
# AliHB Real-Time Cold data Backup

孟庆义(mengqingyi)

# 目录

Content

# HBase Backup State

| | Against Hardware failure | Against User Application error | RPO | RTO |
|---|---|---|---|---|
| Snapshot | NO | YES | N/A | N/A |
| Replication | YES | NO | seconds | seconds |
| HBase Backup Restore | YES | YES | minutes | Increase with data size |
| AliHB Real-Time cold data backup | YES | YES | seconds | minutes |

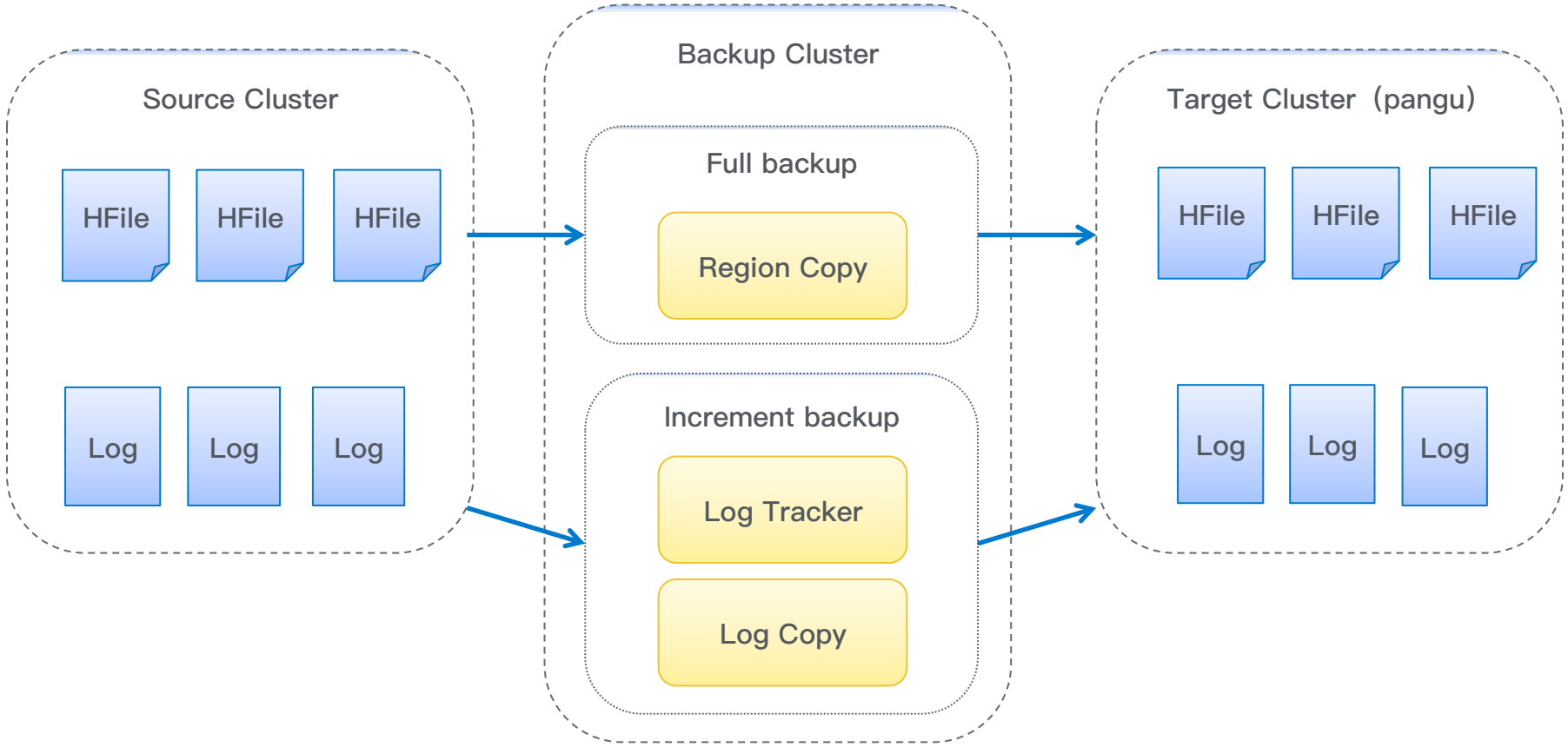
# Alibaba's requirements for Backup

- RPO < 1minutes
- Predictable RTO for PB scale data
- Low Cost
- NO affect on Online service
- Easy Management

