

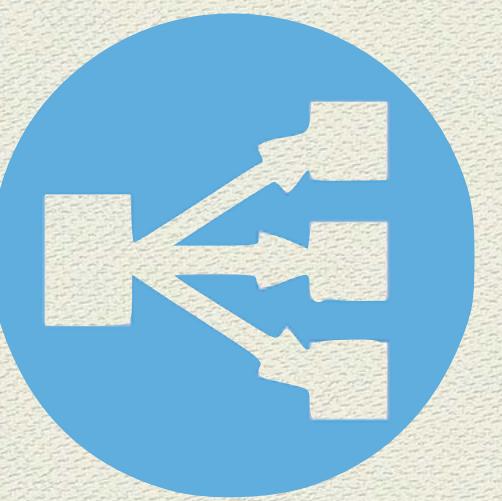
# ES完全入门

## 18. 理解ES中的路由 (Routing)



# 理解路由

- ◆ 插入文档时，ES如何选择一个Shard进行插入？
- ◆ 读取文档时，ES如何定位文档存在哪个Shard上？
- ◆ 路由 -> 将文档映射到某个Shard



# 缺省路由算法

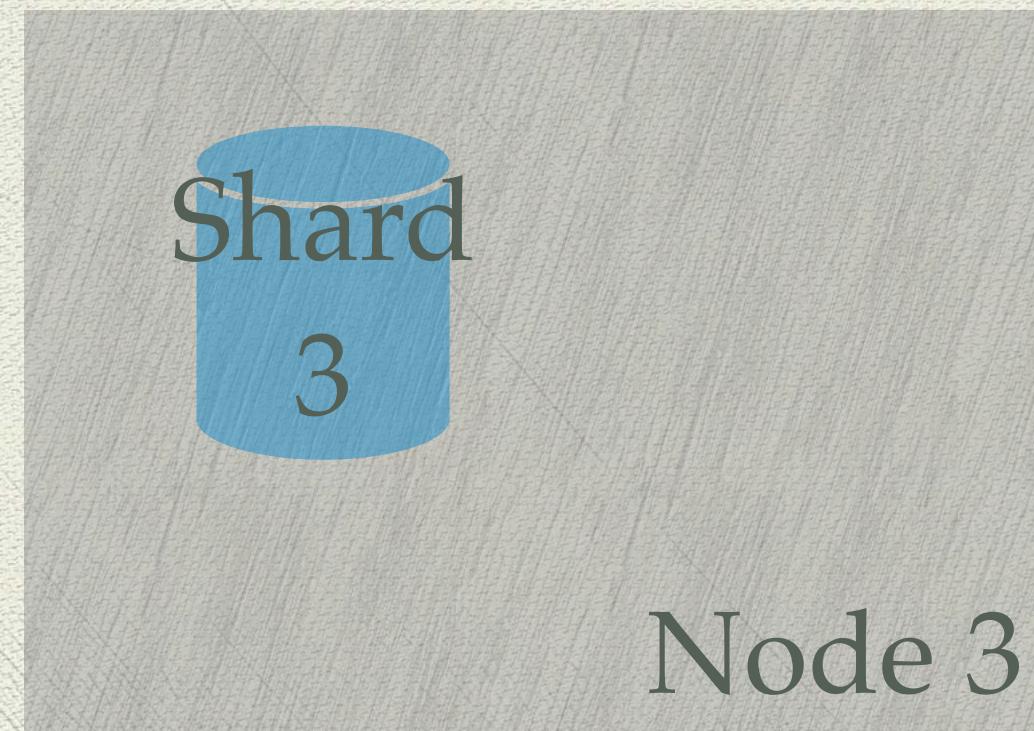
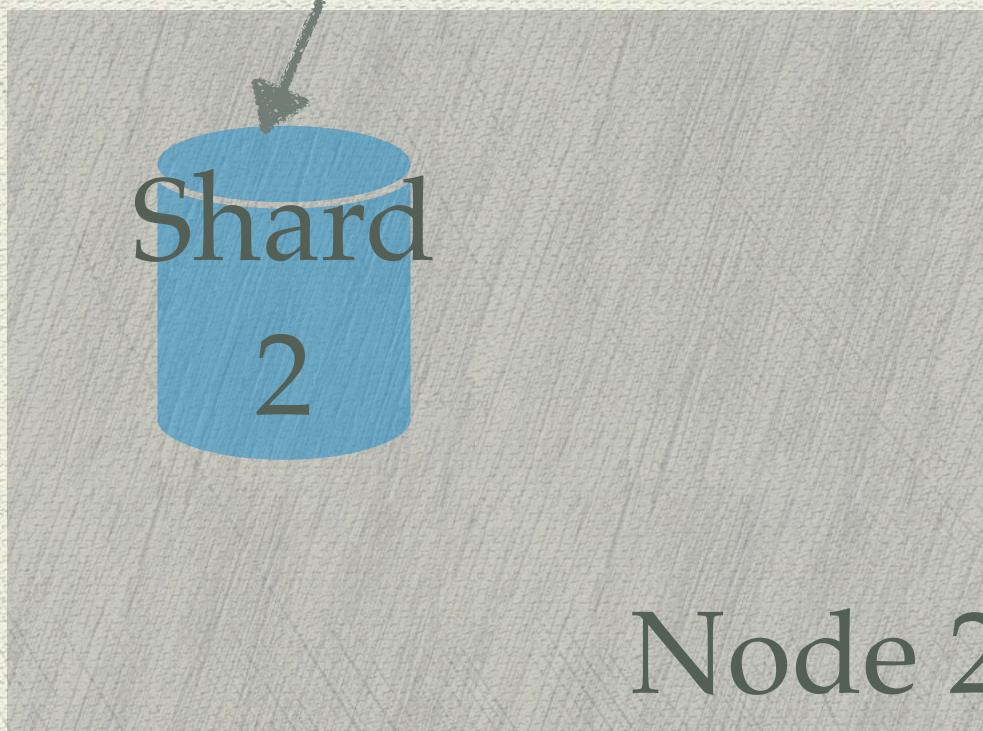
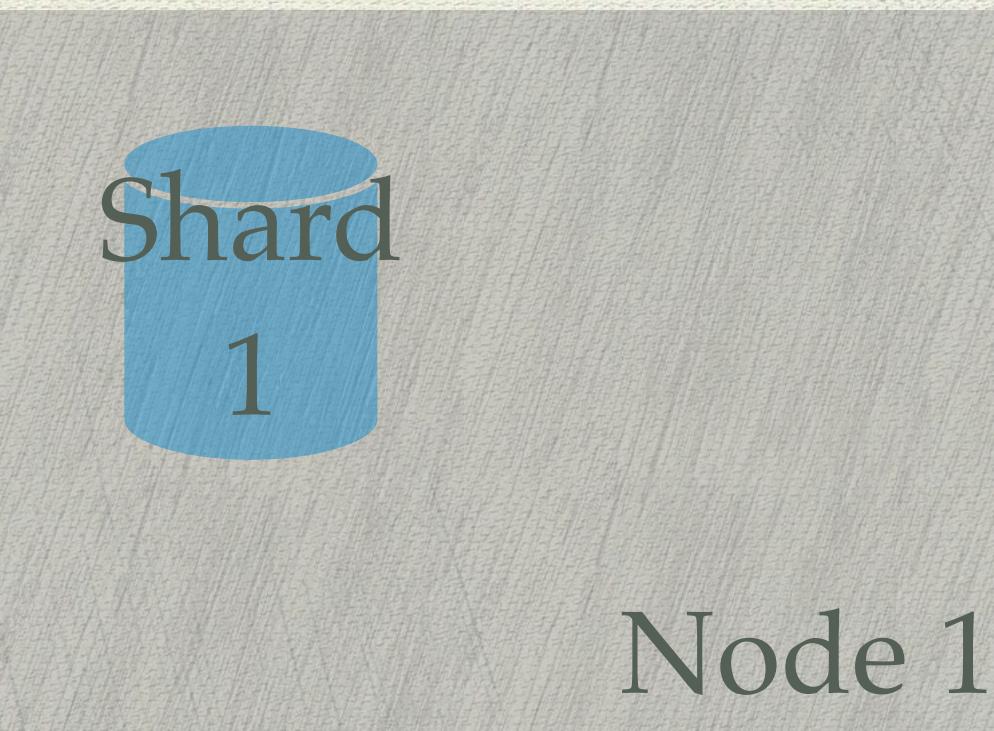
- ◆  $\text{shard\_num} = \text{hash}(\text{doc\_id}) \% \text{num\_of\_primary\_shards}$

