Desperately Seeking ... Optimal Multi-Tier Cache Configurations

12th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage '20)

Tyler Estro[1], Pranav Bhandari[2], Avani Wildani[2], and Erez Zadok[1]

[1]Stony Brook University; [2]Emory University

EMORY UNIVERSITY Stony Brook University

Outline

- **Motivation**
- Cache Analysis
- Multi-Tier Cache Simulation
- Evaluation
- Conclusion

Motivation

- Multi-tier cache configuration space is huge
  - Many tiers, partitions, and arrangements
  - Virtual machines, cloud storage, distributed systems
  - Continuous advances in hardware and firmware
- Workloads and access patterns are evolving
  - Big data and growing working sets
  - Complex access patterns and multi-tenancy
    - Frequency and distribution of data reusage

Key Problems

- Cache analysis practices are short-sighted
  - Performance is often the sole focus
    - e.g., Miss ratio curves (MRCs)
  - Only a single tier of cache is analyzed
- Evaluated metrics are not comprehensive
  - Most common
    - Hit/miss ratio, average throughput, average latency
  - Fail to examine relationship between metrics
    - Is your cache $cost efficient?

Challenges

- Physical experiments are $costly and time-consuming
- Lack of a general, *n*-level I/O simulator
  - Publicly available I/O simulators
    - Simulate a fixed or limited number of tiers
    - Missing key features
    - Limited analysis tools
    - Not easily extendable
  - Proprietary and ad-hoc simulators

We propose best practices in cache research including the analysis of multi-tier configurations and a more comprehensive set of evaluation metrics (e.g., monetary cost)

Outline

- Motivation
- **Cache Analysis**
- Multi-Tier Cache Simulation
- Evaluation
- Conclusion

Cache Analysis Metrics

- Performance is constrained by $cost
  - Best practice: evaluate $cost and performance together
- Relevant metrics
  - Tail latency (P95/P99)
  - Total cost of ownership (TCO)
    - Purchase cost, energy cost, device lifetime, manpower, …
  - Inter-cache & network traffic
- Complex metrics can improve design decisions
  - Throughput/$, throughput/energy, …

Multi-Tier Analysis

- Performance gains do not translate to $cost efficiency
  - Behavior in one tier cascades throughout the system
  - Device characteristics of each tier can widely vary
    - Relative latency between tiers
    - Device durability (e.g., DRAM vs. SSD)
  - Best practice: analyze multi-tier configurations whenever possible
- Visualization
  - Trends in objectives are not obvious in multi-tier
    - Pareto frontier to locate optimal configurations
  - Interactive visual analytics [Cao et al., HotStorage '19]

Outline

- Motivation
- Cache Analysis
- **Multi-Tier Cache Simulation**
- Evaluation
- Conclusion

Multi-Tier I/O Simulator Features

- Write policy
  - Write through, write back, write around, user-defined
- Admission policy
  - Restrict admission of data at any tier(s) based on size or even ML algorithms
  - Inclusive, exclusive, or non-inclusive non-exclusive (NINE)
- Eviction policy
  - Individual layer algorithms, global-aware algorithms, or some mix
    - e.g., Adaptive Level-Aware Caching Algorithm (ALACA) [Cheng et al.]
- Trace sampling
  - Miniature Simulations [Waldspurger et al.]
- Preteching
  - MITHRIL [Yang et al.]