# AliHB Real-Time Cold data Backup

- Real-Time incremental backup
- Independent with HBase
  - No need for snapshot
- Stateless worker node
- Backup in heterogeneous Storage maintained by another team

# Backup Overview

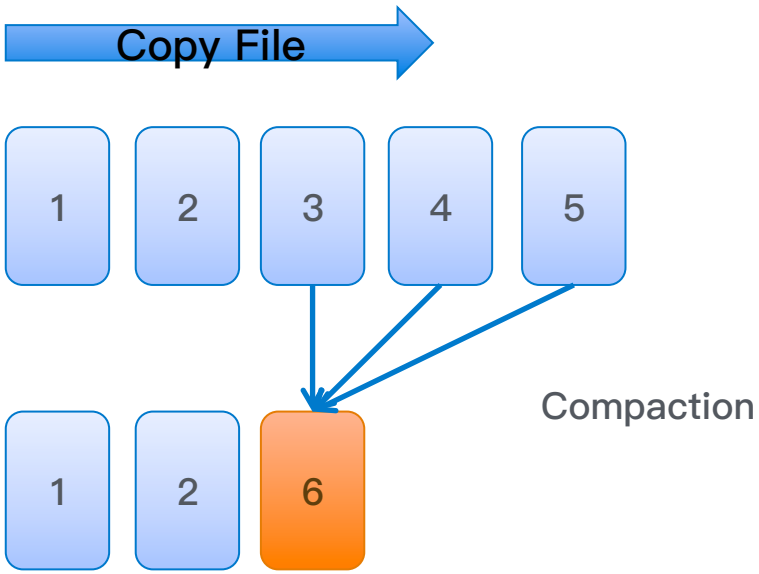


# Full Backup

- Job copy for a table
- Task copy for a region
- Challenge: region's file list keep changing
  - Compaction remove old files
  - Split remove the entire region
  - Merge remove the entire region

# Compaction

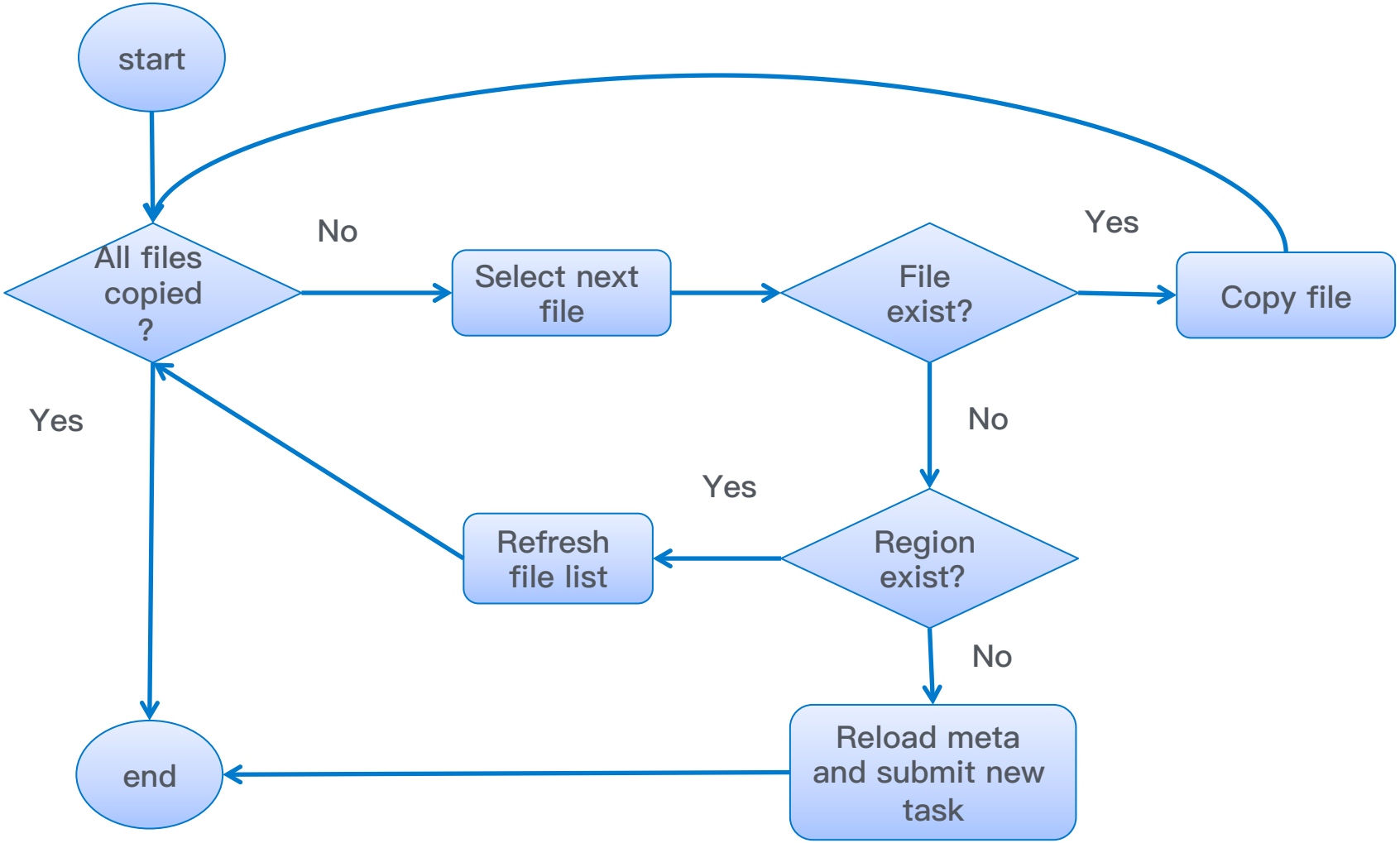
- At first we have file 1,2,3,4,5
- When copy 4, found it missing
- Refresh list we have 1,2,6
- Copy 6



