Northeastern University

UNIVERSITY OF TORONTO

清華大學 Tsinghua University

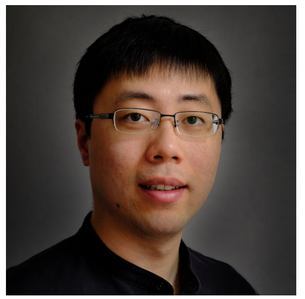
Aeris Filter

Robust range filtering through adaptivity

SIGMOD 2026

Yuvaraj Chesetti, Navid Eslami, Niv Dayan, Huanchen Zhang, Prashant Pandey

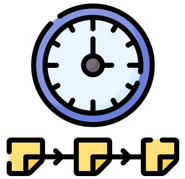


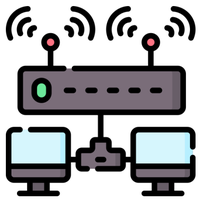

Range queries are ubiquitous

Range Query
Get all keys in range[a, b]

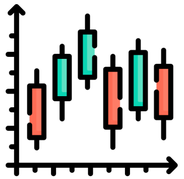
OLAP



Time-series



Networking

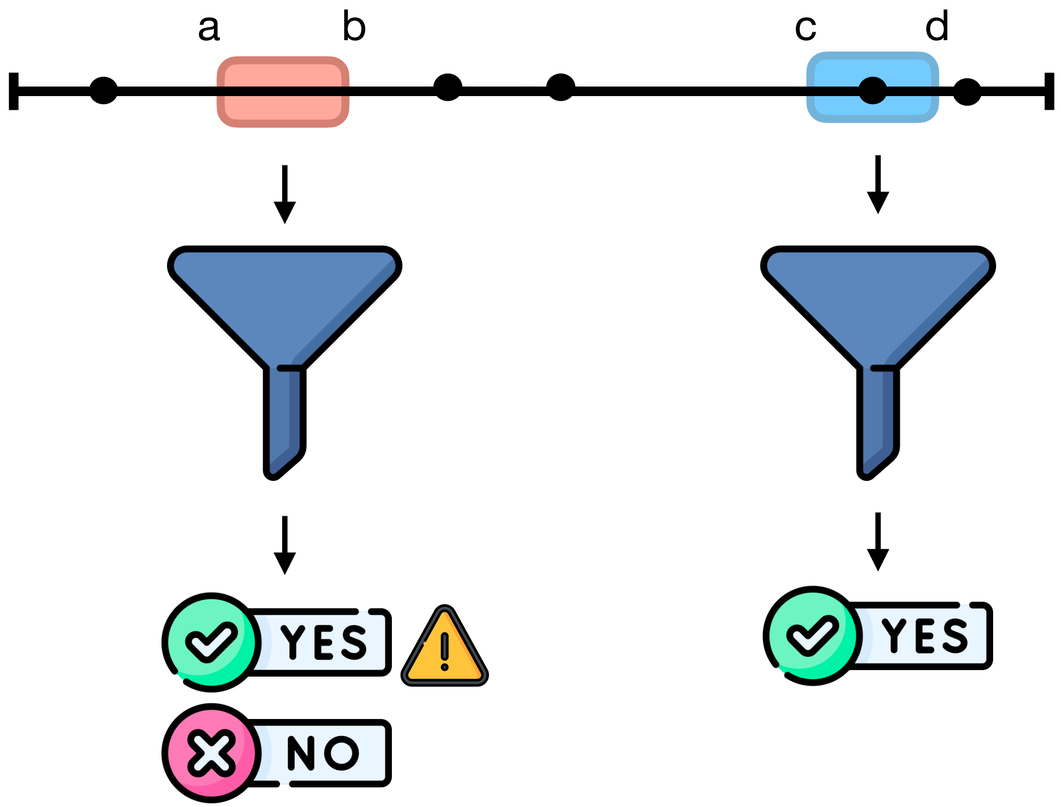


Financial Analysis

Bioinformatics

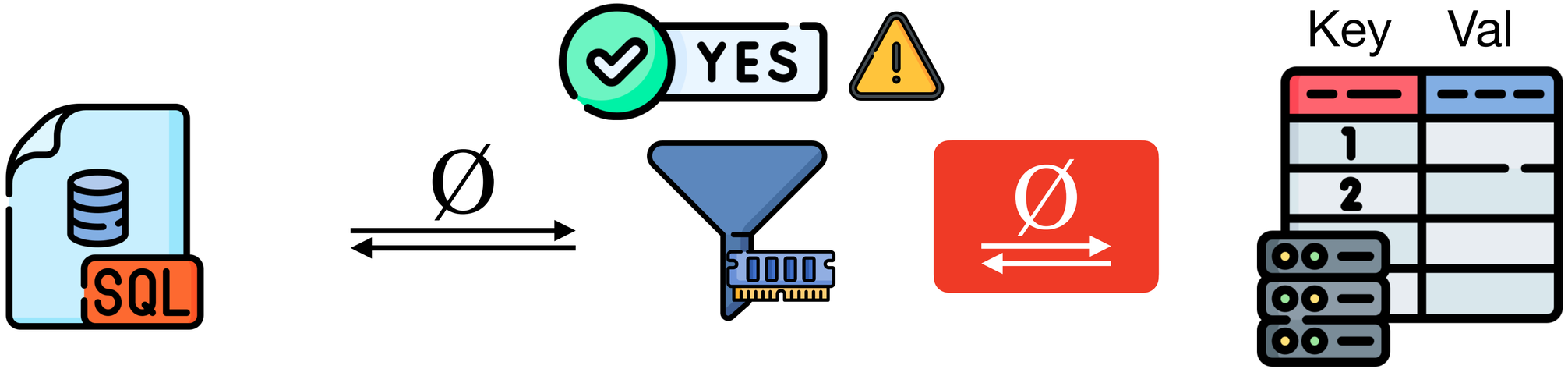
Spatial queries

What is a range filter?

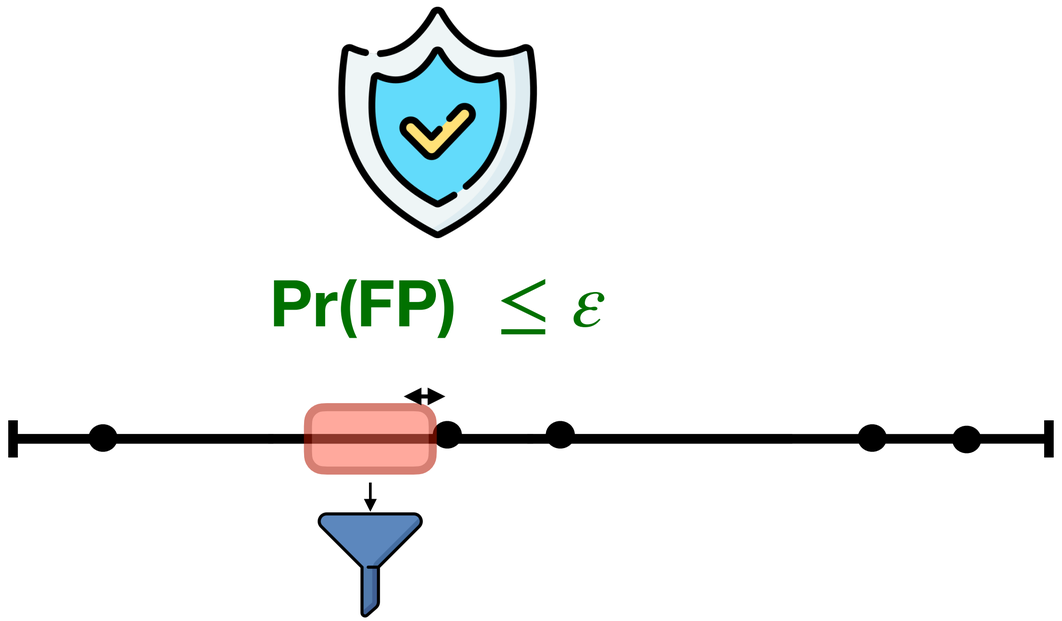


Range emptiness query: Are there any elements in the queried range?