# 插入文档

PUT /books/\_doc/10



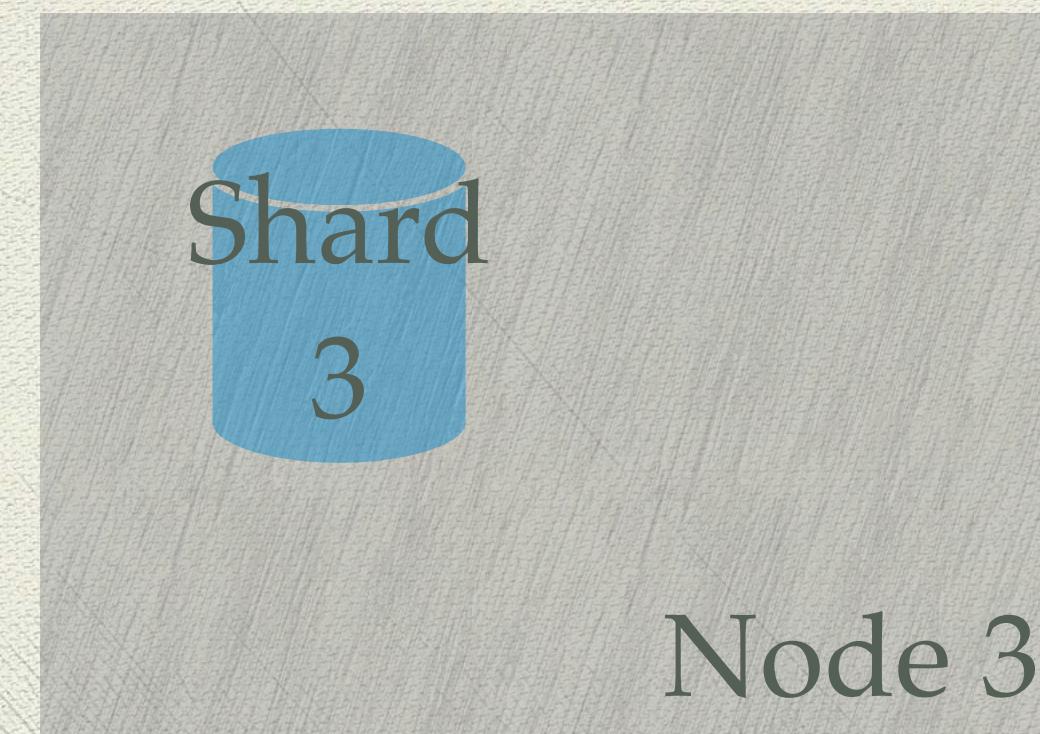
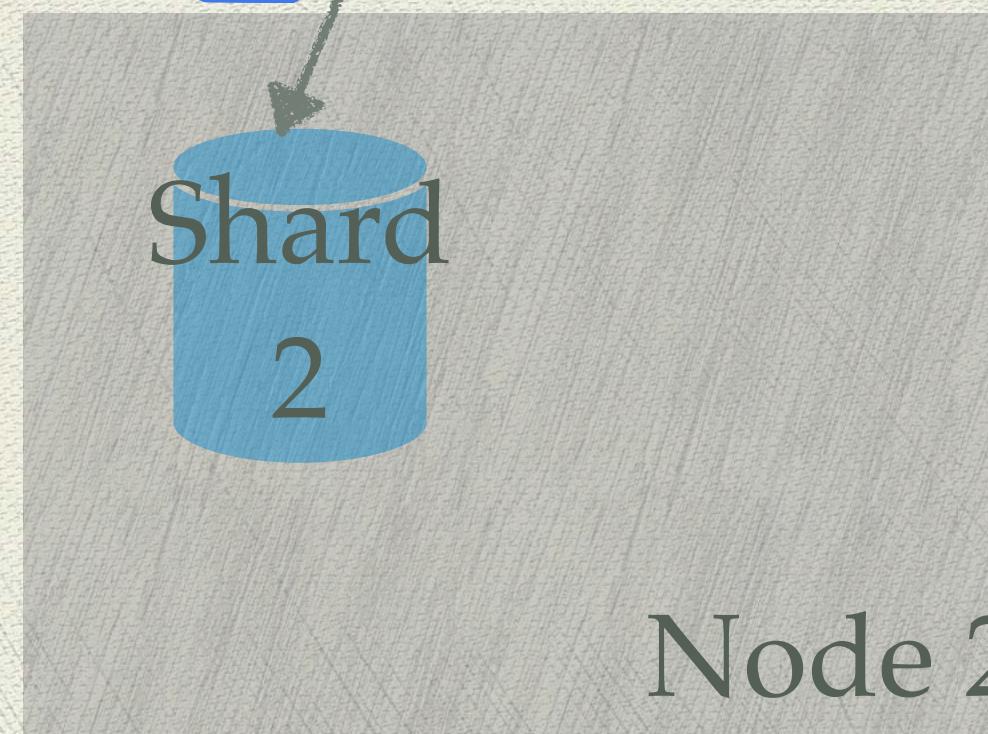