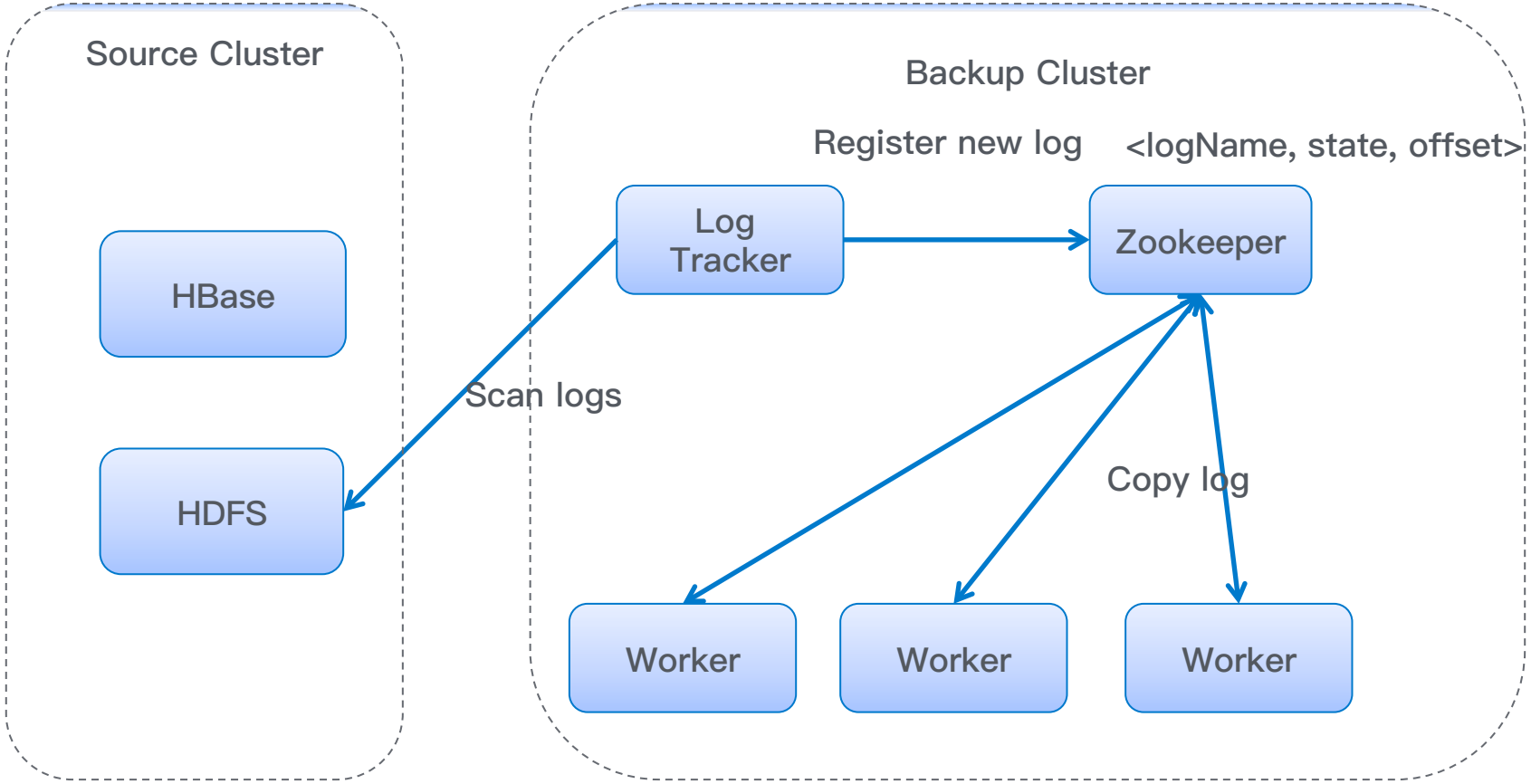
# Split

- We are the parent region
  - Found region missing, reload meta and resubmit tasks
- We are the child region
  - Copy the reference file and it's original file
  - If referenced file missing, refresh the file list and continue
- Merge works like split

# Algorithm


start
All files copied ?
No
Select next file
File exist?
Yes
Copy file
Yes
No
Yes
Refresh file list
Region exist?
No
end
Reload meta and submit new task

# Incremental Backup




Latency < 10 seconds

# Log Lifecycle

- Writing
  - Log Tracker period scan and find new logs
- Closed
  - If not the latest log of the region server or in the “.oldlogs”
