**虚机云盘热迁移（自动稀疏版）**

1. **blockdev-mirror介绍**

blockdev-mirror 的主要功能是将源块设备（source block device）的数据同步到目标块设备（target block device），可以用于以下场景：

* 实时存储迁移：将虚拟机的磁盘镜像从一个存储位置迁移到另一个存储位置，而不中断虚拟机的运行。
* 创建点对点副本：在同步完成后，生成目标设备的点对点副本，用于备份或切换存储。
* 支持复杂存储格式：与 drive-mirror 相比，blockdev-mirror 允许通过 blockdev-add 命令更灵活地配置目标设备，例如支持 LUKS 加密磁盘。

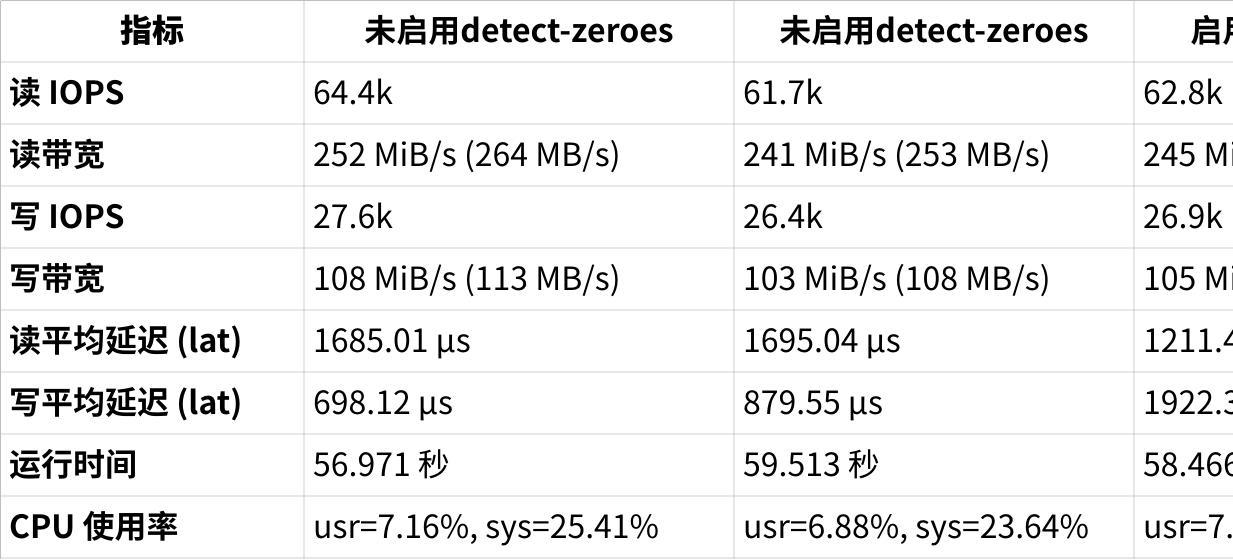
同步完成后，可以选择：

* 取消任务（block-job-cancel）：生成目标设备在取消时刻的点对点副本。
* 完成任务（block-job-complete）：将虚拟机的块设备切换到目标设备，新的写入操作直接写入目标设备。

detect-zeroes=unmap 是 blockdev-mirror 中用于优化零数据区域处理的配置选项。通过检测并丢弃零数据区域，它可以显著减少目标设备的存储占用并提高镜像效率。特别适合需要高效存储或稀疏镜像的场景，如存储迁移或备份。