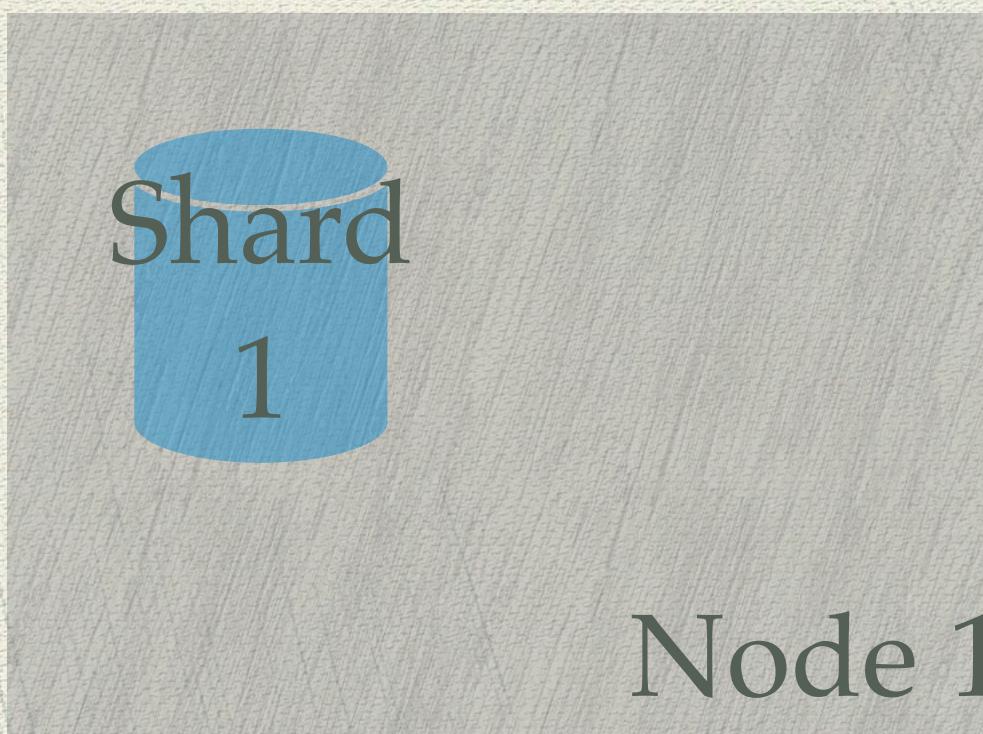
```
shard_num = hash(doc_id) %  
num_primary_shards
```



