Tagged Consistency and Garbage Identification in Deduplication-enabled Storage Systems

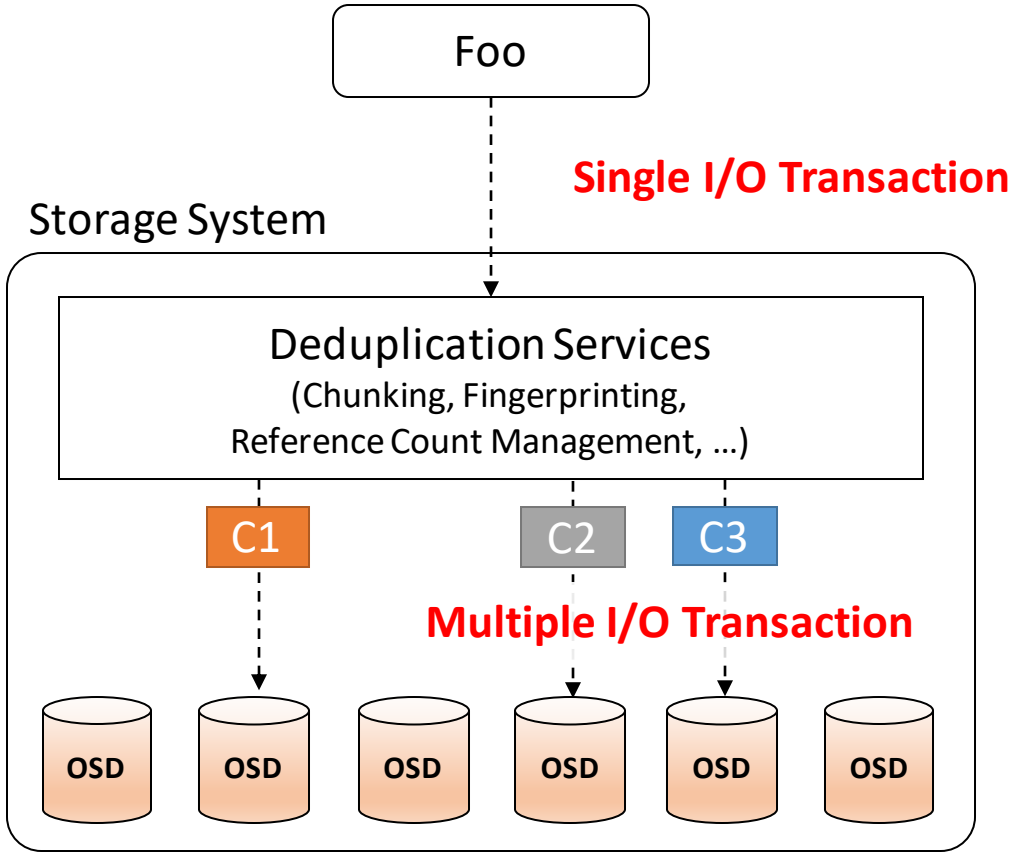
Awais Khan, Changgyu Lee, Sungyong Park, Youngjae Kim

Sogang University, Seoul, Republic of Korea



Sub-Transactions in Deduplication System

- In deduplication-enabled system, single I/O transaction can have multiple sub-transactions.



