Leveraging Value Locality in Optimizing NAND Flash-based SSDs

Aayush Gupta, Raghav Pisolkar,
Bhuvan Urgaonkar and Anand Sivasubramaniam
Computer Systems Lab
The Pennsylvania State University

Agenda

- Relook at Locality
- Another dimension of Locality : Value Locality
  - Value Locality and SSDs
- CA-SSD Design
  - Mapping Structures
  - Metadata Management
- Evaluation
  - CA-SSD vs Traditional SSD

Locality: The pillar of storage

- Temporal Locality
  - If a logical address is accessed now, it is likely to be accessed again in the near future
- Spatial Locality
  - If a logical address is accessed now, there is a high likelihood that its neighboring addresses will be accessed in the near future
- Pervasive : L1/L2 cache, TLB, Buffer Cache, Virtual Memory, Disk Cache, Web Cache …

Another Dimension of Locality

- Value Locality
  - Certain content is accessed preferentially
- Data deduplication using Content Addressable Storage (CAS)
- Use cases of Value Locality (VL)
  - Network traffic reduction
  - Content based Caching
  - Efficient data storage (archival/backup)
    - E.g: Venti, Foundation, EMC Centera, Data Domain Storage Systems

Can we use Value Locality to address the idiosyncrasies of SSDs?

CAS suits SSD

SSD

Writes are a bottleneck

- Read/Write asymmetry
- Block Erases

CAS

Provides Deduplication

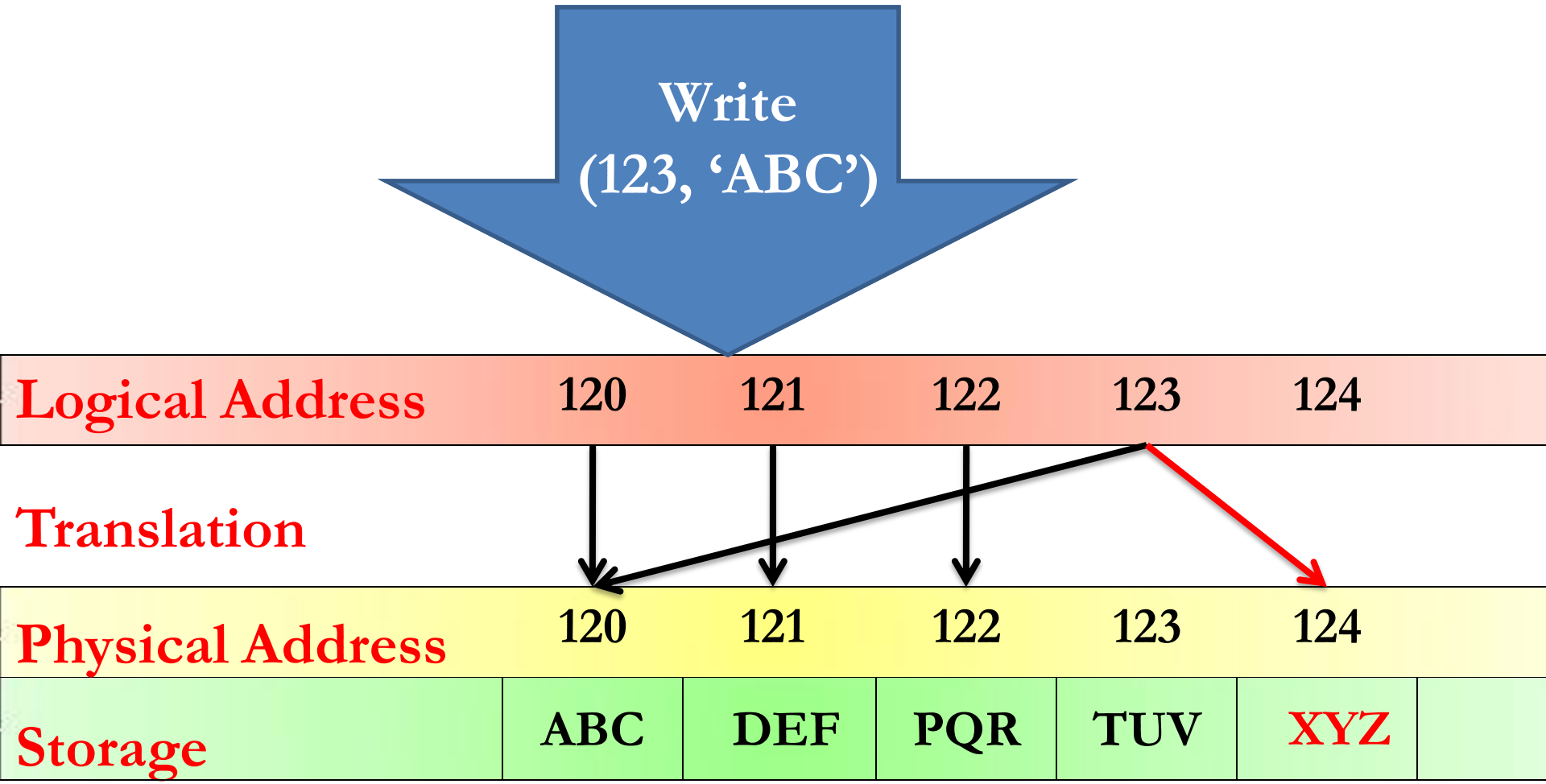
CAS and SSD: Made for each other?

CAS

Out of Place Updates

SSD

Out of Place Updates in CAS



CAS and SSD: Made for each other?

CAS	SSD
Out of Place Updates	Erase before Write

Problem with CAS

CAS	SSD
Loss of Sequentiality	Fast Random Reads

Do real workloads exhibit Value Locality?

Workloads [Koller10]

Workload	Writes (%)	Total Requests (Millions)	Unique Write Requests (%)	Unique Read Requests (%)
web	77.0		42.35	32.05
mail	77.3	3.6	7.83	80.85
homes	96.7	4.4	66.37	80.75

