

HBase Client and Replication

目录

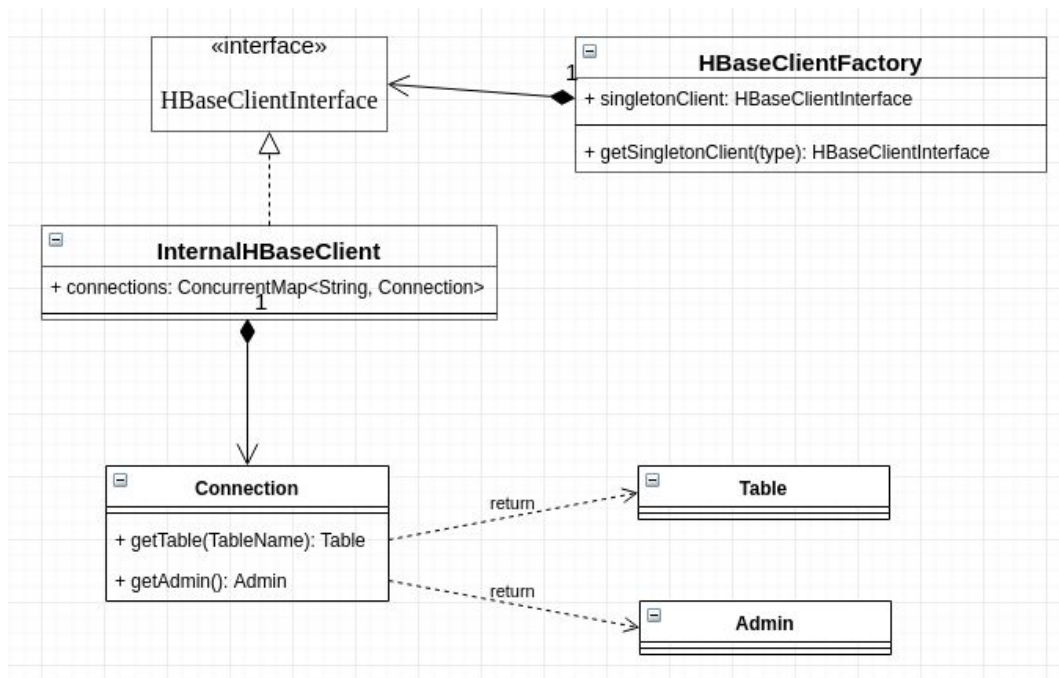
- Client
- Replication

Client

- API, RPC(基本)
- 在Client实现Region Locate
- 在Client实现Replica Read功能等Feature

Client - API

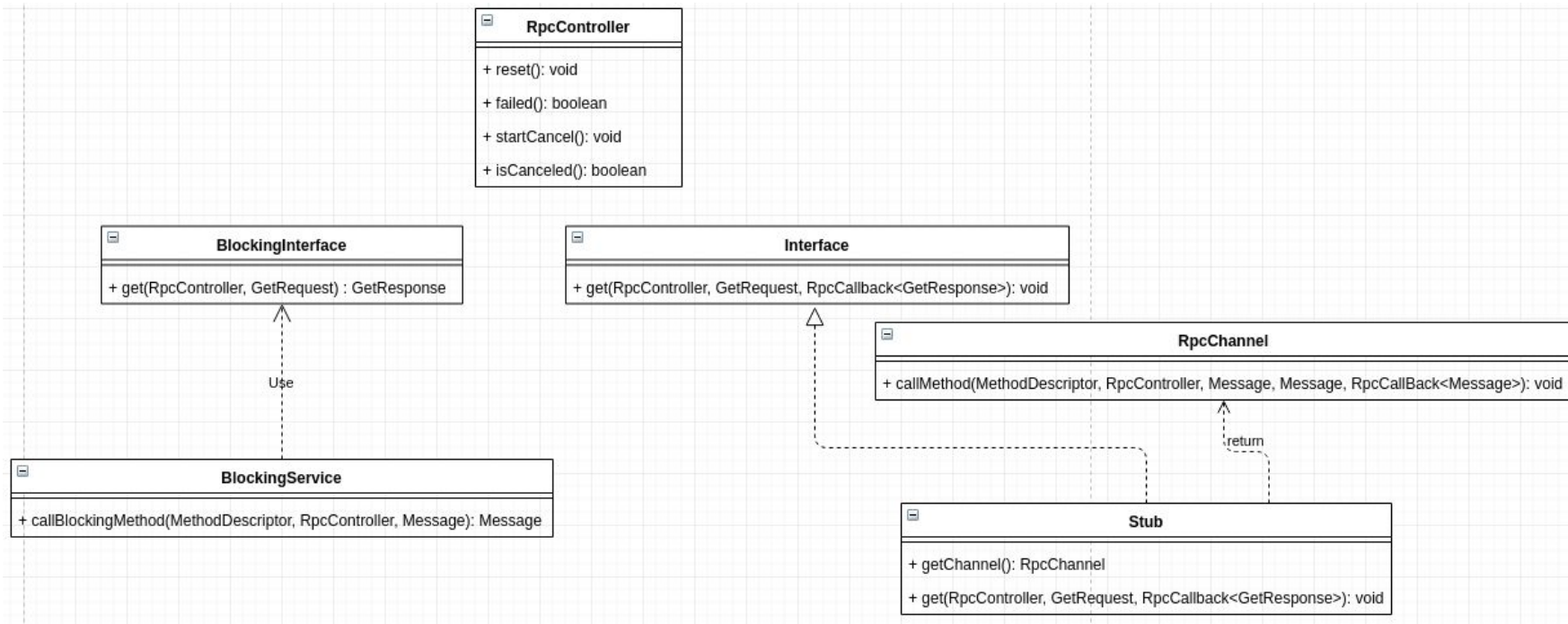
- Connection
 - heavy-weight
 - thread-safe
- Admin
- Table



