capstract:2611.00050v1 [cs.DB] 20 Nov 2026

Read Amplification in the Domestic Junk Drawer: A Log-Structured Cost Model and Field Measurements from 64 Households

Marguerite A. Solheim-Adeyemi
Storage Systems Laboratory, Ellesmere Institute of Technology
Rotterdam, Netherlands
m.solheim-adeyemi@storage.ellesmere.example

Tobias Wrenfield
Storage Systems Laboratory, Ellesmere Institute of Technology
Rotterdam, Netherlands

Priyanka Venkataraghavan
Department of Household Informatics, Kestrel Hill University
Bergen, Norway

Abstract

Households keep a single container, usually a kitchen drawer, for objects with no assigned location. We model this drawer as the top level of a log-structured merge tree, with other storage locations in the dwelling as lower levels, and test three predictions of the model on 64 households. Over eight weeks of overhead video we annotated 7,893 lookups. Positive lookups touched 26.9 items on average (household mean), against 21.9 predicted by the expected search cost of a move-to-front self-organising list. Household beliefs that an item is present behaved like Bloom filters combined by union: the false-positive rate rose from 7.8% in one-person households to 45.6% in households of four or five, and a per-member rate of 11.3% (95% CI 10.2 to 12.5) fitted the data. A twelve-month movement diary kept by 58 households gave a geometric mean size ratio between levels of 2.77 (95% CI 2.62 to 2.92), compatible with e, the minimiser of leveled write amplification. Negative lookups were 17% of lookups and about 39% of search time, and tombstones accounted for 19% of drawer contents. In a trace replay of six maintenance policies, a four-compartment divider insert gave the lowest latency (18.7 s per lookup), and scheduled tiering, which we propose, halved the cost of negative lookups.

CCS Concepts

• Information systems → Data structures; • Information systems → Storage management; • Theory of computation → Caching and paging algorithms.

Keywords

LSM-tree, read amplification, self-organizing lists, Bloom filters, tombstones, storage hierarchy, trace replay

Reference Format:
Marguerite A. Solheim-Adeyemi, Tobias Wrenfield, and Priyanka Venkataraghavan. 2026. Read Amplification in the Domestic Junk Drawer: A Log-Structured Cost Model and Field Measurements from 64 Households. Preprint, November 20, 2026.

1 Introduction

Most households keep one container, usually a kitchen drawer, for objects that have no assigned location: adhesive tape, spare batteries, scissors, takeout menus, keys of uncertain function. In a screening survey of 212 respondents, 91% could identify theirs without prompting. From the point of view of data management it is a key-value store with no schema, no index and no administrator. Items are inserted at the front, retrieved by scanning, and removed rarely and for reasons nobody records.

The storage-engine literature has a settled vocabulary for the cost of such a system. Read, write and space amplification are the standard quantities, and the RUM conjecture states that an access method cannot minimise all three at once [1]. Studies of household storage have measured retrieval time under different arrangements [14, 20], but they report seconds and do not separate the cost of finding an item from the cost of maintaining the container. That separation is what lets a design be compared against another, so we make it here.

We treat the drawer as the top level (L_0) of a log-structured merge tree [17], with the other places in the dwelling where objects are kept as the lower levels. We then test three predictions of the resulting cost model against measurements from 64 households.

1. Retrieval cost in the drawer follows the expected search cost of a move-to-front self-organising list [18] under a Zipf-like request distribution.
2. A household's belief that an item is in the drawer behaves like a Bloom filter [4], and the filters of co-resident members combine by union, so the false-positive rate grows with the number of occupants.
3. The size ratio between adjacent levels is close to e, the value that minimises leveled write amplification.

We recorded 7,893 lookups in 64 households over eight weeks, and followed 58 of them with a movement diary for a further twelve months. Positive lookups touched 26.9 items on average (household mean), against 21.9 predicted for a strict move-to-

front list. The union model fitted the false-positive rates of multi-occupant households with a per-member rate of 11.3% (95% CI 10.2 to 12.5). The geometric mean size ratio was 2.77 (95% CI 2.62 to 2.92). A trace-driven comparison of six maintenance policies found that a commercial divider insert gave the lowest latency, and that scheduled tiering, the policy we propose, gave the lowest cost for negative lookups.

