Pulsar升级回退演练

准备工作

检查命令:

```
bin/pulsar zookeeper-shell
1s /ledgers/cookies
[10.1.32.151:3181, 10.1.32.152:3181, 10.1.32.153:3181]
get /ledgers/cookies/10.1.32.151:3181
bookieHost: "10.1.32.151:3181"
journalDir: "/data/pulsar/bookkeeper/journal"
ledgerDirs: "1\t/data/pulsar/bookkeeper/ledgers"
instanceId: "8dc37d76-5739-40d1-9ad8-ab4fe61f4b67"
get /ledgers/cookies/10.1.32.152:3181
4
bookieHost: "10.1.32.152:3181"
journalDir: "/data/pulsar/bookkeeper/journal"
ledgerDirs: "1\t/data/pulsar/bookkeeper/ledgers"
instanceId: "8dc37d76-5739-40d1-9ad8-ab4fe61f4b67"
get /ledgers/cookies/10.1.32.153:3181
bookieHost: "10.1.32.153:3181"
journalDir: "/data/pulsar/bookkeeper/journal"
ledgerDirs: "1\t/data/pulsar/bookkeeper/ledgers"
instanceId: "8dc37d76-5739-40d1-9ad8-ab4fe61f4b67"
```

zookeeper中查询到bookie的 instanceld保持一致说明集群三个bookie正常

更新前制造一些数据:

```
bin/pulsar-admin brokers list zhoufr_pulsar_test
bin/pulsar-admin tenants create zhoufr
bin/pulsar-admin namespaces create zhoufr/pulsar
bin/pulsar-admin topics create-partitioned-topic zhoufr/pulsar/test-topic -p 3
bin/pulsar-admin namespaces set-persistence zhoufr/pulsar -a 2 -e 3 -w 2 -r 0
bin/pulsar-admin namespaces get-persistence zhoufr/pulsar
bin/pulsar-client produce persistent://zhoufr/pulsar/test-topic -m "hellworld"
-n 1000
```

一、升级zookeeper

将新包解压至目标目录,检查配置,关闭老的进程,启动新的进程:

```
[root@zhoufr-pulsar-node01 scripts]# ./stop_zookeeper.sh
doing stop zookeeper ...
stopping zookeeper
Shutdown is in progress... Please wait...
Shutdown completed.
...
[root@zhoufr-pulsar-node01 scripts]# ./start_zookeeper.sh
doing start zookeeper ...
starting zookeeper, logging to /opt/Apps/pulsar280/logs/pulsar-zookeeper-zhoufr-pulsar-node01.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is switched off due to performance considerations.
```

注:另外两台进行相同的动作,但是不能同时操作,防止脑裂

二、升级bookeeper

首先, 停止老的进程

```
[root@zhoufr-pulsar-node01 pulsar]# /opt/Apps/pulsar/scripts/stop_bookie.sh doing stop bookie ... stopping bookie
Shutdown is in progress... Please wait...
Shutdown completed.
```

然后以只读的形式启动新版:

```
bin/pulsar bookie --readOnly
```

如果成功启动

```
[root@zhoufr-pulsar-node01 pulsar280]# scripts/start_bookie.sh doing start bookie ... starting bookie, logging to /opt/Apps/pulsar280/logs/pulsar-bookie-zhoufr-pulsar-node01.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is switched off due to performance considerations.
```

另外两台做相同的操作, 不要同时操作, 防止集群脑裂

```
[root@zhoufr-pulsar-node02 pulsar280]# /opt/Apps/pulsar/scripts/stop_bookie.sh
doing stop bookie ...
stopping bookie
Shutdown is in progress... Please wait...
Shutdown completed.
[root@zhoufr-pulsar-node02 pulsar280]# pwd
/opt/Apps/pulsar280
[root@zhoufr-pulsar-node02 pulsar280]# bin/pulsar bookie --readOnly
```

