

分 数:	
评卷人:	

华中科技大学

研究生（数据中心技术）课程 实验报告

学 号 M202173487

姓 名 廖子逸

专 业 电子信息

课程指导教师 施展 童薇

院（系、所） 武汉光电国家研究中心

2021 年 1 月 6 日

一 Minio

实验一：系统搭建

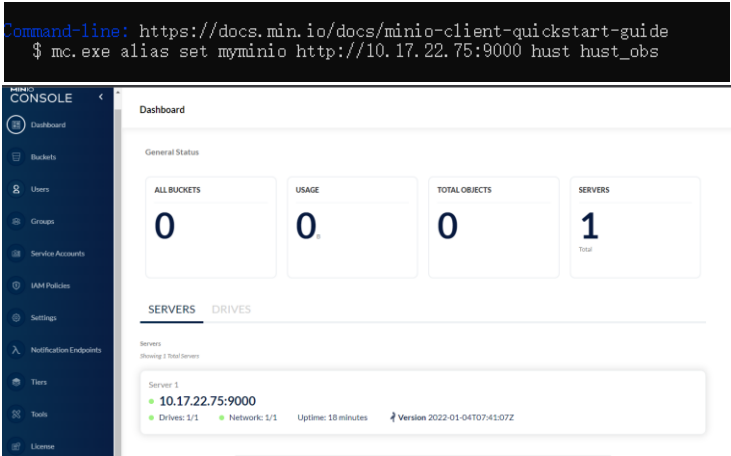


图 1 Minio 搭建 Dashboard 示例图

实验二：性能观测

步骤一：新建桶

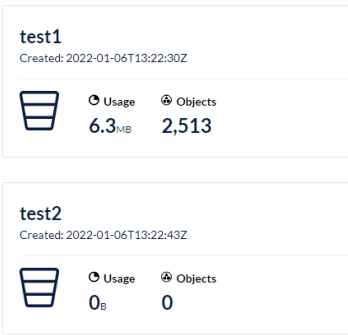


图 2 新建 Bucket 示例图

步骤二：修改参数，观测结果

Config1:

```
s3bench.exe ^
-accessKey=hust ^
-accessSecret=hust_obs ^
-bucket=test2 ^
-endpoint=http://127.0.0.1:9000 ^
-numClients=8 ^
-numSamples=100 ^
-objectNamePrefix=loadgen ^
-objectSize=1024
pause
```

```
Results Summary for Write Operation(s)
Total Transferred: 0.098 MB
Total Throughput: 0.02 MB/s
Total Duration: 4.723 s
Number of Errors: 0

-----
Write times Max: 0.779 s
Write times 99th %ile: 0.779 s
Write times 90th %ile: 0.551 s
Write times 75th %ile: 0.448 s
Write times 50th %ile: 0.371 s
Write times 25th %ile: 0.260 s
Write times Min: 0.105 s

Results Summary for Read Operation(s)
Total Transferred: 0.098 MB
Total Throughput: 1.84 MB/s
Total Duration: 0.053 s
Number of Errors: 0

-----
Read times Max: 0.047 s
Read times 99th %ile: 0.047 s
Read times 90th %ile: 0.002 s
Read times 75th %ile: 0.001 s
Read times 50th %ile: 0.001 s
Read times 25th %ile: 0.001 s
Read times Min: 0.000 s
```

图 3 Config1 下的性能观测

Config2:

```
s3bench.exe ^
  -accessKey=hust ^
  -accessSecret=hust_obs ^
  -bucket=test2 ^
  -endpoint=http://127.0.0.1:9000 ^
  -numClients=8 ^
  -numSamples=100 ^
  -objectNamePrefix=loadgen ^
  -objectSize=10240
pause
```

```
Results Summary for Write Operation(s)
Total Transferred: 0.977 MB
Total Throughput: 0.17 MB/s
Total Duration: 5.742 s
Number of Errors: 0
-----
Write times Max: 0.750 s
Write times 99th %ile: 0.750 s
Write times 90th %ile: 0.628 s
Write times 75th %ile: 0.530 s
Write times 50th %ile: 0.437 s
Write times 25th %ile: 0.360 s
Write times Min: 0.142 s

Results Summary for Read Operation(s)
Total Transferred: 0.977 MB
Total Throughput: 94.05 MB/s
Total Duration: 0.010 s
Number of Errors: 0
-----
Read times Max: 0.002 s
Read times 99th %ile: 0.002 s
Read times 90th %ile: 0.001 s
Read times 75th %ile: 0.001 s
Read times 50th %ile: 0.001 s
Read times 25th %ile: 0.001 s
Read times Min: 0.001 s
```

