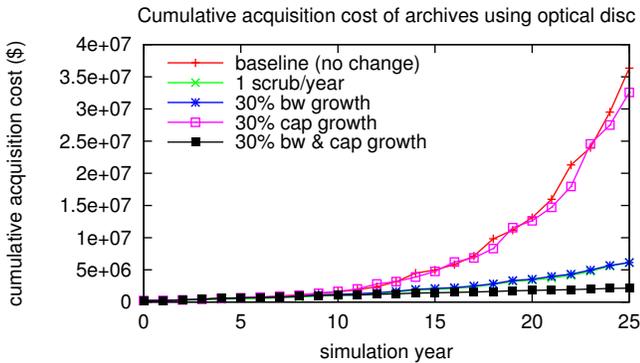


Fig. 5. The cost of optical disc depends on our assumptions about workload and rates of development. Reducing the scrubbing workload can significantly reduce the total acquisition cost.

improving HDD reliability could reduce the TCO for HDD archives so long as the increased cost of manufacturing more reliable hard drives is less than possible cost savings of doing so.

E. Optical Disc Archives

Figure 5 shows the cost of an archival system based on optical disc technology. Disc capacity and throughput grow more slowly than other technologies. As the simulated archival system grows, the minimum throughput needed to scrub the archive’s data 12 times a year increases and quickly requires more drives to increase throughput. Also, new optical discs can only be read by drives of the same generation or newer, and the simulator adds new drives to the archive to support the additional capacity of new optical disc media. In the simulation labeled *No Changes*, we use the same configuration as we did for the devices in figure 5. The cost for optical disc-based archival storage grows exponentially. We found that reducing the workload on the disc archive by reducing the number of annual scrubbing operations from 12 to one results in a cost savings of 80%. Scrubbing is needed to verify the data in an archive, but if discs are sufficiently reliable compared with other storage devices, there will be little or no added risk

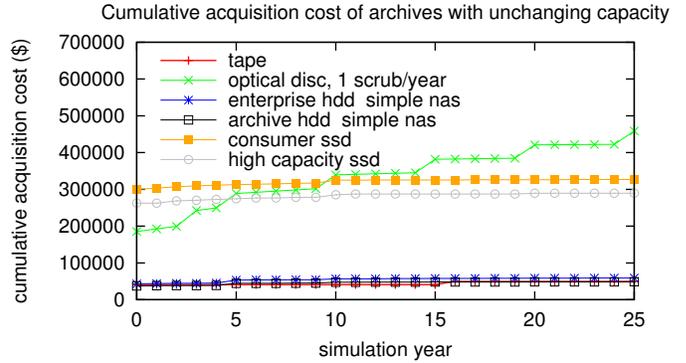


Fig. 6. Replacing media and drives increases the cumulative capital cost of storing 1 PB of data for 25 years, modifying 10% each year, and scrubbing 12 times a year.

of data loss if scrubbing operations occur more infrequently. We also present the effect of optimistic assumptions about the growth rate of disc’s capacity and bandwidth. A 30% CAGR for disc throughput would also result in an 80% cost savings. We did not observe a significant cost savings if only the capacity of optical disc increases by 30% annually. We attribute the importance of bandwidth over capacity to the fact that optical disc media are much cheaper than optical disc drives. The cost of the optical disc archive was in large part compounded by the high price of optical drives. For this reason, reducing the archive’s workload or technological developments to increase optical drive throughput can reduce the number of drives needed in the archival system and result in a cost saving of 80% or more.

Figure 6 shows the capital cost of an archive with each technology if the archive’s capacity does not grow over time. This experiment includes modifying 10% of the archive’s data each year, reading 75%, and scrubbing the archive 12 times. Optical disc, which is a write once read many (WORM) medium, must replace media every time data is modified. The simulator adds a new optical drive when a new generation of optical media comes out because old optical drives do not read new optical discs that have higher capacity. Archives using other storage technologies must replace drives and media as they fail, and the number of drives and media needed decreases as the capacity of each drive or media increases.