Client - RPC

```
message GetRequest {  
  
}  
  
message GetResponse {  
  
}  
  
service ClientService {  
    rpc Get(GetRequest) returns(GetResponse);  
}
```

Client - RPC



Client - Region Locate

- Meta Region Locate

存储在ZK, Client直接从ZK读取

默认结点: {zookeeper.znode.parent}/meta-region-server-{replicaId}

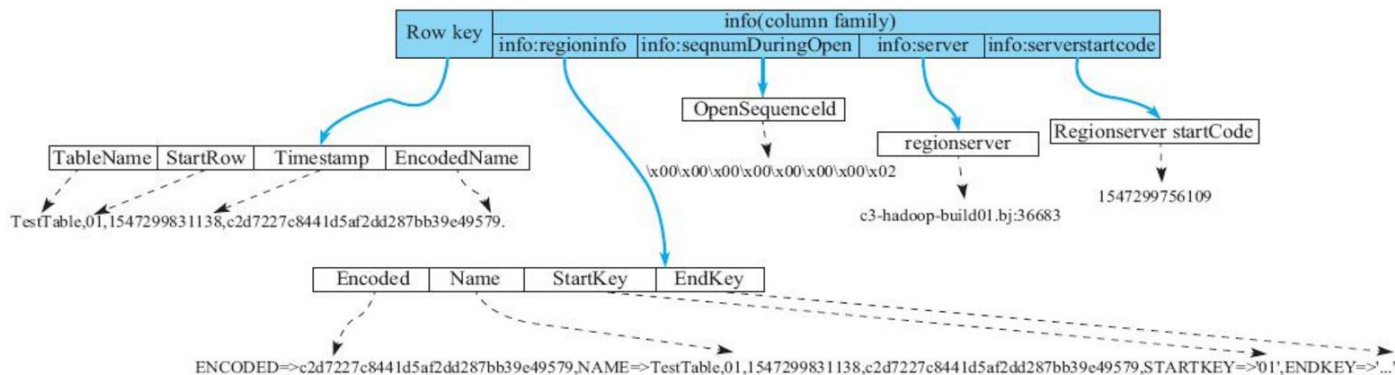
例如: /hbase/c4tst-sx8-1/meta-region-server