2 Background

2.1 Log-structured merge trees

An LSM-tree buffers insertions in a small level L_0 and moves data to progressively larger levels $L_1, \dots, L_k$ by merging [16, 17]. With a size ratio T between adjacent levels, the number of levels for a dataset R times larger than L_0 is $\ln R / \ln T$. In leveled merging each level holds one sorted run, and an entry is rewritten about $T/2$ times per level on its way down [11]. Tiered merging reduces writes at the cost of more runs to probe on a read. A lookup probes levels from the top and stops at the first match, and a lookup for an absent key must probe all of them, which is why implementations attach a Bloom filter to each run [10]. Deletes are written as tombstones that occupy space until a compaction reaches the level holding the original entry, and space amplification, the ratio of stored to live data, is driven largely by them [12]. When L_0 fills faster than compaction can drain it, writes are throttled or stopped [2].

2.2 Self-organising lists

A self-organising list reorders itself on access so that frequently requested items drift towards the front. Under the move-to-front rule an accessed item is relocated to the head [19]. With requests drawn independently with probabilities $p_1, \dots, p_n$, the expected position of the requested item is

$$\mathbb{E}[\mathrm{RA}_{\mathrm{MTF}}] = 1 + 2 \sum_{1 \le i < j \le n} \frac{p_i \, p_j}{p_i + p_j}, \tag{1}$$

a result due to Rivest, who also showed that the transposition rule, which swaps the accessed item with its predecessor, has no larger expected cost in the stationary regime [18]. Ordering by decreasing request frequency is optimal for a known stationary distribution. Move-to-front adapts faster when the distribution changes [3, 15].

2.3 Bloom filters

A Bloom filter answers set membership with one-sided error: a present key is never reported absent, and an absent key is reported present with probability ϵ [4]. Two filters built with the same parameters over different sets can be combined by bitwise OR, and the result answers queries for the union. The false-positive rate of the combined filter is higher than that of either input [7].

2.4 Household storage

Havelock and Ndungu [14] timed retrieval of kitchen utensils under three storage regimes and found no effect of regime on items used daily. Brandt-Okeke [5] modelled the accumulation of objects in home inventories as a stochastic process with a small deletion rate. Tsunoda et al. [20] related the recency of an item's last use to retrieval time in 31 households. None of these studies counted items touched, and none had access to a second level of storage.

3 A Cost Model of the Drawer

3.1 Levels

We number the drawer L_0, with capacity B items. Other storage locations form $L_1, \dots, L_k$ in order of increasing retrieval cost, for example a hall cupboard, a shelf in the utility room, and boxes in a basement, garage or loft. Level L_i has capacity $B\,T^{\,i}$ items for a size ratio T between adjacent levels, which we estimate per household. Insertions enter at the front of L_0. A compaction event relocates items from L_i to L_{i+1} and discards tombstones, and it may cascade.

3.2 Read path

A lookup for a live key present in L_0 is a scan from the front. We define read amplification RA as the number of items touched, meaning lifted, displaced or moved aside, up to and including the target. Participants return a retrieved item to the front of the drawer on most occasions and elsewhere on the rest, so the drawer is a move-to-front list with a leak. If the leak is zero, Eq. (1) gives the expected RA. If an item is returned to a uniformly random position with probability $1 - q$, the expected cost is higher, and we estimate the effect by replay (Section 6).

A lookup for a key that is not in L_0 costs $|L_0|$, the full drawer, because a scan cannot stop early. We call such a lookup negative. Tombstones count towards this cost.

3.3 Belief filters

Before opening the drawer, a household member usually has a belief about whether the item is in it. We model each member's belief as a Bloom filter with false-positive rate f and no false negatives. A household consults the drawer when any member believes the item is there, so the household's filter is the union of its members'. For k occupants whose errors are independent,

$$\epsilon_k = 1 - (1 - f)^k. \tag{2}$$

An unjustified belief produces a negative lookup. We make no claim about how members form their beliefs, and the model needs only that the errors are independent across members.