F. Power Consumption

In Section V-D we proposed an SSD with higher capacity than consumer-oriented SSDs while using the same form factor and interface. In our example, a high-capacity SSD would offer a capacity four times greater than the consumer SSD, and the high-capacity SSD would require fewer drives to store the archives data and fewer NAS devices to attach the drives to a network. Our experiments also demonstrate that a high-capacity SSD uses less total electricity than a consumer-oriented SSD because of the fewer number of NAS devices. Figure 7 shows the annual electricity consumption for an archive with consumer SSDs and high-capacity SSDs and

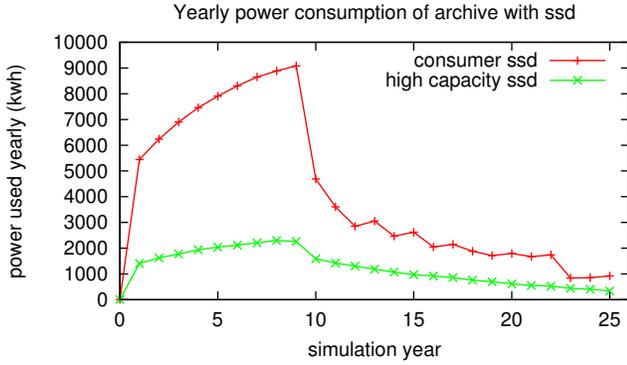


Fig. 7. Total power consumption includes the power used by the drives and NAS devices. The archive uses fewer drives and less networking infrastructure when using high capacity SSDs. The annual power consumption declines dramatically after 10 years because older SSDs are retired and replaced with fewer drives that have much higher capacity.

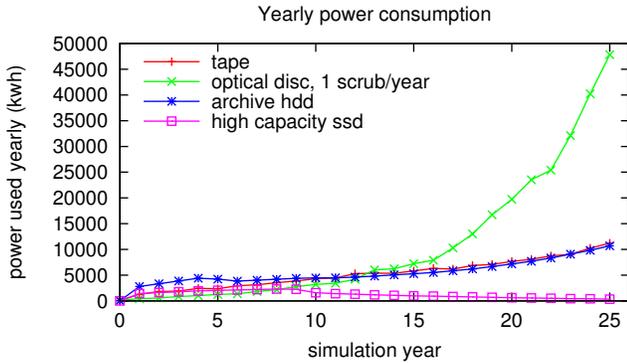


Fig. 8. Archives with removable media use the least power initially. Tape and disc archives, which have less bandwidth than HDDs or SSDs, use more active time to verify the archive’s contents, and therefore they use more power than HDDs or SSDs to scrub the archive’s data each year.

using the low-cost NAS devices described in Section VI-C. The electricity consumption is much lower for high-capacity SSDs due to the efficiency of high-capacity SSDs within an archival storage system. Also, the annual power consumption decreases significantly after ten years when the simulator retires the SSDs that were purchased at the beginning of the simulation and replaces them with fewer and much higher capacity SSDs. The simulations replace the original SSDs with fewer new drives after ten years because, based on our assumptions that the capacity of new SSDs grows by 44% each year [45], [46] and the amount of data in the archive grows by 30% each year [1], the number of SSDs needed to store the archive’s data decreases over time. Our prediction that SSD capacity will grow by 44% annually may prove to be excessively optimistic. Nevertheless, archives using high-capacity SSDs will likely consume less electricity than archives with lower capacity SSDs.

Figure 8 shows the power consumption of archival systems using each of the storage technologies that we have profiled with our simulator. We observe that electrical consumption

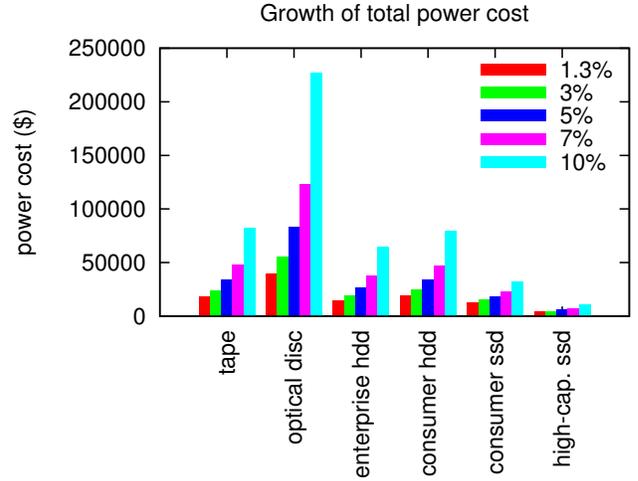


Fig. 9. We used the low-cost NAS for HDD and SSD simulations, and the optical disc archive scrubs once per year. The cost of electricity grows most dramatically in archival systems that spend more time reading, writing, or scrubbing data.

