

# TGIP-CN 025

Pulsar Q&A

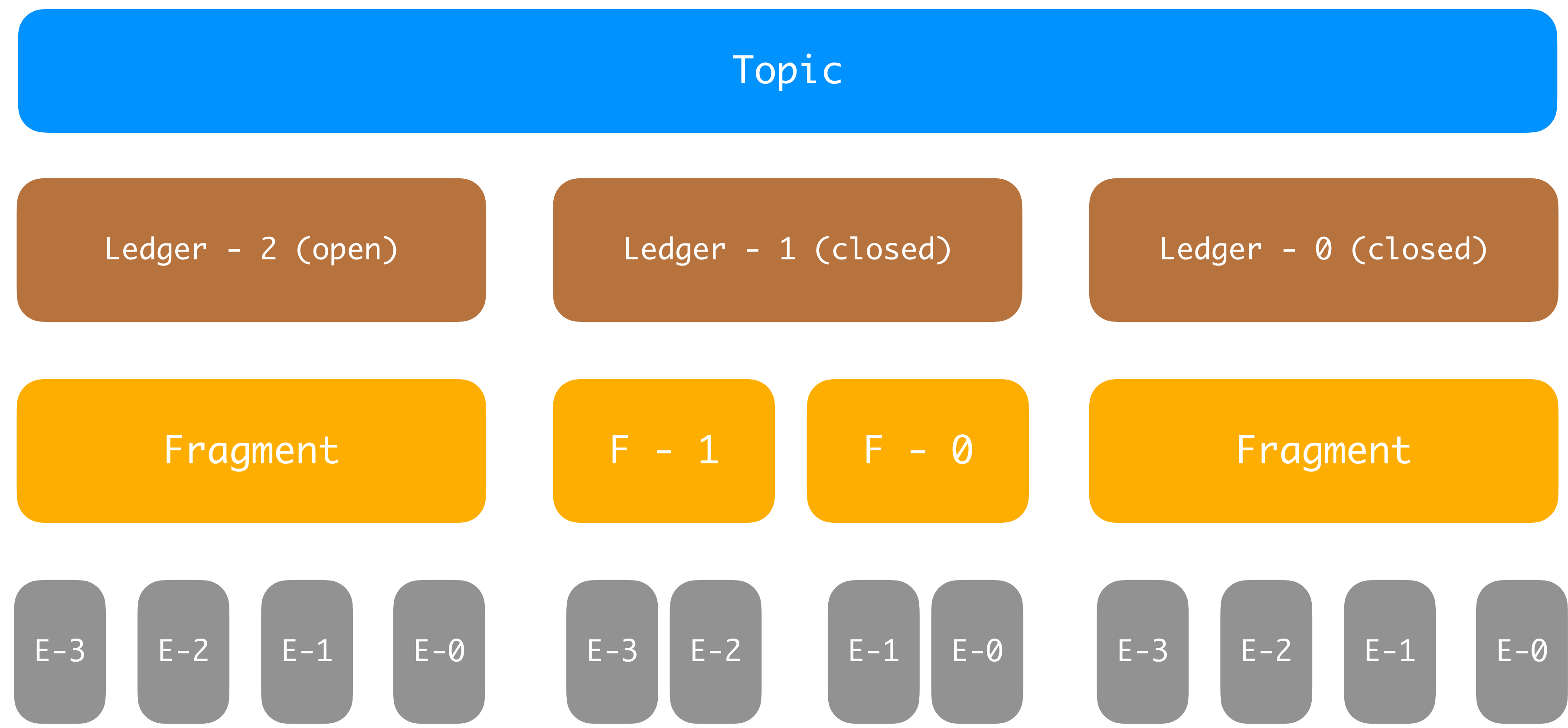
Penghui Li & Xiaolong Ran

## Features and bug fixes

- A. Add ability to specify the subscription init position for functions #7891
- B. Fix IndexOutOfBoundsException in the Key\_Shared subscription when dispatch batch message to consumers #8024
- C. Support set max consumers per subscription on the topic level #8003
- D. Support set max consumers on the topic level #7968
- E. Support set message key in pulsar-perf #7989
- F. Support key based message batching in the CPP Client #7996
- G. Transaction message ack abort on subscriptions #7979