Write dominant

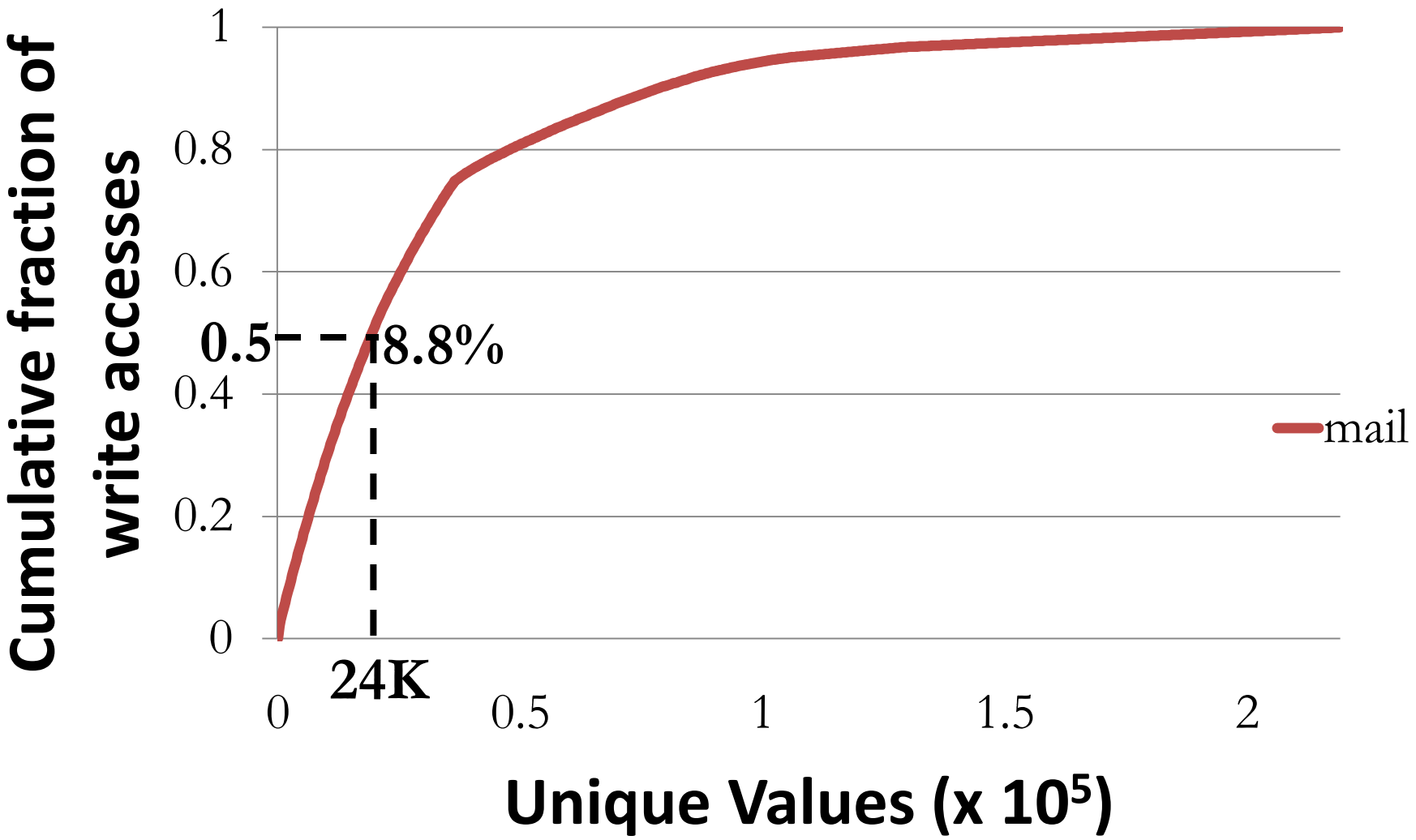
Duplication

[Koller10] Koller, R., and Rangaswami, R. "I/O Deduplication: Utilizing Content Similarity to Improve I/O Performance." (*FAST'10*)

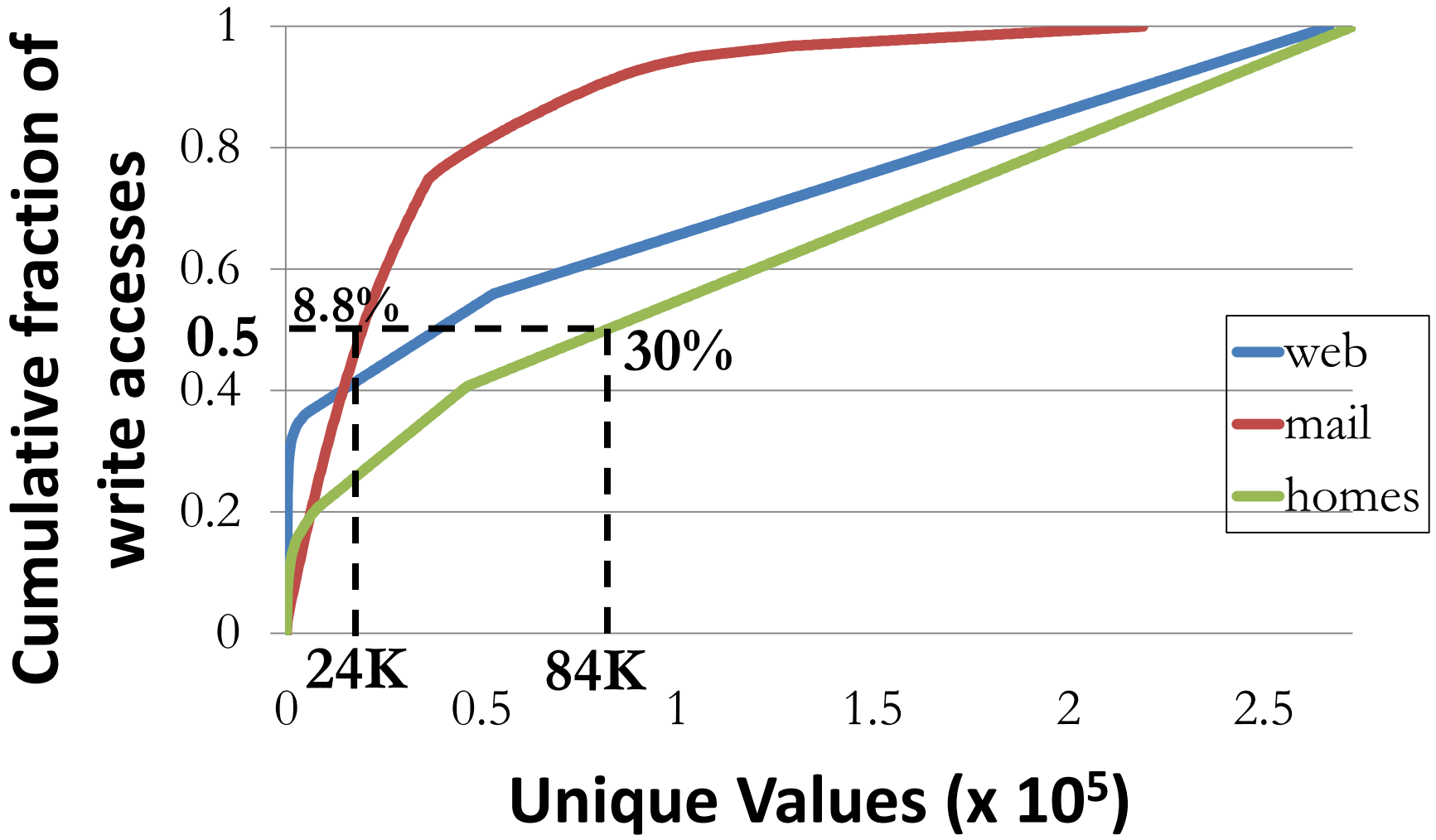
Value Popularity

- **VP represents the number of occurrences of each unique value in a workload**
  - **Signifies potential for deduplication for a workload**