注意日志显示成功启动后方可进行broker的操作

三、升级broker

broker属于内存类型, 可以直接关闭和开启

```
[root@zhoufr-pulsar-node01 pulsar280]# /opt/Apps/pulsar/scripts/stop_broker.sh doing stop broker ... stopping broker
Shutdown is in progress... Please wait...
Shutdown completed.
[root@zhoufr-pulsar-node01 pulsar280]# /opt/Apps/pulsar280/scripts/start_broker.sh doing start broker ... starting broker, logging to /opt/Apps/pulsar280/logs/pulsar-broker-zhoufr-pulsar-node01.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is switched off due to performance considerations.
```

四、检测

```
[root@zhoufr-pulsar-node01 pulsar280]# bin/pulsar-admin broker-stats load-
report
{
    "webServiceUrl" : "http://10.1.32.151:8080",
    "pulsarServiceUrl" : "pulsar://10.1.32.151:6650",
    "persistentTopicsEnabled" : true,
    "nonPersistentTopicsEnabled" : true,
    "cpu" : {
        "usage" : 28.87490539166667,
    }
}
```

```
"limit" : 400.0
},
"memory" : {
 "usage": 501.1755065917969,
 "limit" : 2048.0
},
"directMemory" : {
 "usage" : 128.0,
 "limit" : 4096.0
},
"bandwidthIn" : {
 "usage": 40.49401041666667,
 "limit" : 1.024E7
},
"bandwidthOut" : {
 "usage" : 51.95390625,
 "limit" : 1.024E7
},
"msgThroughputIn" : 0.0,
"msgThroughputOut" : 0.0,
"msgRateIn" : 0.0,
"msgRateOut" : 0.0,
"lastUpdate" : 1662631601242,
"lastStats" : {
  "zhoufr/pulsar/0x80000000_0xc00000000" : {
    "msgRateIn" : 0.0,
    "msgThroughputIn" : 0.0,
    "msgRateOut" : 0.0,
    "msgThroughputOut" : 0.0,
    "consumerCount" : 0,
    "producerCount" : 0,
    "topics" : 1,
   "cacheSize": 0
 }
},
"numTopics" : 1,
"numBundles" : 1,
"numConsumers" : 0,
"numProducers" : 0,
"bundles" : [ "zhoufr/pulsar/0x80000000_0xc00000000" ],
"lastBundleGains" : [],
"lastBundleLosses" : [ ],
"brokerVersionString": "2.8.0",
"protocols" : { },
"advertisedListeners" : { },
"loadReportType" : "LocalBrokerData",
"maxResourceUsage" : 0.24471460282802582,
"bundleStats" : {
  "zhoufr/pulsar/0x80000000_0xc0000000" : {
    "msgRateIn" : 0.0,
    "msgThroughputIn" : 0.0,
    "msgRateOut" : 0.0,
    "msgThroughputOut" : 0.0,
    "consumerCount" : 0,
    "producerCount" : 0,
```

```
"topics" : 1,
    "cacheSize" : 0
}
}
```

五、回退操作

回退操作之前看了一下pulsar的原理图,发现broker和bookkeeper元数据保存在zk当中,然后本身bookkeeper也存储ledger数据,保证数据的一致性很重要,单纯关闭进程,然后拷回数据目录(假设有备份),再启动,这样一台一台的操作存在风险,当其他节点有写入新数据,那数据一致性就会存在问题(我们的ledger的配置是: -a 2 -e 3 -w 2)。想来想去,还是按照升级的方式逆向操作会安全一些。

- Ensemble Size (E) 决定了Pulsar写入Ledger可用的Bookies池的大小
- Write Quorum Size (Qw) Pulsar将要写入的实际的Bookies数量
- Ack Quorum Size (Qa) 是确认写入Bookies的数量

