A Simulation Result of Replicating Data with Another Layout for Reducing Media Exchange of Cold Storage

Satoshi Iwata* and Kensuke Shiozawa
FUJITSU LABORATORIES LTD.

* Temporary transferred to NIFTY Corp.
Cold Storage

- Storage which is slow but cheap
  - Tape and optical disc
  - Cheap because of not only price of media but also reduced electricity usage and long media life
    - No electricity is required unless they are accessed
    - Long media life: 30 years for a tape and 50 years for an optical disc

- Suitable for large, less frequently accessed archival data

- Example use cases
  - Facebook migrates old photos to optical disc storage
    - "50 percent cheaper than using hard disk drives for cold storage, and 80 percent more energy efficient"
      [CES press conf. ’16]
  - ECMWF stores weather information onto tape storage which capacity exceeds 50PB
    [Grawinkel et al. ’15]
Media Exchange Incurs Large Latency

- A reason why they are slow is a large latency incurred by media exchange
- Media and drives are separated
  - In contrast to all-in-one device such as HDD or SSD
- A few minutes to exchange media
  - A robot carries media between shelves and drives inside a library

- Media exchange is norm rather than special [Grawinkel et al. ’15]
  - 9 loadings / minute
  - 231 drives, 32,712 tapes

- Active archive usage must increase for the purpose of big data analysis
  - In contrast to backup and deep archive

- Reducing media exchange must lower a barrier for installing cold storage
Placing Correlated Data Together

- Placing correlated data together is helpful for reducing media exchange

- E.g., A case for log of users’ actions
  - Log is a good example of using cold storage
    - Can become enormous in a large system
    - Can be kept for a long time as evidence
    - It’s rare for all the logs to be accessed frequently

- Logs are typically collected in time order
  - E.g., Discovering a root cause of a failure by checking logs around the time of occurrence

- Should be stored in the order of generation

<table>
<thead>
<tr>
<th>Device config. change</th>
<th>Taking out files</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st day</td>
<td>2nd day</td>
</tr>
</tbody>
</table>
Multi-dimensional Searches Disrupt It

- Logs can be mined in multiple ways
  - E.g., Focusing a particular action
    - Searching for a malicious employee who leaked confidential information by checking only *taking out files*
  - In this case the layout on the right side is more suitable
    - On the contrary, if all the actions are required, the left one is more suitable
    - One layout doesn’t much both requirements

- Not just in the case of log
  - Another example is weather information mined in temporal and spatial manner
Basic Idea

One solution might be …

1. Replicate data
2. Keep both layouts
3. Choose an appropriate layout depending on each query

However, simply doing so reduces capacity efficiency by half
Our proposal

- Meet multi-dimensional searches while achieving small capacity efficiency loss

1. Take relative latency reduction into account
2. Utilize replicas generated for avoiding data loss
### Sample Log

- **Logs of users’ actions on their PCs**
  - In an anonymized format
  - Collected by FUJITSU SOCIAL SCIENCE LABORATORIES LTD.

- **Example log**
  - WIN-HOST,2014/10/6 15:34,user1,G01,Run application,Legal, , Ran [iexplore]

<table>
<thead>
<tr>
<th>Time</th>
<th>User ID</th>
<th>User action</th>
<th>Legality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Two types of search queries**
  1. stats count by “action”
  2. “action”=“Device configuration change” | timechart count by “Legality”
Amount of Logs Varies Among Actions

- 11 actions exist
- 6 filters regarding actions

<table>
<thead>
<tr>
<th>Action</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>File manipulation</td>
<td>21%</td>
</tr>
<tr>
<td>Run application</td>
<td>17%</td>
</tr>
<tr>
<td>Dev. conf. change</td>
<td>3%</td>
</tr>
<tr>
<td>Print out</td>
<td>3%</td>
</tr>
<tr>
<td>Taking out files</td>
<td>3%</td>
</tr>
<tr>
<td>Logoff OR Logon</td>
<td>1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>48%</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Action</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Get window title</td>
<td>45%</td>
</tr>
<tr>
<td>Stop application</td>
<td>7%</td>
</tr>
<tr>
<td>PC shutdown</td>
<td>0%</td>
</tr>
<tr>
<td>PC boot</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>52%</strong></td>
</tr>
</tbody>
</table>
1. Take Relative Latency Reduction into Account

- Smaller amount logs are better candidates for replication

    - Considering ratio of time for reading data and exchanging media
      - 333 minutes for whole LTO Ultrium7 tape and 93 minutes for whole blu-ray disc
    - Action A (50% of a medium)
      - 0.6% decrease from 337 min. to 335 min.
    - Action C (1% of a medium)
      - 18.7% decrease from 10.7 min. to 8.7 min.

- Has another advantage of lowering capacity efficiency loss
  - Action A: 33% capacity efficiency loss
  - Action C: 1% capacity efficiency loss
2. Utilize Replicas Generated for Avoiding Data Loss

- Normal replication layout

- Proposing layout

- Can reduce media exchange without any capacity efficiency loss
Simulation

- Illustrate the efficiency of adding replicas in another layout
  - Based on sample logs
  - Focusing on queries including action-based filters

- Four metrics
  - # media exchange
  - Absolute latency reduction
  - Relative latency reduction
  - Capacity efficiency
Assumptions

- Mean number of media exchange is modeled as $1 + x$
  - $x$ equals the amount of reading data, which is normalized with media size
- One medium holds exactly one day’s logs in the original layout
- Time for reading data increases proportionally to the data amount
  - In other words, exclude seek time from latency
- Simulation based on blu-ray
  - It’s more appropriate for active archive usage, since positioning latency is much less
    - 2 min. for media exchange
    - 93 min. for reading a whole medium
# Media Exchange

- **Original layout**
  - # media exchange are the same among all actions
  - # accessing media is always same as the # collecting days

- **Action-oriented layout**
  - Smaller actions exhibit fewer media exchange
  - One medium can hold more days’ logs with smaller actions

---

**Graphs:**
- **Left Graph:**
  - Title: 
  - X-axis: Days
  - Y-axis: # media exchange
  - Data points showing increasing # media exchange over time.

- **Right Graph:**
  - Title: 
  - X-axis: Days
  - Y-axis: # media exchange
  - Multiple lines indicating different actions (Logoff OR Logon, Print out, Taking out files, Device config. change, Run application, File manipulation).
Absolute Latency Reduction

- Smaller actions exhibit greater reduction
- With the help of fewer media exchange
Differences among actions become much greater

Latency for reading data varies a lot
Finding a Balanced Setting

- Replicating smaller four actions must be the best choice
  - 31% relative latency reduction on average among 6 queries
  - 91% capacity efficiency

- In the case of using replicas for high availability
  - 34% relative latency reduction on average
  - 100% capacity efficiency
Summary

- Cold storage is a good choice for active archiving in the viewpoint of TCO
- Reducing media exchange must lower the barrier for installing such devices
- Proposed two approaches of replication for meeting multi-dimensional searches
  1. Taking relative latency reduction into account
  2. Utilizing replicas prepared for avoiding data loss
- Simulation results based on sample logs
  1. 31% average relative latency reduction with 91% capacity efficiency
  2. 34% average relative latency reduction with 100% capacity efficiency
- Future work
  - Experimentation using a real hardware