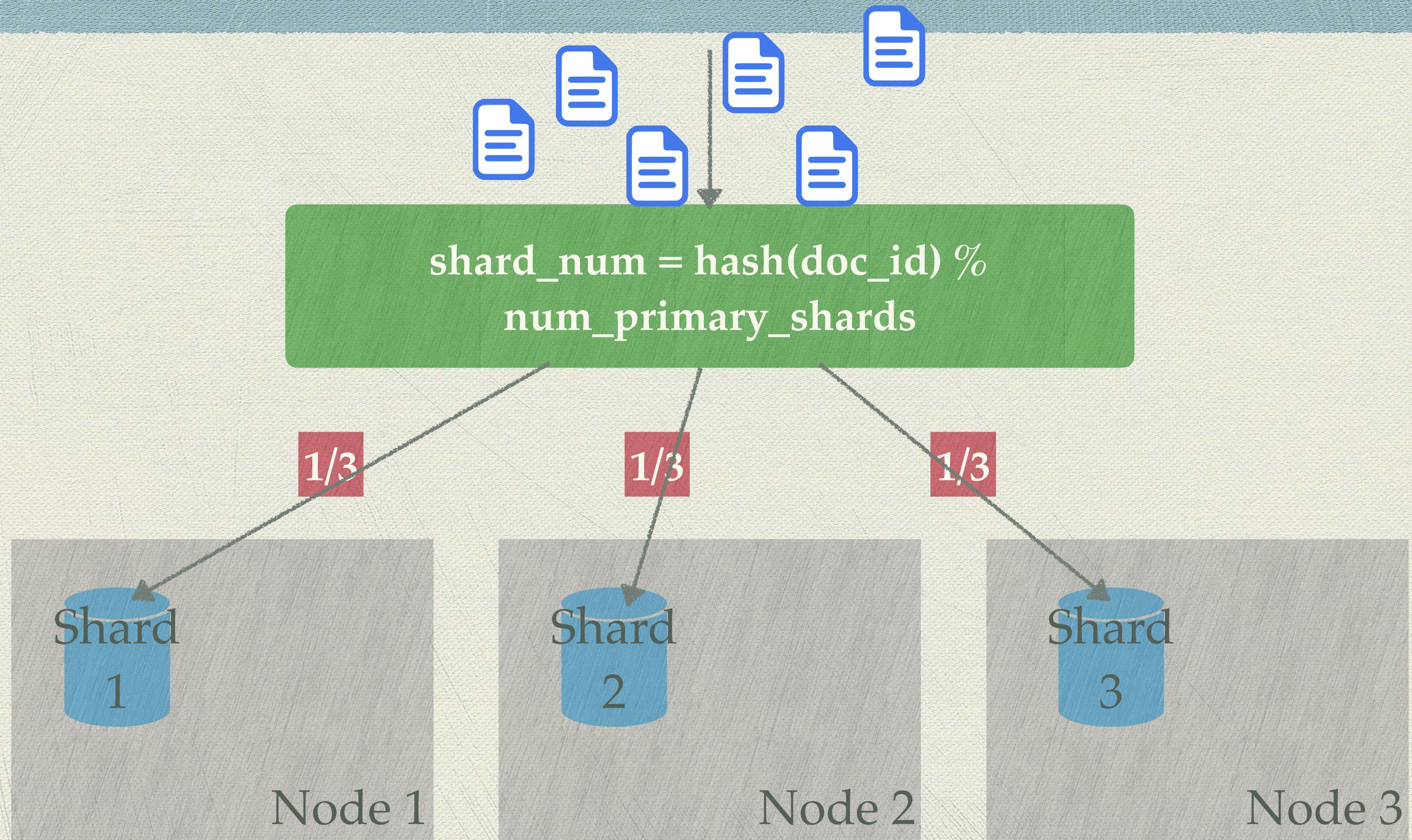
# 读取文档

GET /books/\_doc/10