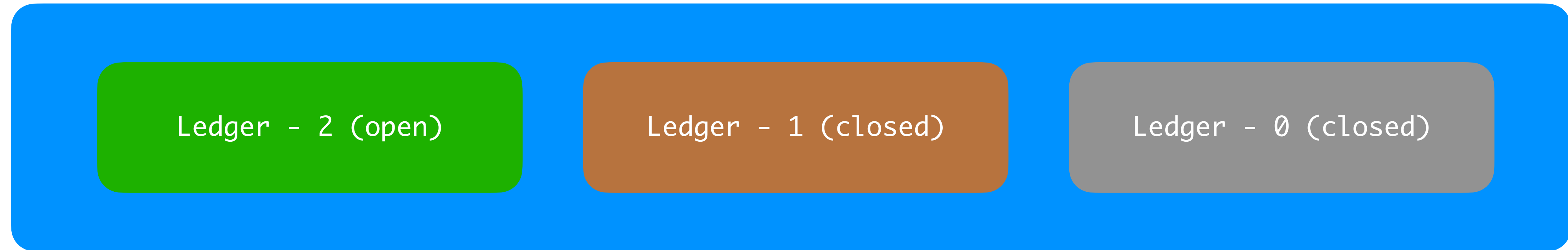
# Data deletion in Pulsar



# Data deletion in Pulsar

Current active in the topic

Removed from the topic



Not removed from the topic



# Data deletion in Pulsar

```
# Max number of entries to append to a ledger before triggering a rollover
# A ledger rollover is triggered on these conditions
# * Either the max rollover time has been reached
# * or max entries have been written to the ledger and at least min-time
#   has passed
managedLedgerMaxEntriesPerLedger=50000

# Minimum time between ledger rollover for a topic
managedLedgerMinLedgerRolloverTimeMinutes=10

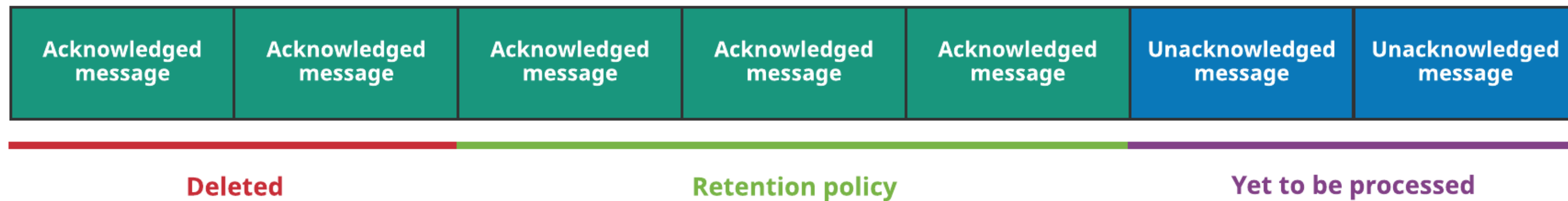
# Maximum time before forcing a ledger rollover for a topic
managedLedgerMaxLedgerRolloverTimeMinutes=240

# Maximum ledger size before triggering a rollover for a topic (MB)
managedLedgerMaxSizePerLedgerMbytes=2048
```