is highest for archives HDDs and lowest for archives with optical disc. As the simulation progresses, the power use of the archive with optical disc increases because it spends more and more time each year reading, writing, and scrubbing data rather than in the low-power idle state. We used the most generous assumption for optical disc archive reliability so that they must verify or scrub the archival data only once each year. Scrubbing once a year reduces the total workload of the archival system, but the experiments for tape, HDDs, and SSDs perform 12 scrubs of the archival data each year. The power consumption of tape increases to match that of hard disks. Similar to optical disc archives, tape and hard disk archives spend increasing amounts of time actively reading, writing, and scrubbing the archives data near the end of the simulation. The power consumption of tape and hard disk archives could be reduced if the number of scrubbing operations could be reduced without compromising the reliability of the archives data. Scrubbing is used to verify that the archives data is still readable from the storage media, and it is necessary because many storage devices, including tapes, HDDs, and SSDs, may lose data over time if not accessed periodically or verified. Improving the reliability of the storage devices could reduce the need for scrubbing operations. SSDs are the least affected by the workload demands of the archive because they have high read and write throughput that grows with the capacity of the drives. SSDs will likely become more attractive compared with other storage technologies in terms of total cost of ownership if the price of electricity increases significantly above the current baseline CAGR of 1.3%.

G. Growth of Power Cost

The cost of electricity varies widely by region. In other simulations we assumed that electricity will grow at a pre-

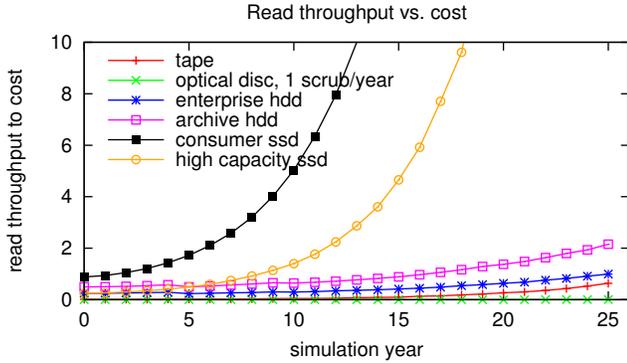


Fig. 10. SSDs are the most expensive storage medium in terms of acquisition cost. SSDs also offer the best value when comparing their performance/cost ratio with other storage technologies.

dictable rate of 1.3% annually. If, however, the cost of energy were to grow more quickly than our baseline assumption, the cost of operating an archival storage system could increase substantially. Figure 9 shows a comparison of the cost of electricity after 25 years of simulation in which electricity costs grow by up to 10% annually. While electricity costs increase for every storage technology, tape and optical disc consume more energy and are therefore more affected by higher energy costs than HDDs or SSDs. Archival systems consume more power when they spend more time reading and writing data. Hard disks consume more energy than a single tape or optical drive with many tapes or optical discs; however, in order to read, write, or scrub the data in the archive, archives with tape and optical disc must spend more time actively reading or writing data compared with HDDs or SSDs. For this reason, we observe that, in addition to the power consumption of individual storage devices and library or NAS devices, read and write throughput affect the total power consumption of archival systems.

H. Performance and Cost

Some archival systems may require high read throughput to support fast data access. Archival systems that are used for data backups may require high read throughput to facilitate a rapid recovery of backed up data. Time-sensitive institutions like banks may prefer archival systems for backup with high throughput instead of the cheapest archival storage available. We calculate the read throughput-to-cost value of an archive by dividing the total read bandwidth of each archive by its total cost of ownership.

We observe in Figure 10 that SSD archives offer the greatest throughput-to-cost value of any candidate archival technology. We assume in our experiments that SSD read and write throughput will continue to increase as SSD capacity increases. The read throughput of each archive is the aggregate total of the storage devices in the archive. The consumer SSD offers better value than the high-capacity SSD because, while its cost per gigabyte of data storage is the same, the archive with consumer SSDs uses more drives than the archive with high-

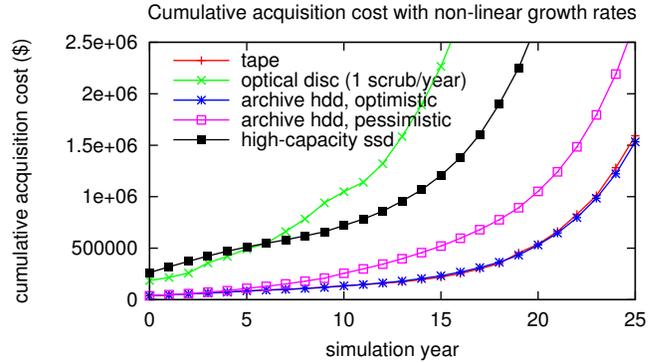


Fig. 11. The total acquisition cost of each archive grows more rapidly as the growth rates for device capacity and performance decrease. In this experiment, tape and optimistic HDD development slow after 10 years, and pessimistic HDD development begins to slow after 18 years. Optical disc and SSD development slow after eight and seven years, respectively.

capacity drives. The archive HDD offers more value than the enterprise HDD for the same reason. Archives with removable media offer the lowest performance-to-cost value because they have fewer drives than media and because the individual drives are much more expensive than SSDs or HDDs.

