



elastic

# 58到家ES服务化实践

张家杰@58到家

Elastic{ON} DevChina – Dec 10, 2016

# 目录

1

ES服务化

2

应用实践

3

未来规划

4

总结

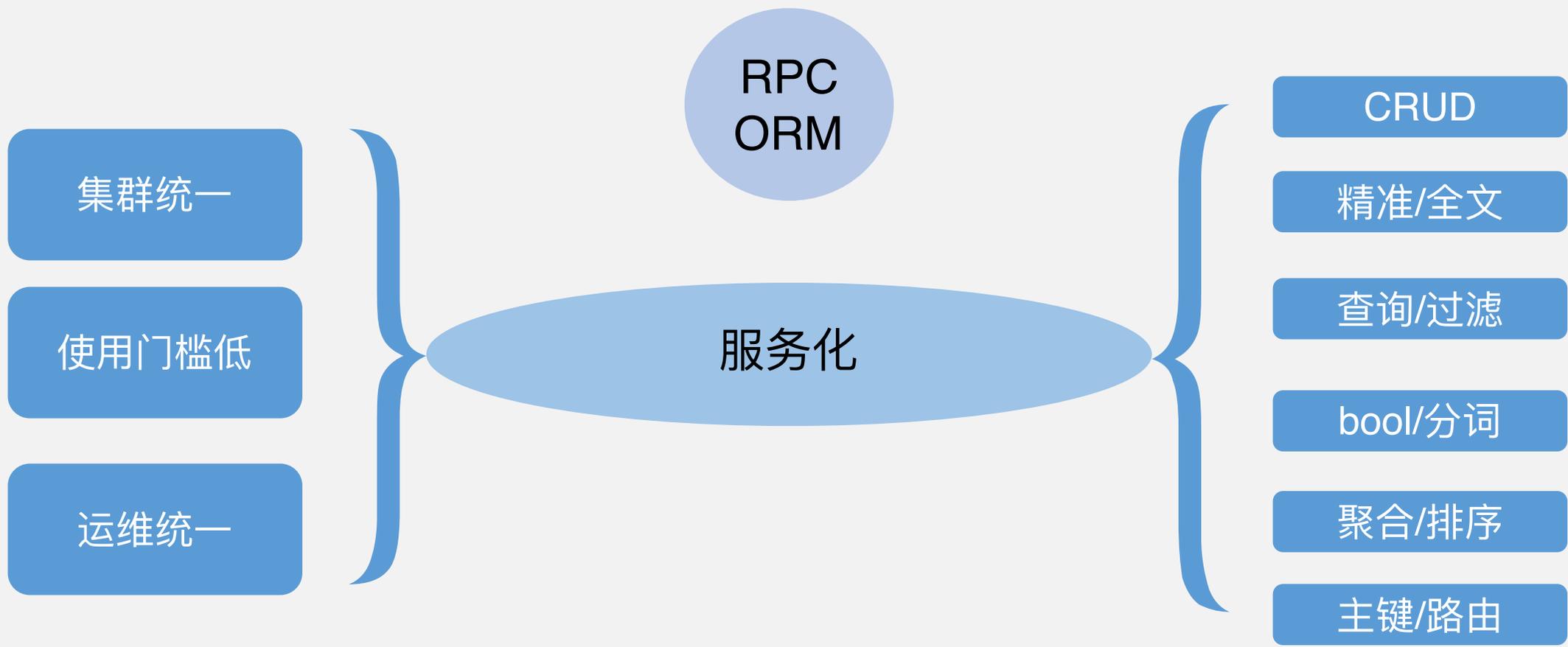