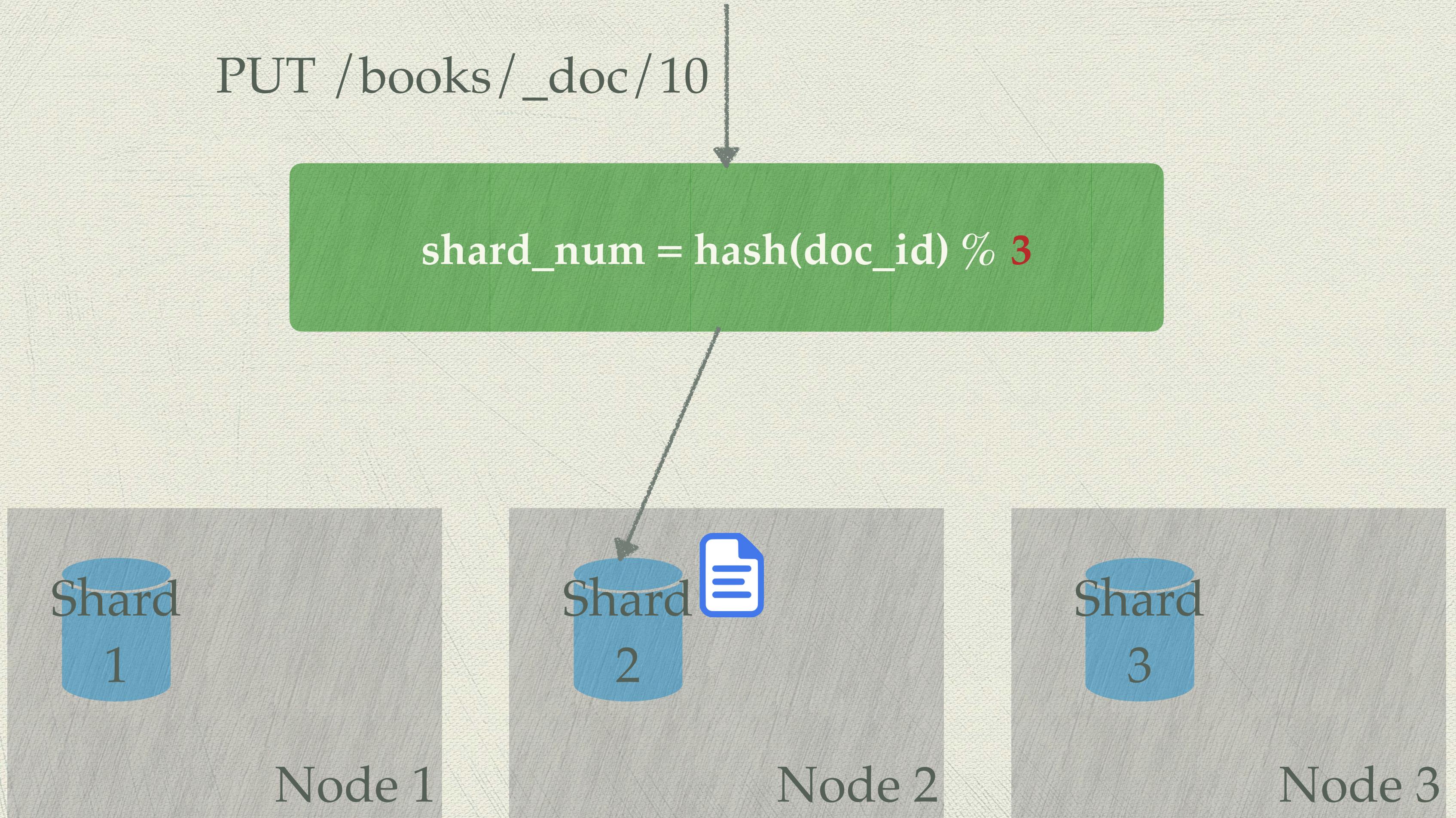
```
shard_num = hash(doc_id) %  
num_primary_shards
```



