The Five-minute Rule Thirty Years Later and its Impact on the Storage Hierarchy

Raja Appuswamy
EURECOM, France
raja.appuswamy@eurecom.fr

Renata Borovica-Gajic
University of Melbourne, AU
renata.borovica@unimelb.edu.au

Goetz Graefe
Google, Madison WI, USA
goetzg@google.com

Anastasia Ailamaki
EPFL, Switzerland
anastasia.ailamaki@epfl.ch

1. INTRODUCTION

The design of data management systems has always been influenced by the storage hardware landscape. In the 1980s, database engines used a two-tier storage hierarchy consisting of Dynamic Random Access Memory (DRAM) and Hard Disk Drives (HDD). Given the disparity in cost between HDD and DRAM, it was important to determine when it made economic sense to cache data in DRAM as opposed to leaving it on the HDD.

In 1987, Jim Gray and Gianfranco Putzolu established the five-minute rule that gave a precise answer to this question–"1KB records referenced every five minutes should be memory resident" [9]. They arrived at this value by using the then-current price–performance characteristics of DRAM and HDD shown in Table 1 for computing the break-even interval at which the cost of holding 1KB of data in DRAM matches the cost of I/O to fetch it from HDD.

Today, enterprise database engines use a three-tier storage hierarchy as depicted in Figure 1. DRAM or NAND Flash Solid State Device (SSD)-based performance tier is used for hosting data accessed by latency-critical, transaction processing and real-time analytics applications. The HDD-based capacity tier hosts data accessed by latency-insensitive batch analytics applications. The archival tier is not used for "online" query processing, but for storing data that is only accessed rarely during regulatory compliance audits or disaster recovery. This tier is primarily based on tape and is extremely crucial as a long-term data repository for several application domains like physics, banking, security, and law enforcement.

In this article, we revisit the five-minute rule three decades after its inception. We recompute break-even intervals for each tier of the modern, multi-tiered storage hierarchy and use guidelines provided by the five-minute rule to identify impending changes in the design of data management engines for emerging storage hardware. We summarize our findings here.

- **HDD is tape.** The gap between DRAM and HDD is increasing as the five-minute rule that was valid for DRAM–HDD case in 1987 is now a four-hour rule. This implies that the HDD-based capacity tier is losing relevance for not just performance sensitive applications, but for all applications with a non-sequential data access pattern.

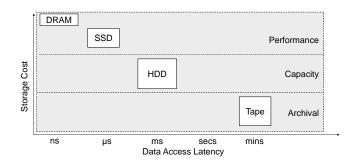- **Non-volatile memory is DRAM.** The gap between DRAM and

Figure 1: Storage tiering for enterprise databases

SSD is shrinking. The original five-minute rule is now valid for the DRAM–SSD case, and the break-even interval is less than a minute for newer Non-Volatile Memory (NVM) devices like 3D-XPoint [22]. This suggests an impending shift from DRAM-based database engines to flash or NVM-based persistent memory engines.

- **Cold storage is hot.** The gap between HDD and tape is also rapidly shrinking for sequential workloads. New cold storage devices that are touted to offer second-long access latency with cost comparable to tape reduce this gap further. This suggests that the HDD-based capacity tier will soon lose relevance even for non-performance-critical batch analytics applications that can be scheduled to run directly over newer cold storage devices.

2. REVISITING THE FIVE-MINUTE RULE

The five-minute rule explores the trade-off between the cost of DRAM and the cost of disk I/O by providing a formula to predict the *break-even interval*–the time window within which data must be re-accessed in order for it to be economically beneficial to be cached in DRAM. The interval is computed as:

$$\frac{PagesPerMBofDRAM}{AccessesPerSecondPerDisk} \times \frac{PricePerDiskDrive}{PricePerMBofDRAM} \quad (1)$$