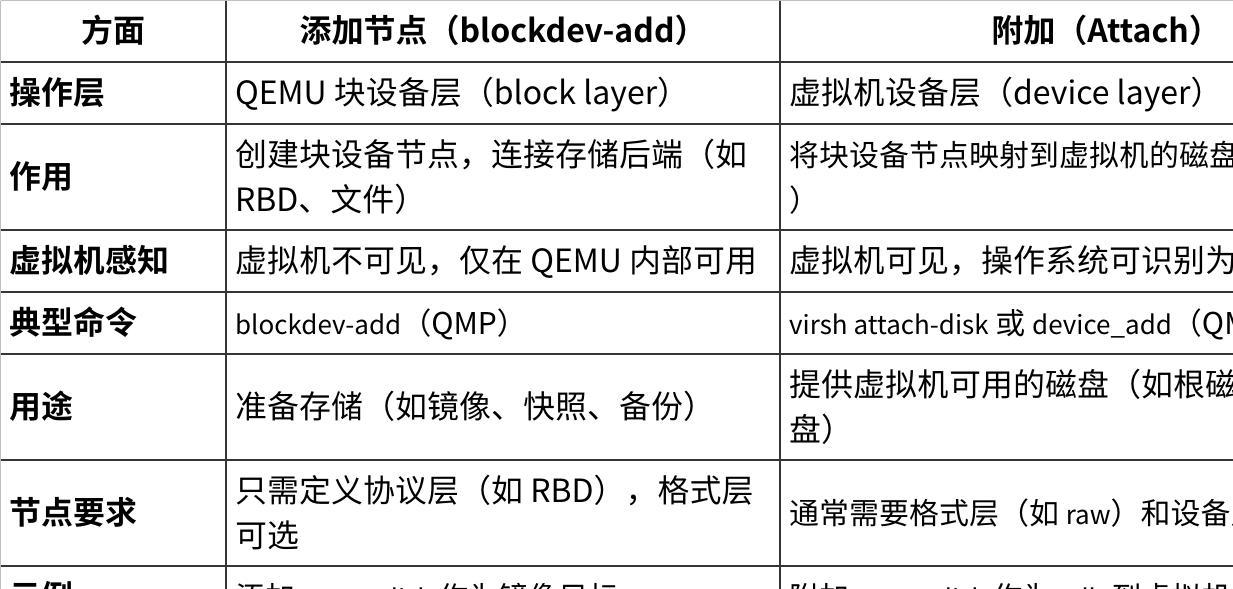
2. **关键性能指标对比**

性能影响：启用 detect-zeroes=unmap 可能增加少量开销，因为 QEMU 需要检测零数据区域，但通常能通过减少数据传输和存储占用带来整体性能提升



**点击图片可查看完整电子表格**

3. **相关概念-添加节点vs附加**



**点击图片可查看完整电子表格**

4. **线上存储稀疏热迁移前置条件**

nova\_libvirt 容器升级，虚机使用新版本qemu进程

artifacts.iflytek.com/docker-private/xfyun/kolla/rocky-source-nova-libvirt:yoga

|  |
| --- |
| Bash (nova-libvirt)[root@rg21-compute138 /]# virsh version Compiled against library: libvirt 9.5.0 Using library: libvirt 9.5.0 Using API: QEMU 9.5.0 Running hypervisor: QEMU 8.0.0  [root@rg21-compute138 ~]# docker ps -a |grep rocky f5f04b9f36a2 artifacts.iflytek.com/docker-private/xfyun/kolla/rocky-source-nova-libvirt:yoga "dumb-init --single-…" 3 hours ago Up 3 hours nova\_libvirt |

5. **开始迁移步骤**

* 准备

|  |
| --- |
| Bash #目标存储集群创建卷 rbd create --size 50G nvme3\_pool/volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23  #需要以下特性，默认是有这些特性得 features: layering, exclusive-lock, object-map, fast-diff  #查看虚机块设备 (nova-libvirt)[root@rg21-compute138 /]# virsh qemu-monitor-command instance-0000359f --hmp 'info block' libvirt-2-format: json:{"driver": "raw", "file": {"pool": "vms", "image": "a24b00f2-0e0c-41ff-8edf-5b91a986f787\_disk.config", "server.0.host": "172.29.224.11", "server.1.host": "172.29.224.12", "server.2.host": "172.29.224.13", "driver": "rbd", "server.0.port": "6789", "server.1.port": "6789", "server.2.port": "6789", "user": "nova"}} (raw, read-only)  Attached to: ide0-0-0  Removable device: locked, tray closed  Cache mode: writeback, direct  libvirt-1-format: json:{"driver": "raw", "file": {"pool": "nvme1\_pool", "image": "volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23", "server.0.host": "172.29.240.11", "server.1.host": "172.29.240.12", "server.2.host": "172.29.240.13", "driver": "rbd", "server.0.port": "6789", "server.1.port": "6789", "server.2.port": "6789", "user": "cinder"}} (raw)  Attached to: /machine/peripheral/virtio-disk0/virtio-backend  Cache mode: writeback, direct  #将目标集群创建的卷添加到qemu virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"blockdev-add","arguments":{"node-name":"target-disk","driver":"rbd","pool":"nvme3\_pool","image":"volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23","server":[{"host":"172.29.224.11","port":"6789"}],"detect-zeroes":"unmap","discard":"unmap"}}' {  "return": {   },  "id": "libvirt-535" }  #可以看到target-disk已经添加qemu virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"query-named-block-nodes"}' |grep target-disk |