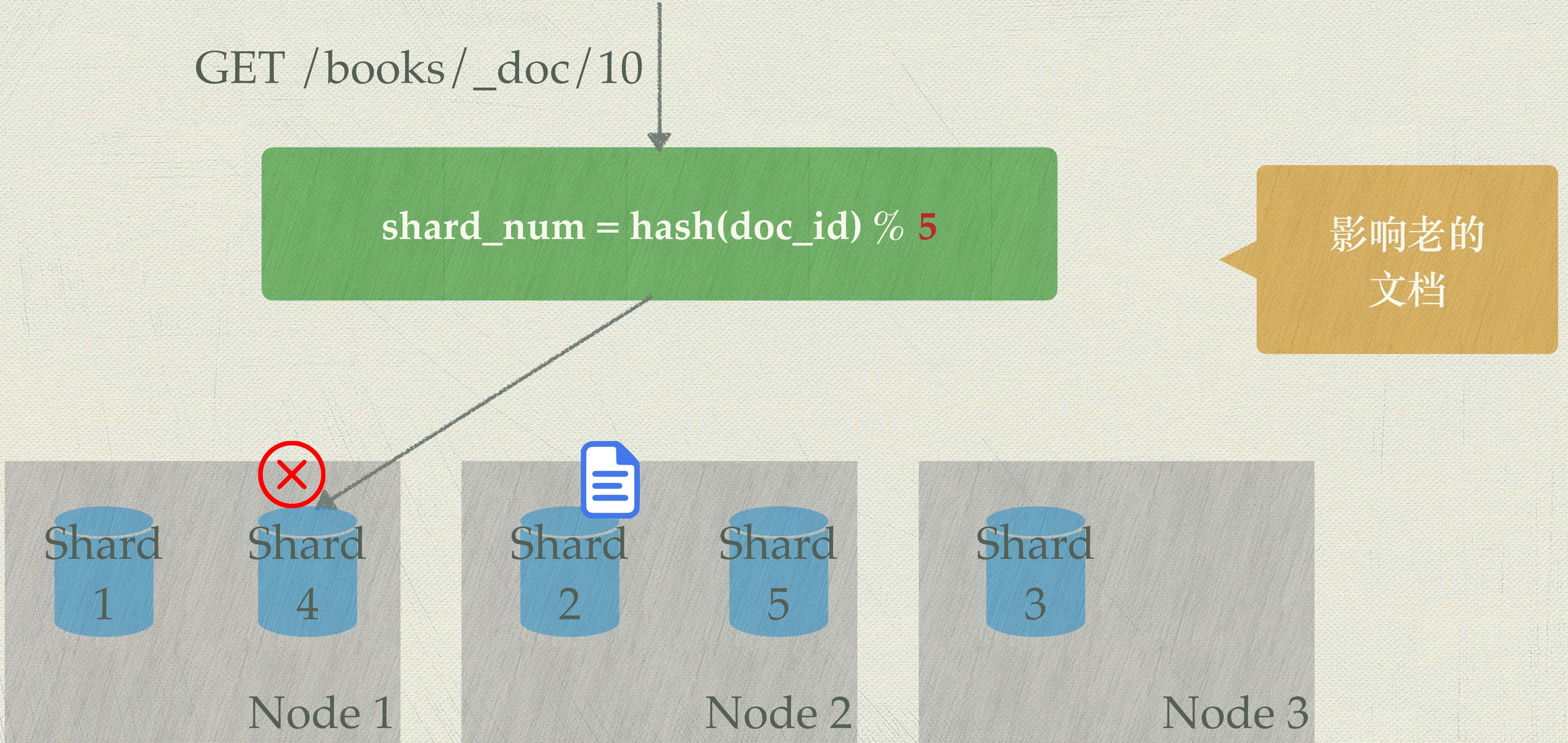
# 均衡分布文档



# 插入文档(Shard数量 = 3)



# 读取文档(Shard数量 = 5)



# 小结

- ◆ 路由解决文档到Shard的映射问题
- ◆ 插入/读取/更新/删除文档时使用同一路由算法
- ◆ 路由算法可以自定义(高级主题)
- ◆ 缺省路由算法可均衡分布文档
- ◆ 索引的Shard数量不能随意更改，因为路由算法会导致找不到之前插入的文档

# 补充Reindex

- ◆ 调整索引Index的Shard数量
- ◆ Setting / Schema变更

```
POST _reindex
{
  "source": {
    "index": "movies"
  },
  "dest": {
    "index": "movies_new"
  }
}
```