False positives result in disk accesses



Spatially robust range filters

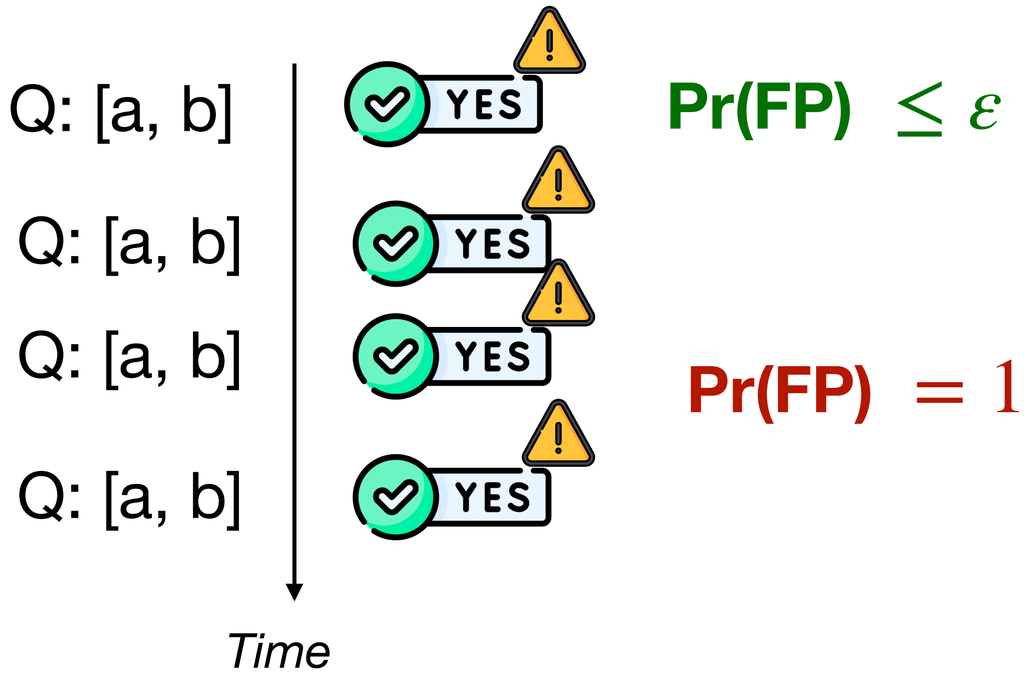


Rosetta	**Grafite**	**Memento**
SIGMOD 2020	SIGMOD 2024	SIGMOD 2025

Spatially robust: Bounded false positive rate, even under spatial skew

Non-robust: Unbounded false positive rate under spatial skew

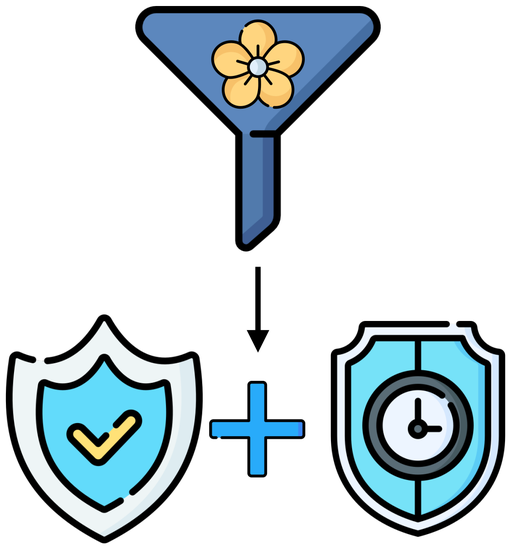
Spatial robustness ≠ Temporal robustness



Spatial robustness only hold for a single query in isolation

Temporal robustness needed to ensure guarantees hold across repeated queries

Aeris Filter

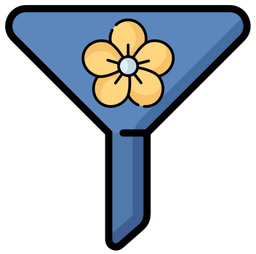


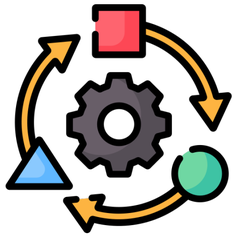
Spatial and **temporal** robustness

10x lower false positives on skewed workloads

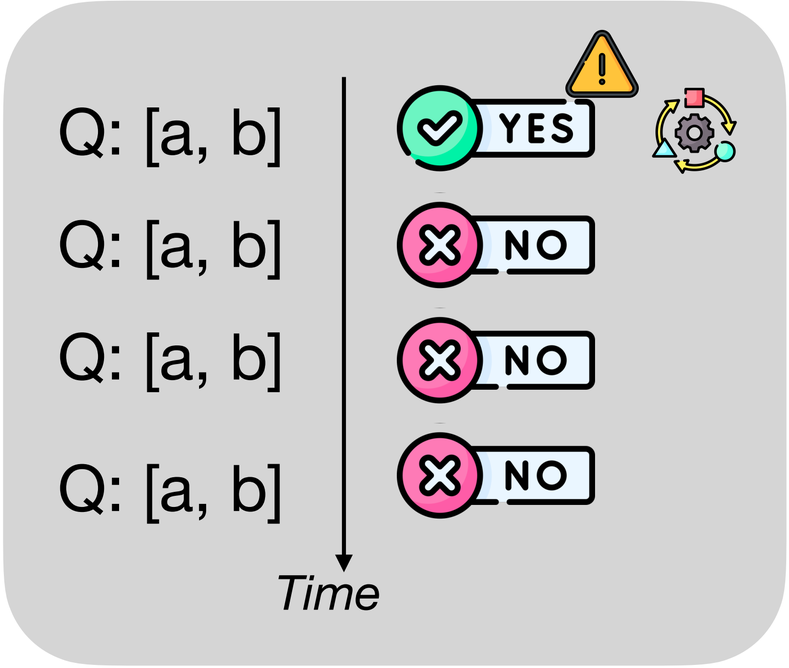
8x higher throughput on adversarial workloads

Based on Memento range filter
[ED, 2024]





Implements adaptivity
[BFG+, 2018]



Spatial + Temporal Robustness:

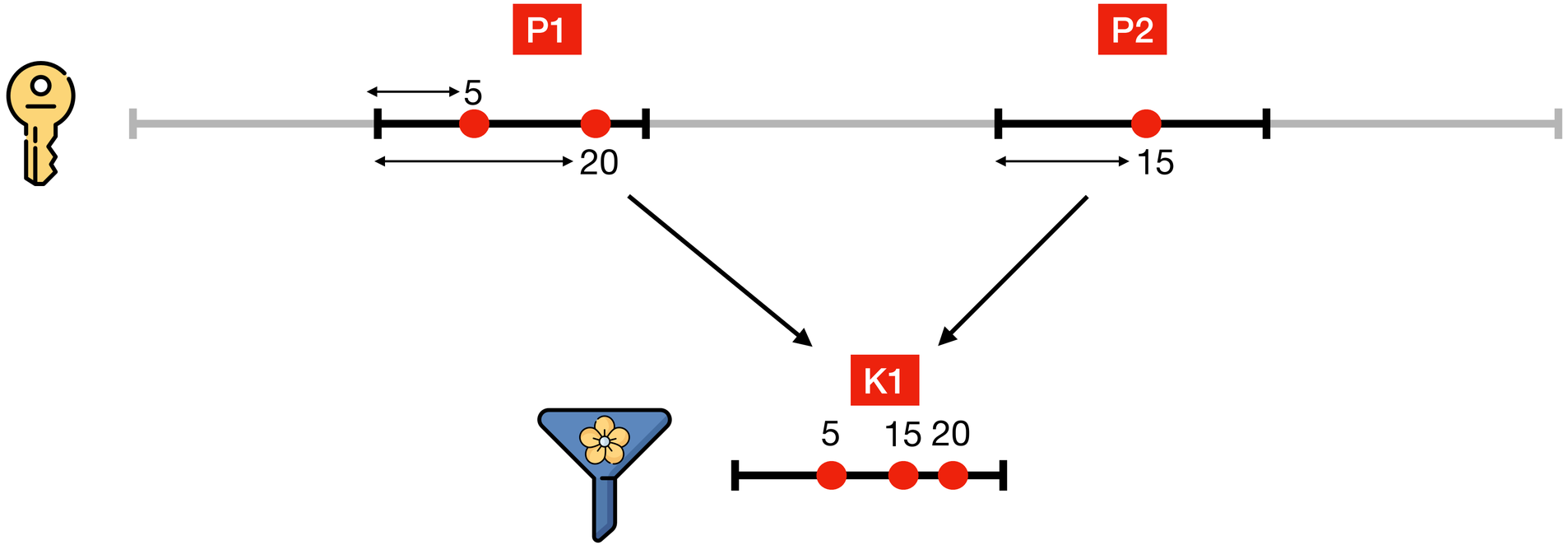
Pr(FP) $\leq \varepsilon$

Negligible space overhead:
1 extra metadata bit per item over Memento

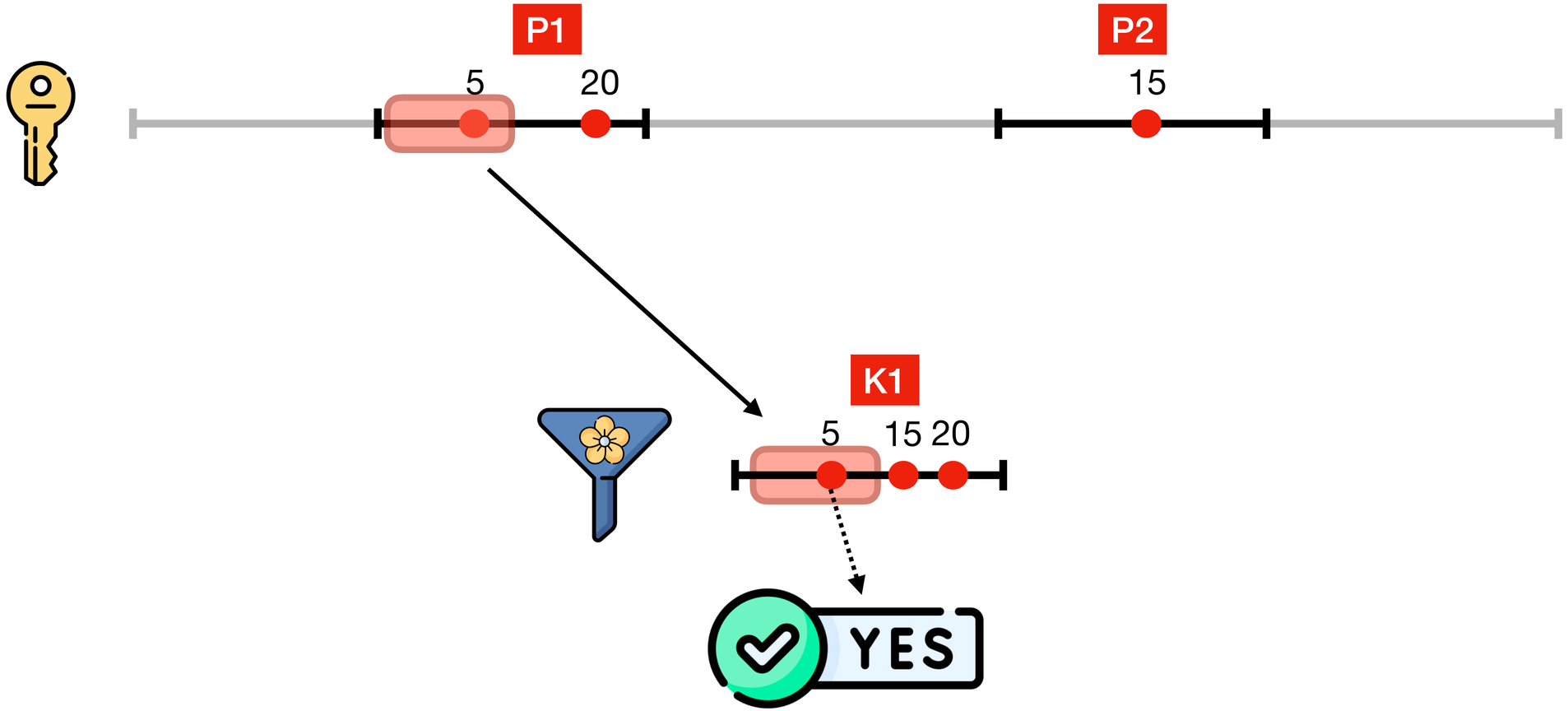
[BFG+2018]: Bender, M. A., Farach-Colton, M., Goswami, M., Johnson, R., McCauley, S., & Singh, S. (2018, October). Bloom filters, adaptivity, and the dictionary problem. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)* (pp. 182-193). IEEE.