/hbase/c4tst-sx8-1/meta-region-server-1

Client - Region Locate

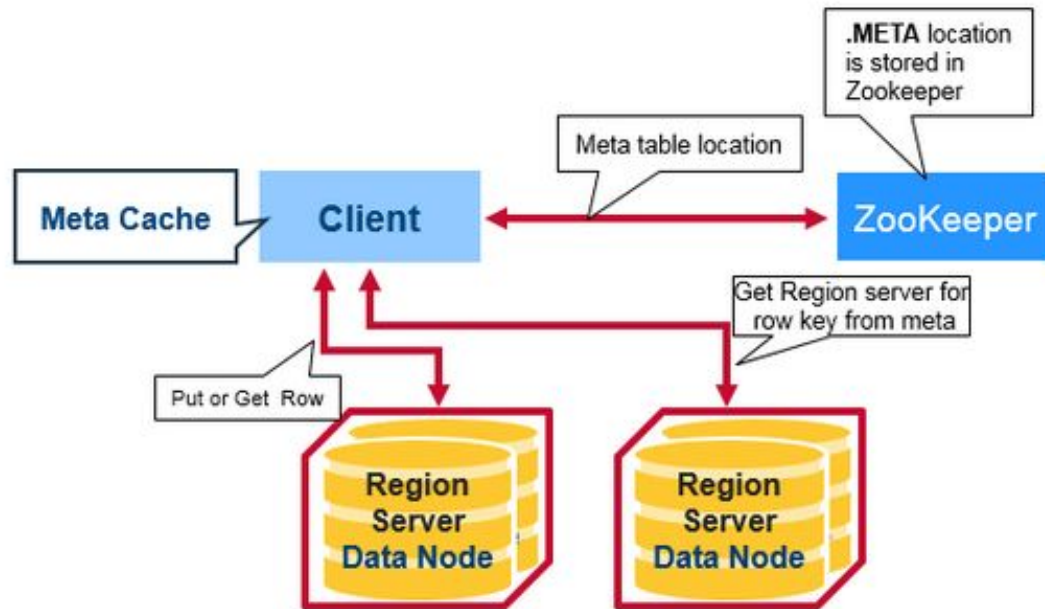
- Non Meta Region Locate

存储在 hbase:meta 表

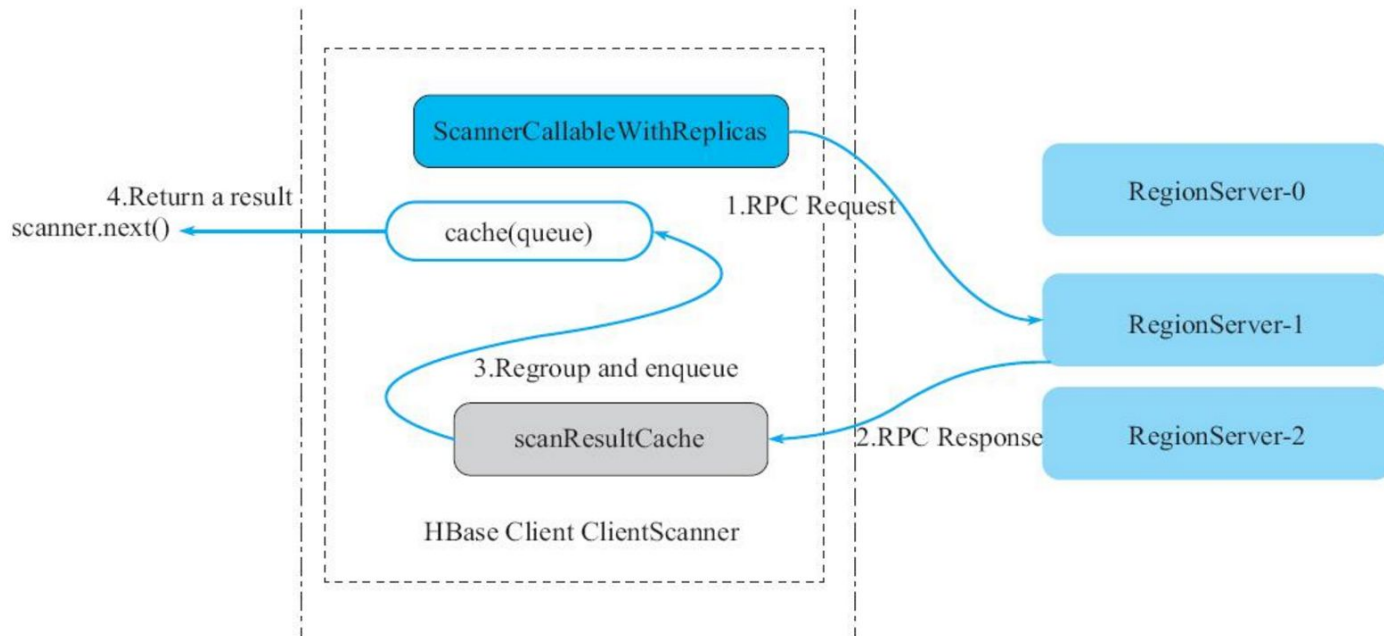


```
scan 'hbase:meta', {STARTROW => 'table,user1111,99999999999999', REVERSED => true, LIMIT => 1}
```