图 4 Config2 下的性能观测

Config3:

```
s3bench.exe ^
  -accessKey=hust ^
  -accessSecret=hust_obs ^
  -bucket=test2 ^
  -endpoint=http://127.0.0.1:9000 ^
  -numClients=30 ^
  -numSamples=100 ^
  -objectNamePrefix=loadgen ^
  -objectSize=1024
pause
```

```
Results Summary for Write Operation(s)
Total Transferred: 0.098 MB
Total Throughput: 0.01 MB/s
Total Duration: 8.555 s
Number of Errors: 0
-----
Write times Max: 4.979 s
Write times 99th %ile: 4.979 s
Write times 90th %ile: 4.413 s
Write times 75th %ile: 3.240 s
Write times 50th %ile: 1.724 s
Write times 25th %ile: 1.591 s
Write times Min: 0.771 s

Results Summary for Read Operation(s)
Total Transferred: 0.098 MB
Total Throughput: 6.10 MB/s
Total Duration: 0.016 s
Number of Errors: 0
-----
Read times Max: 0.009 s
Read times 99th %ile: 0.009 s
Read times 90th %ile: 0.008 s
Read times 75th %ile: 0.005 s
Read times 50th %ile: 0.004 s
Read times 25th %ile: 0.003 s
Read times Min: 0.001 s
```

图 5 Config3 下的性能观测

总结:

对比图 3 和图 4, 配置 2 相比配置 1, 扩大了 10 倍块大小, 由结果得出: 扩大块大小, 在写测试性能上变化不大, 但在读测试上, 带宽明显变大, 延迟明显变小。

对比图 3 和图 5, 配置 3 相比配置 1, 增加了约 4 倍数量的 clients, 导致: 写测试中, 带宽降低 1 倍, 延迟大幅度变大 (增加后的最小延迟等于增加前的最大延迟); 读测试中, 带宽提高 4 倍, 延迟大幅度降低。

实验三: 尾延迟

步骤一: 尾延迟观测

尾部延迟 (也称为高百分比延迟) 是指客户端很少看到的高延迟。例如: “我的服务通常在 10 毫秒左右响应, 但有时需要 100 毫秒左右”。世界上有很多导致尾部等待时间的原因, 包括争用, 垃圾回收, 数据包丢失, 主机故障以及操作系统在后台执行的奇怪操作。

为什么要关注尾延迟? 木桶效应, 短板效应。

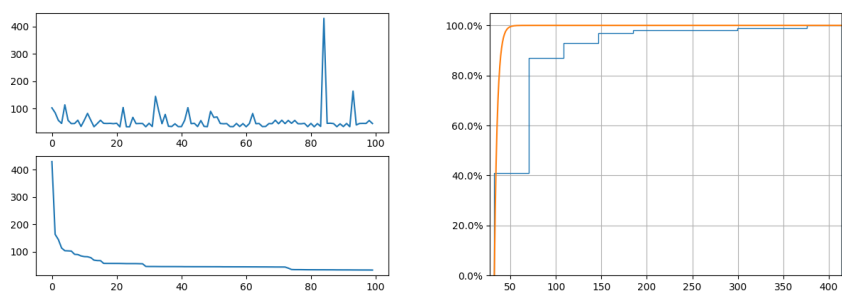


图 6 尾延迟观测

步骤二：对冲请求

原理：当请求在指定的时间间隔后没有返回时，会发起对冲请求。通常对于 95%尾延迟作为阈值，认为 95%尾延迟以上的请求出现问题，然后重新发送请求。

操作：修改 `request_timing()` 函数，当 `system_time` 超过阈值的时候，递归调用一次 `request_timing()` 函数并且返回较小一个 `system_time`。由图 6 观察得到，阈值选取 50ms。

结果：如图 7 所示，尾延迟降低至 110ms（原来为 400ms）。实验验证了对冲请求的有效性。

不足：在 `request_timing()` 函数中穿插时间判断，然后递归调用 `request_timing()` 是一种比较简单的实现方式，但是有可能影响原有进程效率，且不符合实际的多客户访问情况。应该在 `executor.submit()` 函数中设置计时器，超时则重新请求，才是完整端到端测试，但这样需要修改的代码较多，留给以后尝试。