## 服务化构建—背景

- 地理位置搜索成本过高
- 线上搜索业务逐渐增多
- 有些场景DB很难继续扩展支持

# 服务化构建—问题



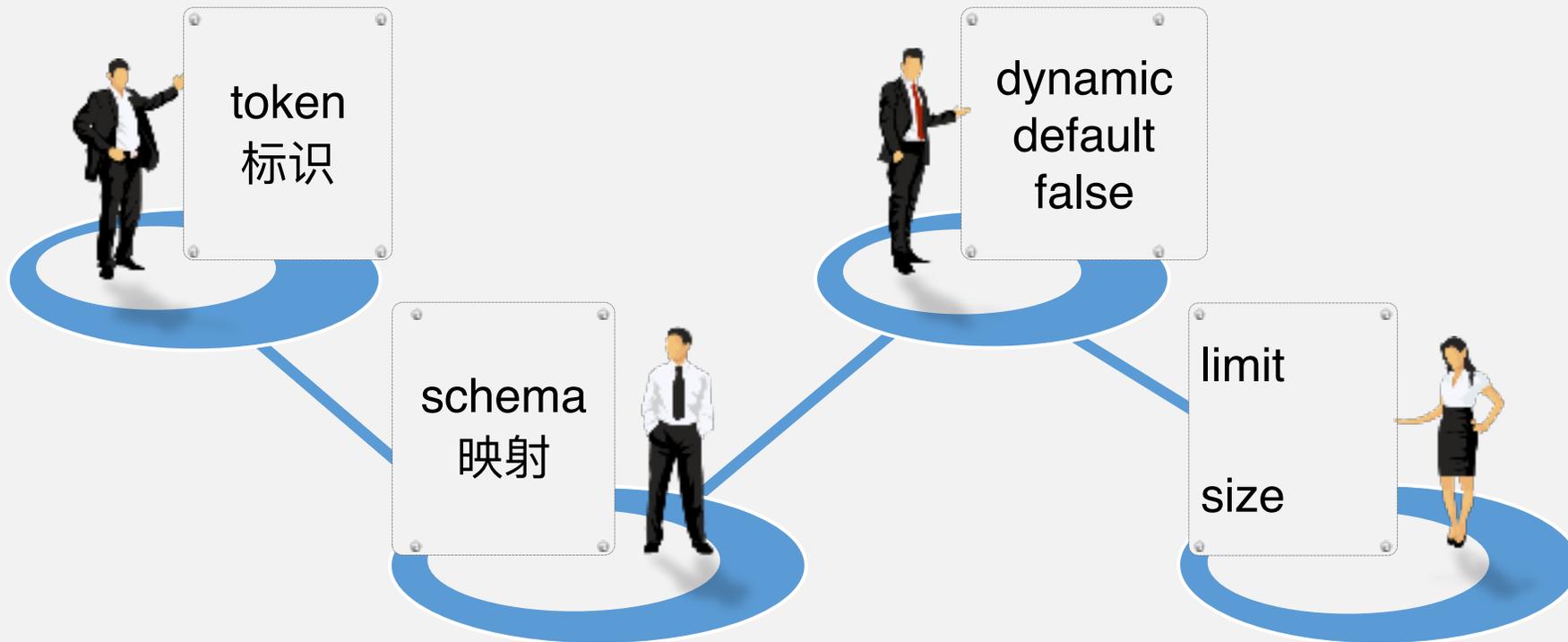
# 服务化构建—目标特点