Client - Region Locate



Client - Region Locate

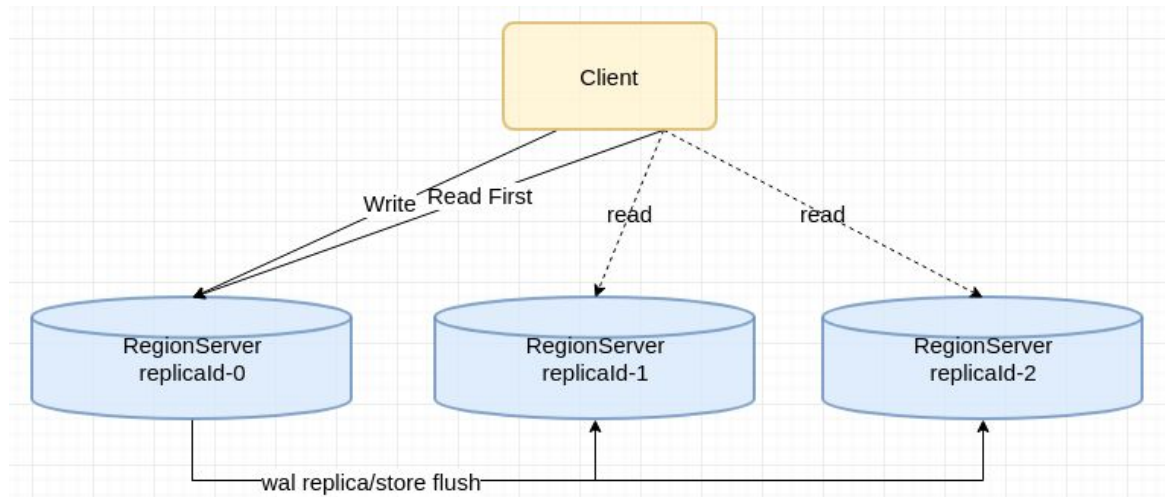


client scan 流程

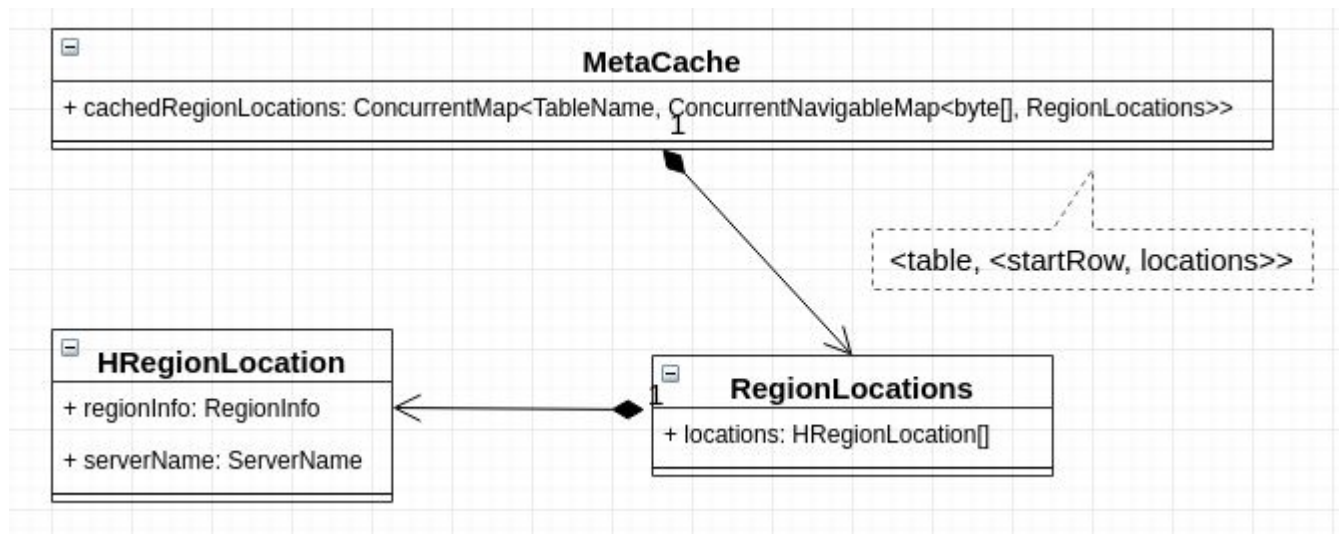
Client - Read Replicas

建表 : `create 'test', 'info', {REGION_REPLICATION => 3}`

查询 : `new Get(get).setConsistency(Consistency.TIMELINE)`