3.4 Space amplification

Let τ be the fraction of items in the drawer that are tombstones: items with no remaining use, such as depleted batteries, dried markers and expired coupons, so judged by the participant at the annual inventory. Space amplification is $\mathrm{SA} = 1 / (1 - \tau)$.

3.5 Write amplification and the size ratio

In a leveled design with size ratio T and R the ratio of total capacity to L_0 capacity, an item passes through $L = \ln R / \ln T$

levels and is rewritten about $T/2$ times at each, so

$$W(T) = \frac{T}{2} \cdot \frac{\ln R}{\ln T}. \tag{3}$$

Proposition 1 (Optimal size ratio)**.** *For fixed $R > 1$, $W(T)$ is minimised over $T > 1$ at $T = e$. Over integer $T \geq 2$ it is minimised at $T = 3$, and $W(2) = W(4)$.*

Proof. Write $W(T) = \frac{1}{2}\ln R \cdot g(T)$ with $g(T) = T/\ln T$. Then $g'(T) = (\ln T - 1)/\ln^2 T$, which is negative for $1 < T < e$ and positive for $T > e$, so g has its minimum at e. For integers, $g(2) = 2.885$, $g(3) = 2.731$ and $g(4) = 2.885$, and g increases for $T \geq 4$. The equality $g(2) = g(4)$ follows from $4/\ln 4 = 4/(2\ln 2) = 2/\ln 2$. ∎

The minimum is shallow. With $R = 24.3$, the ratio of mean level capacities we observed in four-level households (2,191 to 90 items), $W(e) = 4.34$ and $W(2) = W(4) = 4.60$, and W exceeds 5.0 only for T below 1.75 or above 5.1.

4 Study Design

4.1 Households and drawers

We recruited 64 households through a housing cooperative mailing list, from 212 that expressed interest. Households had a median of 2 occupants (1 to 5; 9, 31, 15, 8 and 1 households with 1, 2, 3, 4 and 5 occupants). We asked each household to name the drawer where they keep “the things that don’t go anywhere else”, and studied that drawer. Between March and September 2025 we photographed its contents at enrolment. The median drawer held 72.5 items (range 25 to 344).

4.2 Instrumentation and annotation

A fixed overhead camera, mounted over the drawer and angled to show hands and contents but not faces, recorded whenever the drawer was open for eight weeks. Two annotators independently marked each lookup with the requested item class, the number of items touched before the target was found, whether the target was found, and where the retrieved item was returned. On a 10% subsample (789 lookups) agreement on the items-touched count within one item gave Cohen’s $\kappa = 0.91$ [8]. Disagreements were resolved by a third annotator. Insertions were logged when an item entered the drawer. A lookup was positive if the target was found in the drawer. We analysed 7,893 lookups, of which 6,549 were positive and 1,344 negative (17.0%), and 497 insertions.

We fitted retrieval time to items touched by ordinary least squares and obtained $t = 3.1 + 0.62\,r$ seconds ($R^2 = 0.79$, $n = 7{,}893$). We use this relation to convert items touched into latency in Section 6. Retrieving an item from L_1 took a median of 41 s (IQR 29 to 66) in the 212 retrievals for which a participant carried one out while the camera was running. This is the cost we assign to an L_1 hit.

4.3 Belief probes

At the end of the eight weeks, each household answered 40 probes of the form “Is there a [item] in the drawer?”. Twenty probe items were present and twenty were absent, as verified by the annotators, and we drew probes from the 60 most frequently requested item classes in the sample. A household answered “yes” if any member said so, with members asked in sequence and the answer recorded as soon as one said yes. This protocol yields the household-level filter that Eq. (2) describes.