The first ratio in the equation was referred to as the *technology ratio*, as random I/O access capability of the secondary storage device, and the page size used by the database engine for performing I/O, both directly depend on the hardware technology used for secondary storage. The second ratio, in contrast, is referred to as the *economic ratio* as pricing is determined by factors other than just hardware technology. Rearranging the formulation by swapping

Metric	DRAM				HDD				SATA Flash SSD	
	1987	1997	2007	2018	1987	1997	2007	2018	2007	2018
Unit price($)	5k	15k	48	80	30k	2k	80	49	1k	415
Unit capacity	1MB	1GB	1GB	16GB	180MB	9GB	250GB	2TB	32GB	800GB
$/MB	5k	14.6	0.05	0.005	83.33	0.22	0.0003	0.00002	0.03	0.0005
Random IOPS	-	-	-	-	15	64	83	200	6.2k	67k (r)/20k (w)
Sequential b/w (MB/s)	-	-	-	-	1	10	300	200	66	500 (r)/460 (w)

Table 1: The evolution of DRAM, HDD, and Flash SSD properties

	1987	1997	2007	2018
Break-even (4KB page)	100s	9m	1.5h	4h
Page size (5-minute interval)	1KB	8KB	64KB	512KB

Table 2: The evolution of the page size for which the five-minute rule holds across four decades based on appropriate price, performance, and page size values

Tier	1987	1997	2007	2018
DRAM–SSD	-	-	15m	5m
SSD–HDD	-	-	2.25h	1.5d

Table 3: The evolution of the break-even interval across four decades based on appropriate price, performance, and page size values

the denominators provides the intuition behind the five-minute rule, as it reduces the equation to price-per-disk-access-per-second normalized by the price-per-page of DRAM. This term directly compares the cost of performing I/O to fetch a page from disk versus the cost of caching it in DRAM.

Table 1 shows the price, capacity, and performance of DRAM, HDD, and NAND flash-based SSDs across four decades. The values shown for 1987, 1997, and 2007 are those reported by previous revisions of the five-minute rule [6, 8, 9]. The values listed for 2018 are performance metrics listed in vendor specifications, and unit price quoted by www.newegg.com as of March 1, 2018, for DRAM, SSD, and HDD components specified in a recent TPC-C report [24].

DRAM–HDD. Table 2 presents both the break-even interval for 4KB pages and the page sizes for which the five-minute rule is applicable across four decades. In 1987, the break-even interval was 400 seconds for 1KB pages. This was rounded down to five minutes, thus, lending the name for the rule. For 4KB pages, the break-even interval was 100 seconds. When the study was repeated in 1997, the break-even interval had increased to 9 minutes for 4KB pages, and the five-minute rule was determined to hold only for 8KB pages. Between 1997 and 2007, DRAM and HDD prices dropped further resulting in the economic ratio increasing from 133 ($2k/$15) to 1700 ($80/$0.05). However, the technology ratio did not drop proportionately due to a lack of improvement in HDD random access latency. As a result, the break-even interval for 4KB pages increased $10\times$, from nine minutes to 1.5 hours. The five-minute rule was applicable only for 64KB pages in 2007. Continuing this trend, the break-even interval for DRAM–HDD case today is four hours for 4KB pages. The five-minute rule is valid today for 512KB pages. The break-even interval trend indicates that it is more economical to store most data in DRAM instead of the HDD, and the page size trend indicates that even rare accesses to HDD should be performed in large granularities.

DRAM–SSD. SSDs are being increasingly used as the storage medium of choice in the latency-critical performance tier due to their superior random access capability compared to HDDs. Thus, the five-minute rule can be used to compute a break-even interval for the case where DRAM is used to cache data stored in SSDs. Table 3 shows the interval in 2007, when SSDs were in the initial stages of adoption, and today, based on metrics listed in Table 1.