I. Non-Linear Growth

Figures 11 and 12 compare the costs of archival systems when the growth of capacity and performance decreases over time. We assume that each technology will continue to develop linearly until the device reaches the capacity promised in each its roadmap as described in Section V-F. We also assume that the demand for archival storage capacity will continue to grow by 30% annually.

The pace at which each technology develops determines the long-term cost of using it in an archival storage system that grows over time. We observe that technologies like tape and hard disk may provide better value in archival systems than optical disc and SSD in part because they are poised to continue developing rapidly throughout the next decade. Optical disc and SSD, on the other hand, may begin to exhibit slower development in less than ten years. In the optimistic case of 30% annual increases in capacity, hard disks will remain cost-competitive with tape for archival storage; however, if hard disks develop at the pessimistic rate of 15% each year, we would expect the cost of using hard disks in archival storage systems to grow more rapidly. Even though the slower pace of 15% annual capacity increase may last for up to 18 years, which is longer than the optimistic alternative, the annual increase of 30% in the amount of data in the archival system results in a steeper curve for the pessimistic cost prediction of hard disks.

Figure 12 shows that the annual acquisition cost of each technology begins to grow dramatically soon after the rate of development for the technology begins to decrease. We suggest that archival storage system designers should consider the roadmap of future developments for each candidate storage technology when designing an archival storage system.

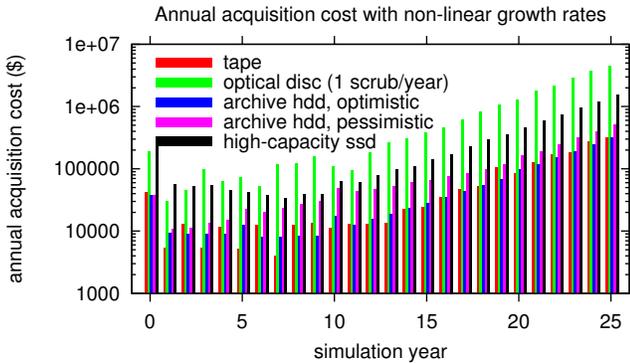


Fig. 12. The annual acquisition cost for each archive remains mostly constant while the individual storage devices develop at their historical rate. Slower rates of development result in higher acquisition cost as the archive grows. The y-axis is logarithmic to show the exponential growth of cost.

VII. RELATED WORK

Earlier research presented models to optimize large storage systems. The Disk Array Designer (DAD) searched for optimal data placement within a disk-based storage system [62]. DAD used the features of hard disks to predict total storage system performance and to optimize the placement of data within the system. Storage device performance and cost characteristics were also used to compare solid state disks and hard disks for data centers [63]. As of 2009, SSDs were more expensive than HDDs in data centers, notwithstanding the lower power consumption of SSDs. Later research discovered increasing value of SSDs in applications where power cost contributes significantly to a storage system’s long-term TCO.

Capital and power costs contribute to the total cost of ownership for archival storage systems, and the growth rates of storage device capacity and performance impact the long-term cost of using a storage technology within a growing archive. The importance of storage device capacity and performance to archival storage systems motivates the need to compare storage devices by measuring their total cost of ownership in an archival system [2].

Gupta et al. compared the projected costs of HDD and NAND flash in an archival system over 100 years. The cost per byte of storage on HDD is declining more slowly than the cost per byte of storage on SSD, and this trend may continue for years into the future. As SSDs become more cost-competitive with HDDs, SSDs will become increasingly suitable for archival storage. Also, the longer service life of SSDs compared with HDDs enhances the value of an SSD-based archive. Furthermore, the high power consumption of HDDs limits their economic advantages over SSDs in long-term storage.

The design of computation and storage systems can affect their capital costs and long-term operating costs. Previous work has introduced storage system designs that mitigate the power consumption of large systems. The cost of electricity motivated the designs of FAWN, a Fast Array of Wimpy Nodes, and FAWN-DLi [64], [65]. Energy efficiency also

motivates the design of storage systems intended specifically for archives.

Earlier research has evaluated the suitability of solid state disks for archival storage systems. DAWN, a Durable Array of Wimpy Nodes, proposed storage class memory for an archival system to target low power consumption and durability [52]. While NAND flash and other solid state storage devices incur higher initial capital costs than tape or disc, the higher performance-per-watt and lower active power consumption of SSDs can result in lower long-term operating costs.