4.4 Movement diary

Compaction events and relocations are too rare to observe in eight weeks. Fifty-eight of the 64 households therefore kept a movement diary for twelve months after the camera period, recording each relocation of an item between locations and each discard, and the reason, and completed an inventory of all locations at the end. Six households did not complete the diary. The inventory gave the capacity of each level, from which we estimated the size ratio T per household as the exponent of a log-linear regression of level size on level index. Households had 3 levels (13) or 4 levels (45). The ethics board of the host institution approved the study (protocol 2024-0417), and all adult members of each household gave written consent.

4.5 Statistics

Intervals are 95% percentile bootstrap intervals over households with 2,000 resamples [13], unless stated. Pooled quantities weight each lookup equally, and household means weight each household equally.

5 Results

5.1 Workload

Requests were concentrated on a few item classes (Table 1). The Zipf exponent [6], estimated per household by maximum likelihood on item-level request counts, had mean 0.95 (SD 0.12), within the range typical of key-value benchmarks [9]. The ratio of lookups to insertions was 15.9 to 1. Negative lookups were 17% of the total. Of the 1,344 negative lookups, 54 (4.0%) were followed within seven days by the insertion of an item of the same class. We counted these as duplicate versions of an existing key. Lighters had a mean of 3.4 live copies in the 38 drawers that held any, tape measures 2.2 copies (47 drawers) and scissors 1.9 (61 drawers).

Table 1: Most frequently requested item classes among the 6,549 positive lookups, with items touched (pooled mean).

Item class	Lookups	Share (%)	Items touched
Scissors	642	9.8	9.1
Adhesive tape	596	9.1	14.8
Batteries (AA, AAA)	570	8.7	19.7
Pens and markers	537	8.2	18.2
Lighters and matches	432	6.6	22.4
Screwdrivers	354	5.4	31.0
Tape measures	308	4.7	36.9
Keys (labelled)	269	4.1	28.3
Other (128 classes)	2,841	43.4	34.1
All	6,549	100.0	26.3

5.2 Read amplification

Positive lookups touched a mean of 26.3 items (pooled; 26.9 as a household mean, 95% CI 23.3 to 31.1), a median of 15 and a 90th percentile of 62. The mean of the closed form in Eq. (1) over the 64 households, using each household's fitted distribution and drawer size, was 21.9. Strict move-to-front replay of the same traces gave 22.6 (19.8 to 26.0), 3% above the closed form. The excess of the observed value over the replay is 4.3 items. In 22% of retrievals the item was returned to a place other than the front, and a replay that reproduces this leak at the observed rate recovers the observed mean (Section 6).

Figure 1 shows items touched by popularity decile. The most requested decile, which accounts for 51% of positive lookups, cost 17.3 items observed against 9.8 under strict move-to-front. The two curves agree from the fourth decile onward, so the leak costs the most where the list is shortest.

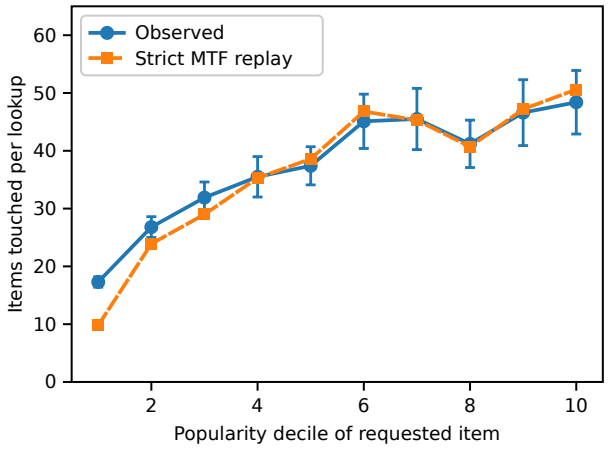


Figure 1: Items touched per positive lookup by popularity decile of the requested item (1 = most requested), pooled over 6,549 lookups. Bars show 95% intervals for the observed values. The dashed line is a replay of the same traces under strict move-to-front.

5.3 Negative lookups and belief filters

A negative lookup cost a mean of 93.9 items touched, which is the full drawer, or 61 s under the time model. Negative lookups were 17.0% of lookups and used about 39% of the total search time.