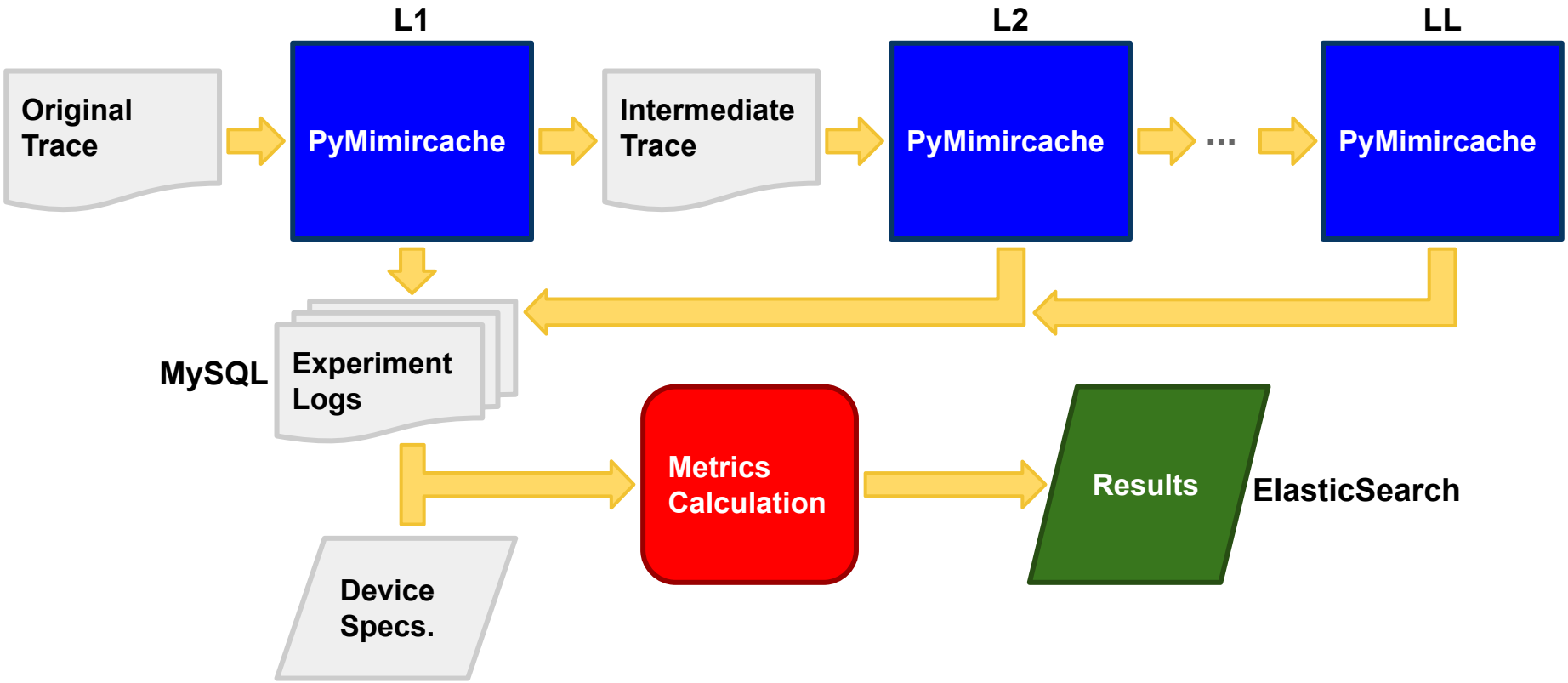
Reverse Miniature Simulations

- Multi-tier caching requires large datasets
  - Entire working set should not easily fit into DRAM
- Simulating large traces is extremely time-consuming
  - Use Miniature Simulations to sample the trace [Waldspurger et al.]
- What if our trace is too small?
  - Treat an original trace as if it is a scaled down version of a much larger trace

PyMimircache

- PyMimircache
  - Single layer cache simulator
  - Python framework with an efficient C backend
  - Supports numerous eviction algorithms
- Multi-tier PyMimircache extension
  - Write-through and write-back
  - MySQL and ElasticSearch
  - Simulate *n* cache tiers serially