We see that the interval has dropped from 15 minutes to five minutes for 4KB pages. Thus, the five-minute rule is valid for SSDs

today. This is in stark contrast with the DRAM–HDD case, where the interval increased $2.7\times$ from 1.5 hours to four hours. In both DRAM–HDD and DRAM–SSD cases, the drop in DRAM cost/MB dominated the economic ratio. However, unlike the $2.5\times$ improvement in random I/Os-per-second (IOPS) with HDDs, SSDs have managed to achieve an impressive $11\times$ improvement (67k/6.2k). Thus, the increase in economic ratio was overshadowed by the decrease in technology ratio with SSDs, resulting in the interval shrinking.

SSD–HDD. As SSDs can also be used as a cache for HDD, the same formula can also be used to estimate the break-even interval for the SSD–HDD case. From Table 3, we see that the break-even interval for this case has increased by a factor of $10\times$ from 2.25 hours in 2007 to 1.5 days in 2018. The SSD–HDD interval is nine times longer than the DRAM–HDD interval of four hours.

Implications. There are two important consequences of these results. First, in 2007, the turn-over time in the DRAM–HDD case was six times higher than the DRAM–SSD case(1.5h/15m). In 2018, it is nearly $50\times$ higher (4h/5m). Thus, in systems tuned using economic considerations, one should replace HDD with SSD, as it would not only improve performance, but also reduce the amount of DRAM required for caching data. Second, given the four-hour DRAM–HDD and one day SSD–HDD intervals, it is important to keep all active data in the DRAM or SSD-based performance tier and relegate the HDD-based capacity tier to storing only infrequently accessed data. The growing gap between performance and capacity tiers also implies that SSD vendors should optimize for $/IOPS, and HDD vendors, in contrast, should optimize for $/GB. In the next two sections, we will highlight recent changes in performance and capacity tiers that indicate that such targeted optimizations are already underway.

3. THE PERFORMANCE TIER

NAND flash. NAND Flash-based solid-state storage has been steadily inching its way closer to the CPU over the past two decades. When NAND flash was introduced in the early 2000s, solid-state storage was dominated by DRAM-based SSD products. By the mid 2000s, improvements in performance and reliability of NAND flash resulted in flash-based Serial AT Attachment (SATA) SSDs gaining popularity in niche application domains. The late 2000s witnessed the emergence of a new breed of Peripheral Component Interconnect Express (PCIe) flash SSDs that could deliver two orders of magnitude higher throughput than their SATA counterparts. Since

then, a rapid increase in capacity, drop in pricing, and new low-overhead interfaces like Non-Volatile Memory Express (NVMe), have all resulted in PCIe flash SSDs displacing their SATA counterparts as server accelerators of choice.

Table 4 (first row) shows the price/performance characteristics of a representative, state-of-the-art PCIe SSD. Comparing this with Table 1, we can see that the PCIe SSD offers five times higher read IOPS and sequential access bandwidth than its SATA counterpart.

Device	Capacity	Price($)	IOPS(k)	B/w (GB/s)
Intel 750	800GB	589	460	2.5
Intel P4800X	480GB	617	550	2.5

Table 4: Price/performance metrics for the NAND-based Intel 750 PCIe SSD and 3D-XPoint-based Intel Optane P4800X PCIe SSD

NVDIMM. As SSD vendors continue to improve throughput and capacity, the bottleneck in the storage subsystem has shifted from the device itself to the PCIe bus that is used to interface with the SSD. Thus, over the past few years, NAND flash has started transitioning once again from storage devices that are interfaced via the high-latency, bandwidth-limited PCIe bus into Non-Volatile Memory (NVM) devices that are interfaced via the low-latency, high-bandwidth memory bus. These devices, also referred to as Non-Volatile DIMMs (NVDIMM), use a combination of DRAM and flash storage media packaged together as a Dual Inline Memory Module (DIMM).