Deduplication Metadata

Object Map

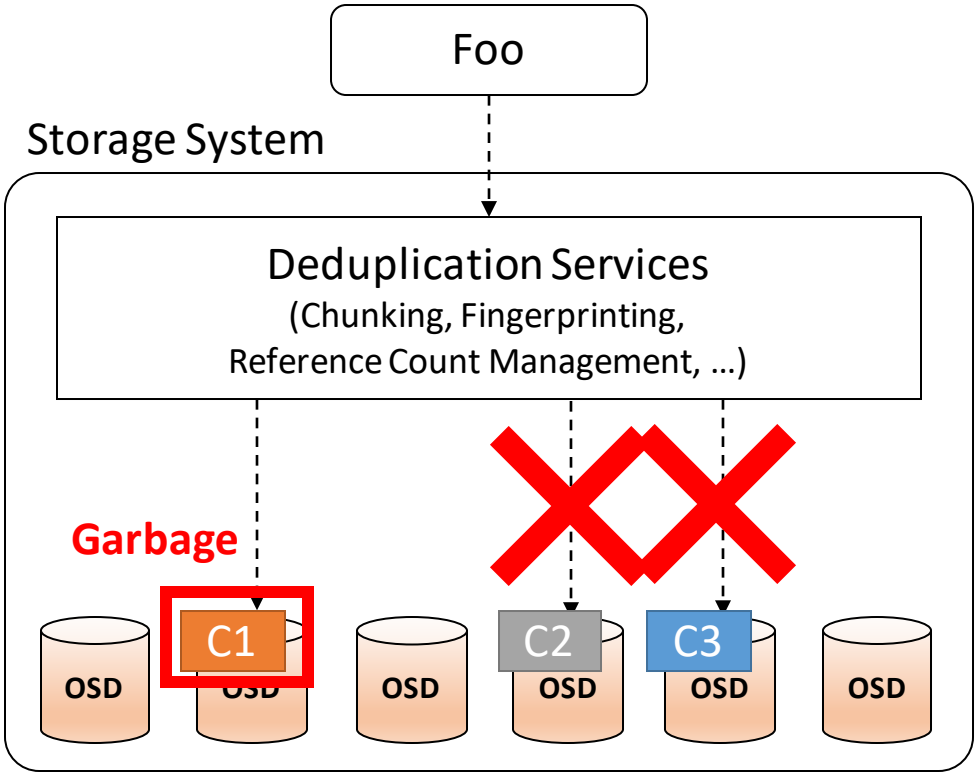
Object Name	Fingerprint	Chunk List
Foo	0H2x42	C1, C2, C3

Chunk Information

Chunk	Reference Count
C1	1
C2	1
C3	1

Sub-Transaction Failure

- If some of sub-transaction failed, subsequent transaction will cause system inconsistency.