PyMimircache Extension



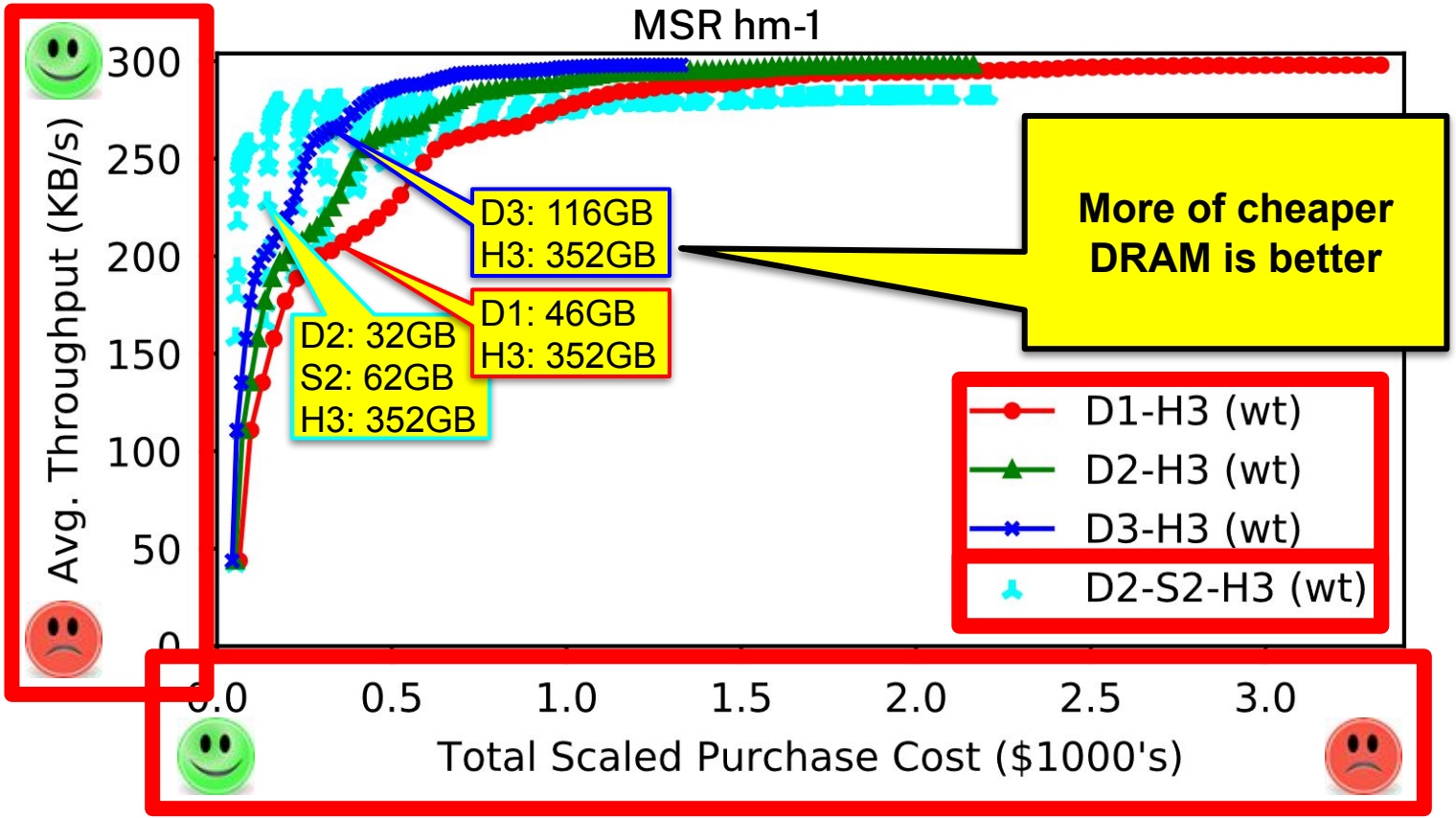
Outline

- Motivation
- Cache Analysis
- Multi-Tier Cache Simulation
- **Evaluation**
- Conclusion

