

4.7. 性能管家 profile 和 top

对于性能监控,glusterfs 是提供了一些工具的,主要是三个命令 profile、top 和 statedump,这三者面向是不一样的,其中 statedump 主要是排查一些文件句柄泄漏这样的问题,并且可以获取到申请的系统内存等信息。而 profile 和 top 命令,则是通用的性能检测工具,其中 profile 命令默认是不打开的,需要手动启动该命令,适合用在排查当前的 volume 运行情况,哪些 fop 在运行并且延迟怎样; top 命令则适合查看某些文件的句柄连接和打开情况等。下面先来感受一下 profile 命令的使用。

```
1. root@gfs03:~# gluster volume profile test-shd-num info
2. Profile on Volume test-shd-num is not started
3.
4. root@gfs03:~# gluster volume profile test-shd-num start
5. Starting volume profile on test-shd-num has been successful
```

这里需要手动打开 profile 命令,然后下面挂载 volume,使用脚本创建大量的文件。

```
1. root@gfs01:/mnt/test-shd-num# cat test.sh
2. #!/bin/bash
3.
4.
5. for((i=1;i<=10000;i++));
6. do
7.     sudo touch $i.txt
8.     sudo date > $i.txt
9. done
```

这里默认创建一万个小文件,并且文件写入时间,接着执行脚本,并且使用 profile 命令来查看情况。

```
1. root@gfs03:~# gluster volume profile test-shd-num info
2. Brick: 10.0.12.9:/glusterfs/test-shd-num
```

```

3. -----
4. Cumulative Stats:
5.   Block Size:      32b+      64b+      4096b+
6.   No. of Reads:    0          0          0
7.   No. of Writes:   5838      1          1
8.
9.   %-latency Avg-latency Min-Latency Max-Latency No.of
      calls   Fop
10.  -----
11.   0.00   0.00 us   0.00 us   0.00 us   5680   FORGET
12.   0.00   0.00 us   0.00 us   0.00 us   11684  RELEASE
13.   0.00   0.00 us   0.00 us   0.00 us   82     RELEASEDIR
14.   0.03   34.08 us  28.46 us  39.69 us   2      GETXATTR
15.   0.04   45.86 us  25.16 us  66.56 us   2      OPENDIR
16.   2.48   15.42 us  9.22 us   68.12 us  328    ENTRYLK
17.   3.46   21.50 us  9.36 us   131.09 us 328    FINODELK
18.   4.52   56.15 us  38.82 us  243.58 us 164    SETATTR
19.   4.84   60.14 us  47.06 us  155.08 us 164    TRUNCAT
    E
20.   5.33   16.57 us  6.63 us   210.16 us 656    FLUSH
21.   5.38   66.90 us  49.67 us   172.31 us 164    WRITE
22.   5.93   18.44 us  9.63 us   1674.93 us 656    INODELK
23.   6.68   83.05 us  46.11 us   166.93 us 164    OPEN
24.   11.87  73.79 us  35.11 us   287.36 us 328    XATTROP
25.   14.59  90.70 us  29.52 us   7414.06 us 328    FXATTRO
    P
26.   14.75  183.39 us 127.94 us   745.42 us 164    CREAT
    E
27.   20.10  62.03 us  20.71 us   884.99 us 661    LOOKUP
28.
29.   Duration: 5141 seconds
30.   Data Read: 0 bytes
31. Data Written: 191004 bytes
32.
33. Interval 1 Stats:
34.   Block Size:      32b+
35.   No. of Reads:      0
36. No. of Writes:     164

```

```

37. %-latency Avg-latency Min-Latency Max-Latency No.of calls
    Fop
38. -----
39.  0.00  0.00 us  0.00 us  0.00 us  328  RELEASE
40.  0.00  0.00 us  0.00 us  0.00 us   2  RELEASEDIR
41.  0.03  34.08 us  28.46 us  39.69 us   2  GETXATTR
42.  0.04  45.86 us  25.16 us  66.56 us   2  OPENDIR
43.  2.48  15.42 us  9.22 us  68.12 us  328  ENTRYLK
44.  3.46  21.50 us  9.36 us  131.09 us  328  FINODELK

45.  4.52  56.15 us  38.82 us  243.58 us  164  SETATTR

46.  4.84  60.14 us  47.06 us  155.08 us  164  TRUNCAT
    E
47.  5.33  16.57 us  6.63 us  210.16 us  656  FLUSH
48.  5.38  66.90 us  49.67 us  172.31 us  164  WRITE
49.  5.93  18.44 us  9.63 us  1674.93 us  656  INODEL
    K
50.  6.68  83.05 us  46.11 us  166.93 us  164  OPEN
51. 11.87  73.79 us  35.11 us  287.36 us  328  XATTRO
    P
52. 14.59  90.70 us  29.52 us  7414.06 us  328  FXATTR
    OP
53. 14.75 183.39 us 127.94 us 745.42 us  164  CREAT
    E
54. 20.10 62.03 us 20.71 us 884.99 us  661  LOOKUP

55.
56.  Duration: 22 seconds
57.  Data Read: 0 bytes
58. Data Written: 5248 bytes
59.
60. Brick: 10.0.12.2:/glusterfs/test-shd-num
61. -----
62. Cumulative Stats:
63.  Block Size:    32b+          64b+          4096b+
64. No.of Reads:    0            0            0
65. No.of Writes:   5838         1            1
66. %-latency Avg-latency Min-Latency Max-Latency No.of calls
    Fop
67. -----
68.  0.00  0.00 us  0.00 us  0.00 us  5680  FORGET
69.  0.00  0.00 us  0.00 us  0.00 us 11684  RELEASE
70.  0.00  0.00 us  0.00 us  0.00 us   82  RELEASEDIR