Storage systems can be challenging to model accurately. Several studies have proposed techniques to improve both the accuracy of the performance model and the ease of configuring distributed storage systems [66], [67]. IRONModel [68] and Flamingo [69] enable accurate modeling to improve predictions about storage system design. Storage system modeling relies on descriptions of the storage device features, storage needs, and the relative importance of each parameter within the storage system. Within the context of some storage systems, cost is the most important variable to select between alternative storage devices; however, a utility function that includes the weighted values of a broad set of parameters can also improve decisions about storage system design [70]. The utility function predicts the value of a set of storage devices within the context of a storage system, and different sets of storage devices can be compared based on their utility value.

VIII. CONCLUSION

Developments in storage technology affect the long-term total cost of ownership for archival systems. We have designed a simulator for archival storage that compares the relative cost of different technologies in an archival system. We found that the growth rates of performance and capacity for different storage technologies predict the cost of using them in archival systems. Hard disks, which require more electricity than other storage devices, offer a competitive solution to tape archival systems, particularly if the archived data must be accessed frequently. Solid state drives, which are more expensive for archival storage than tape or hard disk in terms of capital cost, require less power while offering more throughput than other storage technologies. We observed that the slow pace of development for optical disc technology will cause disc-based archives to become more expensive than other technologies; however, optical disc will remain a viable archival technology if its capacity and throughput increase more rapidly than they have in the past. We observed that the long-term prospect for development varies for different types of technology. Hard disks will likely remain competitive with tape for archival storage systems for years to come notwithstanding the prospect that hard disk capacity will increase more slowly than it has in the past.

ACKNOWLEDGMENTS

We would like to acknowledge Ashwin Balaji for his research contributions and thoughtful insights.