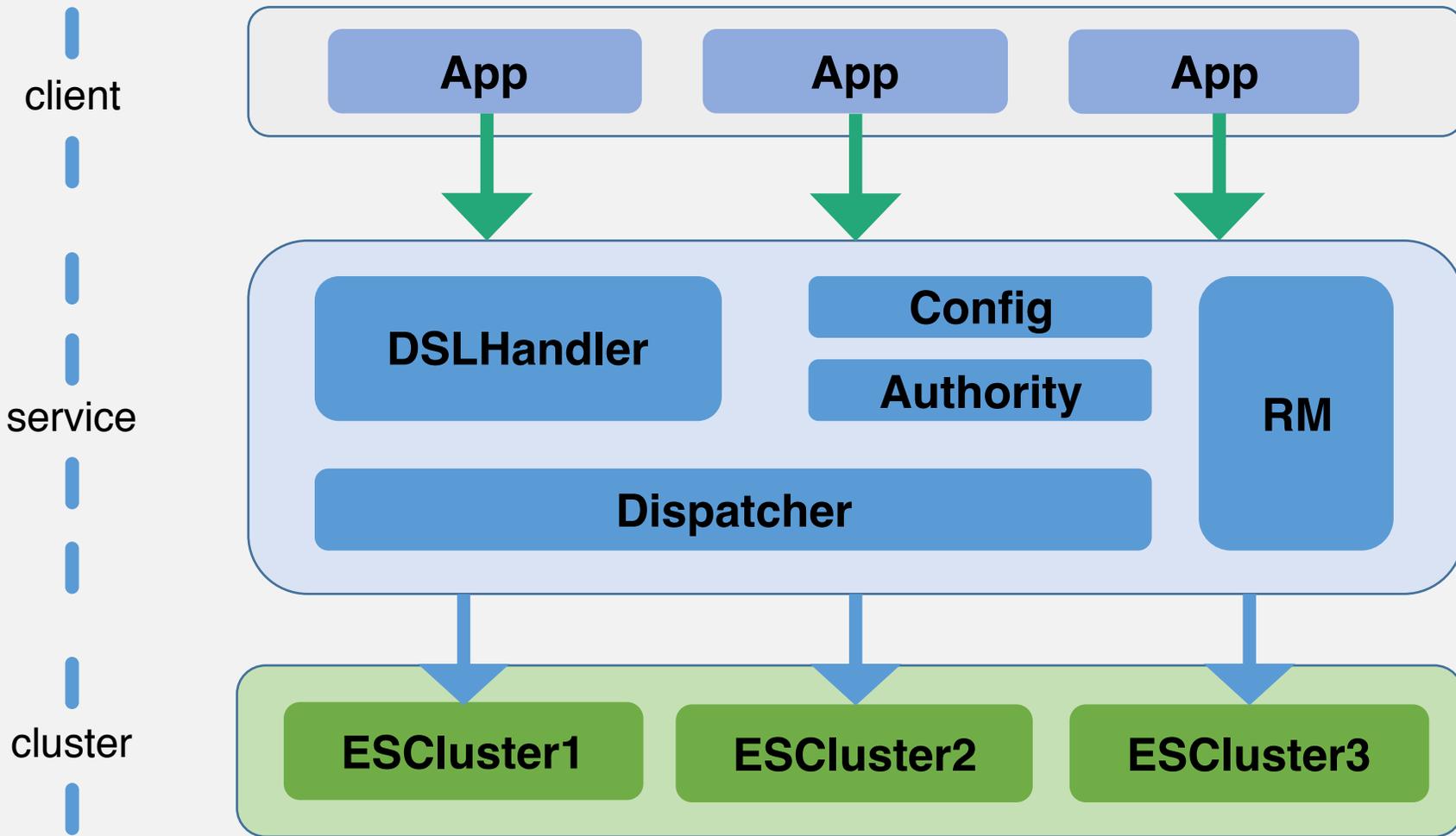
## 服务化构建—基础

DB	ES
DataBase	Index
Table	Type
Schema	Mapping
Index	All Index
SQL	DSL/Interface
Select	GET
Update	PUT