NVM. Today, NVDIMMs are niche accelerators compared to PCIe SSDs due to a high cost/GB. Unlike these NVDIMM technologies that rely on NAND flash, new NVM technologies that are touted to have better endurance, higher throughput, and lower latency than NAND flash are being actively developed. Table 4 (second row) shows the characteristics of Intel Optane DC P4800X–a PCIe SSD based on 3D XPoint, a new phase-changed-media-based NVM technology. The cost/GB of 3D XPoint is higher than NAND flash today as the technology is yet to mature. However, preliminary studies have found that 3D XPoint provides predictable access latencies that are much lower than several state-of-the-art NAND flash devices even under severe load [22].

3.1 Break-even interval and implications

When we apply the five-minute rule formula using metrics given in Table 4, we get a break-even interval of 1 minute for 4KB pages in both the DRAM–NAND Flash PCIe SSD and DRAM–3D XPoint cases. Comparing these results with Table 2, we see that the break-even interval is $10\times$ shorter when PCIe SSDs or new PM technologies are used as the second tier instead of SATA SSDs. This can be attributed to the drop in technology ratio caused by the improvement in random IOPS.

Implications. Today, in the era of in-memory data management, several database engines are designed based on the assumption that all data is resident in DRAM. However, the dramatic drop in break-even interval computed by the five-minute rule challenges this trend of DRAM-based in-memory data management due to three reasons. First, recent projections indicate that flash density is expected to increase 40% annually over the next five years [5]. DRAM, in contrast, is doubling in capacity every three years [15]. As a result, the cost of NAND flash is likely to drop faster than DRAM. This, in turn, will result in the economic ratio dropping further leading to a reduction in the break-even interval.

Second, modern PCIe SSD are highly parallel devices that can provide very high random I/O throughput by servicing multiple

	1997	2018
Tape library cost ($)	10,000	11,000
Number of drives	1	4
Number of slots	14	10
Max capacity per tape	35GB	15TB
Transfer rate per drive (MB/s)	5	750
Access latency	30s	65s

Table 5: Price/performance characteristics of tape

outstanding I/Os concurrently. New non-volatile memory technologies like 3D XPoint promise further improvements in both throughput and access latencies over NAND flash. With interfaces like NVMe, the end-to-end latency of accessing data from PCIe 3D XPoint SSDs is just tens of μs. Thus, further improvements in non-volatile solid-state storage media will result in a drop in technology ratio, thereby reducing the break-even interval further.

Third, SSDs consume substantially lower power than DRAM. The Intel 750 SSD consumes 4W of power when idle and 22W when active. In contrast, 1TB of DRAM in a server would consume 50W when idle and 100W when active [1]. It is also well known that DRAM power consumption increases non-linearly with capacity, as high-density DRAM consumes substantially more power than their low-density counterparts. A recent study that focuses on power consumption in main memory databases showed that in a server equipped with 6TB of memory, the idle power of DRAM would match that of four active CPUs [1]. Such a difference in power consumption between SSD and DRAM directly translates into higher Operational Expenses (OPEX), and hence, higher Total Cost of Ownership (TCO), for DRAM-based database engines.

Given these three factors, the break-even interval from the five-minute rule seems to suggest an inevitable shift from DRAM-based data management engines to NVM-based persistent-memory engines. In fact, this change is already well underway, as state-of-the-art database engines are being updated to fully exploit the performance benefits of PCIe NVMe SSDs [26]. Researchers have recently highlighted the fact that *data caching systems* that trade off performance for price by reducing the amount of DRAM are gaining market share over in-memory database engines [16].

4. THE CAPACITY TIER

HDD. Traditionally, HDDs have been the primary storage media used for provisioning the capacity tier. For several years, areal density improvements enabled HDDs to increase capacity at Kryder's rate (40% per year), outstripping Moore's law. However, over the past few years, HDD vendors have hit walls in scaling areal density with conventional Perpendicular Magnetic Recording (PMR) techniques resulting in annual areal density improvement of only around 16% instead of 40% [17].