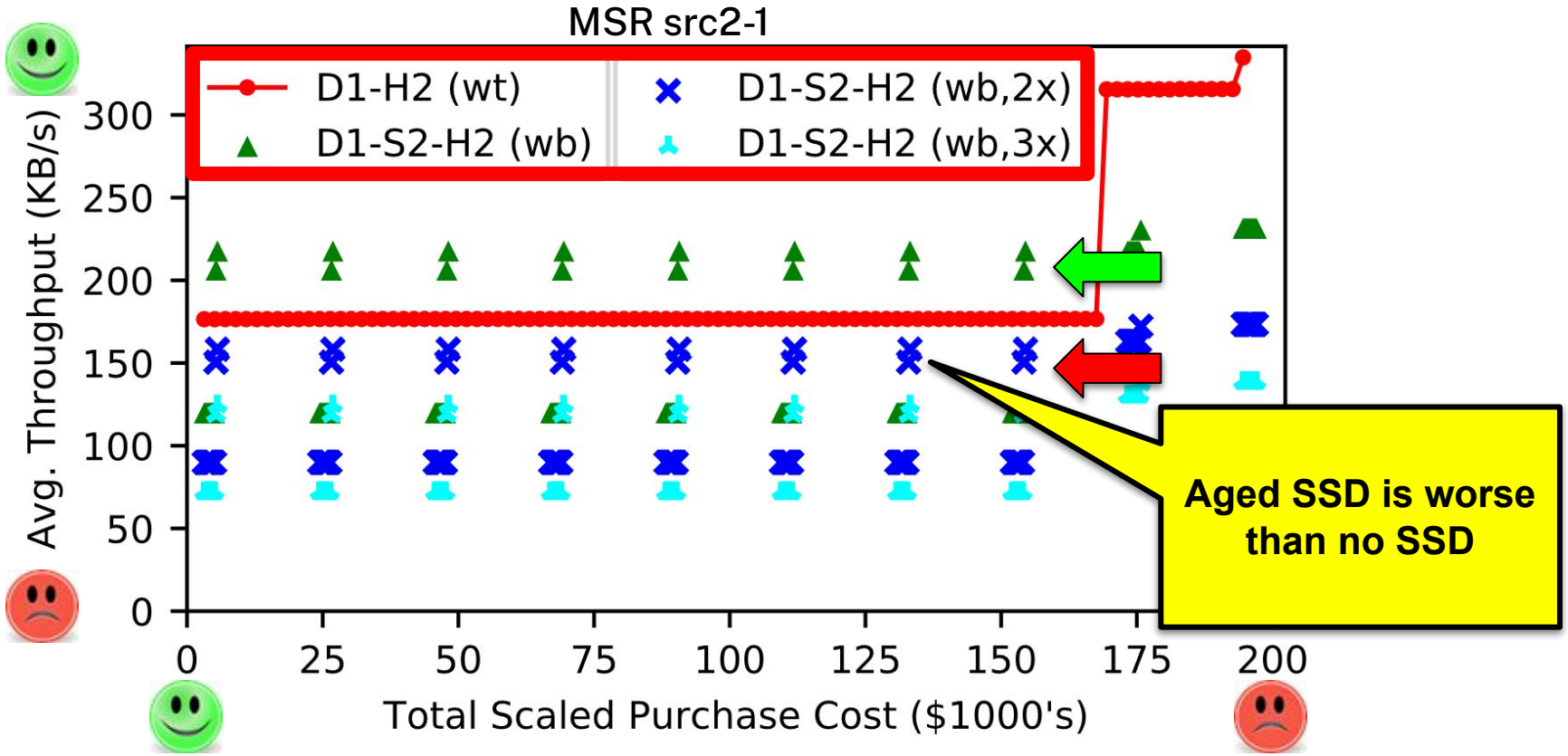
Experimental Setup

- Microsoft Research (MSR) Cambridge block traces
  - 13 enterprise data center servers, 36 volumes [Narayanan et al.]
- Device information
  - DRAM (D), SSD (S), HDD (H)
  - High-end (1), mid (2), low (3)
- Reverse miniature simulation scaled 1000x
  - MSR traces access <1GB of data

Evaluation 1: DRAM Performance



Evaluation 2: SSD Degradation



Outline

- Motivation
- Cache Analysis
- Multi-Tier Cache Simulation
- Evaluation
- **Conclusion**

Conclusion

- Proposed best practices in cache research
  - Analysis of multi-tier systems
  - Evaluation of more comprehensive metrics
- Counter-intuitive simulation results
  - Demonstrated the need for multi-tier analysis and simulation
- Development of an *n*-level I/O cache simulator
- Visual Analytics

Desperately Seeking ... Optimal Multi-Tier Cache Configurations

Tyler Estro, Pranav Bhandari, Avani Wildani, and Erez Zadok

{testro, ezk}@cs.stonybrook.edu, pranav.bhandari@emory.edu, avani@mathcs.emory.edu

Thank You

Q&A

EMORY UNIVERSITY Stony Brook University

Desperately Seeking ... Optimal Multi-Tier Cache Configurations

Tyler Estro, Pranav Bhandari, Avani Wildani, and Erez Zadok

{testro, ezk}@cs.stonybrook.edu, pranav.bhandari@emory.edu, avani@mathcs.emory.edu

Thank You
Q&A

Discussion

1. Multi-tier simulator features.
2. Is monetary cost the primary metric?
3. Cache reconfiguration.
4. Other best practices in cache research.
5. Algorithmic multi-tier solutions.
6. Existing I/O simulators.

EMORY UNIVERSITY Stony Brook University