Table 2 gives the probe results by household size. The false-positive rate rose from 7.8% in one-person households to 45.6% in households of four or five. No household of three or more answered "no" for a present item. In one-person households the true-positive rate was 94.4%, with 10 of 180 present items missed. The maximum-likelihood estimate of the per-member rate f in Eq. (2), fitted to all 1,280 probes jointly, was 0.113 (profile-likelihood 95% CI 0.102 to 0.125). Figure 2 shows the fitted curve against the observed rates. The four-and-five-occupant group is above the fitted line (45.6% against 38.9%), although the intervals overlap, and it has few households.

Table 2: Belief probes by household size. Each household answered 20 absent-item probes. FPR is the share of absent items reported present, with the half-width of the normal-approximation 95% interval in parentheses. Predicted FPR uses Eq. (2) with $f = 0.113$.

Occupants	Households	Absent probes	False positives	FPR (%)	Predicted FPR (%)
1	9	180	14	7.8 (3.9)	11.3
2	31	620	126	20.3 (3.2)	21.3
3	15	300	89	29.7 (5.2)	30.2
4–5	9	180	82	45.6 (7.3)	38.9

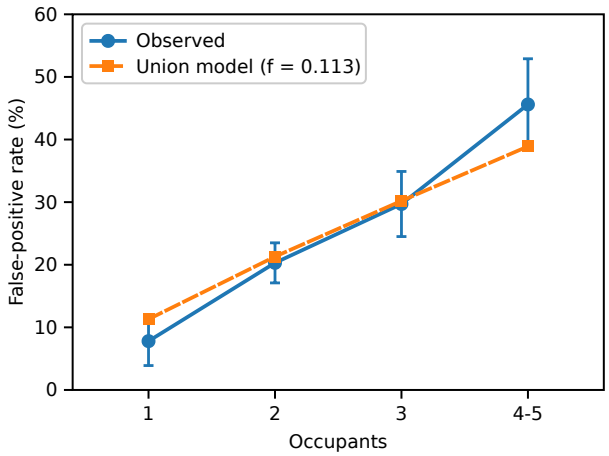


Figure 2: False-positive rate of the household belief that an absent item is in the drawer, by number of occupants. Bars show 95% intervals. The dashed line is the union model with per-member rate f = 0.113.

5.4 Tombstones and space amplification

The mean fraction of tombstones in the drawer was $\tau = 0.190$ (SD 0.067), giving $\mathrm{SA} = 1.23$. The largest classes were depleted batteries, dried markers and pens, expired coupons and menus, and twist ties without a bag. A drawer with the mean tombstone fraction wastes 17.8 of the 93.9 items touched in a negative lookup, about 11 s. Tombstone fractions were higher in drawers that had gone longer since a compaction (Spearman $\rho = 0.52$, $n = 58$).

5.5 Compaction

The 58 diary households recorded 71 compaction events in twelve months (1.22 per household; 11 households had none). We classified each by its trigger, defined as the circumstance reported to have started it (Table 3). In 34 events the drawer would no longer close. In the other 37 the trigger came from outside the drawer. A mean of 205 relocations took place per event, counting cascades into lower levels.

Table 3: Triggers of compaction events in the 58 diary households.

Trigger	Events	Share (%)
Drawer would not close	34	47.9
External (visitors, inspection, move)	22	31.0
Calendar (season, holiday)	15	21.1
All	71	100.0

The estimated size ratio per household had a geometric mean of 2.77 (95% CI 2.62 to 2.92), a median of 2.81 and a range of 1.58 to 4.51. The interval contains $e = 2.718$. Fifty-one of 58 households (88%) had T between 2 and 4. Mean capacities for the 45 four-level households were 90, 259, 714 and 2,191 items.