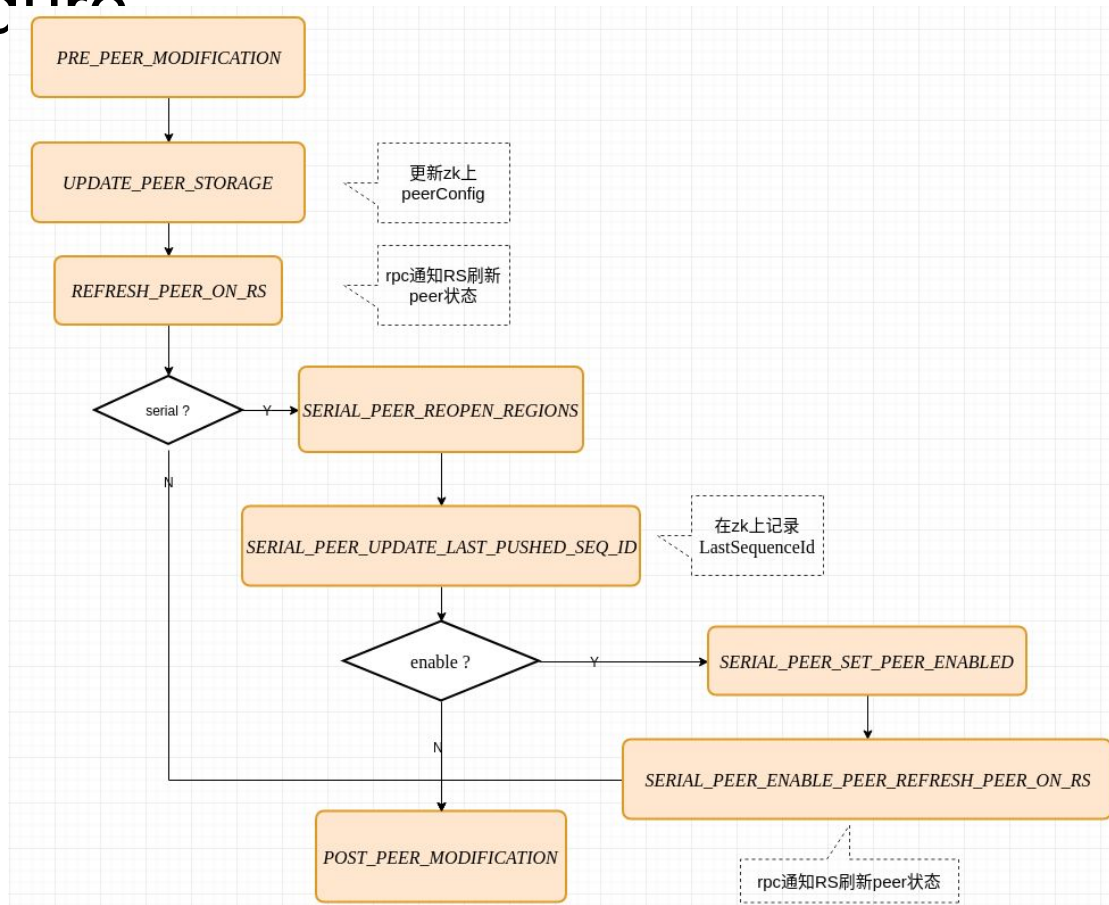
Client - Read Replicas



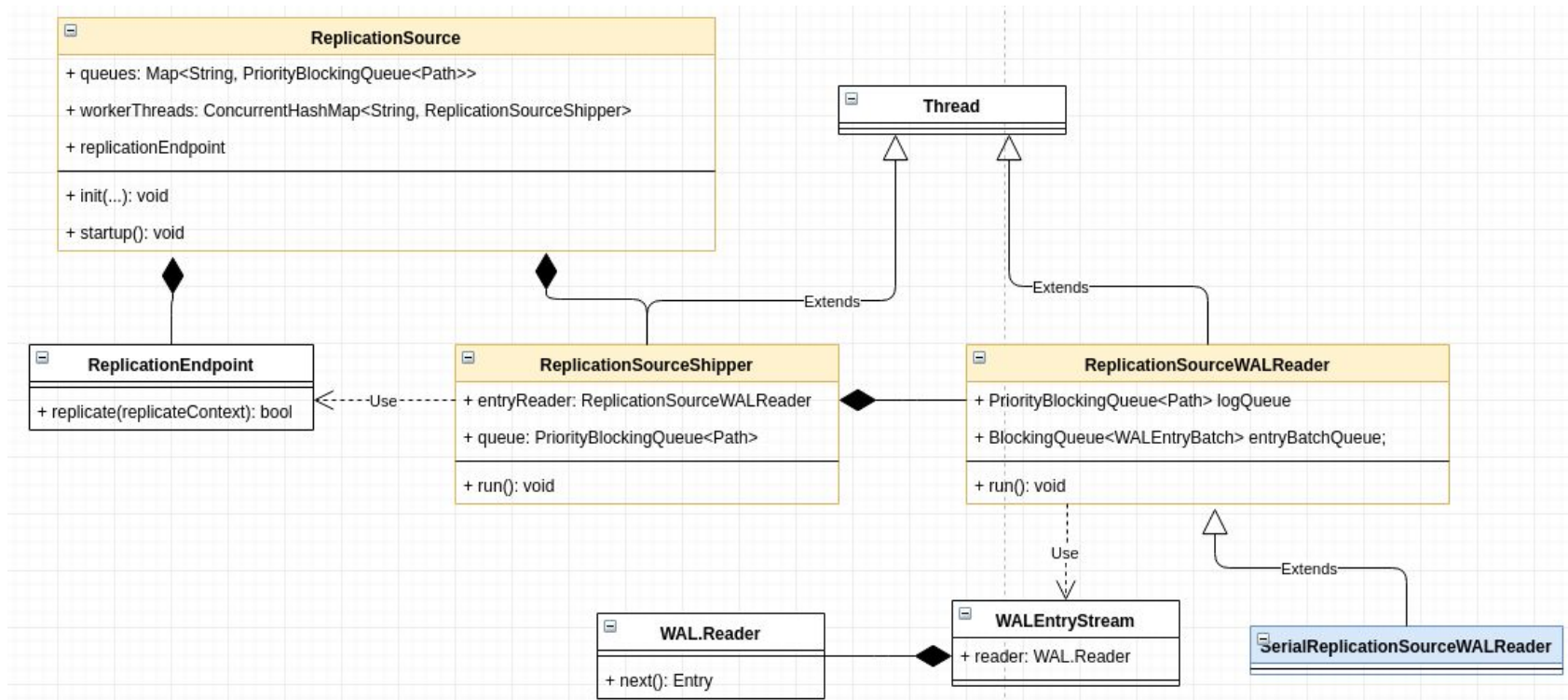
支持read replicas的MetaCache实现

Replication - Procedure

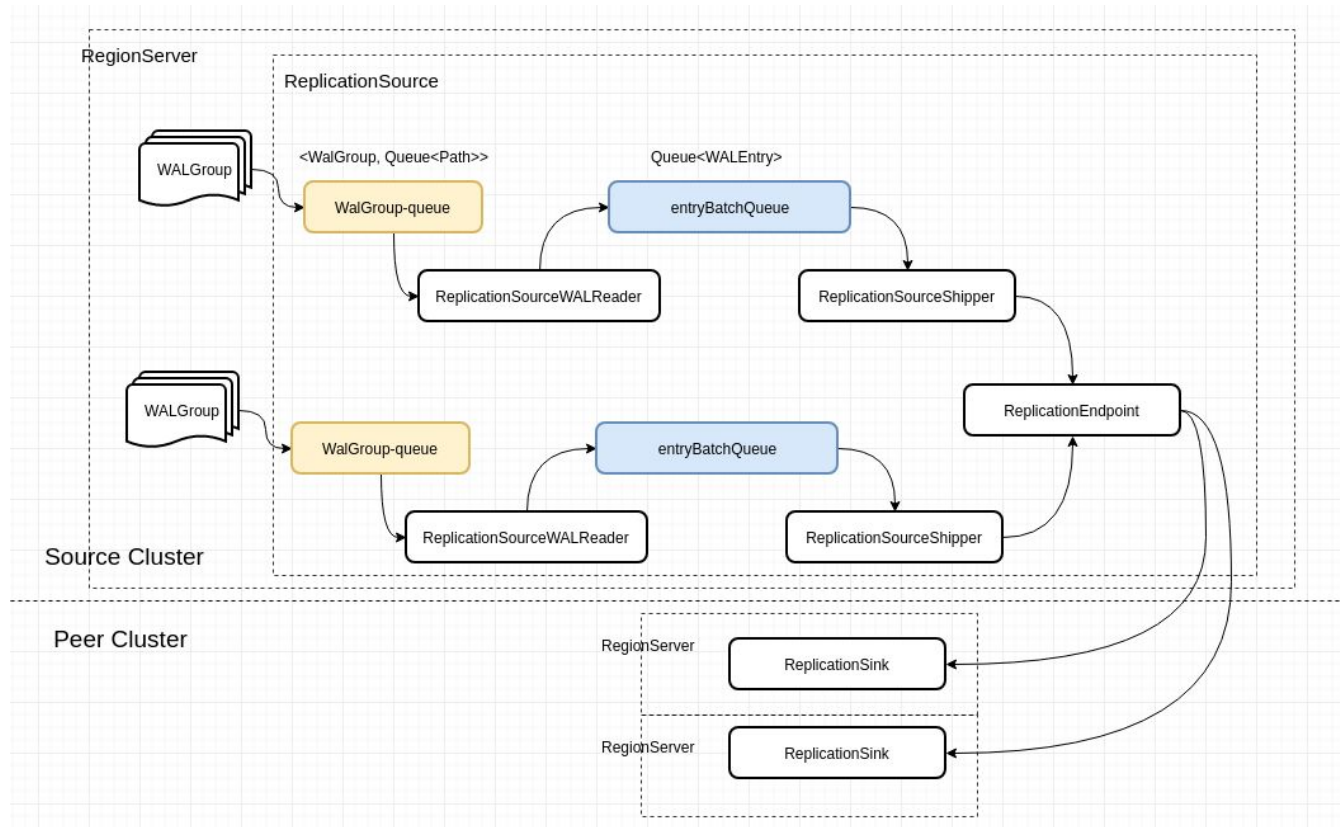
- 复杂状态机
- 拆分子任务
- 持久化在HDFS上
- 可重试、可回滚
- 要求操作幂等性