[ED, 2024]: Eslami, N., & Dayan, N. (2024). Memento filter: A fast, dynamic, and robust range filter. *Proceedings of the ACM on Management of Data*, *2*(6), 1-27.

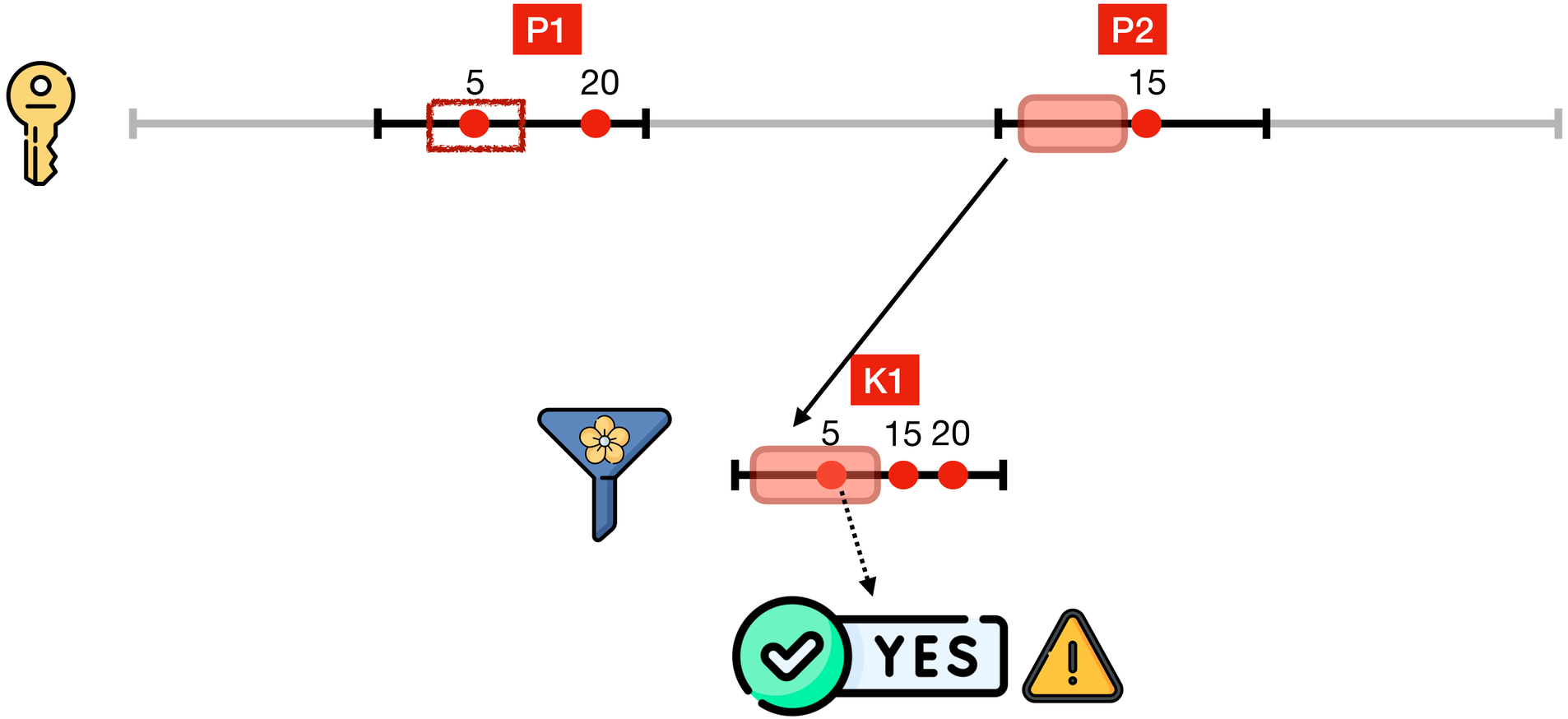
Memento combines partition key offsets



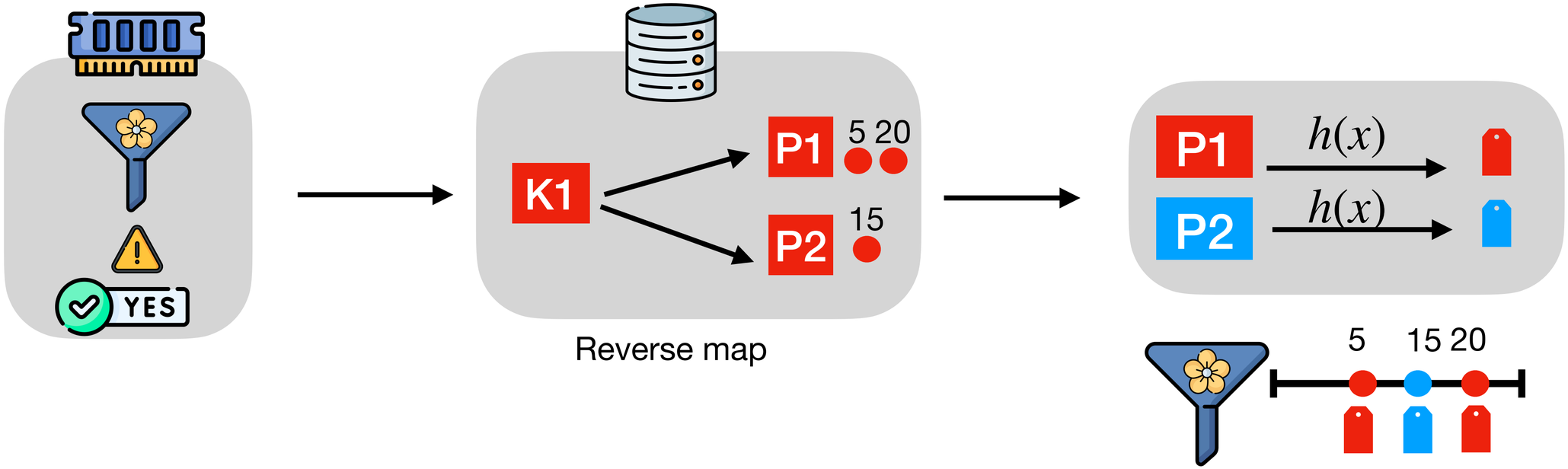
Queries are answered in filter space



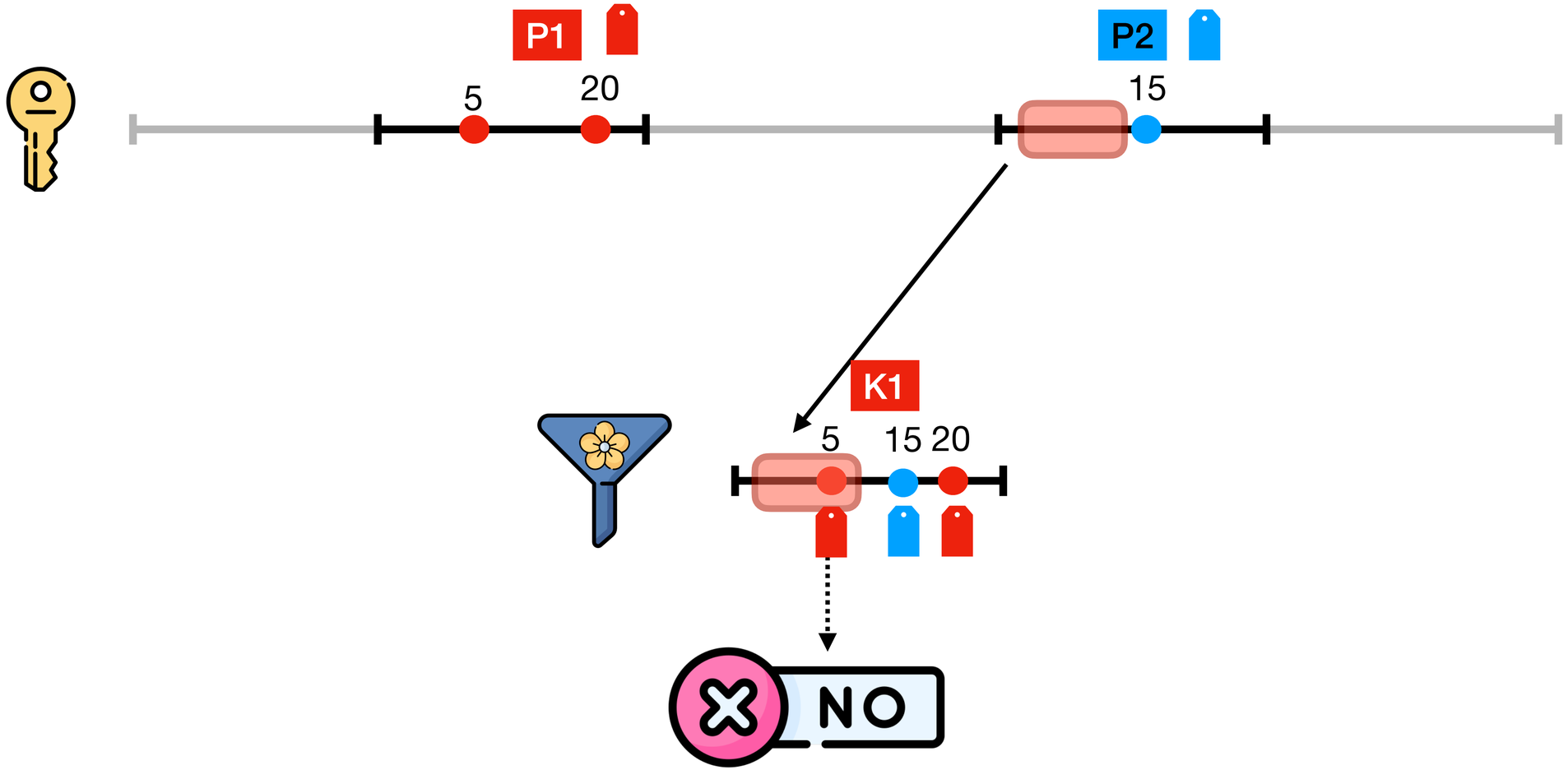