HDDs also present another problem when used as the storage medium of choice for building a capacity tier, namely, high idle power consumption. Although enterprises gather vast amounts of data, as one might expect, not all data is accessed frequently. Recent studies estimate that as much as 80% of enterprise data is "cold", meaning infrequently accessed, and that cold data is the largest growing segment with a 60% Cumulative Annual Growth Rate (CAGR) [10, 11, 23]. Unlike tape, which consumes no power once unmounted, HDDs consume a substantial amount of power even while idle. Such power consumption translates to a proportional increase in TCO.

Tape. The areal density of tape has been increasing steadily at a rate of 33% per year and roadmaps from the Linear Tape Open

Metric	DRAM	HDD	Tape
Unit capacity	16GB	2TB	10×15TB
Unit cost ($)	80	50	11,000
Latency	100ns	5ms	65s
Bandwidth	100 GB/s	200 MB/s	4×750MB/s
Kaps	9,000,000	200	0.02
Maps	10,000	100	0.02
Scan time	0.16s	3hours	14hours
$/Kaps	9e-14	5e-09	8e-03
$/Maps	9e-12	8e-09	8e-03
$/TBscan	8e-06	0.003	0.03
$/TBscan (97)	0.32	4.23	296

Table 6: Price/performance metrics of DRAM, HDD, and tape

Figure 2: Break-even interval asymptotes for DRAM–HDD and DRAM–tape cases

consortium (LTO) [25] and the Information Storage Industry Consortium (INSIC) [4] project a continued increase in density for the foreseeable future.

Table 5 shows the price/performance metrics of tape storage both in 1997 and today. The 1997 values are based on the corresponding five-minute rule paper [8]. The 2018 values are based on a SpectraLogic T50e tape library [21] using LTO-7 tape cartridges.

With individual tape capacity increasing $200\times$ since 1997, the total capacity stored in tape libraries has expanded from hundreds of gigabytes to hundreds of petabytes today. Further, a single LTO-7 cartridge is capable of matching, or even outperforming a HDD, with respect to sequential data access bandwidth as shown in Table 6. As modern tape libraries use multiple drives, the cumulative bandwidth achievable using even low-end tape libraries is 1–2GB/s. High-end libraries can deliver well over 40GB/s. These benefits have made tape the preferable media of choice in the archival tier both on-premise and in the cloud, for several applications ranging from natural sciences, like particle physics and astronomy, to movies archives in the entertainment industry [13, 18]. However, random access latency of tape is still $1000\times$ higher than HDD (minutes vs ms) due to the fact that tape libraries need to mechanically load and wind tape cartridges before data can be accessed.

4.1 Break-even interval and implications

Using metrics from Tables 1, 5 to compute the break-even interval for the DRAM–tape case results in an interval of over 300 years for a page size of 4KB! Jim Gray referred to tape drives as the "data motel" where data checks in and never checks out [7], and this is certainly true today. Figure 2 shows the variation in break-even interval for both HDD and tape for various page sizes. We see that the interval asymptotically approaches one minute in the DRAM–HDD case and ten minutes in the DRAM–tape case. The HDD asymptote is reached at a page size of 100MB and the tape asymptote is reached at a size of 100GB. This clearly shows that: i) randomly accessing data on these devices is extremely expensive, and ii) data transfer sizes with these devices should be large to amortize the cost of random accesses.

However, the primary use of the capacity tier today is not supporting applications that require high-performance random accesses. Rather, it is to reduce the cost/GB of storing data over which latency-insensitive batch analytics can be performed. Indeed, Gray and Graefe noted that metrics like KB-accesses-per-second (Kaps) are less relevant for HDD and tape as they grow into infinite-capacity resources [8]. Instead, MB-accesses-per-second (Maps) and time to scan the whole device are more pertinent to these high-density storage devices. Table 6 shows these new metrics and their values for DRAM, HDD, and tape. In addition to Kaps, Maps, and scan time, the table also shows $/Kaps, $/Maps, and $/TBscan, where
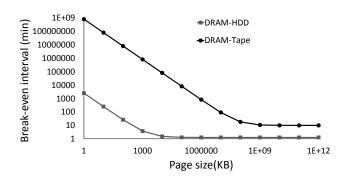
costs are amortized over a three-year time frame as proposed by Gray and Graefe [8].