REFERENCES

- [1] D. Reinsel, J. Gantz, and J. Rydning, "Data age 2025: The evolution of data to life-critical," April 2017. [Online]. Available: <https://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf>
- [2] D. S. H. Rosenthal, D. C. Rosenthal, E. L. Miller, I. F. Adams, M. W. Storer, and E. Zadok, "The economics of long-term digital storage," in *The Memory of the World in the Digital Age: Digitization and Preservation*, Sep. 2012.
- [3] Blue Ribbon Task Force on Sustainable Digital Preservation and Access, "Sustainable economics for a digital planet: Ensuring long-term access to digital information," Blue Ribbon Task Force on Sustainable Digital Preservation and Access, Tech. Rep., 2010. [Online]. Available: http://blueribbontaskforce.sdsu.edu/biblio/BRTF_Final_Report.pdf
- [4] Energy Information Administration. (2017, Nov.) Electric power monthly. [Online]. Available: https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_5_6_a
- [5] I. F. Adams, M. W. Storer, and E. L. Miller, "Analysis of workload behavior in scientific and historical long-term data repositories," *ACM Transactions on Storage*, vol. 8, no. 2, 2012.
- [6] C. Walter, "Kryder's law," *Scientific American*, Aug. 2005. [Online]. Available: <https://www.scientificamerican.com/article/kryders-law/>
- [7] P. Gupta, A. Wildani, D. Rosenthal, E. L. Miller, I. Adams, C. Strong, and A. Hospodor, "An economic perspective of disk vs. flash media in archival storage," in *Proceedings of the 22nd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS '14)*, Sep. 2014. [Online]. Available: <http://www.ssrc.ucsc.edu/Papers/gupta-mascots14.pdf>
- [8] B. Madden, I. Adams, J. Frank, and E. L. Miller, "Analyzing user behavior: A trace analysis of the near archival storage system," University of California, Santa Cruz, Tech. Rep. UCSC-SSRC-ssrcr-12-02, Mar. 2012.
- [9] I. F. Adams, B. Madden, J. Frank, M. W. Storer, and E. L. Miller, "Usage behavior of a large-scale scientific archive," in *Proceedings of the 2012 International Conference for High Performance Computing, Networking, Storage and Analysis (SC12)*, Nov. 2012. [Online]. Available: <http://www.ssrc.ucsc.edu/Papers/adams-sc12.pdf>
- [10] A. Wildani and I. F. Adams, "A case for rigorous workload classification," in *2015 IEEE 23rd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, Oct 2015, pp. 146–149.
- [11] Oracle Corporation, "StorageTek SL150 modular tape library," 2015. [Online]. Available: <http://www.oracle.com/us/products/servers-storage/storage/tape-storage/sl150-modular-tape-library/oracle-stk-sl150-ds-1665043.pdf>
- [12] Panasonic Corporation, "Optical data archive," 2017. [Online]. Available: ftp://ftp.panasonic.com/datastorage/intelligentarchive_referencearchitecture_brochure.pdf
- [13] HGST. (2017) Ultrastar HS14. [Online]. Available: <https://www.hgst.com/sites/default/files/resources/Ultrastar-Hs14-DS.pdf>
- [14] Seagate Technology. (2014) Archive HDD data sheet. [Online]. Available: <https://www.seagate.com/www-content/product-content/hdd-fam/seagate-archive-hdd/en-us/docs/archive-hdd-ds1834-3-1411us.pdf>
- [15] Micron Technology. (2015) Micron 1100 2.5-inch and m.2 SATA NAND flash SSD. [Online]. Available: https://www.micron.com/~/media/documents/products/data-sheet/ssd/1100_ssd.pdf
- [16] Oracle Corporation, "StorageTek SL150 power calculator," 2018. [Online]. Available: <http://www.oracle.com/us/products/servers-storage/sun-power-calculators/calcul/sl150-power-calculator-1954625.html>
- [17] —, "StorageTek LTO tape drives," 2017. [Online]. Available: <http://www.oracle.com/us/products/servers-storage/storage/tape-storage/033631.pdf>
- [18] —, "California state price list," Jul. 2013. [Online]. Available: <http://www.oracle.com/us/corporate/pricing/systems-hw-sw-pricelist-1863546.pdf>
- [19] Dell EMC, "Dell powervault LTO tape drives," 2010. [Online]. Available: http://www.dell.com/downloads/global/products/pvaul/en/tbu_drives_spec.pdf
- [20] R. Costea. (2017) Oracle launches new StorageTek LTO8 tape drive. [Online]. Available: <https://blogs.oracle.com/oraclepartners/oracle-launches-new-storage-tek-lto8-tape-drive>