```

71.	0.03	32.32 us	27.99 us	36.64 us	2	GETXATTR
72.	0.04	41.98 us	26.41 us	57.56 us	2	OPENDIR
73.	2.19	14.77 us	8.53 us	181.93 us	328	ENTRYLK
74.	3.66	24.66 us	9.71 us	796.68 us	328	FINODELK
75.	4.10	13.82 us	8.88 us	78.91 us	656	INODELK
76.	4.26	57.36 us	38.09 us	136.71 us	164	SETATTR
77.	4.84	65.21 us	45.30 us	367.69 us	164	TRUNCATE
78.	5.07	68.35 us	50.81 us	212.95 us	164	WRITE
79.	5.83	78.50 us	44.37 us	144.30 us	164	OPEN
80.	6.43	21.64 us	6.32 us	1814.42 us	656	FLUSH
81.	8.39	56.54 us	28.91 us	167.37 us	328	FXATTROP
82.	11.80	79.46 us	35.23 us	1428.97 us	328	XATTROP
83.	18.47	61.74 us	20.05 us	1442.66 us	661	LOOKUP
84.	24.89	335.33 us	148.63 us	3005.57 us	164	CREATE
85.						
86.	Duration: 5141 seconds					
87.	Data Read: 0 bytes					
88.	Data Written: 191004 bytes					
89.						
90.	Interval 1 Stats:					
91.	Block Size:		32b+			
92.	No. of Reads:		0			
93.	No. of Writes:		164			
94.	%-latency	Avg-latency	Min-Latency	Max-Latency	No.of calls	
95.	-----	-----	-----	-----	-----	
96.	0.00	0.00 us	0.00 us	0.00 us	328	RELEASE
97.	0.00	0.00 us	0.00 us	0.00 us	2	RELEASEDIR
98.	0.03	32.32 us	27.99 us	36.64 us	2	GETXATTR
99.	0.04	41.98 us	26.41 us	57.56 us	2	OPENDIR
100.	2.19	14.77 us	8.53 us	181.93 us	328	ENTRYLK
101.	3.66	24.66 us	9.71 us	796.68 us	328	FINODELK

102.	LK	4.10	13.82 us	8.88 us	78.91 us	656	INODE
103.	TTR	4.26	57.36 us	38.09 us	136.71 us	164	SETA
104.	CATE	4.84	65.21 us	45.30 us	367.69 us	164	TRUN
105.	TE	5.07	68.35 us	50.81 us	212.95 us	164	WRI
106.	N	5.83	78.50 us	44.37 us	144.30 us	164	OPE
107.	SH	6.43	21.64 us	6.32 us	1814.42 us	656	FLU
108.	TROP	8.39	56.54 us	28.91 us	167.37 us	328	FXAT
109.	TROP	11.80	79.46 us	35.23 us	1428.97 us	328	XAT
110.	OKUP	18.47	61.74 us	20.05 us	1442.66 us	661	LO
111.	REATE	24.89	335.33 us	148.63 us	3005.57 us	164	C
112.							
113.	Duration: 22 seconds						
114.	Data Read: 0 bytes						
115.	Data Written: 5248 bytes						
116.							
117.	Brick: 10.0.12.12:/glusterfs/test-shd-num						
118.	-----						
119.	Cumulative Stats:						
120.	Block Size:		32b+	64b+	4096b+		
121.	No. of Reads:		0	2	0		
122.	No. of Writes:		5838	1	1		
123.							
124.	%-latency lls Fop		Avg-latency	Min-Latency	Max-Latency	No.of ca	
125.	-----	-----	-----	-----	-----	----	
126.	0.00	0.00 us	0.00 us	0.00 us	5680	FORG	
127.	0.00	0.00 us	0.00 us	0.00 us	11684	RELEA	
128.	0.00	0.00 us	0.00 us	0.00 us	82	RELEASE	
129.	0.12	526.14 us	28.92 us	1023.36 us	2	OPEN	
	DIR						

