SmartDedup: Optimizing Deduplication for Resource-constrained Devices

Qirui Yang, Runyu Jin, Ming Zhao

Arizona State University

http://visa.lab.asu.edu

Resource Management on Edge and IoT

- Limited on-device storage
  - Limited I/O performance
  - Limited capacity
  - Limited endurance
- Deduplication can
  - Eliminate redundant I/Os → improve performance
  - Reduce flash writes → improve endurance
  - Remove redundant data → improve utilization
- But
  - Is there enough data duplication in device workloads?
  - How to exploit it using limited resources on the device?

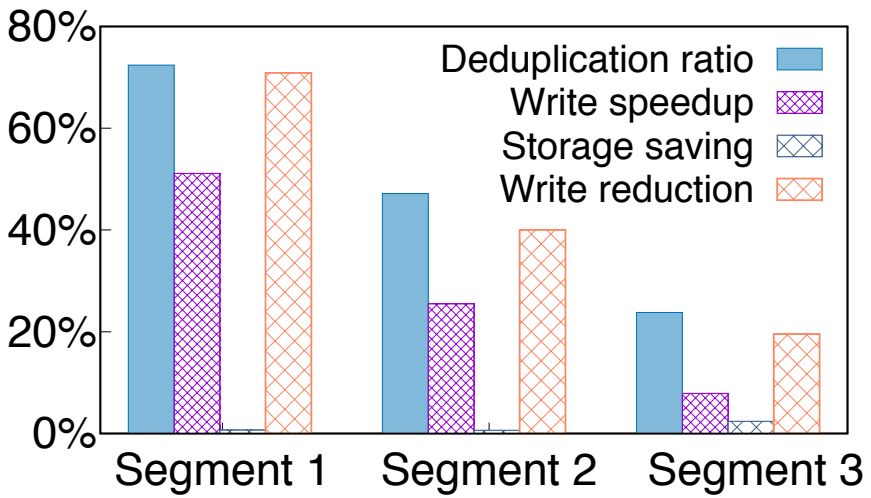
SmartDedup—A Smart Deduplication Solution for Smart Devices

- Cohesively designed in-memory and on-disk fingerprint stores
  - On-disk store complements the small in-memory store
- Synergistically integrated in-line and out-of-line deduplication
  - Out-of-line deduplicates data skipped by in-line
- Adaptive deduplication according to resource and battery availability
  - Dynamically enabled/disabled, dynamic processing rate

Trace Replay (on Nexus)

- Real-world device workloads have a good level of duplicates
- SmartDedup achieves
  - Up to 51.1% write speedup
  - Up to 70.9% write reduction

Segment	Write (GB)	Duplication ratio (%)	Read/write ratio
1	17.2	75.8	1.5
2	12.4	47.9	2.2
3	9.1	26.4	6.8



SmartDedup compared to EXT4

Welcome to Our Talk

SmartDedup

Thursday 3:50 pm

Track II: Deduplication Session