Replication

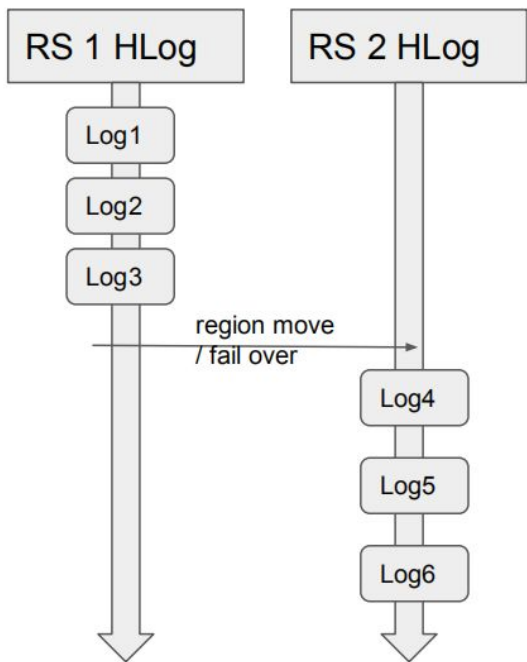


Replication - Async Replication

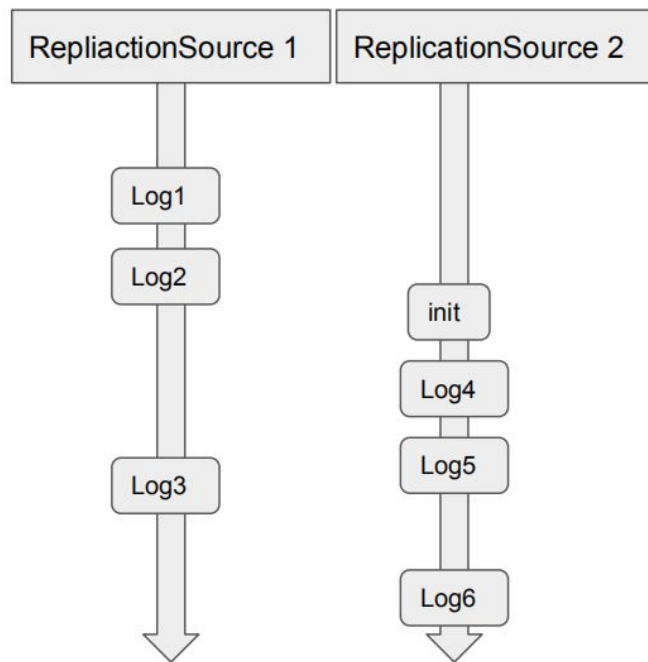


Replication - Serial Replication

Before HBASE-9465(<= 1.3.x):