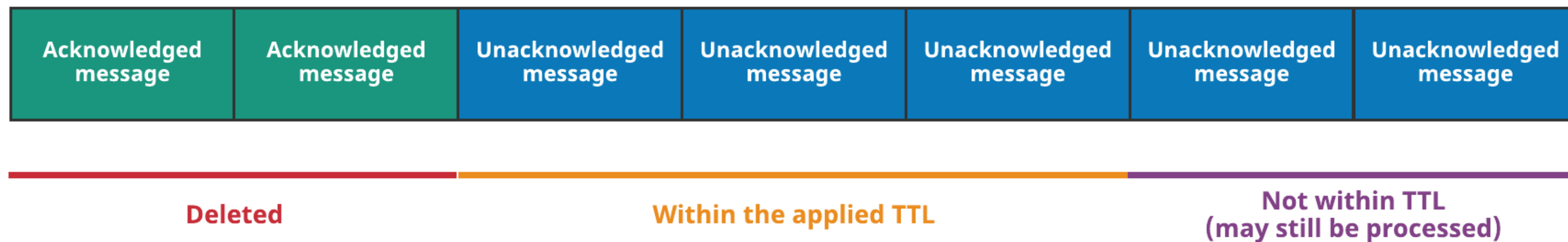


# Data deletion in Pulsar

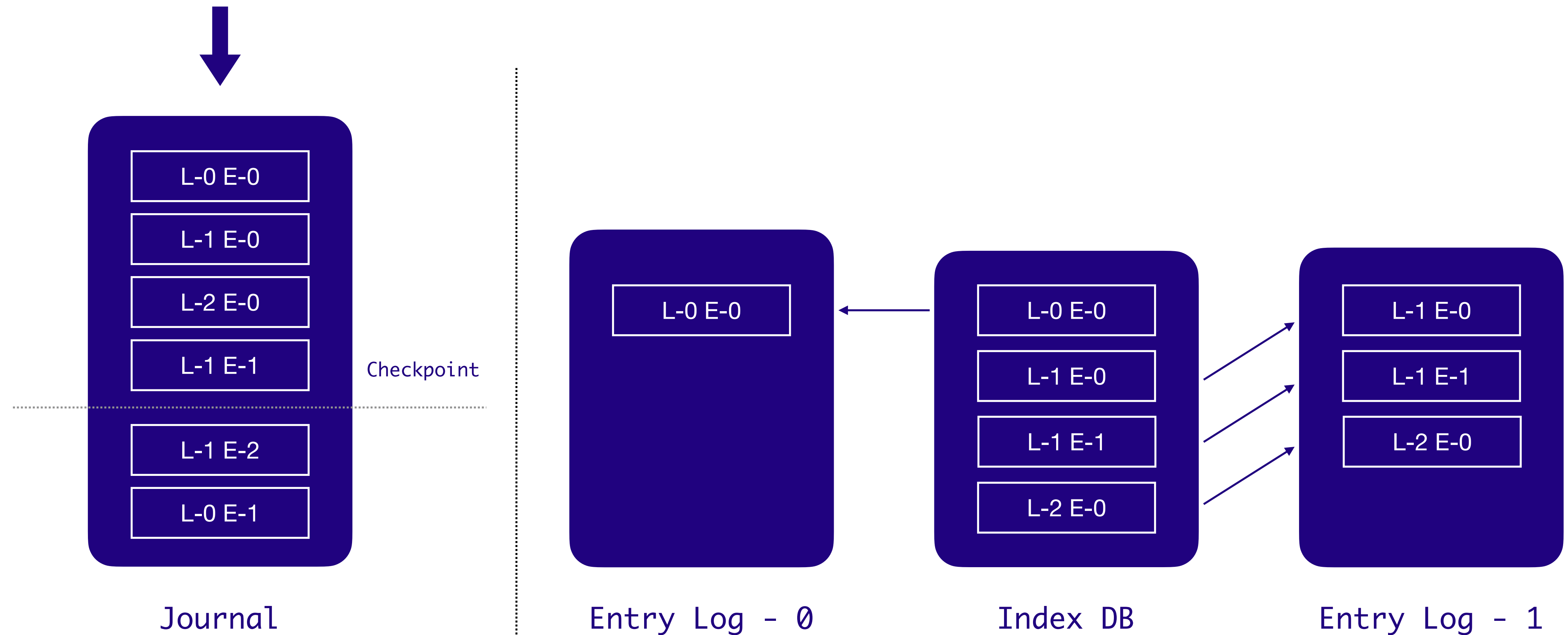
## Message retention



## Message expiry



# Data deletion in Pulsar



# Data deletion in Pulsar

`minorCompactionThreshold=0.2`

`majorCompactionThreshold=0.5`

`minorCompactionInterval=3600`

`majorCompactionInterval=86400`





# Question 1

最近刚刚接触pulsar项目，之前一直关注confluence的kafka项目，对pulsar很感兴趣。用pulsar对数据做永久保存怎么样，是否是个好主意？我看到对接presto时候是直接去bookie里面取历史数据，那么当做批处理的存储层也可以吧，会对broker的消息吞吐产生影响么？



## Question 2

最请问有没有pulsar性能压测的指导，不知道咋测试pulsar的性能，求大佬指导指导，谢谢！

<https://github.com/apache/pulsar/tree/master/pulsar-testclient>

<https://github.com/openmessaging/openmessaging-benchmark>



# Question 3

通过bookkeeper shell metaformat 操作后，是不是整个bookie集群的元数据都格式化了？  
而不是单个bookie节点。然后broker消费不了数据（生产没问题），异常ERROR  
org.apache.pulsar.broker.service.persistent.PersistentDispatcherMultipleConsumers - [persistent://public/default/topic\_aep\_order /  
aep\_order\_consumer\_group] Error reading entries at 110260:190 : No such  
ledger exists, Read Type Normal - Retrying to read in 55.281 seconds。请问  
这种情况怎么快速恢复？ topic比较多



# Question 4

Pulsar 如何保证 Bookie 节点的负载均衡呢 比如原来有三个 Bookie 节点，现新加了一个 bookie节点，在写入数据时会提高新加 Bookie 节点的选中率来实现均衡吗

```
# Enable/disable disk weight based placement. Default is false  
bookkeeperDiskWeightBasedPlacementEnabled=false
```