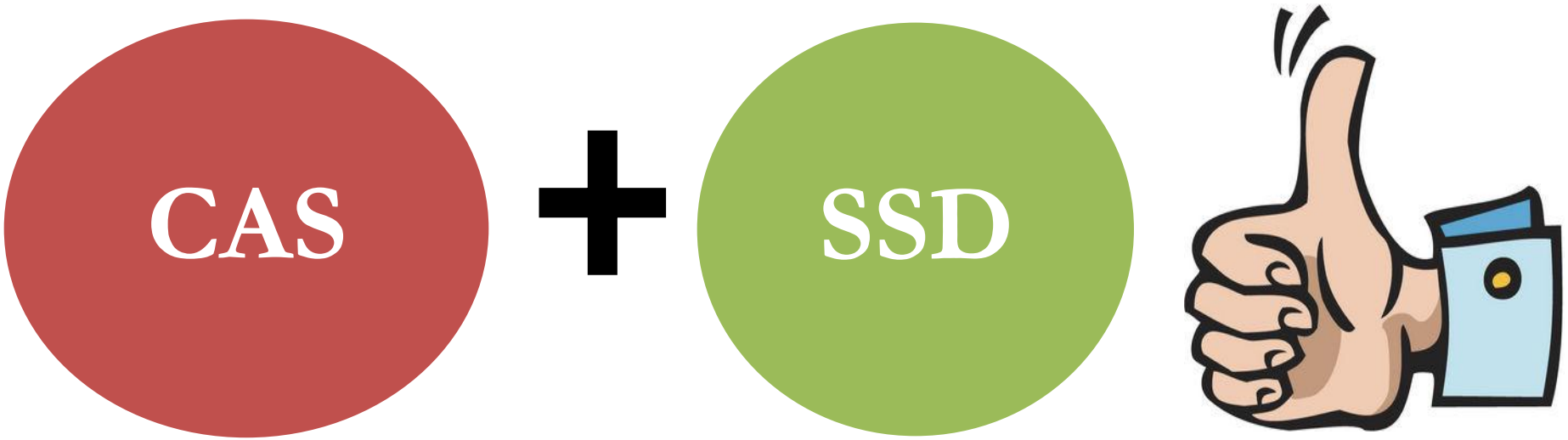
Some Values are Very Popular



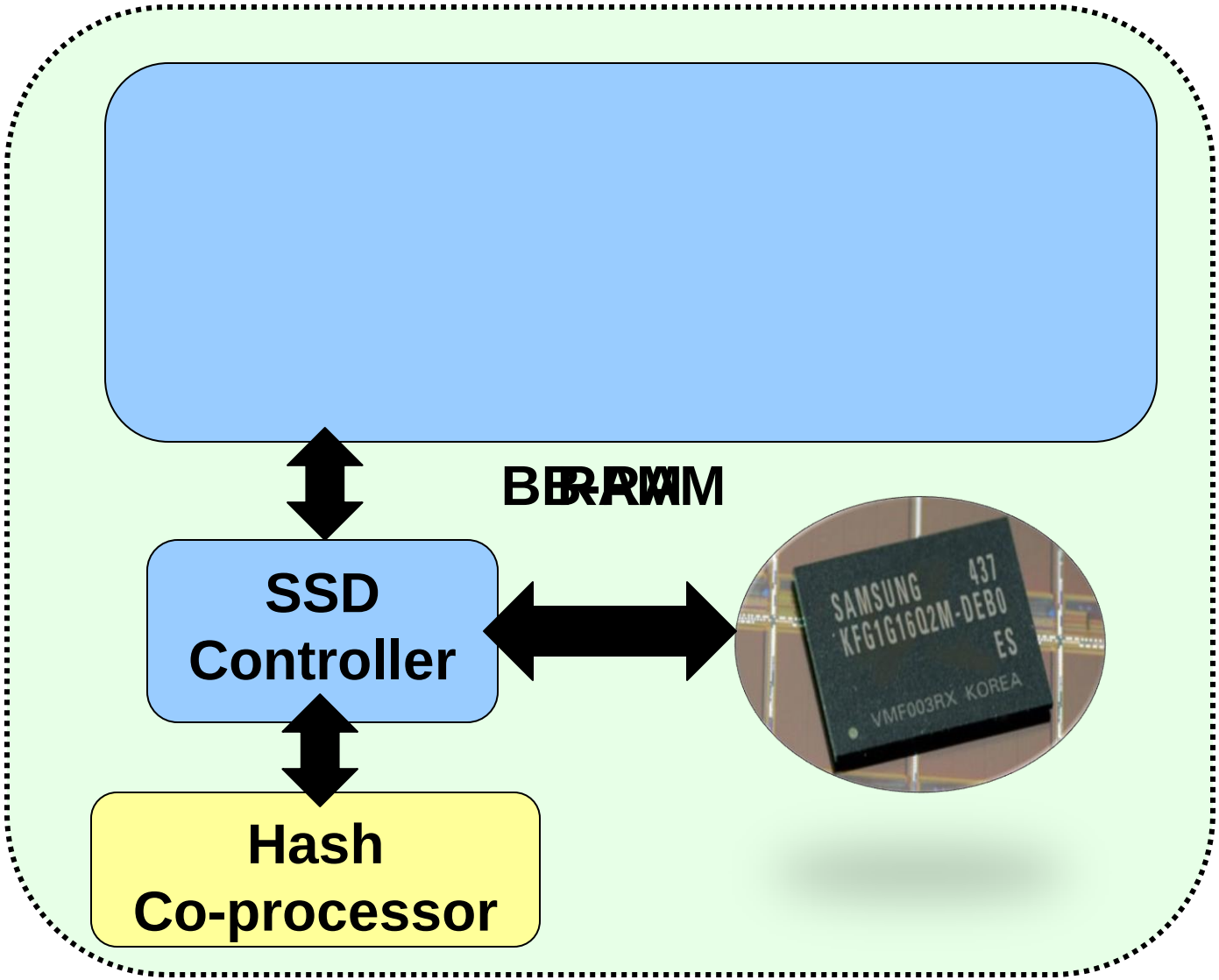
Some Values are Very Popular



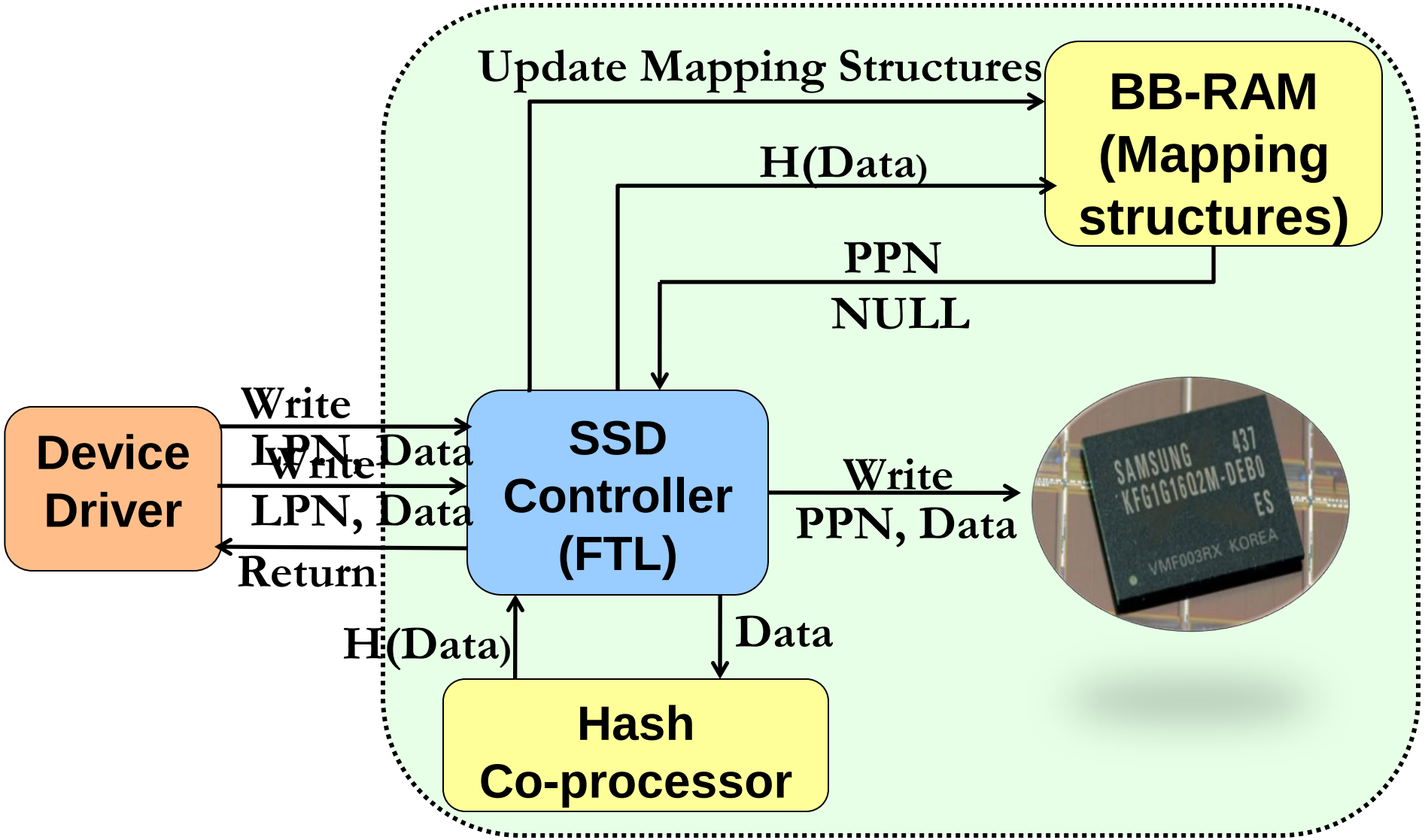
CA-SSD



CA-SSD Design



CA-SSD Design



Mapping Structures: LPT & HPT

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

Mapping Structures: iLPT

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

iLPT	
PPN	LPN
P1	L1, L3
P2	L2
P3	INV
P4	L4

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

Mapping Structures: iLPT & iHPT