* 开始迁移

|  |
| --- |
| Bash #开始迁移 (nova-libvirt)[root@rg21-compute138 /]# virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"blockdev-mirror","arguments":{"device":"libvirt-1-format","target":"target-disk","sync":"full","job-id":"mirror-job"}}' {  "return": {   },  "id": "libvirt-549" }  #查看迁移进度 (nova-libvirt)[root@rg21-compute138 /]# virsh qemu-monitor-command instance-0000359f --hmp 'info block-jobs' Type mirror, device mirror-job: Completed 53689188352 of 53689188352 bytes, speed limit 0 bytes/s  #完成迁移 (nova-libvirt)[root@rg21-compute138 /]# virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"block-job-complete","arguments":{"device":"mirror-job"}}' {  "return": {   },  "id": "libvirt-574" } |

* 迁移后验证

|  |
| --- |
| Bash #再次查看虚机块设备 libvirt-1-format 被替换成了target-disk (nova-libvirt)[root@rg21-compute138 /]# virsh qemu-monitor-command instance-0000359f --hmp 'info block' libvirt-2-format: json:{"driver": "raw", "file": {"pool": "vms", "image": "a24b00f2-0e0c-41ff-8edf-5b91a986f787\_disk.config", "server.0.host": "172.29.224.11", "server.1.host": "172.29.224.12", "server.2.host": "172.29.224.13", "driver": "rbd", "server.0.port": "6789", "server.1.port": "6789", "server.2.port": "6789", "user": "nova"}} (raw, read-only)  Attached to: ide0-0-0  Removable device: locked, tray closed  Cache mode: writeback, direct  target-disk: json:{"pool": "nvme3\_pool", "image": "volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23", "server.0.host": "172.29.224.11", "driver": "rbd", "server.0.port": "6789"} (rbd)  Attached to: /machine/peripheral/virtio-disk0/virtio-backend  Cache mode: writeback  Detect zeroes: unmap  **#使用du查看ceph集群卷,可以看到没有膨胀** [root@rg21-control001 ~]# rbd du nvme3\_pool/volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23  NAME PROVISIONED USED  volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23 50 GiB 24 GiB      修改openstack数据库 ，硬重启或者热迁移后xml文件自动更新 |

|  |
| --- |
| Bash #原集群卷从qemu删除，否则老存储集群还有watcher 无法删除卷 virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"blockdev-del","arguments":{"node-name":"libvirt-1-format"}}' virsh qemu-monitor-command instance-0000359f --pretty '{"execute":"blockdev-del","arguments":{"node-name":"libvirt-1-storage"}} |

* 热迁移验证

|  |
| --- |
| Bash #其他相关测试验证 热迁移虚机测试  可以看到热迁移后unmap消失 libvirt-1-format 的pool和host为目标集群信息 (nova-libvirt)[root@rg21-compute114 /]# virsh qemu-monitor-command instance-0000359f --hmp 'info block' libvirt-2-format: json:{"driver": "raw", "file": {"pool": "vms", "image": "a24b00f2-0e0c-41ff-8edf-5b91a986f787\_disk.config", "server.0.host": "172.29.224.11", "server.1.host": "172.29.224.12", "server.2.host": "172.29.224.13", "driver": "rbd", "server.0.port": "6789", "server.1.port": "6789", "server.2.port": "6789", "user": "nova"}} (raw, read-only)  Attached to: ide0-0-0  Removable device: locked, tray closed  Cache mode: writeback, direct  libvirt-1-format: json:{"driver": "raw", "file": {"pool": "nvme3\_pool", "image": "volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23", "server.0.host": "172.29.224.11", "server.1.host": "172.29.224.12", "server.2.host": "172.29.224.13", "driver": "rbd", "server.0.port": "6789", "server.1.port": "6789", "server.2.port": "6789", "user": "cinder"}} (raw)  Attached to: /machine/peripheral/virtio-disk0/virtio-backend  Cache mode: writeback, direct  I/O throttling: bps=838860800 bps\_rd=0 bps\_wr=0 bps\_max=0 bps\_rd\_max=0 bps\_wr\_max=0 iops=8000 iops\_rd=0 iops\_wr=0 iops\_max=0 iops\_rd\_max=0 iops\_wr\_max=0 iops\_size=0 group=/machine/peripheral/virtio-disk0/virtio-backend    老存储集群watcher查看 [root@rg21-control001 cinder-volume]# rbd status nvme1\_pool/volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23 -c cephB.conf Watchers: none 新存储集群watcher查看 [root@rg21-control001 ~]# rbd status nvme3\_pool/volume-d7e4a44e-63bc-4748-8419-3c0dc6eb7b23 Watchers:  watcher=172.29.224.196:0/1116629098 client.352163930 cookie=140526565228816   **热迁移也没问题** |

* 相关截图