Where false positives come from



How Aeris filter adapts to false positives



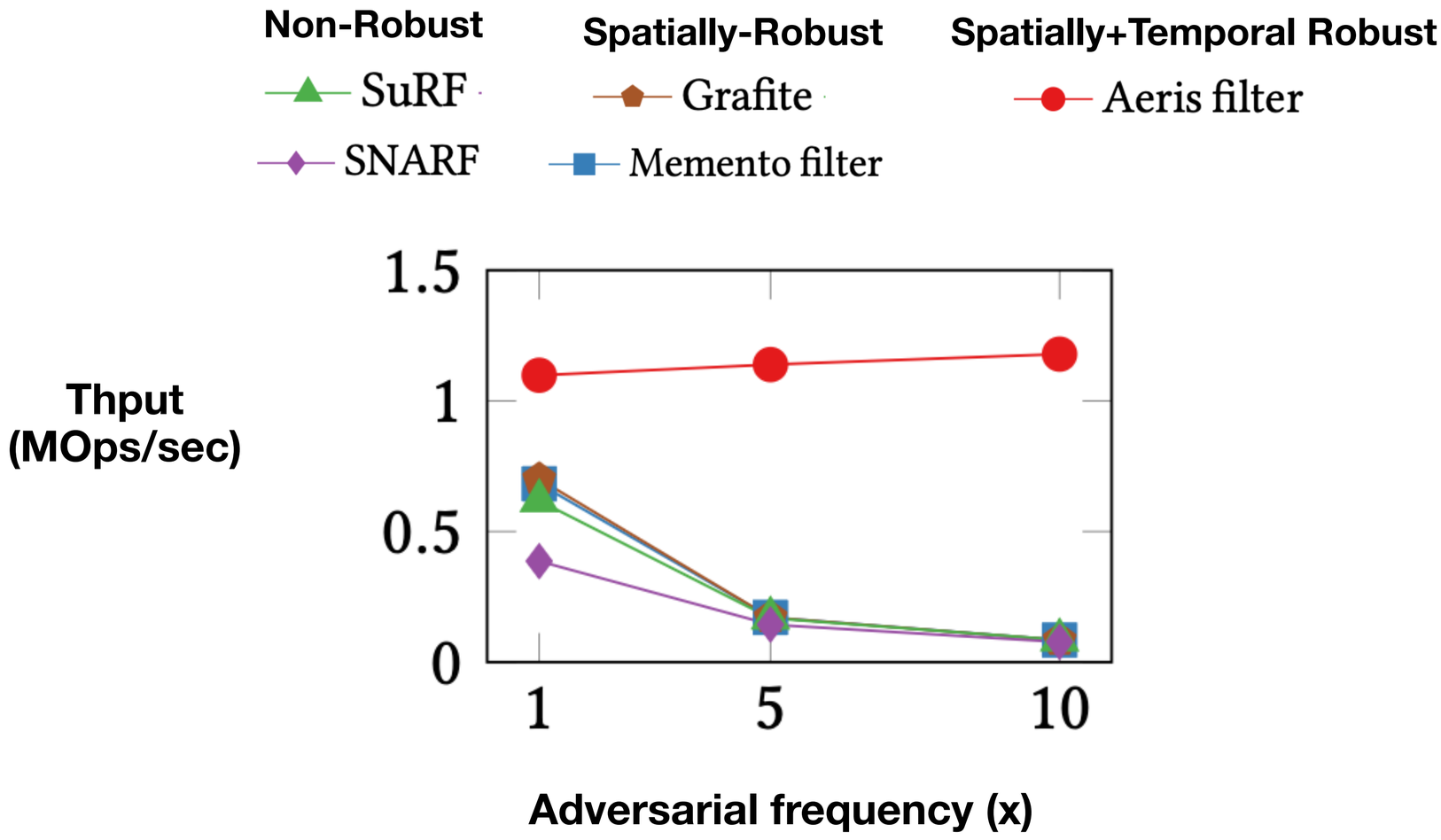
Disambiguating offsets resolves false positives



