QEMU BACKUP STATUS

Vladimir Sementsov-Ogievskiy
October 2018
QEMU Backup

• Incremental
• External API
• Architecture & performance
Incremental backup

disk

0 1 1 1 0 0 1 0

dirty bitmap

target
incremental backup

persistent (qcow2)

migration
suspend/resume

dirty bitmaps

NBD export

snapshots
Bitmaps: persistent (2.10)

qmp: block-dirty-bitmap-add (persistent=true)

persistent bitmap in Qemu

persistent bitmap in Qcow2

open()
close()
Bitmaps: migration (2.12)

qmp: migrate-set-capabilities (capabilities=

  [{capability: dirty-bitmaps, state: True}])
qmp: migrate
qmp: migrate-start-postcopy

source Qemu

01110010

target Qemu

0000000000
01110010

merge

postcopy
Bitmaps: migration through storage (2.10)

qmp: migrate

source Qemu

01110010

target Qemu

01110010

downtime

shared Qcow2

01110010
## Bitmaps: snapshots

<table>
<thead>
<tr>
<th>make a snapshot</th>
<th>switch to the snapshot</th>
</tr>
</thead>
<tbody>
<tr>
<td>do nothing</td>
<td>all bitmaps are invalid</td>
</tr>
<tr>
<td></td>
<td>• mark them all dirty</td>
</tr>
<tr>
<td></td>
<td>• drop all the bitmaps</td>
</tr>
<tr>
<td></td>
<td>• <strong>error</strong></td>
</tr>
</tbody>
</table>
## Bitmaps: snapshots

<table>
<thead>
<tr>
<th>Make a snapshot</th>
<th>Switch to the snapshot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do nothing</td>
<td>All bitmaps are invalid</td>
</tr>
<tr>
<td></td>
<td>• Mark them all dirty</td>
</tr>
<tr>
<td></td>
<td>• Drop all the bitmaps</td>
</tr>
<tr>
<td></td>
<td>• Error</td>
</tr>
<tr>
<td>Enable <strong>dirty-bitmaps migration capability</strong></td>
<td>Drop all the bitmaps and load them from vmstate</td>
</tr>
</tbody>
</table>
Bitmaps: snapshots

Legend:
snapshot
active state
Bitmaps: snapshots

Legend:
- snapshot
- active state
- incremental backup
Bitmaps: snapshots

Legend:
snapshot
active state
incremental backup

[Diagram showing a tree with nodes representing transitions between states. The nodes are colored to indicate different states, with the legend providing a key for snapshot, active state, and incremental backup.]
Bitmaps: snapshots

Legend:
snapshot
active state
incremental backup

01110010
Bitmaps: snapshots

Legend:
- snapshot
- active state
- incremental backup

01110010

00000000
Bitmaps: NBD Export (3.0)

Negotiation phase: select metadata contexts

- Commands:
  - NBD_OPT_LIST_META_CONTEXT
  - NBD_OPT_SET_META_CONTEXT
- base:allocation - common block status (2.12)
- qemu:dirty-bitmap:<bitmap-export-name> (3.0)

Transmission phase: get metadata

- NBD_CMD_BLOCK_STATUS
Bitmaps: QMP API

basic (2.4)
- block-dirty-bitmap-add
  - persistent (2.10)
  - x-disabled (3.0)
- block-dirty-bitmap-remove
- block-dirty-bitmap-clear

new (3.0)
- x-block-dirty-bitmap-enable
- x-block-dirty-bitmap-disable
- x-block-dirty-bitmap-merge
  - node
  - dst_name
  - src_name
- x-nbd-server-add-bitmap
  - name
  - bitmap
  - bitmap-export-name
External backup API: image fleecing (3.0)
External backup API: image fleecing (next)
External backup API: image fleecing (next)

Diagram:

- Device
- Root
- Backup-top
- Active disk
- Target
- Temp qcow2
- Backup
- Far target

Backings:
- Backing
- Backing
External backup API: image fleecing (next)

- device
  - root
  - backup-top
  - active disk

  backing

  target

  temp qcow2

  backup

  far target

  — extra CBW operations
  — extra save/load to temp
Backup architecture and performance

Guest write hangs
  • don’t wait for write to far target (fleecing to temp)
  • don’t wait for write at all if possible
  • priority of CBW
Total number of iops
  • don’t write to temp if possible
  • share reads
Decrease number of CBW
  • lazy discards
  • copy from temp at last (or at first, optionally?)
  • predict guest writes
Optimize copying loop
  • asynchronous copying like in mirror instead of sequential
Backup architecture and performance

device

root

backup-top

backing

active disk

backup reads

target

RAM cache

backup writes

far target

temp qcow2

backing
Qemu backup status

Merged
• QCOW2 bitmaps (2.10)
• Migration (2.12) ⇒ suspend/resume & snapshots
• Fleecing (3.0)
• NBD BLOCK_STATUS (3.0)
• QMP API (3.0, still with x- prefix)

Plans
• Backup architecture
Qemu backup status

Merged
• Qcow2 bitmaps (2.10)
• Migration (2.12) ⇒ suspend/resume & snapshots
• Fleecing (3.0)
• NBD BLOCK_STATUS (3.0)
• QMP API (3.0, still with x- prefix)

Plans
• Backup architecture

Vladimir Sementsov-Ogievskiy
vsementsov@virtuozzo.com