Looking at $/Kaps, we see that DRAM is five orders of magnitude cheaper than HDD, which, in turn, is six orders of magnitude cheaper than tape. This is expected given the huge disparity in random access latencies and is in accordance with the five-minute rule that favors using DRAM for randomly accessed data. Looking at $/Maps, we see that the difference between DRAM and HDD shrinks to roughly $1,000\times$. This is due to the fact that HDDs can provide much higher throughput for sequential data accesses over random ones. However, HDD continue to be six orders of magnitude cheaper than tape even for MB-sized random data accesses. This, also, is in accordance with the HDD/tape asymptote shown in Figure 2. Finally, $/TBscan paints a very different picture. While DRAM remains $300\times$ cheaper than HDD, the difference between HDD and tape shrinks to $10\times$.

Comparing the $/TBscan values with those reported in 1997, we can see two interesting trends. First, the disparity between DRAM and HDD is growing over time. In 1997, it was $13\times$ cheaper to use DRAM for a TBscan than HDD. Today, it is $300\times$ cheaper. This implies that even for scan-intensive applications, unsurprisingly, optimizing for performance requires avoiding using HDD as the storage medium. Second, the difference between HDD and tape is following the opposite trend and shrinking over time. In 1997, HDD was $70\times$ cheaper than tape. However, today, it is only $10\times$ cheaper. Unlike HDD, sequential data transfer bandwidth of tape is predicted to double for the foreseeable future. Hence, this difference is likely to shrink further. Thus, in the near future, it might not make much of a difference whether data is stored in a tape or HDD with respect to the price paid per TBscan.

Implications. Today, all data generated by an enterprise has to be stored twice, once in the traditional HDD-based capacity tier for enabling batch analytics, and a second time in the tape-based archival tier for meeting regulatory compliance requirements. The shrinking difference in $/TBscan between HDD and tape suggests that it might be economically beneficial to merge the capacity and archival tiers into a single *Cold Storage Tier* [2]. However, with such a merger, the Cold Storage Tier would no longer be a near-line tier that is used rarely during disaster recovery, but an online tier that is used for running batch analytics applications. Recent hardware and application trends indicate that it might be feasible to build such a Cold Storage tier.

On the hardware front, storage vendors have recently started building new *Cold storage devices* (CSD) for storing cold data. Each CSD is an ensemble of HDDs grouped in a Massive Array of Idle Disks (MAID) setup where only a small subset of disks are active at any given time [3, 19, 27]. For instance, Pelican CSD pro-

vides 5PB of storage using 1,152 SMR disks packed as a 52U rack appliance [19]. However, only 8% of disks can be spun up simultaneously due to cooling and power restrictions enforced by hardware. Access to data in any of the spun-up disks can be done with latency and bandwidth comparable to that of the traditional capacity tier. For instance, Pelican, OpenVault Knox, and ArticBlue provide between 1–2 GB/s of throughput for reading data from spun-up disks [19, 20, 27]. However, accessing data on a spun-down disk takes several seconds, as the disk has to be spun up before data can be retrieved. Thus, CSDs form a perfect middle ground between HDD and tape with respect to both cost/GB and access latency.

On the application front, there is a clear bifurcation in demand between latency-sensitive interactive applications and latency insensitive batch applications. As interactive applications are isolated to the performance tier, the Cold Storage Tier only has to cater to the bandwidth demands of latency-insensitive batch analytics applications. Nearline storage devices like tape libraries and CSD are capable of providing high-throughput access for sequentially accessed data. Thus, researchers have recently started investigating extensions to batch processing frameworks for enabling analytics directly over data stored in tape archives and CSD. For instance, Nakshatra implements prefetching and I/O scheduling extensions to Hadoop so that map–reduce jobs can be scheduled to run directly on tape archives [12]. Skipper is a query processing framework that uses adaptive query processing techniques in combination with customized caching and I/O scheduling to enable query execution over CSD [2]. Skipper even shows that for long-running batch queries, using CSD results in query execution time increasing by only 35% compared to a traditional HDD despite the long disk spin-up latency. With such frameworks, it should be possible for installations to switch from the traditional three-tier hierarchy to a two-tier hierarchy consisting of just a performance tier with DRAM and SSDs, and a Cold Storage tier with CSDs.