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

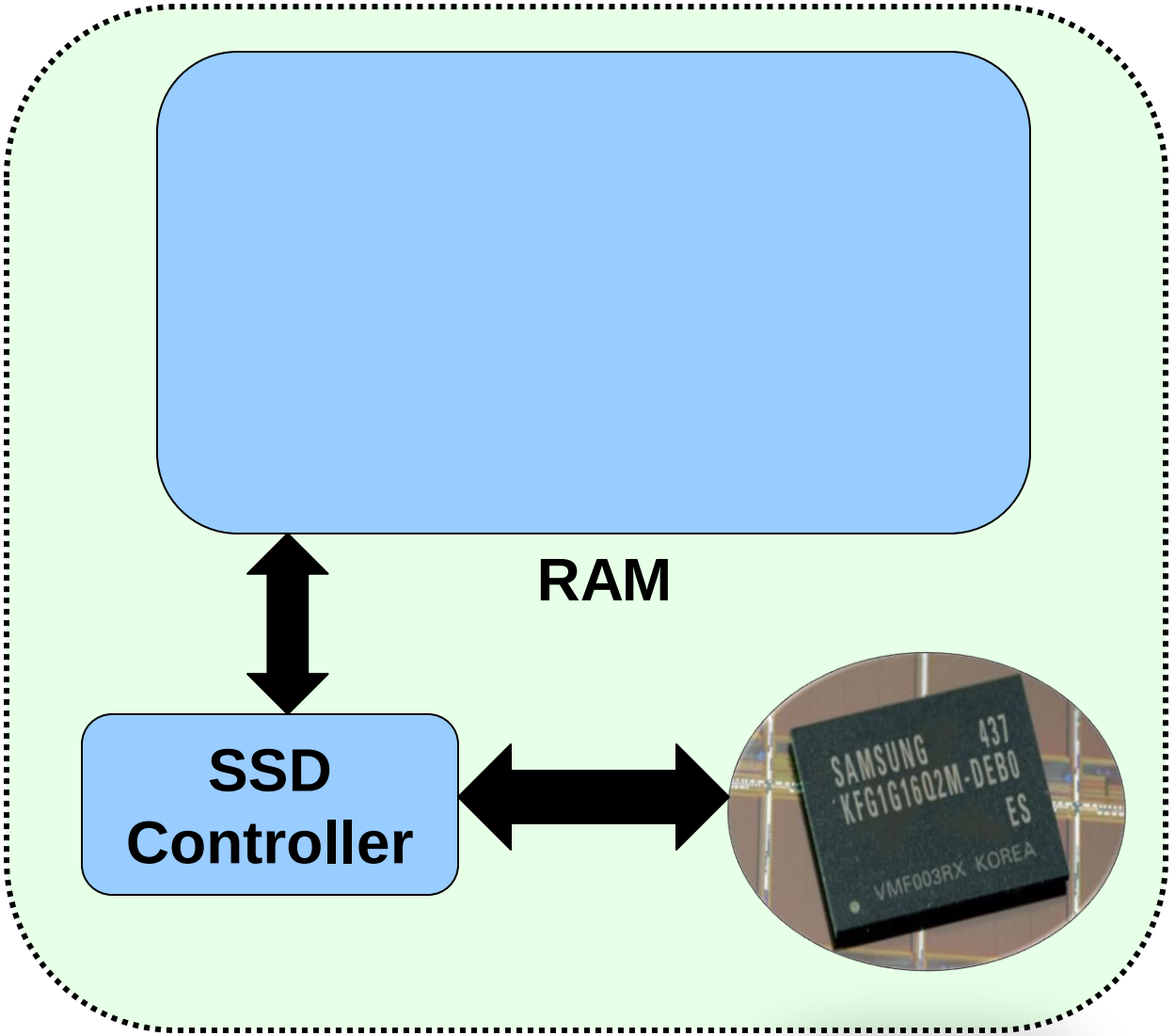
Remove

iLPT	
PPN	LPN
P1	L1, L3
P2	L2
P3	INV
P4	L4

iHPT	
PPN	Hash
P1	H1
P2	H2
P3	H3
P4	H4

Metadata: Traditional SSD

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P3
L4	P4



Optimal Large SRAM

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

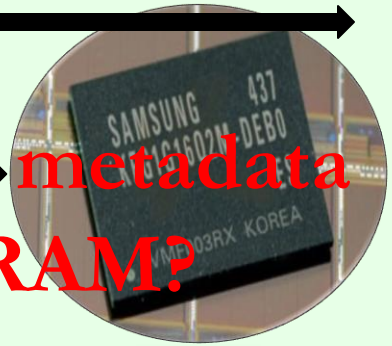
iLPT	
PPN	LPN
P1	L1,L3
P2	L2
P3	INV
P4	L4

iHPT	
PPN	Hash
P1	H1
P2	H2
P3	H3
P4	H4

BB-RAM

BB-RAM

SSD Controller



Not Scalable!!