为了更好的验证回退,我们将2.8.0再做一次升级至2.8.2,然后我再进行性能测试,在性能测试过程中进行回退,来模拟在业务没有中断的情况,了解其影响程度。

```
"msgThroughputOut" : 31307.3588674461,
"consumerCount" : 28,
"producerCount" : 28,
"cacheSize" : 0
},
"numTopics" : 57,
"numBundles" : 2,
"numConsumers" : 56,
"numProducers" : [ */2.001/pulsar/0x4000000_0x80000000", "zhoufr/pulsar/0x0000000_0x40000000"],
"lastBundleGains" : [ ],
"lastBundleGains" : [ ],
"lastBundleGains" : [ 2.8.2",
"producelsage" : [ 2.8.2",
"producers" : [ 3.944109380245209,
"msgThroughputTor : 0.3944109380245209,
"bundleStats" : [
"msgRateIn" : 28.883561084301846,
"msgThroughputTor" : 31424.28362609384,
"msgThroughputTor" : 31424.28362609384,
"consumerCount" : 28,
"producerCount" : 28,
"producerCount" : 29,
"cacheSize" : 0
},
"numCount : 29,
"cacheSize" : 0
},
"numCount : 31307.35891682774,
"msgThroughputIn" : 31307.35891682774,
"msgThroughputOut" : 31307.35891682774,
"msgThroughputOut" : 31307.3588674461,
"consumerCount" : 28,
"producerCount" : 28,
"msgThroughputOut" : 31307.3588674461,
"consumerCount" : 28,
"producerCount" : 28,
"produce
```

生产:

bin/pulsar-perf produce -t 100 persistent://zhoufr/pulsar/test-topic

```
15:42:46.567 [main] INFO org.apache.pulsar.testclient.PerformanceProducer - Throughput produced: 999tt: 9.685 - 99.9pct: 14.675 - 99.99pct: 15.08 - 99.9pct: 14.675 - 99.99pct: 15.08 - 99.9pct: 16.08 - 99.9pct: 15.08 - 99.9pct:
```

消费:

bin/pulsar-perf consume -t 100 persistent://zhoufr/pulsar/test-topic

```
15:42:52.395 [main] IMFO org apache.pulsar.testclient.PerformanceConsumer - Throughput received:

Latency: mean: 0.000 ms - med: 0 - 95pct: 0 - 99pct: 0 - 99.9pct: 0 - 99.9pc
```