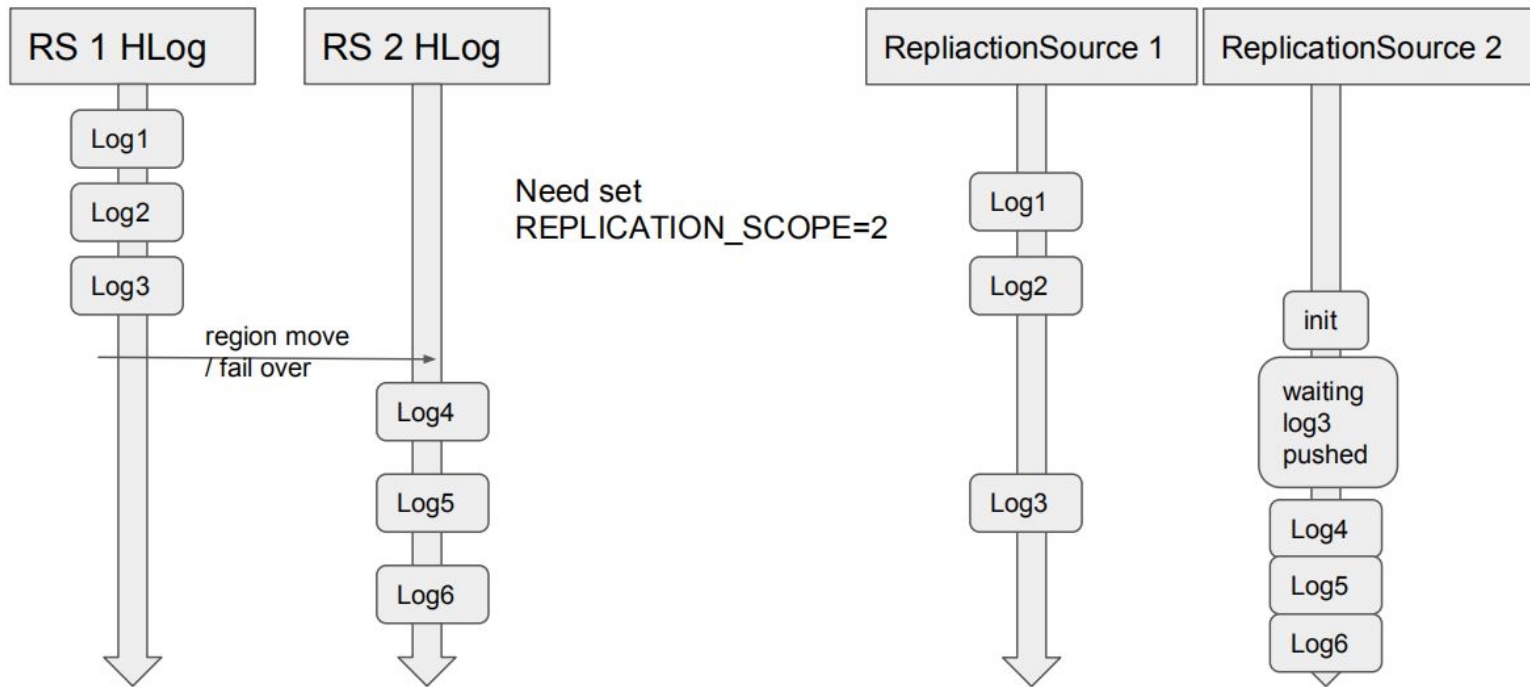


Log4/5 is pushed before log3

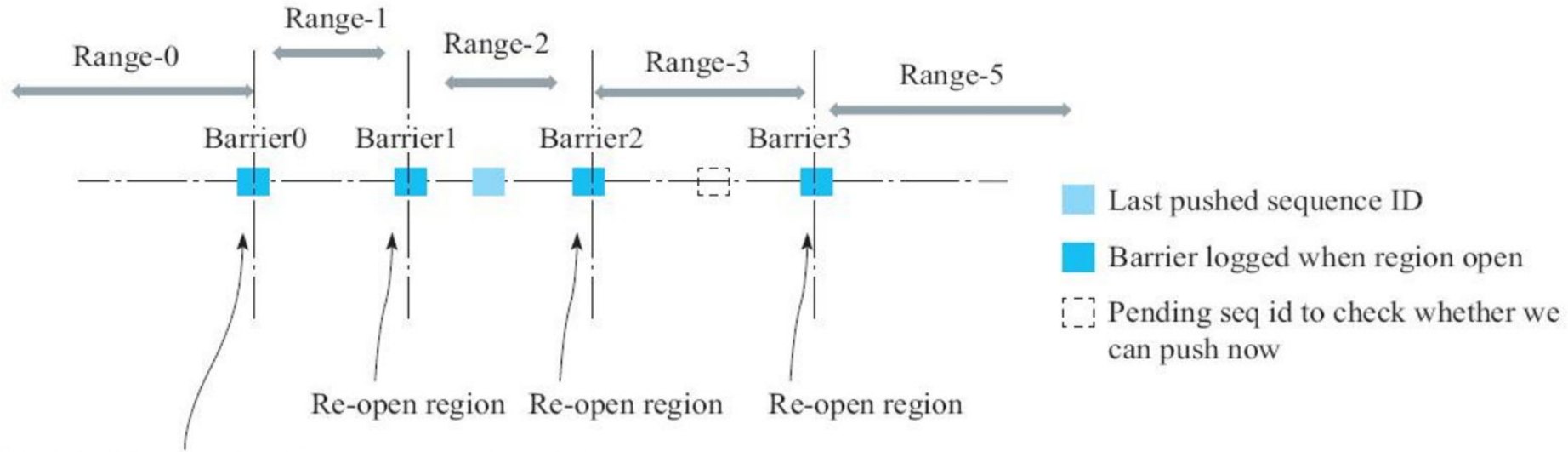


Replication - Serial Replication

After HBASE-9465(>= 1.4.0/2.0.0):