Evaluation setup

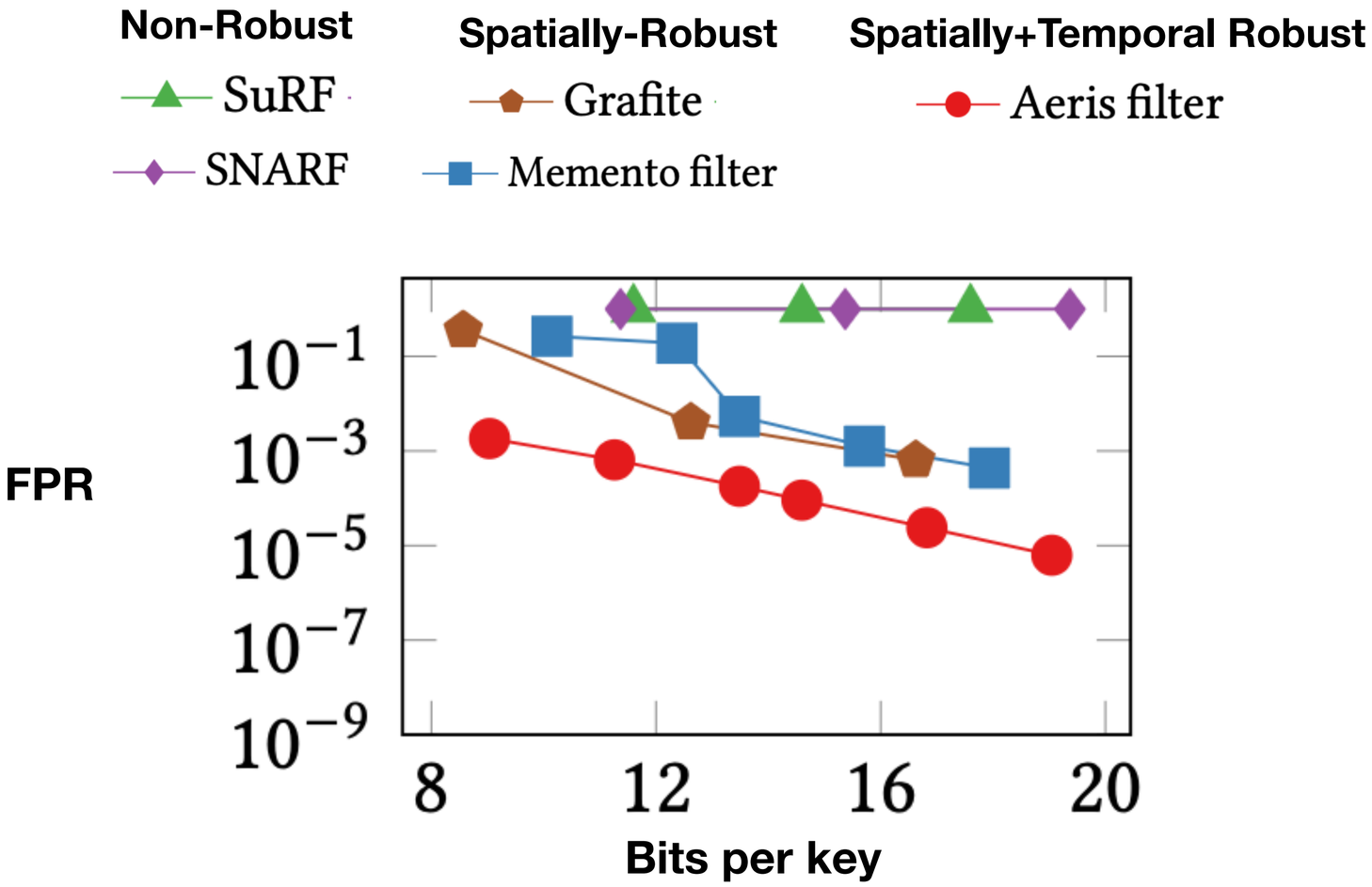
- WiredTiger database (100 million pairs of 64-bit keys, 512-bit values)
- Tested for point queries (R=1), short range queries (R=32) and long range queries (R=1024)
- All filters tested on 100 million queries with configured false positive rate $\epsilon = 2^{-9}$
- Filters
  - **Non robust:** SNARF, SuRF
  - **Spatially robust:** Grafite, Memento
  - **Spatially and temporally robust**: Aeris

System throughput on adversarial workload



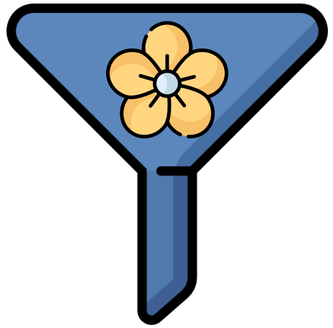
(Y-Axis: Linear scale, higher is better)

Temporal robustness at negligible space cost



(Y-Axis: Log scale, lower is better)

Takeaways



- **If your workload has any skew, Aeris is the right filter!**
- Provides both **spatial and temporal robustness**
  - **10x lower FPR** on skewed workloads,
  - **8x higher throughput** on adversarial workloads
  - Deamortizes filter expansions for **3x faster expansions**
  - Negligible space overhead
- **Code**: https://github.com/saltsystemslab/AerisFilter