回退broker

下面,我们开始回退broker:

```
[root@zhoufr-pulsar-node01 pulsar282]#
/opt/Apps/pulsar282/scripts/stop_broker.sh
[root@zhoufr-pulsar-node01 pulsar282]#
/opt/Apps/pulsar280/scripts/start_broker.sh
doing start broker ...
starting broker, logging to /opt/Apps/pulsar280/logs/pulsar-broker-zhoufr-pulsar-node02.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is switched off due to performance considerations.
```

```
Creating producer on cnx [id: 0x52359fee, L:/10.1.32.151:40916 - R:10.1.32.153/10.1.32.153:6650]
387 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-16] [zhoufr_pulsar_
Created producer on cnx [id: 0x52359fee, L://10.1.32.151:40916 - R:10.1.32.153/10.1.32.153:6650]
388 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-16] [zhoufr_pulsar_
Re-Sending 98 messages to server
689 [pulsar-timer-6-1] INFO org.apache.pulsar.client.impl.ConnectionHandler - [persistent://zhoufr/pulsar/test-topic-43] [zhoufr_pulsar_
                         set-6-6] Re-Sending 98 messages to server
5:48:55.689 [pulsar-timer-6-1] INFO org.apache.pulsar.client.impl.ConnectionHandler - [persistent://zhoufr/pulsar/test-topic-43] [zhoufr_pulsar_test-6-8] Reconnecting after connection was closed
5:48:55.692 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-43] [zhoufr_pulsar_test-6-8] Creating producer on cnx [id: 0x61227961, L:/10.1.32.151:40912 - R:10.1.32.153/10.1.32.153:6650]
5:48:55.697 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-43] [zhoufr_pulsar_test-6-8] Created producer on cnx [id: 0x61227961, L:/10.1.32.151:40912 - R:10.1.32.153/10.1.32.153:6650]
5:48:55.698 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-43] [zhoufr_pulsar_test-6-8] Re-Sending 99 messages to server
5:48:55.6061 [pulsar-timer-6-1] INFO org.apache.pulsar.client.impl.ConnectionHandler - [persistent://zhoufr/pulsar/test-topic-98] [zhoufr_pulsar_test-6-10] Reconnecting after connection was closed
5:48:55.6064 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-98] [zhoufr_pulsar_test-6-10] Creating producer on cnx [id: 0xdbf250c3, L:/10.1.32.151:40936 - R:10.1.32.153/10.1.32.153:6650]
5:48:55.6069 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-98] [zhoufr_pulsar_test-6-10] Created producer on cnx [id: 0xdbf250c3, L:/10.1.32.151:40936 - R:10.1.32.153/10.1.32.153:6650]
5:48:55.6069 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-98] [zhoufr_pulsar_test-6-10] Created producer on cnx [id: 0xdbf250c3, L:/10.1.32.151:40936 - R:10.1.32.153/10.1.32.153:6650]
5:48:55.6069 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-98] [zhoufr_pulsar_test-6-10] Created producer o
                         /s --- failure 0.0 msg/s --- Latency: mean: 3.936 ms - med. 5.000 sppt. 4.804 st.4804 st.49:27.712 [main] INFO org.apache.pulsar.testclient.PerformanceProducer - Throughput produced: 79927 msg --- 100.0 msg/s --- 0.8 Mbi /s --- failure 0.0 msg/s --- Latency: mean: 4.063 ms - med: 3.889 - 95pct: 5.065 - 99pct: 9.179 - 99.9pct: 15.158 - 99.99pct: 15.68 - Max: 15.668 - Max: 15.668 st.49:37.736 [main] INFO org.apache.pulsar.testclient.PerformanceProducer - Throughput produced: 80930 msg --- 100.1 msg/s --- 0.8 Mbi /s --- failure 0.0 msg/s --- Latency: mean: 4.062 ms - med: 3.870 - 95pct: 5.243 - 99pct: 9.045 - 99.9pct: 13.994 - 99.99pct: 14.09 - Max: 14.409
15:52:56.129 [pulsar-client-io-1-1] INFO org.apache.pulsar.client.impl.ConsumerImpl - [persistent://zhoufr/pulsar/test-topic-71][sub] Subscribin g to topic on cnx [id: 0x7c119ea8] L:/10.1.32.152:46028 - R:10.1.32.151/10.1.32.151/6650] consumerId 50 [id: 0x7c119ea8] L:/10.1.32.151/10.1.32.151/6650] consumerId 50 [id: 0x7c119ea8] L:/10.1.32.151/6650] consumerId 50 [id: 0x7c119ea8] [id: 0x7c119ea8] L:/10.1.32.151/6650] consumerId 50 [id: 0x7c119ea8] [id: 0x7c119ea
```

此时我们再次检测pulsar的版本:

另外两台进行相同的操作,注意不要在同一时间操作

回退bookie

回退bookkeeper需要关注日志,并且也是滚动回退,防止集群脑裂, 同样会发现生产和消费的窗口会重启:

```
[root@zhoufr-pulsar-node01 pulsar282]# /opt/Apps/pulsar282/scripts/stop_bookie.sh
[root@zhoufr-pulsar-node01 pulsar282]#
/opt/Apps/pulsar280/scripts/start_bookie.sh
doing start bookie ...
starting bookie, logging to /opt/Apps/pulsar280/logs/pulsar-bookie-zhoufr-pulsar-
node01.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging
event is flushing to disk immediately. The default behavior is switched off due
to performance considerations.
```