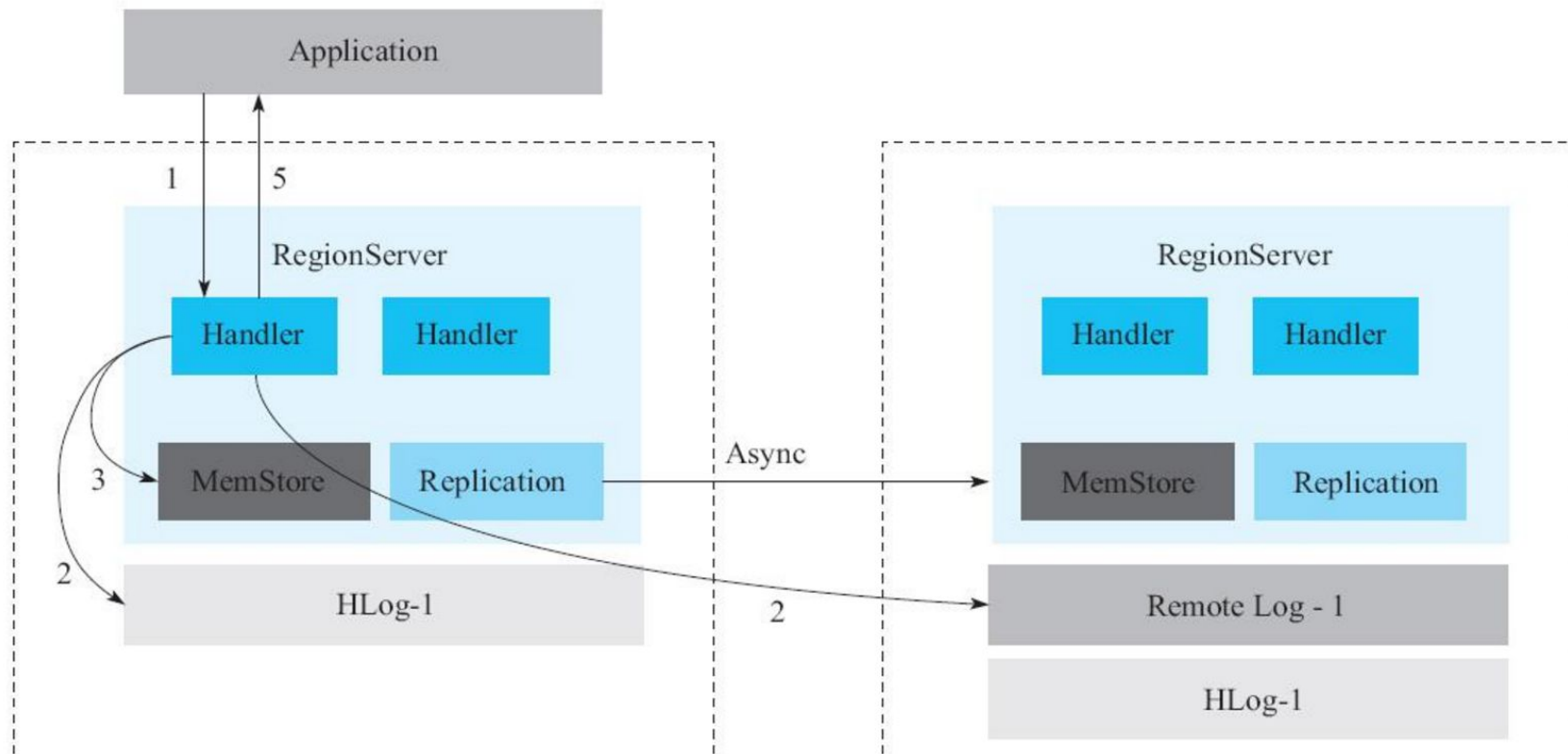


Replication - Serial Replication



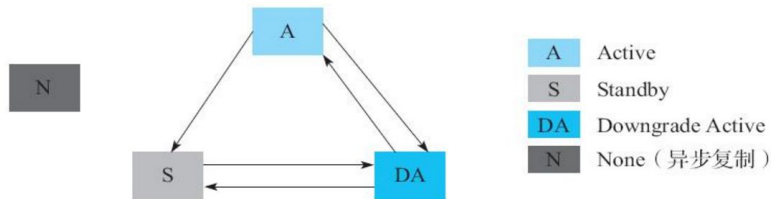
The initialized barrier when set peer to be serial.

Replication - Sync Replication



Replication - Sync Replication

集群同步复制的四种状态



- Active: 这种状态的集群将在远程集群上写RemoteWAL日志, 同时拒绝来自其他集群的复制数据。
- Downgrade Active: 这种状态的集群将跳过写RemoteWAL流程, 同时拒绝来自其他集群的复制数据。
- Standby: 这种状态的集群不容许Peer内的表被客户端读写, 只接收来自其他集群的复制数据。
- None: 表示没有开启同步复制。

对比项	Active	Downgrade Active	Standby	None
是否写 RemoteWAL	是	否	否	否
是否容许客户端读写集群	是	是	否	是
是否接收异步复制请求	否	否	是	是
是否能复制 Entry 到其他集群	是	是	否	是

Replication

对比项	异步复制	同步复制
读路径	无影响	无影响
写路径	无影响	多写一份RemoteWAL
网络带宽	1倍带宽	2倍带宽
存储空间	无需额外空间	RemoteWAL多占用1备WAL的存储空间
最终一致性	无法保证	总能保证
可用性	主集群故障，业务不可用	主集群故障，备集群回放RemoteWAL便可提供服务，可用性更高
运维复杂度	运维操作简单	操作复杂，需要手动切换主备集群

Thanks