- Finished
  - If worker has copied the whole closed Log
- Deleted
  - If Log Tracker can not find it in HBase and it’s finished on backup, then delete the log record on backup system

# Data Consistence

- Full comparison
  - Do sample comparison
  - Sample on every region
  - Balanced sample, use index of the largest file for each region
- Incremental comparison
  - Compare recent logs

# Restore Scenes

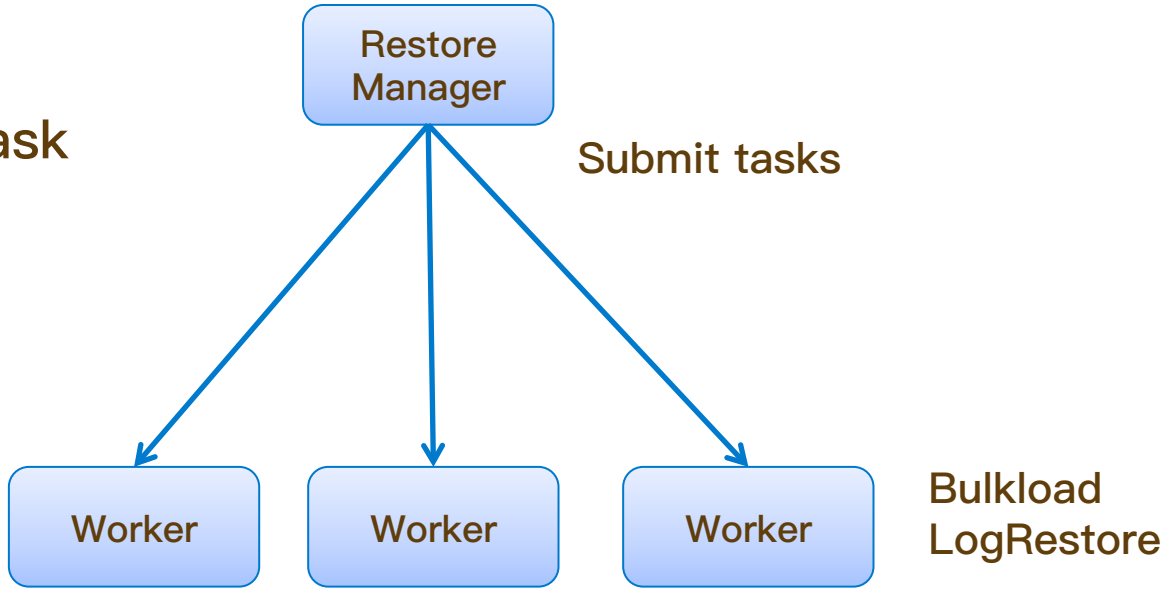
- Cluster Level
  - Restore the whole cluster
- Table Level
  - Restore one or list of tables
- Region Level
  - Restore ranged data of some table
- Restore to given time point

