

# 2020-04-01

时间: 15:00~16:00

与会人员: 陈耿阳 田晨江 赵文祺 孙逸伦

记录人: 赵文祺

## 1. 尽量建立覆盖索引

publicationName_contentType						
contentType	publicationName	REGULAR	933.9 KB	4 since Wed Apr 01 2020	COMPOUND	

## 2. 是否分片?

#### 至少使用3个数据节点的复制集

MongoDB的建议最小部署是3个数据节点构成的复制集。复制集可以提供以下优点：

- 系统99.999% 高可用
- 自动故障切换
- 数据冗余
- 容灾部署
- 读写分离

#### 不用太早分片

分片可以用来扩展你系统的读写能力，但是分片也会带来不少新的挑战比如说管理上的复杂度，成本的增加，选择合适片键的挑战性等等。一般来说，你应该先穷尽了其他的性能调优的选项以后才开始考虑分片，比如说，索引优化，模式优化，代码优化，硬件资源优化，IO优化等。

#### 选择合适的分片数

分片的一些触发条件为：

- 数据总量太大，无法在一台服务器上管理
- 并发量太高，一台服务器无法及时处理
- 磁盘IO压力太大
- 单机系统内存不够大，无法装下热数据
- 服务器网卡处理能力达到瓶颈
- 多地部署情况下希望支持本地化读写

取决于你分片的触发条件，你可以按照总的需求 然后除以每一台服务器的能力来确定所需的分片数。

### 3. 需要阅读服务器的日志，解决慢查询问题（服务器请求和本地请求速度差异过大）

---

## 4. mtools

---

搜索关键词高亮：目前倾向于只传输和接受一个特定的HTML子集，预防XSS攻击

首页排名的问题，机构排名比其他几块都要长，以什么样的形式呈现？

- 省略 + 展开
- 上面放四个，下面放一个
- .....?

柱状图、饼状图

加入需求以外的限定条件：比如关键词接口中加入时间范围的限定条件（近五年）

搜索接口可能需要同时返回author的Id（在现有的返回name的基础上）

## 5. 已知的情报：

---

1. 白嫖GKE 8 core 57 G RAM