- [21] Oracle Corporation. (2017) StorageTek LTO tape drives: Overview and frequently asked questions. [Online]. Available: <https://www.oracle.com/assets/lto-tape-drives-overview-faq-1906830.pdf>
- [22] Hewlett Packard Enterprise, "HPE LTO ultrium cartridges," 2018. [Online]. Available: <https://psnow.ext.hpe.com/doc/PSN34648USEN.pdf>
- [23] R. Cook, "How to estimate the lifespan of LTO tapes," May 2007. [Online]. Available: <http://searchdatabackup.techtarget.com/tip/How-to-estimate-the-lifespan-of-LTO-tapes>
- [24] Imation, "Expected usage life of imation media," 2010. [Online]. Available: https://web.archive.org/web/20160108131038/http://support2.imation.com/downloads/imm/LTO/Usage_Life_Imation_Media.pdf
- [25] LTO Consortium, "LTO program outlines generation 8 specifications and extends technology roadmap to 12th generation," Oct. 2017. [Online]. Available: <https://www.lto.org/2017/10/lto-program-outlines-generation-8-specifications-extends-technology-roadmap-12th-generation/>
- [26] M. Rouse, "LTO-1 (linear tape-open 1)," *TechTarget*, Oct. 2015. [Online]. Available: <http://searchdatabackup.techtarget.com/definition/LTO-1-Linear-Tape-Open-1>
- [27] A. Woodie, "LTO roadmap extended to gen 10," *ITJungle*, Sep. 2014. [Online]. Available: <https://www.itjungle.com/2014/09/15/tfh091514-story05/>
- [28] LTO Consortium, "The LTO program announces upcoming generation 7 specifications for licensing," Sep. 2015. [Online]. Available: <https://www.lto.org/2015/09/>
- [29] J. Niccolai, "Panasonic to commercialize facebook's blu-ray cold storage systems," Jan. 2016. [Online]. Available: <https://www.pcworld.com/article/3019413/panasonic-to-commercialize-facebooks-blu-ray-cold-storage-systems.html>
- [30] Sony Corporation and Panasonic Corporation, "White paper: Archival disc technology," Tech. Rep., Jul. 2015. [Online]. Available: https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf
- [31] H. Bennett, "Understanding CD-R and CD-RW," Optical Storage Technology Association, Tech. Rep., Jan. 2003. [Online]. Available: http://www.osta.org/technology/pdf/cdr_cdrw.pdf
- [32] —, "Understanding recordable and rewritable DVD," Optical Storage Technology Association, Tech. Rep., Apr. 2004. [Online]. Available: <http://www.osta.org/technology/dvdqa/pdf/dvdqa.pdf>
- [33] Blu-ray Disc Association, "Blu-ray disc format," Tech. Rep., Aug. 2015. [Online]. Available: http://www.blu-raydisc.com/Assets/Downloadablefile/White_Paper_General_4th_20150817_clean.pdf
- [34] Panasonic Corporation, "Panasonic optical data archiver," May 2016. [Online]. Available: <https://na.industrial.panasonic.com/sites/default/pidsa/files/downloads/files/panasonic-optical-data-archiver-freeze-ray-datasheet.pdf>
- [35] —, "Panasonic FreezeRay optical data archive system SRP cost catalog," 2016. [Online]. Available: https://s3-us-west-2.amazonaws.com/naspoaluepoint/1498745693_Panasonic%20FreezeRay%20SRP%20Cost%20Catalog.pdf
- [36] E. Dahl, "PC drive reaches 500GB," *PCWorld*, Mar. 2005. [Online]. Available: <https://www.pcworld.com/article/120102/article.html>
- [37] M. J. Perenson, "Hitachi introduces 1-terabyte hard drive," *PCWorld*, Jan. 2007. [Online]. Available: <https://www.pcworld.com/article/128400/article.html>
- [38] D. Murphy, "Western digital launches world-first 2TB hard drive," *PCWorld*, Jan. 2009. [Online]. Available: https://www.pcworld.com/article/158374/Western_Digital_Launch.html
- [39] S. Anthony, "Hitachi ships world's first 4TB hard drive, sticks it to thor," *ExtremeTech*, Dec. 2011. [Online]. Available: <https://www.extremetech.com/computing/108665-hitachi-ships-worlds-first-4tb-hard-drive-sticks-it-to-thor>
- [40] R. Whitwam, "Western digital launches world's first 14TB hard drive," *ExtremeTech*, Oct. 2017. [Online]. Available: <https://www.extremetech.com/computing/256961-western-digital-launches-worlds-first-14tb-hard-drive>
- [41] A. Klein, "Hard drive cost per gigabyte," *Backblaze*, Jul. 2017. [Online]. Available: <https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/>
- [42] E. Brewer, L. Ying, L. Greenfield, R. Cypher, and T. T'so, "Disks for data centers," Google, Tech. Rep., 2016.
- [43] C. Mallor. (2010) Drive suppliers hit capacity increase difficulties. [Online]. Available: https://www.theregister.co.uk/2010/07/12/adding_platters/