Observed write amplification, defined as relocations per inserted item over the twelve months, had a mean of 4.89 (SE 0.12). The model of Eq. (3) evaluated at each household's own T and $R = 24.3$ gives a mean of 4.44, so observed values were 10% above the model. Figure 3 plots the binned observations against $W(T)$. Households with $T < 2.2$ had the highest write amplification (5.46, SE 0.48), and households with T between 2.2 and 2.6 the lowest (4.52, SE 0.21). The data are too sparse above $T = 3.5$ to say much about the right arm of the curve.

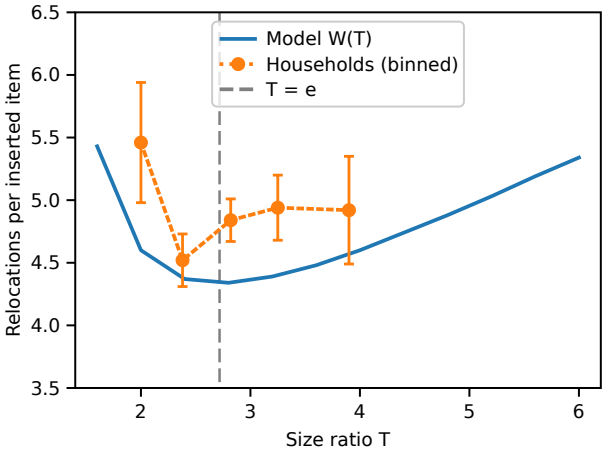


Figure 3: Write amplification against size ratio T. The curve is W(T) from the leveled model with R = 24.3. Points are household means in bins of T with one standard error (58 households). The dashed vertical line marks T = e.

6 Maintenance Policies by Trace Replay

6.1 Policies

We replayed the 7,893 lookups under six policies, each applied to the same request sequence and starting drawer. Before the measured trace, each replay ran 300 synthetic lookups drawn from the household's fitted request distribution, so that the drawer order had reached its stationary behaviour. Insertions were not replayed. The policies are:

- **Observed.** Move-to-front with a return to the front with probability $q = 0.78$ and to a uniformly random position otherwise. This reproduces the field behaviour.
- **Strict MTF.** Move-to-front with $q = 1$.
- **Transpose.** Swap the retrieved item with its predecessor [18].
- **Frequency order.** Items sorted by true request probability. This needs knowledge that households do not have and is reported as a reference for single-list policies.
- **Dividers.** A four-compartment insert. Each item is assigned a compartment, and a lookup scans its own compartment first. With probability 0.15 an item is in the wrong compartment, and the scan then first covers half of the other compartments. Within a compartment, order is move-to-front.
- **Scheduled tiering (ST).** Every 14 days, tombstones are discarded and all but the 48 most recently used live items are demoted to L_1. A lookup that misses in L_0 is served from L_1 at 41 s extra. Retrieved items return to the front (Algorithm 1).

Algorithm 1 Scheduled tiering (ST) of the drawer

Require: drawer D kept from the front backwards, hot-set size K, period P, level L_1
every P days **do**
  discard every tombstone of D
  demote all items of D beyond position K into L_1
end
procedure Lookup(x)
  scan D from the front ▷ cost = items touched
  if $x \notin D$ **then**
    fetch x from L_1
  end if
  place x at the front of D
  return x
end

6.2 Results

Table 4 gives items touched for positive and negative lookups and latency under the time model. The replay of the observed policy gave a mean positive cost of 26.9 and a latency of 26.4 s (23.2 to 30.1), matching the field values by construction of the leak parameter.

Table 4: Trace replay of 7,893 lookups under six policies. Positive and negative columns give items touched (household mean). Latency is pooled, in seconds, with 95% bootstrap intervals over households. The last column is the paired difference from the observed policy. Best in bold.