How do we fit the metadata in CA-SSD's RAM?

Hash Co-processor

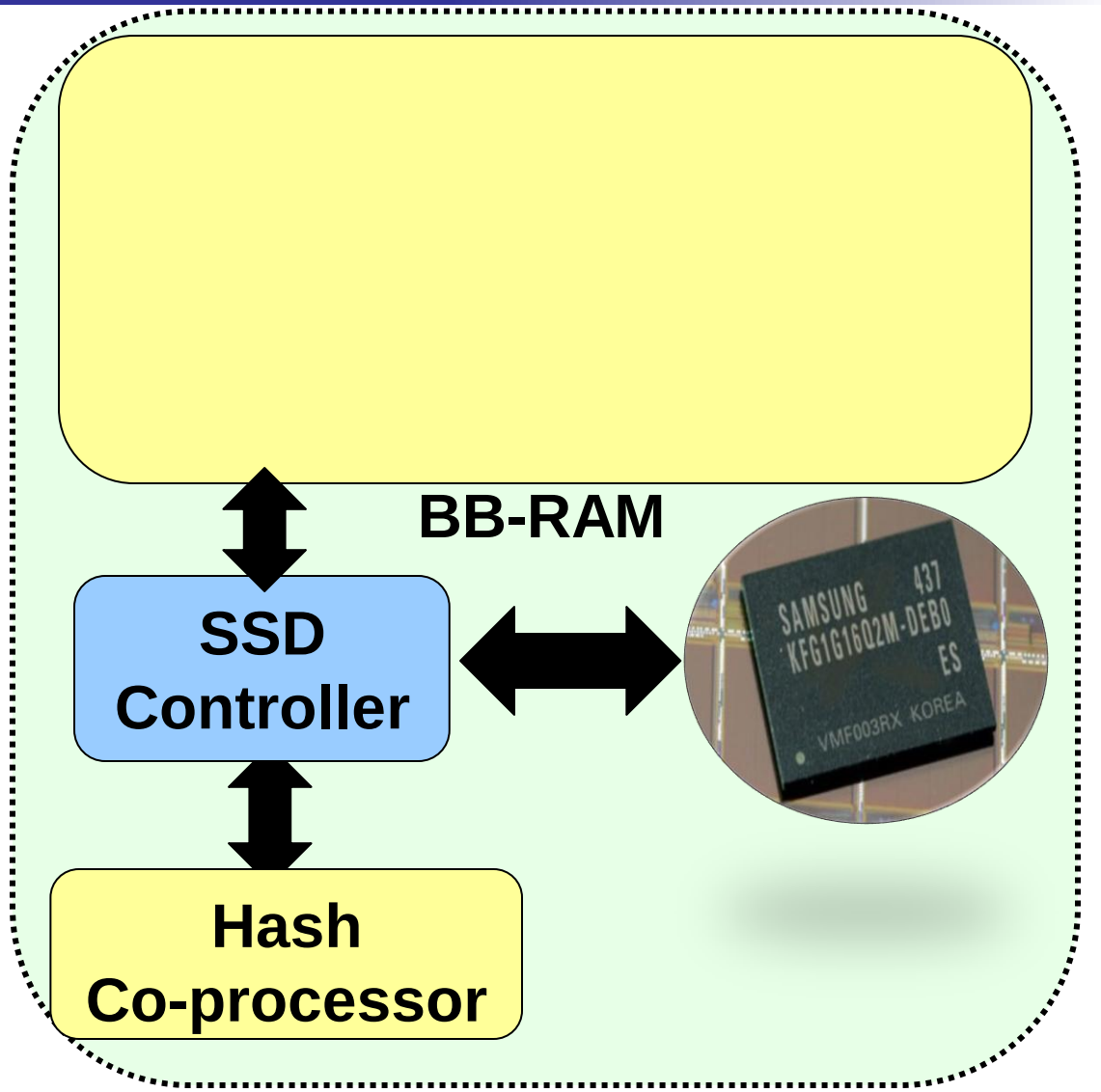
Option 2 : Shrink Metadata

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

iLPT	
PPN	LPN
P1	L1,L3
P2	L2
P3	INV
P4	L4

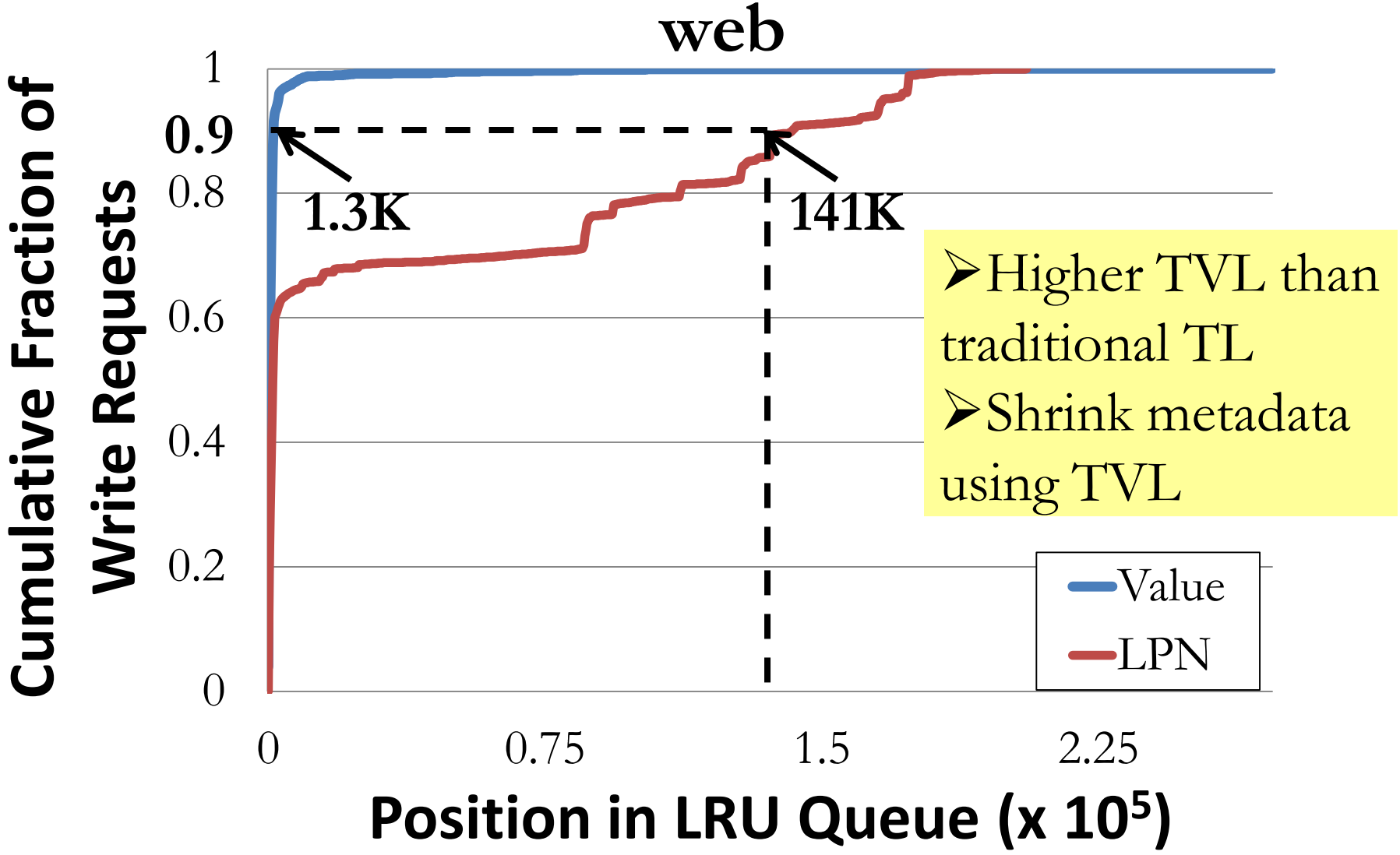
iHPT	
PPN	Hash
P1	H1
P2	H2
P3	H3
P4	H4



Temporal Value Locality

- **TVL implies that if a certain *value* is accessed now, it is likely to be accessed again in the near future *not necessarily from the same address***