5. CONCLUSION AND FUTURE WORK

Modern database engines use a three-tier storage hierarchy across four primary storage media (DRAM, SSD, HDD, and tape) with widely varying price–performance characteristics. In this article, we revisited the five-minute rule in the context of this modern storage hierarchy and used it to highlight impending changes based on recent trends in the hardware landscape.

In the performance tier, NAND flash is inching its way closer to the CPU resulting in dramatic improvements in both access latency and bandwidth. For state-of-the-art PCIe SSDs, the break-even interval predicted by the five-minute rule is one minute for 4KB pages. Going forward, further improvements in NAND flash and the introduction of new NVM technologies will likely result in this interval dropping further. As the data reuse window shrinks, it will soon be economically more valuable to store most, if not all, data on solid state storage devices instead of DRAM. This will invariably necessitate revisiting several techniques pioneered by traditional HDD-based database engines, but eschewed by in-memory engines, like buffer caching, on-disk storage layout, and index persistence, to name a few, for these new low-latency, high-bandwidth storage devices.

Traditionally, HDDs have been used for implementing the capacity tier. However, our analysis showed that the difference between HDD and tape is shrinking when $/TBScan is used as the metric. Given the latency-insensitive nature of batch analytics workloads, it is economically beneficial to merge the HDD-based capacity tier and the tape-based archival tier into a single Cold Storage tier as demonstrated by recent research [2]. However, several open questions still need to answered in order for the Cold Storage tier to be feasible in practice.

Over the past few years, several other systems have been built to reduce the cost of storing cold data using alternative storage media. For instance, DTStore [14] uses LTFS tape archive for reducing the TCO of online multimedia streaming services by storing cold data in tape drives. ROS [28] is a PB-sized, rack-scale cold storage library built using thousands of optical discs packed in a single 42U Rack. Today, it is unclear as to how these alternative storage options fare with respect to HDD-based CSD as the storage media of choice for storing cold data. Furthermore, in order for the Cold Storage Tier to be realized in practice, an ideal cold storage media needs to support batch analytics workloads. CSD, tape, and optical media are all primarily used today for archival storage where data is rarely read. Further research is required to understand the reliability implications of using these storage devices under batch analytics workloads.

Finally, with wide-spread adoption of cloud computing, the modern enterprise storage hierarchy not only spans several storage devices, but also different geographic locations from direct-attached low-latency devices, through network-attached storage servers, to cloud-hosted storage services. The price–performance characteristics of these storage configurations vary dramatically depending not only on the storage media used, but also on other factors like the total capacity of data stored, the frequency and granularity of I/O operations used to access the data, the read–write ratio, the duration of data storage, and the cloud service provider used, to name a few. Given the multitude of factors, determining the break-even interval for cloud storage is a complicated problem that we did not consider in this work. Thus, another interesting avenue of future work is extending the five-minute rule to such a distributed cloud storage setting.

6. REFERENCES

[1] R. Appuswamy, M. Olma, and A. Ailamaki. Scaling the memory power wall with dram-aware data management. In *DAMON*, 2015.

[2] R. Borovica-Gajic, R. Appuswamy, and A. Ailamaki. Cheap data analytics using cold storage devices. *PVLDB*, 9(12), 2016.

[3] D. Colarelli and D. Grunwald. Massive arrays of idle disks for storage archives. In *Conference on Supercomputing*, 2002.

[4] I. S. I. Consortium. International magnetic tape storage roadmap. http://www.insic.org/news/2015roadmap/15index.html.