- [44] A. Klein, "Backblaze hard drive stats for 2017," Feb. 2018. [Online]. Available: <https://www.backblaze.com/blog/hard-drive-stats-for-2017/>
- [45] C. Mellor, "SSD price premium over disk faaaalling," *The Register*, May 2017. [Online]. Available: https://www.theregister.co.uk/2017/05/22/ssd_price_premium_over_disk_falling/
- [46] L. Mearian, "SSD prices plummet again, close in on HDDs," *Computerworld*, Mar. 2016. [Online]. Available: <https://www.computerworld.com/article/3040694/data-storage/ssd-prices-plummet-again-close-in-on-hdds.html>
- [47] R. Micheloni, A. Marelli, and S. Commodaro, *NAND Overview: From Memory to Systems*. Springer, 2010, ch. 2, pp. 19–53.
- [48] L. M. Grupp, J. D. Davis, and S. Swanson, "The bleak future of NAND flash memory," in *Proceedings of the 10th USENIX Conference on File and Storage Technologies*, ser. FAST'12. Berkeley, CA, USA: USENIX Association, 2012. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2208461.2208463>
- [49] Hewlett Packard Enterprise, "HPE MSA 2050 SAN storage - specification," 2018. [Online]. Available: https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-a00017539en_us
- [50] —, "QucikSpecs, HPE MSA 2050 storage," 2018. [Online]. Available: <https://h20195.www2.hpe.com/v2/GetDocument.aspx?docname=a00008276enw>
- [51] Sinovoip. (2018) BPI-M2 berry quad-core development board. [Online]. Available: <http://www.banana-pi.org/m2ub.html>
- [52] I. Adams, E. L. Miller, and D. S. Rosenthal, "Using storage class memory for archives with DAWN, a durable array of wimpy nodes," University of California, Santa Cruz, Tech. Rep. UCSC-SSRC-11-07, Oct. 2011.
- [53] M. W. Storer, K. M. Greenan, E. L. Miller, and K. Voruganti, "Pergamum: Replacing tape with energy efficient, reliable, disk-based archival storage," in *Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST)*, Feb. 2008. [Online]. Available: <http://www.ssrc.ucsc.edu/Papers/storer-fast08.pdf>
- [54] S. Balakrishnan, R. Black, A. Donnelly, P. England, A. Glass, D. Harper, S. Legtchenko, A. Ogus, E. Peterson, and A. Rowstron, "Pelican: A building block for exascale cold data storage." 11th USENIX Symposium on Operating Systems Design and Implementation (OSDI '14), October 2014. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/pelican-a-building-block-for-exascale-cold-data-storage/>
- [55] Sony Corporation. (2016, Apr.) Optical disc archive, generation 2: White paper. [Online]. Available: <http://assets.pro.sony.eu/Web/ngp/pdf/optical-disc-archive-generation-two.pdf>
- [56] A. Klein. (2017, Dec.) What is HAMR and how does it enable the high-capacity needs of the future? [Online]. Available: <https://www.backblaze.com/blog/hamr-hard-drives/>
- [57] R. Bauer. (2018, Mar.) HDD vs SSD: What does the future for storage hold? part 2. [Online]. Available: <https://www.backblaze.com/blog/hdd-vs-ssd-in-data-centers/>
- [58] M. Re. (2015, Aug.) Tech talk on HDD areal density. [Online]. Available: https://www.seagate.com/www-content/investors/_shared/docs/tech-talk-mark-re-20150825.pdf
- [59] A. Moscaritolo. (2018, Mar.) Nimbus Data's 100TB ExaDrive is the world's largest SSD. [Online]. Available: <https://www.pcmag.com/news/359941/nimbus-datas-100tb-exadrives-is-the-worlds-largest-ssd>
- [60] T. Coughlin. (2016, Feb.) Flash memory areal densities exceed those of hard drives. [Online]. Available: <https://www.forbes.com/sites/tomcoughlin/2016/02/03/flash-memory-areal-densities-exceed-those-of-hard-drives/#53dd240a7c72>
- [61] L. Mearian. (2016, Feb.) Flash memory's density surpasses hard drives for first time. [Online]. Available: <https://www.computerworld.com/article/3030642/data-storage/flash-memorys-density-surpasses-hard-drives-for-first-time.html>
- [62] E. Anderson, S. Spence, R. Swaminathan, M. Kallahalla, and Q. Wang, "Quickly finding near-optimal storage designs," *ACM Trans. Comput. Syst.*, vol. 23, no. 4, pp. 337–374, Nov. 2005. [Online]. Available: <http://doi.acm.org/10.1145/1113574.1113575>
- [63] D. Narayanan, E. Thereska, A. Donnelly, S. Elnikety, and A. Rowstron, "Migrating server storage to SSDs: Analysis of tradeoffs," in *Proceedings of the 4th ACM European Conference on Computer Systems*, ser. EuroSys '09. New York, NY, USA: ACM, 2009, pp. 145–158. [Online]. Available: <http://doi.acm.org/10.1145/1519065.1519081>
- [64] D. G. Andersen, J. Franklin, M. Kaminsky, A. Phanishayee, L. Tan, and V. Vasudevan, "FAWN: A fast array of wimpy nodes," in *Proceedings of the 22nd ACM Symposium on Operating Systems Principles (SOSP '09)*, Big Sky, MT, Oct. 2009, pp. 1–14. [Online]. Available: <http://www.ssrc.ucsc.edu/PaperArchive/andersen-sosp09.pdf>
- [65] J. Gluck, "FAWN-DLI: A data library for a fast array of wimpy nodes," 2012. [Online]. Available: <https://www.sccs.swarthmore.edu/users/12/jgluck/resources/FAWNDLI.pdf>
- [66] D. M. Eyers, R. Routray, R. Zhang, D. Willcocks, and P. Pietzuch, "Towards a middleware for configuring large-scale storage infrastructures," in *Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and e-Science*, ser. MGC '09. New York, NY, USA: ACM, 2009, pp. 3:1–3:6. [Online]. Available: <http://doi.acm.org/10.1145/1657120.1657123>
- [67] S. Gopisetty, E. Butler, S. Jaquet, M. Korupolu, T. K. Nayak, R. Routray, M. Seaman, A. Singh, C.-H. Tan, S. Uttamchandani, and A. Verma, "Automated planners for storage provisioning and disaster recovery," *IBM J. Res. Dev.*, vol. 52, no. 4, pp. 353–365, Jul. 2008. [Online]. Available: <http://dx.doi.org/10.1147/rd.524.0353>
- [68] E. Thereska and G. R. Ganger, "Ironmodel: Robust performance models in the wild," in *Proceedings of the 2008 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '08. New York, NY, USA: ACM, 2008, pp. 253–264. [Online]. Available: <http://doi.acm.org/10.1145/1375457.1375486>
- [69] S. Legtchenko, X. Li, A. Rowstron, A. Donnelly, and R. Black, "Flamingo: Enabling evolvable hdd-based near-line storage," in *Proceedings of the 14th Usenix Conference on File and Storage Technologies*, ser. FAST'16. Berkeley, CA, USA: USENIX Association, 2016, pp. 213–226. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2930583.2930599>
- [70] J. D. Strunk, E. Thereska, C. Faloutsos, and G. R. Ganger, "Using utility to provision storage systems," in *Proceedings of the 6th USENIX Conference on File and Storage Technologies*, ser. FAST'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 21:1–21:16. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1364813.1364834>