Deduplication Metadata

Object Map

Object Name	Fingerprint	Chunk List
Foo	0H2x42	C1, C2, C3

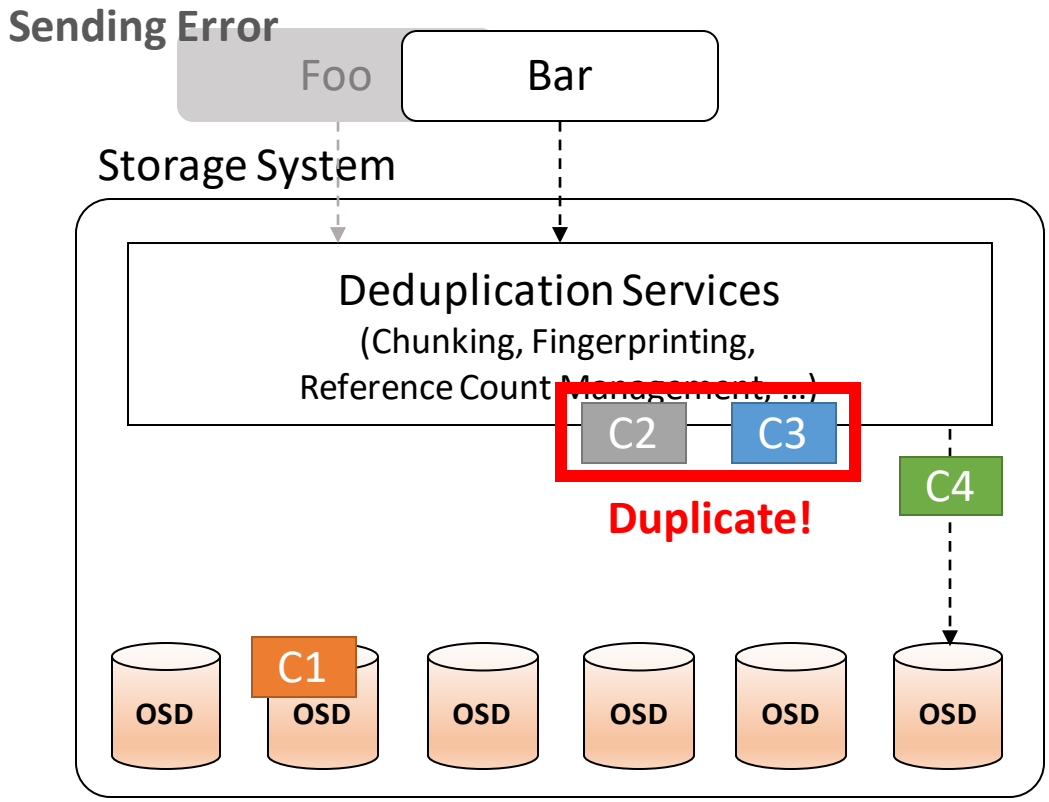
Chunk Information

Chunk	Reference Count
C1	1
C2	1
C3	1

Invalid Reference Count

Sub-Transaction Failure

- If some of sub-transaction failed, subsequent transaction will cause system inconsistency.



Deduplication Metadata

Object Map

Object Name	Fingerprint	Chunk List
Foo	0H2x42	C1, C2, C3
Bar	1Z2F09	C2, C3, C4

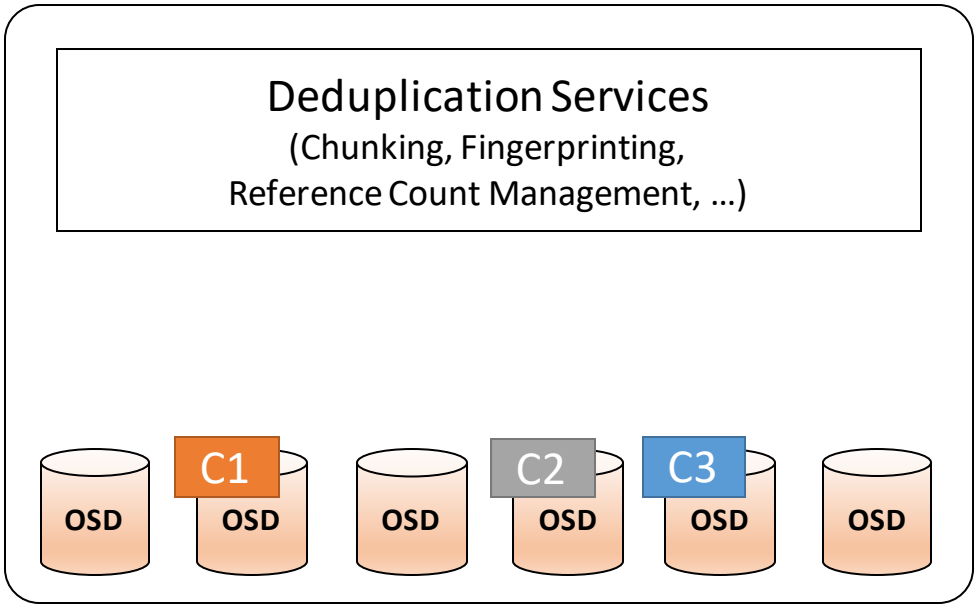
Chunk Information

Chunk	Reference Count
C1	1
C2	1 -> 2
C3	1 -> 2
C4	1

Invalid Reference Count

Tagged Consistency

- In our proposed solution, each entry has a consistency flag.
  - If flag is False, entry can be invalid.
  - Else entry is valid.