# Question 5

要验证pulsar支持百万级topic数量，请问各位有什么建议吗？需要多少个租户/名称空间？创建百万topic后，如何判断集群的稳定性？



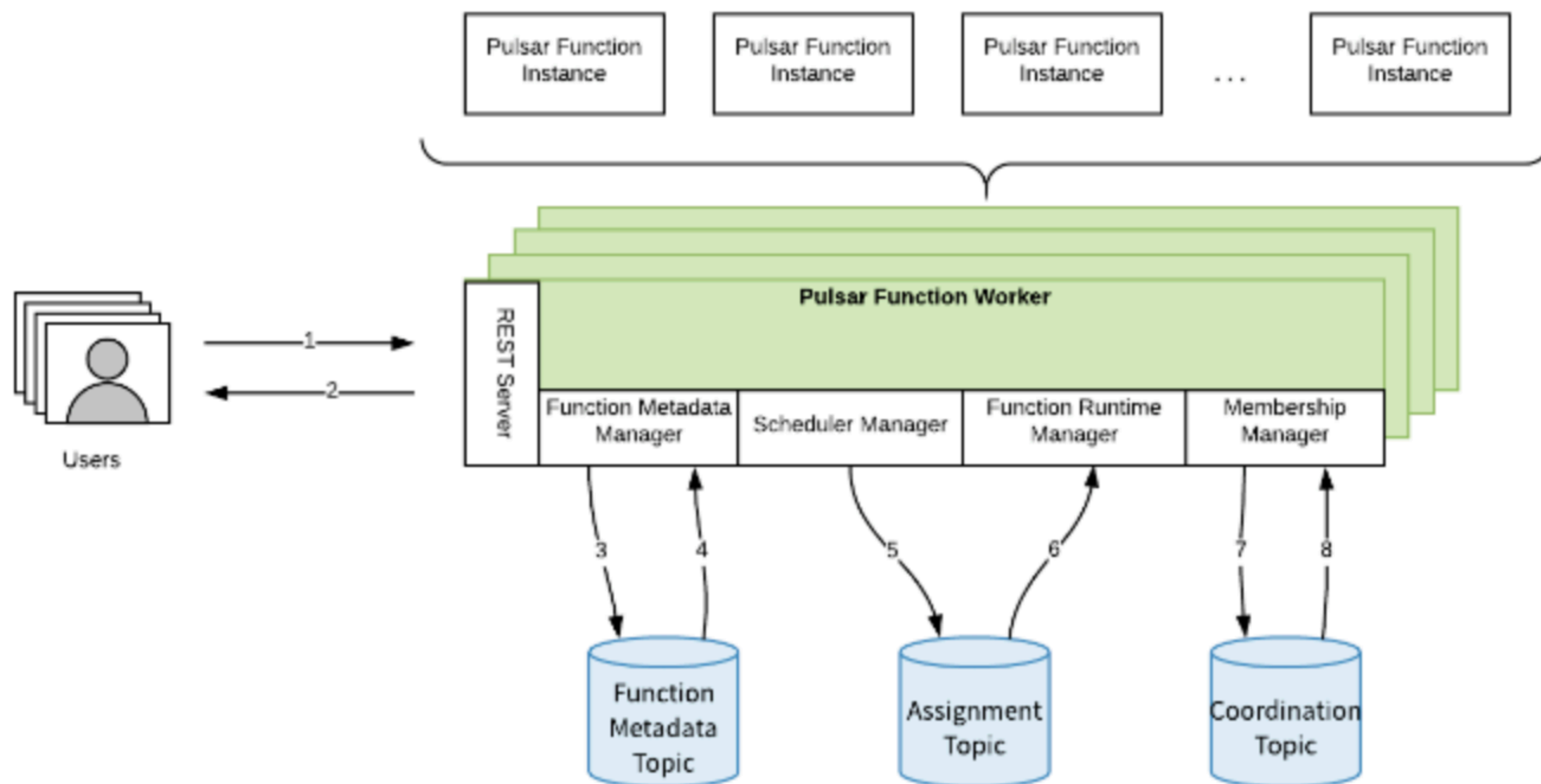
# Question 6

大家有尝试过在function中动态创建topic，并根据不同参数指定不同分区数的经验么？除了用pulsar-admin，还有其他便利的方式么？现在在function中有啥姿势可以尝试？之前我再issue上提了，但回复是用pulsar-admin接口。觉得有点重，不知道有没有更便捷的方式？



# Question 7

请问下，新增了几个functionWorker 节点，怎么让任务做一次 rebalance ？



<https://github.com/apache/pulsar/pull/7388>

<https://github.com/apache/pulsar/pull/7449>





扫码入群！

Pulsar 中国社区等你加入

添加好友后，回复“进群”即可





# Thanks