Policy	Positive	Negative	Latency (s)	Difference (s)
Observed (leaky MTF)	26.9	93.9	26.4 [23.2, 30.1]	—
Strict MTF	22.6	93.9	24.3 [21.4, 27.6]	−2.1 [−2.6, −1.6]
Transpose	35.2	93.9	30.7 [26.1, 35.5]	+4.2 [2.8, 5.8]
Frequency order (oracle)	16.3	93.9	21.2 [18.9, 24.0]	−5.2 [−6.2, −4.3]
Dividers (4 compartments)	**11.3**	93.9	**18.7** [16.4, 21.3]	**−7.7** [−9.0, −6.6]
Scheduled tiering (ST)	18.8	**46.9**	21.0 [19.2, 22.9]	−5.5 [−8.0, −3.5]

Dividers gave the lowest latency and the lowest positive cost, below the frequency-ordered oracle because a divider also restricts the scan to a subset of the contents, which no single list can do. Transpose was worse than the observed policy. In the stationary regime it is the better rule [18], but 300 warm-up lookups per household were not enough for it to converge, and the field traces average 123 lookups per household over the whole study. Strict move-to-front was 2.1 s faster than the field policy. ST halved the cost of negative lookups (46.9 against 93.9 items) and had the narrowest latency interval of the six policies. 10.0% of its positive lookups (655 of 6,549) were served from L_1 and paid the 41 s penalty. Dividers do nothing for negative lookups, since every compartment must be scanned.

7 Discussion

Closed-form read amplification from Eq. (1) came within 3% of a strict-MTF replay, and the field value was 23% above it. The gap is consistent with the 22% of items that participants returned to non-front positions, and replay puts the cost of that habit at 2.1 s per lookup. We did not test whether participants can be made to stop.

The union model fitted the one-, two- and three-occupant rates well. It underestimated the four-and-five-occupant group, where the observed rate was 6.7 points above the prediction. Positively correlated errors among members would lower the union rate below the independence formula, which is the wrong direction for this group. We have no explanation, and the group contains nine households.

The size-ratio result needs care. $W(T)$ is flat near its minimum: any T from 2 to 4 is within 6.1% of $W(e)$, and 88% of households were in that range. The finding that the geometric mean is statistically compatible with e is therefore weak evidence for a design principle and does not require one. The binned write-amplification values were lowest at $T \approx 2.4$, not at e, with intervals that overlap those of neighbouring bins. We do not claim that households optimise T. We note only that no household in the diary reported a rule for sizing levels.

Compaction was mostly reactive. In 47.9% of events the drawer would no longer close, which is the stop-writes condition of an LSM-tree [2]. The remaining triggers came from outside. Scheduled tiering sidesteps both by demoting on a fixed period, and its main benefit in replay was on negative lookups, where the tombstone purge and the 48-item cap together halved the scan. It did not beat the divider insert on latency, and we have not modelled the loss of drawer volume to the dividers or the cost of keeping compartments correct, so the comparison may favour dividers. A divider insert and periodic tiering are compatible, and we did not test their combination.

7.1 Limitations

The sample was recruited from one cooperative and may differ from the general population in drawer contents and household

composition. Eight weeks of video captures the steady state of the read path but not compaction, which depends on a twelve-month diary that participants completed from memory and at their own initiative. Under-reporting of small relocations would bias write amplification downward. The latency model is linear in items touched and was fitted on the same lookups it is applied to. The replay fixes the request sequence, so the effect of a policy on later requests is not captured. Divider mis-filing at 0.15 and the ST parameters were chosen once and not tuned. Probes were drawn from common item classes, and the false-positive rate for rarer items may be higher.

8 Conclusion

Read cost in the drawer was close to the move-to-front prediction, and household beliefs about its contents combined across occupants as the union model says they should. The size ratio between levels was compatible with e, but the optimum is flat enough that this tells us little. Negative lookups were the costliest operation per request, mostly because of tombstones. Scheduled tiering halved their cost in replay and a divider insert gave the lowest overall latency. Combinations of the two should be tested in place, and insertion order should be recorded, which we did not do.

Data availability

Annotated lookup traces, probe responses and the replay code are available from the corresponding author.

References

[1] Manos Athanassoulis, Michael S. Kester, Lukas M. Maas, Radu Stoica, Stratos Idreos, Anastasia Ailamaki, and Mark Callaghan. 2016. Designing Access Methods: The RUM Conjecture. In *Proceedings of the 19th International Conference on Extending Database Technology*. 461–466.