Deduplication Metadata

Object Map

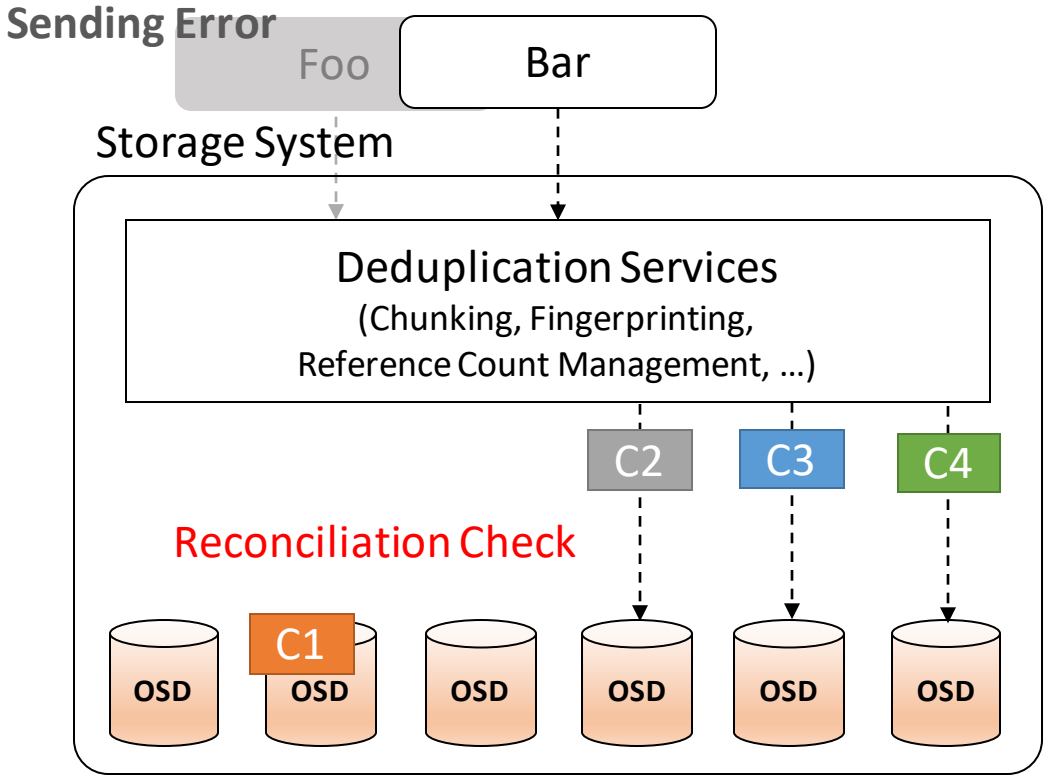
Object Name	Fingerprint	Chunk List	Flag
Foo	0H2x42	C1, C2, C3	False

Chunk Information

Chunk	Reference Count	Flag
C1	1	False
C2	1	False
C3	1	False

Reconciliation Check

- If flag is *ZERO* while checking reference count, Reconciliation Check will be initiated.
- It will directly check OSD for existence of chunk and fix consistency flag.