130.	0.30	1273.69 us	46.65 us	2500.74 us	2	GETX ATTR
131.	1.50	39.03 us	9.81 us	2227.32 us	328	ENTR YLK
132.	1.81	94.05 us	44.32 us	1424.93 us	164	SET ATTR
133.	2.45	126.94 us	56.05 us	3559.47 us	164	TRU NCATE
134.	2.57	133.25 us	53.76 us	4854.92 us	164	WR ITE
135.	2.58	66.91 us	12.21 us	1911.90 us	328	FINO DELK
136.	2.78	36.02 us	7.80 us	1778.86 us	656	FLU SH
137.	2.85	36.95 us	10.73 us	3242.02 us	656	INO DELK
138.	3.19	165.32 us	68.88 us	1437.64 us	164	O PEN
139.	3.73	96.79 us	33.39 us	1454.73 us	328	FXAT TROP
140.	4.74	10077.98 us	15.21 us	40048.82 us	4	REA DDIRP
141.	5.82	151.06 us	39.16 us	3738.27 us	328	XAT TROP
142.	23.35	300.52 us	24.16 us	15041.37 us	661	L OOKUP
143.	42.20	2188.84 us	189.06 us	20331.92 us	164	CREATE
144.						
145.	Duration: 5141 seconds					
146.	Data Read: 184 bytes					
147.	Data Written: 191004 bytes					
148.						
149.	Interval 1 Stats:					
150.	Block Size: 32b+					
151.	No. of Reads: 0					
152.	No. of Writes: 164					
153.	%-latency lls	Avg-latency Fop	Min-Latency	Max-Latency	No.of ca	
154.	-----	-----	-----	-----	----	
155.	0.00	0.00 us	0.00 us	0.00 us	328	RELEAS E
156.	0.00	0.00 us	0.00 us	0.00 us	2	RELEASE DIR

157.	0.12	526.14 us	28.92 us	1023.36 us	2	OPEN DIR
158.	0.30	1273.69 us	46.65 us	2500.74 us	2	GETXA TTR
159.	1.50	39.03 us	9.81 us	2227.32 us	328	ENTR YLK
160.	1.81	94.05 us	44.32 us	1424.93 us	164	SETA TTR
161.	2.45	126.94 us	56.05 us	3559.47 us	164	TRU NCATE
162.	2.57	133.25 us	53.76 us	4854.92 us	164	WR ITE
163.	2.58	66.91 us	12.21 us	1911.90 us	328	FINO DELK
164.	2.78	36.02 us	7.80 us	1778.86 us	656	FLU SH
165.	2.85	36.95 us	10.73 us	3242.02 us	656	INO DELK
166.	3.19	165.32 us	68.88 us	1437.64 us	164	O PEN
167.	3.73	96.79 us	33.39 us	1454.73 us	328	FXAT TROP
168.	4.74	10077.98 us	15.21 us	40048.82 us	4	RE ADDIRP
169.	5.82	151.06 us	39.16 us	3738.27 us	328	XAT TROP
170.	23.35	300.52 us	24.16 us	15041.37 us	661	L OOKUP
171.	42.20	2188.84 us	189.06 us	20331.92 us	164	CREATE
172.						
173.	Duration: 22 seconds					
174.	Data Read: 0 bytes					
175.	Data Written: 5248 bytes					

这里在写入文件的过程中,可以看到 REaddir 和 CREATE 这两个操作耗时是比较多,其中 CREATE 是创建文件需要申请空间,而 REaddir 操作慢,则是因为对于 glusterfs 来说,这里并没有对元数据信息缓存,因此在有很多文件的时候,哪怕执行 ls 耗费的时间也会比较多,因此关于这部分内容,未来也是一个很重要的优化内容。

那么下面也可以使用 top 命令来查看一下 volume 的情况。

```
1. root@gfs03:~# gluster volume top test-shd-num write list-c
   nt 5
2. Brick: 10.0.12.9:/glusterfs/test-shd-num
3. Count    filename
4. =====
5. 1        /773.txt
6. 1        /772.txt
7. 1        /771.txt
8. 1        /770.txt
9. 1        /769.txt
10. Brick: 10.0.12.2:/glusterfs/test-shd-num
11. Count    filename
12. =====
13. 1        /773.txt
14. 1        /772.txt
15. 1        /771.txt
16. 1        /770.txt
17. 1        /769.txt
18. Brick: 10.0.12.12:/glusterfs/test-shd-num
19. Count    filename
20. =====
21. 1        /773.txt
22. 1        /772.txt
23. 1        /771.txt
24. 1        /770.txt
25. 1        /769.txt
26.
27.
28. root@gfs03:~# gluster volume top test-shd-num read list-c
   nt 5
29. Brick: 10.0.12.9:/glusterfs/test-shd-num
30. Brick: 10.0.12.2:/glusterfs/test-shd-num
31. Brick: 10.0.12.12:/glusterfs/test-shd-num
32. Count    filename
33. =====
34. 1        /test.sh
35. 1        /test.sh
```

这里在运行的时候可以看到当前的 volume 读写文件的情况，这里因为只是创建文件然后就关闭，因此这里显示的都是 1，如果想测试，可以自行做

一些其他的。

章节语:

这一章主要关注了 glusterfs 中的一些特性和工具的使用，从快照到 quota 容量限制，还有扩缩容等，这些内容在日常系统使用中会比较频繁遇到，因此有必要平时多测试和进行一些破坏性测试，例如对于快照，可以手动把 lvm 快照删掉或者移除，导致 glusterfs 快照丢失异常等。多进行这样的测试，以便在遇到生产环境问题时能更好地排查问题，平时项目维护的时候规避部分问题等。