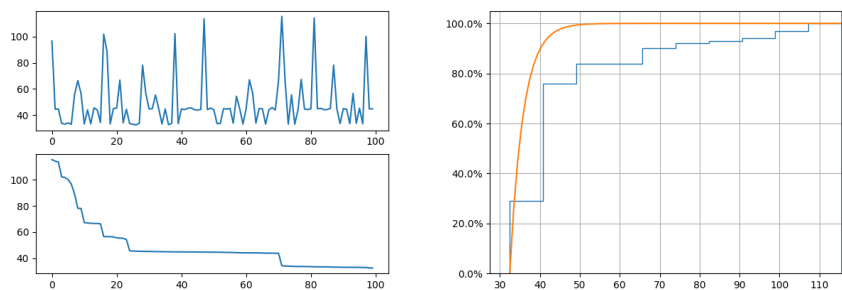


图 7 对冲请求后的尾延迟观测

二 Ceph

实验一：系统搭建

系统使用 VirtualBox 建立三台 Ubuntu 虚拟机搭建 Ceph 集群。

具体步骤为：

- 1 使用 VirtualBox 安装三台虚拟机，系统版本为：Ubuntu20.04，每台虚拟机添加一个虚拟硬盘做 OSD。
- 2 分配 IP 段，连接 Xshell 操作虚拟机，主机和 IP 分配如表 1 所示。
- 3 安装 Cephadm，使用 docker 容器部署 Ceph 集群。
- 4 Ceph 搭建完成。如图 8 所示。

5 配置 Ceph RGW（对象存储网关），用来配合 S3 的负载脚本使用（此步骤一直 BUG，rgw 进程报 error，如图 9。Ceph 实验就差这一步!!! 所以才做了个 Minio 保底实验，后续准备重装系统，重做 Ceph 实验）。

Cluster » Hosts

Hosts List	Overall Performance			
<div>+ Create</div>				
Hostname		Services	Labels	Version
ceph1		mgr.ceph1.sgnujc, mon.ceph1, osd.0		15.2.13
ceph2		mgr.ceph2.lwxdaw, mon.ceph2, osd.1		15.2.13
ceph3		mon.ceph3, osd.2		15.2.13
0 selected / 3 total				

图 8 Ceph Dashboard 示意图

表 1 Ceph Host IP

Host	IP
Ceph1	10.0.0.101
Ceph2	10.0.0.102
Ceph3	10.0.0.103

NAME	HOST	STATUS	REFRESHED	AGE	VERSION	IMAGE NAME	IMAGE ID	CONTAINER ID
crash.ceph1	ceph1	running (8h)	10m ago	5h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	8665cabdd3ac
crash.ceph2	ceph2	running (8h)	10m ago	4h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	68b81cc17748
crash.ceph3	ceph3	running (8h)	10m ago	8h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	ef10d12d04f5
grafana.ceph2	ceph2	running (8h)	10m ago	4h	6.7.4	docker.io/ceph/ceph-grafana:6.7.4	557c83e11646	9cf353f14e6d
mgr.ceph1.sgnujc	ceph1	running (8h)	10m ago	5h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	08e19d785ee9
mgr.ceph2.lwxdaw	ceph2	running (8h)	10m ago	4h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	a911ce401839
mon.ceph1	ceph1	running (8h)	10m ago	5h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	2d1889897eb5
mon.ceph2	ceph2	running (8h)	10m ago	8h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	1a3782251356
mon.ceph3	ceph3	running (8h)	10m ago	8h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	b4522fd66ddb
node-exporter.ceph2	ceph2	running (8h)	10m ago	4h	0.18.1	docker.io/prom/node-exporter:v0.18.1	e5a616e4b9cf	301f78b53638
node-exporter.ceph3	ceph3	running (8h)	10m ago	8h	0.18.1	docker.io/prom/node-exporter:v0.18.1	e5a616e4b9cf	f814a180f15f
osd.0	ceph1	running (7h)	10m ago	7h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	f14c7532ded3
osd.1	ceph2	running (7h)	10m ago	7h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	1f725f3c080a
osd.2	ceph3	running (7h)	10m ago	7h	15.2.13	docker.io/ceph/ceph:v15	2cf504fded39	5faeb6b9b2b1
prometheus.ceph2	ceph2	running (8h)	10m ago	4h	2.18.1	docker.io/prom/prometheus:v2.18.1	de242295e225	e0467455ead2
rgw.rgw.ceph1.qvkkgo	ceph1	error	10m ago	4h	<unknown>	docker.io/ceph/ceph:v15	<unknown>	<unknown>
rgw.rgw.ceph2.wybdfr	ceph2	error	10m ago	4h	<unknown>	docker.io/ceph/ceph:v15	<unknown>	<unknown>
rgw.rgw.ceph3.glnlon	ceph3	error	10m ago	4h	<unknown>	docker.io/ceph/ceph:v15	<unknown>	<unknown>

图 9 Ceph RGW 进程 error