Deduplication Metadata

Object Map

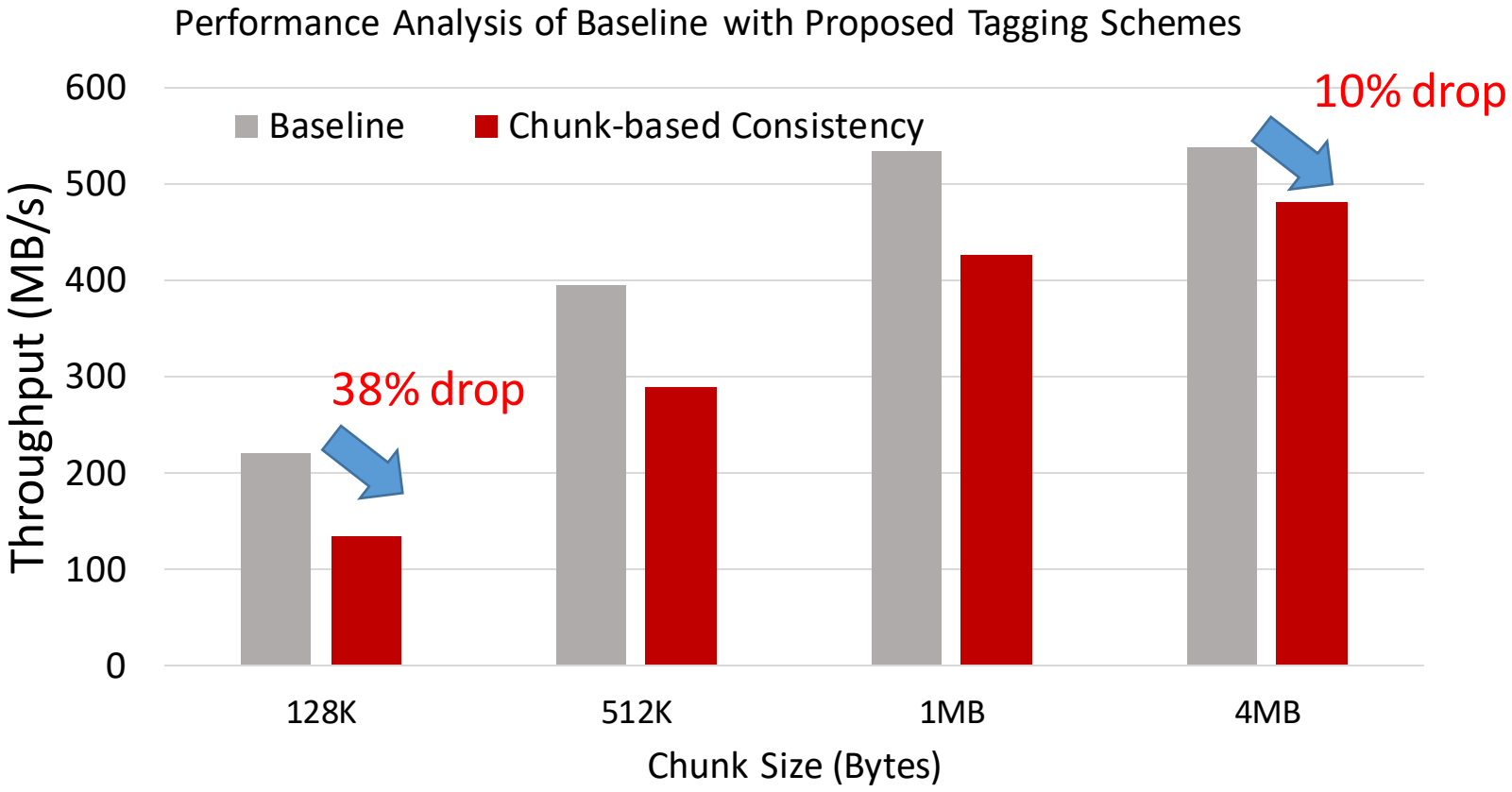
Object Name	Fingerprint	Chunk List	Flag
Foo	0H2x42	C1, C2, C3	False
Bar	1Z2F09	C2, C3, C4	False

Chunk Information

Chunk	Reference Count	Flag
C1	1	False
C2	2	True
C3	2	True
C4	1	False

Preliminary Evaluation

- Ceph Cluster
  - 1 MDS, 2 OSD Servers (4 SSDs), 1 Client,
  - Fio Benchmark, Queue Depth=16, No. of Jobs=16



Q&A

SOGANG UNIVERSITY

Contact:

Awais Khan (awais@sogang.ac.kr)

Department of Computer Science and Engineering
Sogang University, Seoul, Republic of KOREA

Changgyu Lee (changgyu@sogang.ac.kr)

Department of Computer Science and Engineering
Sogang University, Seoul, Republic of KOREA