Temporal Value Locality: Writes



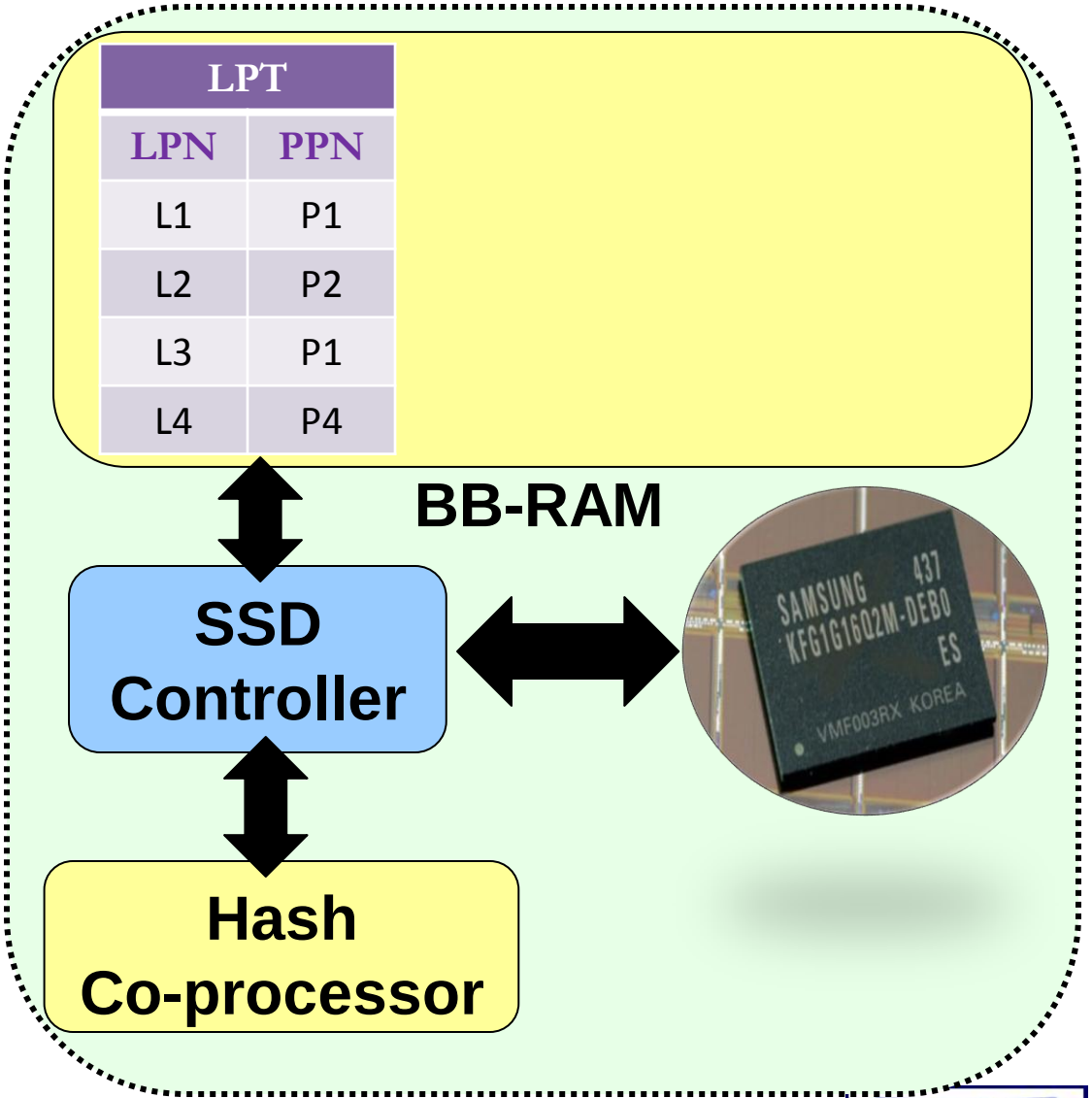
- Higher TVL than traditional TL
- Shrink metadata using TVL