当一个节点发生重置bookkeeper进程时, 日志和测试工具都有相应的错误输出, 如下截图所示

:50.910 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-16] [zhoufr_pulsa

```
106:50.910 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-16] [zhoufr_pulsar_sit-6-6] Re-Sending 20 messages to server
106:50.910 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-33] [zhoufr_pulsar_sit-8-17] Re-Sending 19 messages to server
106:50.911 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-93] [zhoufr_pulsar_sit-7-13] [rotated producer on cnx [id: 0x90deb763, i./10.1.32.151:52228 - R:10.1.32.151/10.1.32.151:6650]
106:50.912 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-93] [zhoufr_pulsar_sit-7-13] [rotated producer on cnx [id: 0x70deb763, i./10.1.32.151:52228 - R:10.1.32.151/10.1.32.151:6650]
106:50.912 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-39] [zhoufr_pulsar_sit-7-22] [rotated producer on cnx [id: 0x70ex]obsp. i./10.1.32.151:52232 - R:10.1.32.151/10.1.32.151/6050]
106:50.912 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-39] [zhoufr_pulsar_sit-7-22] [rotated producer on cnx [id: 0x70e4]014, l./10.1.32.151:52238 - R:10.1.32.151/10.1.32.151/6050]
106:50.913 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-97] [zhoufr_pulsar_sit-7-32] [rotated producer on cnx [id: 0x70e4]014, l./10.1.32.151:52238 - R:10.1.32.151/10.1.32.151/6050]
106:50.913 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-51] [zhoufr_pulsar_sit-7-32] [rotated producer on cnx [id: 0x70e4]014, l./10.1.32.151/52238 - R:10.1.32.151/10.1.32.151/01.32.151/001.790147]

106:50.913 [pulsar-client-io-2-1] INFO org.apache.pulsar.client.impl.ProducerImpl - [persistent://zhoufr/pulsar/test-topic-51] [zhoufr_pulsar_sit-7-32] [rotated_prod
```

三台重启之后,测试工具仍然可生产消费,说明进程切换正常

切换zookeeper

采用相同的方式关闭zookeeper, 并重新启动老的zookeeper, 注意: 不要同一时间内关闭和启动, 重置 zookeeper对性能测试工具无影响

```
[root@zhoufr-pulsar-node01 pulsar282]#
/opt/Apps/pulsar282/scripts/stop_zookeeper.sh
doing stop zookeeper ...
stopping zookeeper
Shutdown is in progress... Please wait...
Shutdown completed.
[root@zhoufr-pulsar-node01 pulsar282]#
/opt/Apps/pulsar280/scripts/start_zookeeper.sh
doing start zookeeper ...
starting zookeeper, logging to /opt/Apps/pulsar280/logs/pulsar-zookeeper-zhoufr-pulsar-node01.log
Note: Set immediateFlush to true in conf/log4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is switched off due to performance considerations.
```

完成后检查服务器进程如下:

```
Note: Set immediateFlush to true in conflog4j2.yaml will guarantee the logging event is flushing to disk immediately. The default behavior is sw inched off due to performance considerations.

[root@zhoufr-pulsar-node01 pulsar282]# ps -ef | grep java | grep -v grep
dominos+ 6295 | 13 15.47 pts/ 0 00:07:14 /usr/local/javaj/gkl.8.0 112/bin/java -Dlog4j.shutdownHookEnabled=false -cp /opt/Apps/pulsar280/conf::/opt/Apps/pulsar280/lib/: -Dlog4j.configurationFileelog4j2.yaml -Djute.maxbuffer=10485760 -Djava.net.preferIPv4Stack=true -Dpulsar.allocat or.exit on cometrue -Dio.netty.recycler.maxCapacity.default=1000 -Dio.netty.recycler.linKapacity=1024 -Xms2g -Xms
```