# Restore Tools

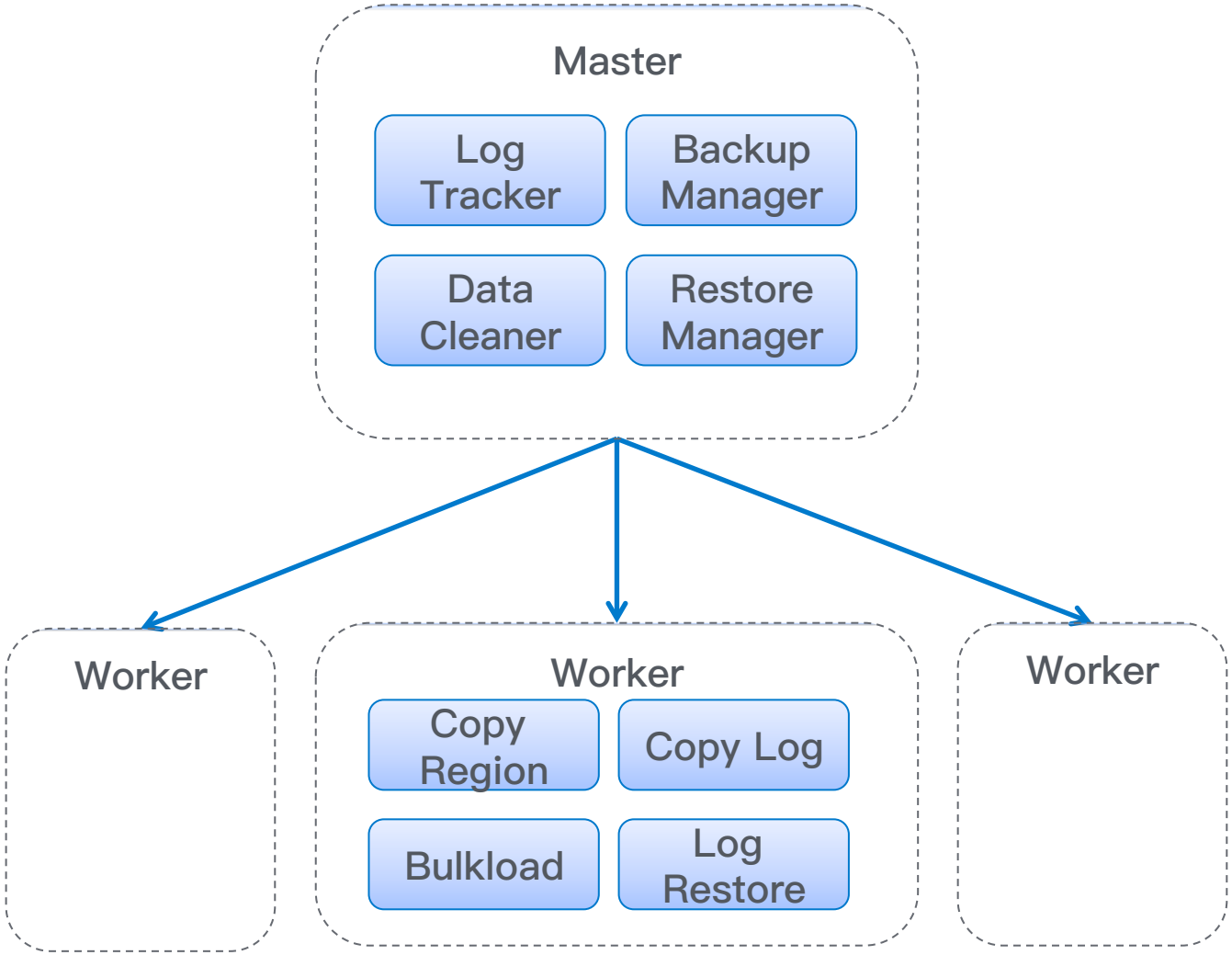
- Bulkload the full backup
  - Filter hfiles by table name and range
- Use LogRestore tool to restore logs
  - Filter by table name
  - Filter by range
  - Filter by timestamp