Metadata Management: TVL

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3
H4	P4

iLPT	
PPN	LPN
P1	L1,L3
P2	L2
P3	INV
P4	L4

iHPT	
PPN	Hash
P1	H1
P2	H2
P3	H3
P4	H4



Metadata Management: TVL

HPT	
Hash	PPN
H1	P1
H2	P2
H3	P3

MRU

LRU

Discard

LPT	
LPN	PPN
L1	P1
L2	P2
L3	P1
L4	P4

BB-RAM

SSD Controller

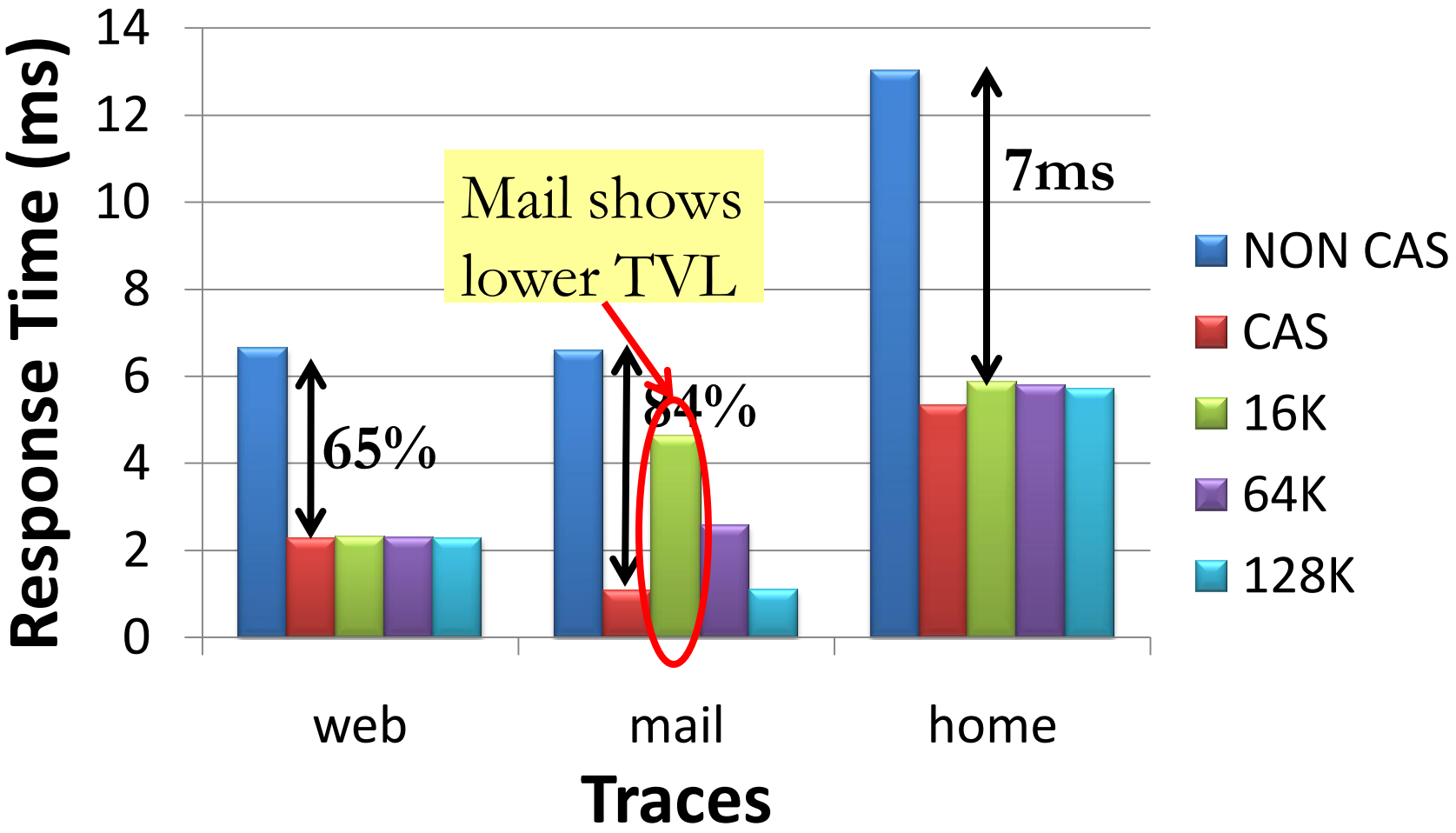
Hash Co-processor

iLPT	
PPN	LPN
P1	L1,L3
P2	L2
P3	INV
P4	L4

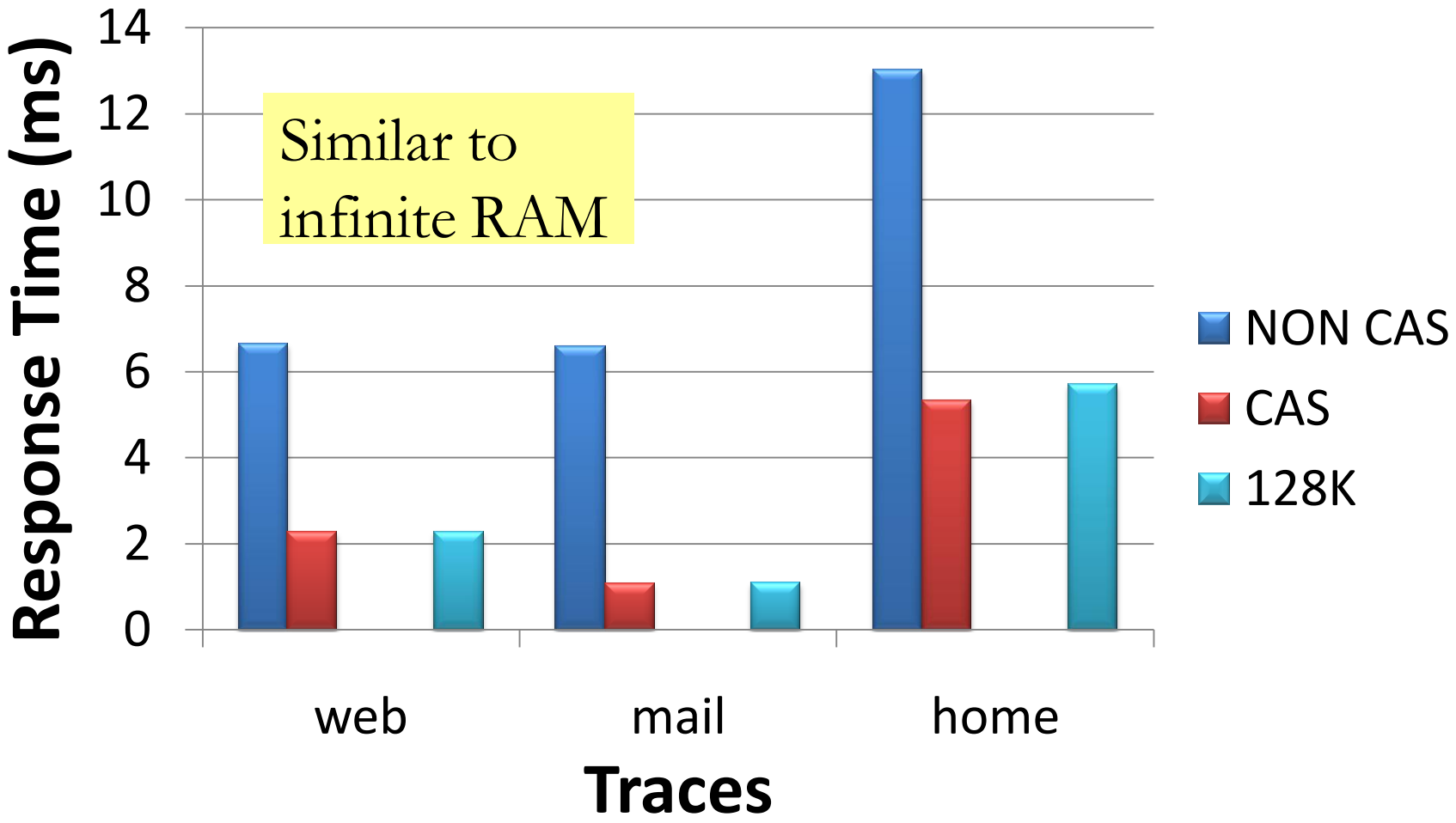
iHPT	
PPN	Hash
P1	H1
P2	H2
P3	H3

How does CA-SSD perform compared to Traditional SSDs?

