Storage Tiering with Deduplication

Nalini Kumari Nallamalli
Santosh Kalekar
Veritas Technologies LLC
Agenda

- Deduplication
- Tiering
- Tiering w/o Deduplication
- Tiering with Deduplication
- Summary
Deduplication

- Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data.
- Used to improve storage utilization and to transfer less data across the network.
### Deduplication Contd..

<table>
<thead>
<tr>
<th>File1</th>
<th>A B C D E</th>
</tr>
</thead>
<tbody>
<tr>
<td>File2</td>
<td>Z Y A B E</td>
</tr>
<tr>
<td>File3</td>
<td>Y A B Z D</td>
</tr>
<tr>
<td>File4</td>
<td>Z A B X Z</td>
</tr>
</tbody>
</table>

**W/O Deduplication**

**With Deduplication**
Tiering

- Progression or demotion of data across different tiers (types) of storage devices and media.
- Assigned to the related media according to performance and capacity requirements.
- User define rules and policies that dictate if and when data can be moved between the tiers.
Tiering Policies

- mtime or atime (Aging)
- I/O or Access temperature (Activity)
Tiering w/o Deduplication

File1

File2

File3

File4

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering w/o Deduplication

File 1 became cold

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering w/o Deduplication

File1
File2
File3
File4

Tier 0 (Hot Data)

YABZE
ZABXZ

Tier 1 (Cold Data)

ABCDE
ZYABE

File2 became cold
Tiering w/o Deduplication

File1
File2
File3
File4

Tier 0 (Hot Data)

Tier 1 (Cold Data)

ZABXZ

A B C D E
Z Y A B E
Y A B Z E

File3 became cold
Tiering w/o Deduplication

Tier 0 (Hot Data)

File1
File2
File3
File4

Tier 1 (Cold Data)

File2 became hot

2019 Storage Developer Conference India © All Rights Reserved.
Tiering w/o Deduplication

Tier 0 (Hot Data)

File1
ZYABE
ZABXZ
XYTAZ

Tier 1 (Cold Data)

File2
ABCDE

File3

File4

File5

X Y T A Z

Y A B Z E
Tiering with Deduplication - Challenges

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering with Deduplication - Challenges

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering with Deduplication - Solution

### Tier 0 (Hot Data)
- File1: ABCDE
- File2: ZYABE
- File3: YABZE
- File4: ABXZ

### Tier 1 (Cold Data)
- Ref: 4
- Ref: 4
- Ref: 1
- Ref: 1
- Ref: 3
- Ref: 1
- Ref: 2
- Ref: 3
- CRef: 0
- CRef: 0
- CRef: 0
- CRef: 0
- CRef: 0
- CRef: 0
- CRef: 0
- CRef: 0

2019 Storage Developer Conference India © All Rights Reserved.
Tiering with Deduplication

Tier 0 (Hot Data)
- File 1: ABCDE
- Ref: 4
- CRef: 1

- File 2: ZYABE
- Ref: 3
- CRef: 1

Tier 1 (Cold Data)
- File 3: YABZD
- Ref: 1
- CRef: 1

- File 4: ABXZ
- Ref: 1
- CRef: 1
Tiering with Deduplication

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering with Deduplication

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Tiering with Deduplication

Tier 0 (Hot Data)

File1
ABCDE

Ref: 2
CRef: 1

Ref: 2
CRef: 1

Ref: 1
CRef: 0

Ref: 1
CRef: 0

Ref: 1
CRef: 0

Ref: 2
CRef: 1

Ref: 2
CRef: 1

Tier 1 (Cold Data)

File3
YABZE

Ref: 1
CRef: 1

Ref: 1
CRef: 1

Ref: 1
CRef: 1

Ref: 0
CRef: 0

Ref: 1
CRef: 0

Ref: 1
CRef: 0

Ref: 1
CRef: 0
Tiering with Deduplication

File1
ABCDE

File3
YABZE

File5
XTAZ

Tier 0 (Hot Data)

Tier 1 (Cold Data)
Summary

- Deduplication
- Tiering
- Challenges with Tiering
- Tiering with Deduplication
Questions?