

1、内部执行流程

1. 客户端发送请求到任意一个 node，成为 coordinating node
2. coordinating node 对 document 进行路由，将请求转发到对应的 node，此时会使用 round-robin 随机轮询算法，在 primary shard 以及将所有的 replica 中随机选择一个，让读请求负载均衡
3. 接受请求的 node 返回 document 给 coordinating node
4. coordinating node 返回 document 给客户端
5. 特殊情况：document 如果还在建立索引过程中，可能只有 primary shard 有，任何一个 replica shard 都没有，此时可能会导致无法读取到 document，但是 document，但是 document 完成索引建立之后，primary shard 和 replica shard 就都有了

2、图解执行流程