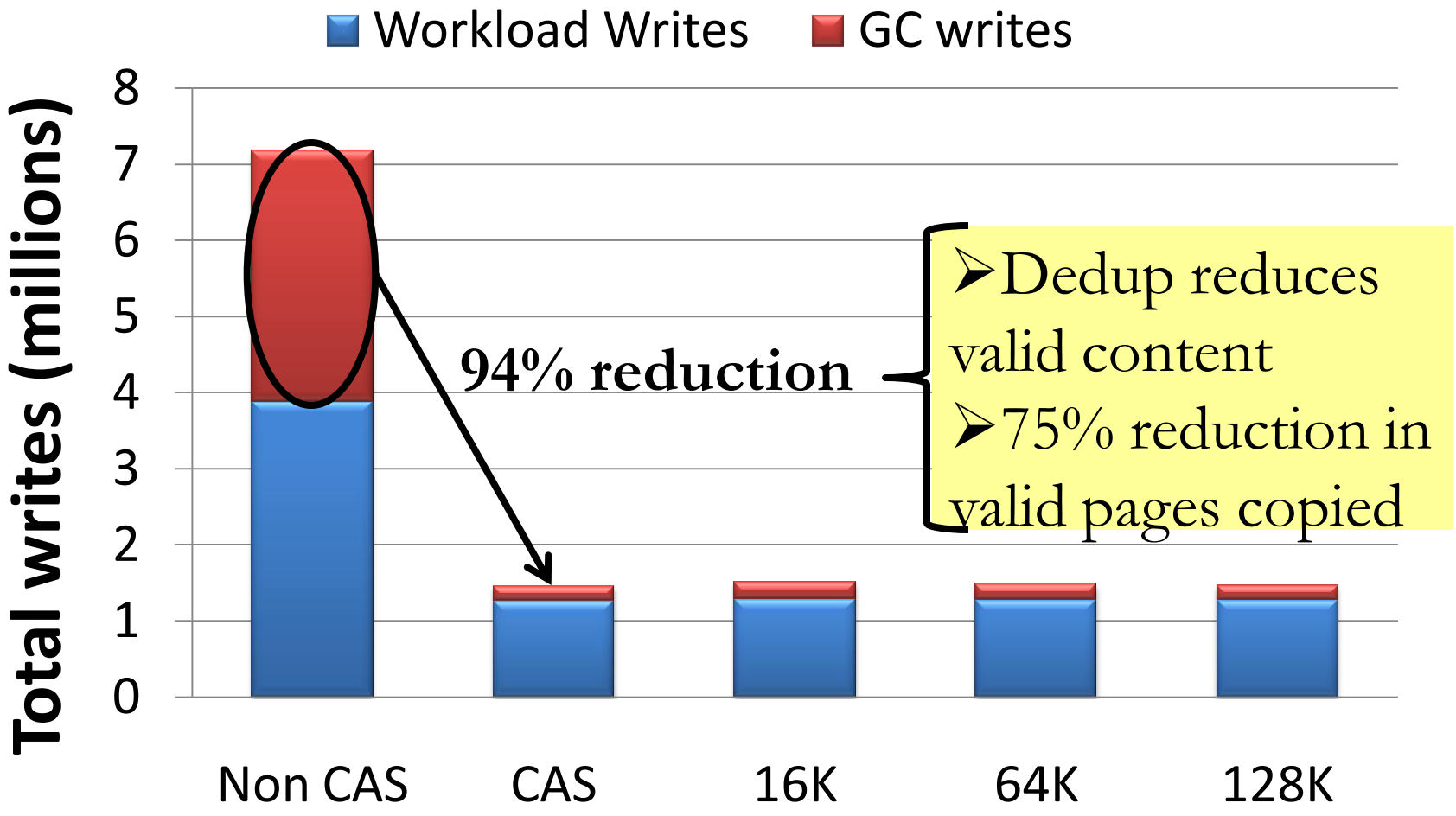
Evaluation : Response Time



Evaluation : Response Time

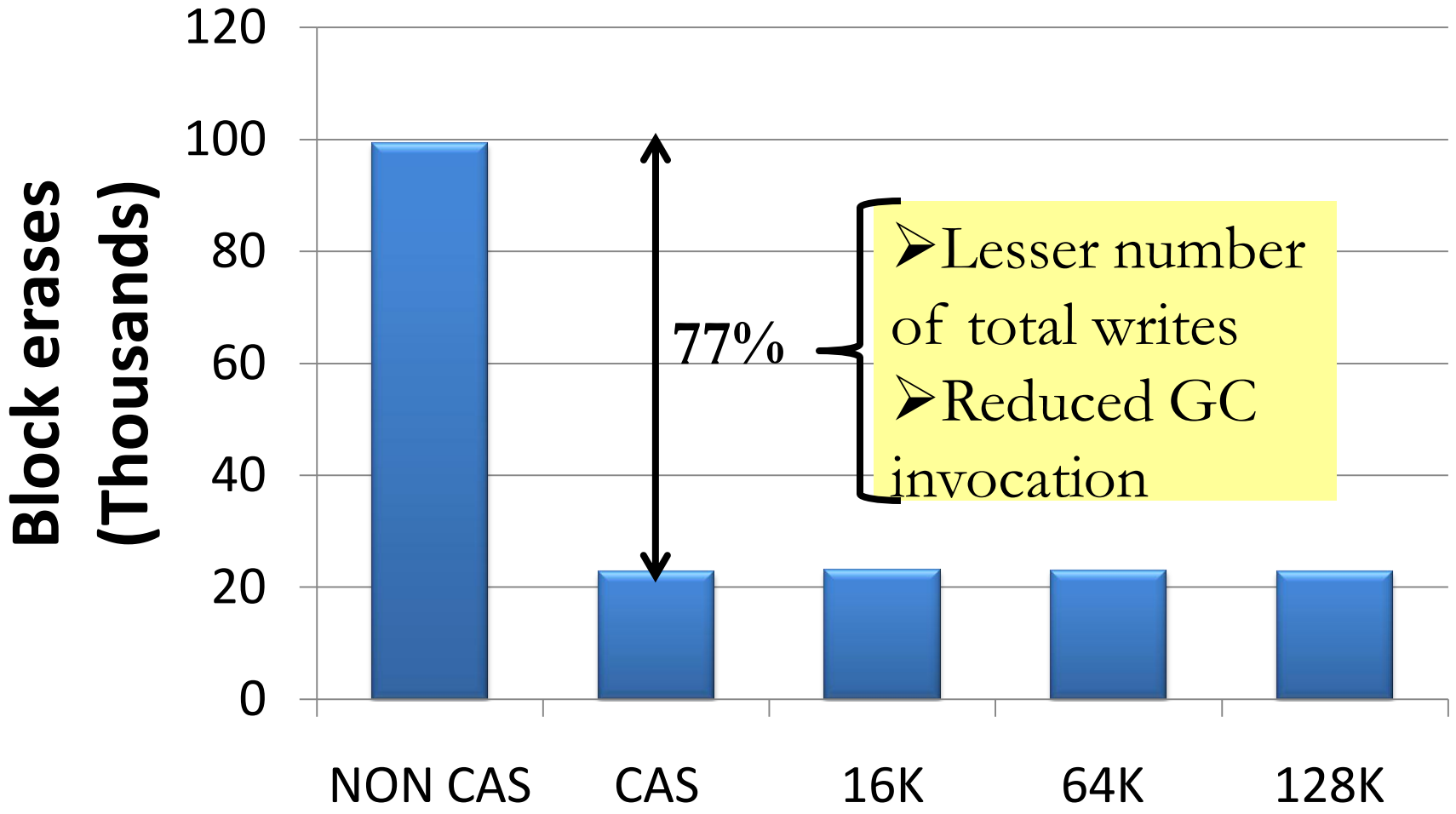


Total Writes : *web*



- Dedup reduces valid content
- 75% reduction in valid pages copied

Total Erases : *web*



Conclusions

- Workloads exhibit significant value locality
  - Characterization of Value Popularity and Temporal Value Locality
- CAS and SSDs complement each other
- Certain implementation challenges need to be addressed
  - Mapping structures
  - Metadata Management

Thank You

Questions???