# Restore Runtime

- HFiles
  - Split by region, one region one task
- Logs
  - Each log is a task



# Real-Time Cold data Backup



# WEB UI

## cluster backup

clusterName:
sourceClusterKey:
sinkFsURI:
sinkDir:
backupTables:

Submit Job

## cluster restore

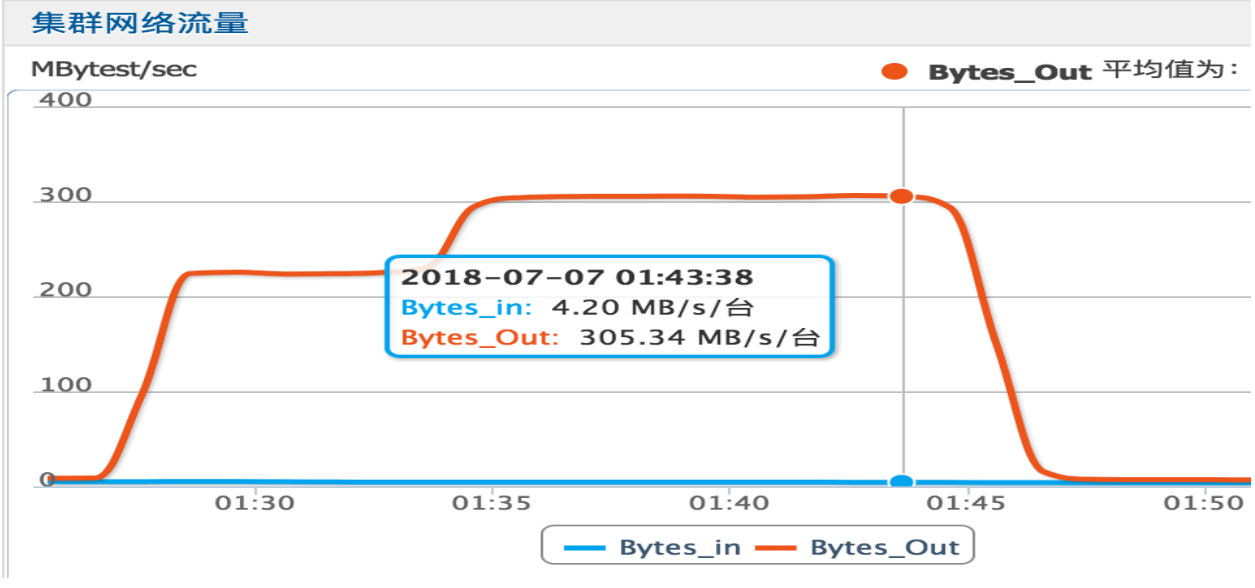
clusterName:
sourceFsURI:
sourceHFileDir:
sourceLogDir:
sinkClusterKey:
logStartTs:
logEndTs:
restoreTables:

Submit Job

# Performance

| sinkTableName | copyCost | copyDataSize(MB) | copyFileNum | bulkLoadCost | message | Oper |
|---|---|---|---|---|---|---|
| [redacted] | 0 Days 0 Hours 7 Minutes 9 Seconds | 31743208 | 33316 | 0 Days 0 Hours 1 Minutes 3 Seconds | null | delete |
| [redacted] | 0 Days 0 Hours 21 Minutes 59 Seconds | 119544362 | 221457 | 0 Days 0 Hours 53 Minutes 14 Seconds | null | delete |

Backup System
200Nodes
110TB data
backup 22minutes
Restore 53minutes




HBase 377Nodes

# Conclusion

- AliHB Real-time Cold data backup
  - Realtime incremental backup keep the latency in seconds
  - Scale out ability to obtain more power on restore
  - Use less resources on normal backup
  - Independent with HBase, easy to deploy and upgrade

# Future works

- Incremental Restore
  - Recognize Hot / Cold Data
  - Resume the hbase service after Restore hot data
  - Access the cold data through reference file
  - Background restore cold data
- Put log lifecycle manage on HBase
  - Period scan on .oldlogs cause pressure on NN
  - Keep only the necessary logs on zookeeper
- Compact hlogs to Hfile
  - Save storage space
  - Speed up restore

# 谢谢观看

Thanks

# 欢迎加入HBase中文社区

- HBase中文技术社区 http://www.hbase.group/


技术社区微信公众号

钉钉技术交流群