[5] T. Coughlin. Flash memory areal densities exceed those of hard drives. https://www.forbes.com/sites/tomcoughlin/2016/02/03/flash-memory-areal-densities-exceed-those-of-hard-drives/#113504fe7c72.

[6] G. Graefe. The five-minute rule 20 years later (and how flash memory changes the rules). *Commun. ACM*, 52(7), 2009.

[7] J. Gray. The five-minute rule. research.microsoft.com/en-us/um/people/gray/talks/fiveminuterule.ppt.

[8] J. Gray and G. Graefe. The five-minute rule ten years later, and other computer storage rules of thumb. *SIGMOD Rec.*, 26(4), 1997.

[9] J. Gray and F. Putzolu. The 5 minute rule for trading memory for disc accesses and the 10 byte rule for trading memory for cpu time. In *SIGMOD*, 1987.

[10] IDC. Technology assessment: Cold storage is hot again âĂŤ finding the frost point. http://www.storiant.com/

resources/Cold-Storage-Is-Hot-Again.pdf.

[11] Intel. Cold Storage in the Cloud: Trends, Challenges, and Solutions. http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/cold-storage-atom-xeon-paper.pdf, 2013.

[12] A. Kathpal and G. A. N. Yasa. Nakshatra: Towards running batch analytics on an archive. In *MASCOTS*, 2014.

[13] M. Lantz. Why the future of data storage is (still) magnetic tape. https://spectrum.ieee.org/computing/hardware/why-the-future-of-data-storage-is-still-magnetic-tape.

[14] J. Lee, J. Ahn, C. Park, and J. Kim. Dtstorage: Dynamic tape-based storage for cost-effective and highly-available streaming service. In *CCGRID*, 2016.

[15] K. Lim, J. Chang, T. Mudge, P. Ranganathan, S. K. Reinhardt, and T. F. Wenisch. Disaggregated memory for expansion and sharing in blade servers. In *ISCA*, 2009.

[16] D. Lomet. Cost/performance in modern data stores: How data caching systems succeed. In *DAMON*, 2018.

[17] F. Moore. Storage outlook 2016. https://horison.com/publications/storage-outlook-2016.

[18] M. Perlmutter. The lost picture show: Hollywood archivists cannot outpace obsolescence. https://spectrum.ieee.org/computing/it/the-lost-picture-show-hollywood-archivists-cant-outpace-obsolescence, 2017.

[19] R. B. Shobana Balakrishnan, 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. In *OSDI*, 2014.

[20] Spectra. Arcticblue deep storage disk. Product, https://www.spectralogic.com/products/arcticblue/.

[21] SpectraLogic. Spectralogic t50e. http://www.pcm.com/p/Spectralogic-Tape-Drives/product~dpno~8247348~pdp.gchecjj.

[22] StorageReview. Intel optane memory review. http://www.storagereview.com/intel_optane_memory_review.

[23] H. I. Strategies. Tiered storage takes center stage. Report, 2015.

[24] TPC-C. Dell-microsoft sql server tpc-c executive summary. http://www.tpc.org/tpcc/results/tpcc_result_detail.asp?id=114112501, 2014.

[25] L. Ultrium. Lto ultrium roadmap. http://www.ltoultrium.com/lto-ultrium-roadmap/.

[26] K. Umamageswaran and G. Goindi. Exadata: Delivering memory performance with shared flash. http://www.oracle.com/technetwork/database/exadata/exadatadeliveringmemoryperformance-3518310.pdf.

[27] M. Yan. Open compute project: Cold storage hardware v0.5. http://www.opencompute.org/wp/wp-content/uploads/2013/01/Open_Compute_Project_Cold_Storage_Specification_v0.5.pdf, 2013.

[28] W. Yan, J. Yao, Q. Cao, C. Xie, and H. Jiang. Ros: A rack-based optical storage system with inline accessibility for long-term data preservation. In *EUROSYS*, 2017.