[2] Oana Balmau, Florin Dinu, Willy Zwaenepoel, Karan Gupta, Ravishankar Chandhiramoorthi, and Diego Didona. 2019. SILK: Preventing Latency Spikes in Log-Structured Merge Key-Value Stores. In *Proceedings of the 2019 USENIX Annual Technical Conference*. 753–766.

[3] Jon L. Bentley and Catherine C. McGeoch. 1985. Amortized Analyses of Self-Organizing Sequential Search Heuristics. *Communications of the ACM* 28 , 4 (1985), 404–411.

[4] Burton H. Bloom. 1970. Space/Time Trade-offs in Hash Coding with Allowable Errors. *Communications of the ACM* 13 , 7 (1970), 422–426.

[5] Sigrid Brandt-Okeke. 2019. Clutter Accumulation as a Stochastic Process: Evidence from Home Inventories. In *Proceedings of the Conference on Household Informatics*. 44–53.

[6] Lee Breslau, Pei Cao, Li Fan, Graham Phillips, and Scott Shenker. 1999. Web Caching and Zipf-like Distributions: Evidence and Implications. In *Proceedings of IEEE INFOCOM '99*. 126–134.

[7] Andrei Broder and Michael Mitzenmacher. 2004. Network Applications of Bloom Filters: A Survey. *Internet Mathematics* 1 , 4 (2004), 485–509.

[8] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20 , 1 (1960), 37–46.

[9] Brian F. Cooper, Adam Silberstein, Erwin Tam, Raghu Ramakrishnan, and Russell Sears. 2010. Benchmarking Cloud Serving Systems with YCSB. In *Proceedings of the 1st ACM Symposium on Cloud Computing*. 143–154.

[10] Niv Dayan, Manos Athanassoulis, and Stratos Idreos. 2017. Monkey: Optimal Navigable Key-Value Store. In *Proceedings of the 2017 ACM International Conference on Management of Data*. 79–94.

[11] Niv Dayan and Stratos Idreos. 2018. Dostoevsky: Better Space-Time Trade-Offs for LSM-Tree Based Key-Value Stores via Adaptive Removal of Superfluous Merging. In *Proceedings of the 2018 International Conference on Management of Data*. 505–520.

[12] Siying Dong, Mark Callaghan, Leonidas Galanis, Dhruba Borthakur, Tony Savor, and Michael Stumm. 2017. Optimizing Space Amplification in RocksDB. In *Proceedings of the 8th Biennial Conference on Innovative Data Systems Research.*.

[13] Bradley Efron. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 7 , 1 (1979), 1–26.

[14] Rosalind Havelock and Amos Ndungu. 2014. Retrieval Times for Kitchen Utensils under Three Storage Regimes. *Journal of Domestic Ergonomics* 9 , 2 (2014), 101–118.

[15] Donald E. Knuth. 1998. *The Art of Computer Programming, Volume 3: Sorting and Searching*. Addison-Wesley, Reading, MA.

[16] Chen Luo and Michael J. Carey. 2020. LSM-Based Storage Techniques: A Survey. *The VLDB Journal* 29 , 1 (2020), 393–418.

[17] Patrick O'Neil, Edward Cheng, Dieter Gawlick, and Elizabeth O'Neil. 1996. The Log-Structured Merge-Tree (LSM-Tree). *Acta Informatica* 33 , 4 (1996), 351–385.

[18] Ronald Rivest. 1976. On Self-Organizing Sequential Search Heuristics. *Communications of the ACM* 19 , 2 (1976), 63–67.

[19] Daniel D. Sleator and Robert E. Tarjan. 1985. Amortized Efficiency of List Update and Paging Rules. *Communications of the ACM* 28 , 2 (1985), 202–208.

[20] Kenji Tsunoda, Lise Eriksen, and Mustafa Demirci. 2021. What the Drawer Remembers: Item Recency and Retrieval Time in Household Storage. *Journal of Domestic Ergonomics* 16 , 1 (2021), 23–40.