# 服务化构建—约束



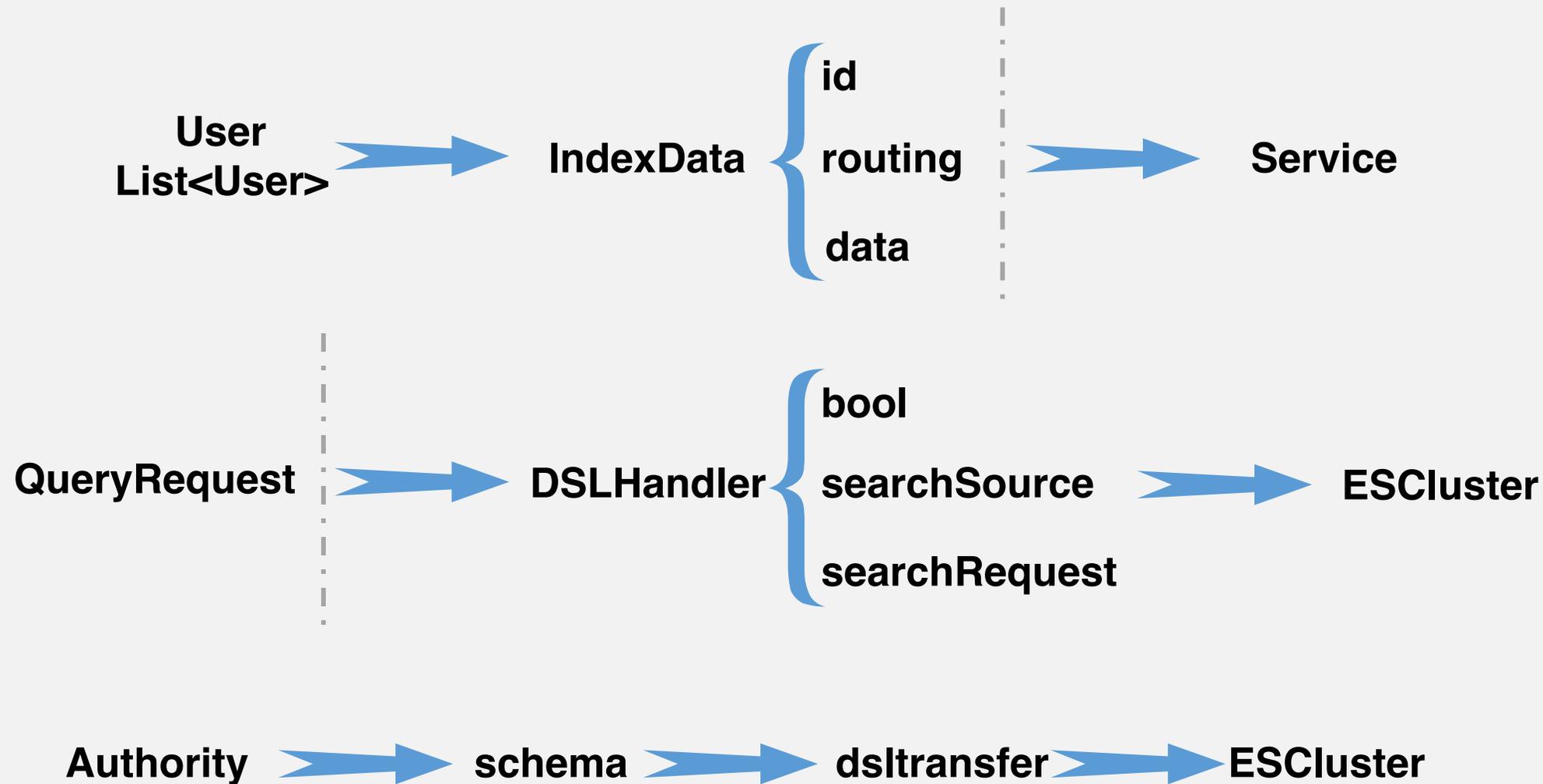
# 服务化构建一平台



## 服务化构建—中间件

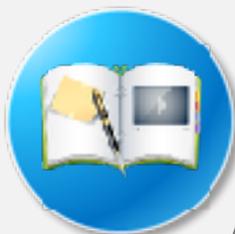
- index/update/upsert/delete (id)
- bulkIndex/bulkUpdate/bulkUpsert
- bulkDelete (ids) /delete (type)
- CommonQuery/GeoQuery/AggregationQuery
- @Id
- @Routing
- @Type

# 服务化构建—中间件



# 服务化构建—tips

filter 和 query  
选择优化



基本类型  
基本对象



选择transport  
client



版本冲突控制

## 应用实践—1个case



## 应用实践—集群分类

- 索引隔离
- 按需优化



## 应用实践—位置搜索

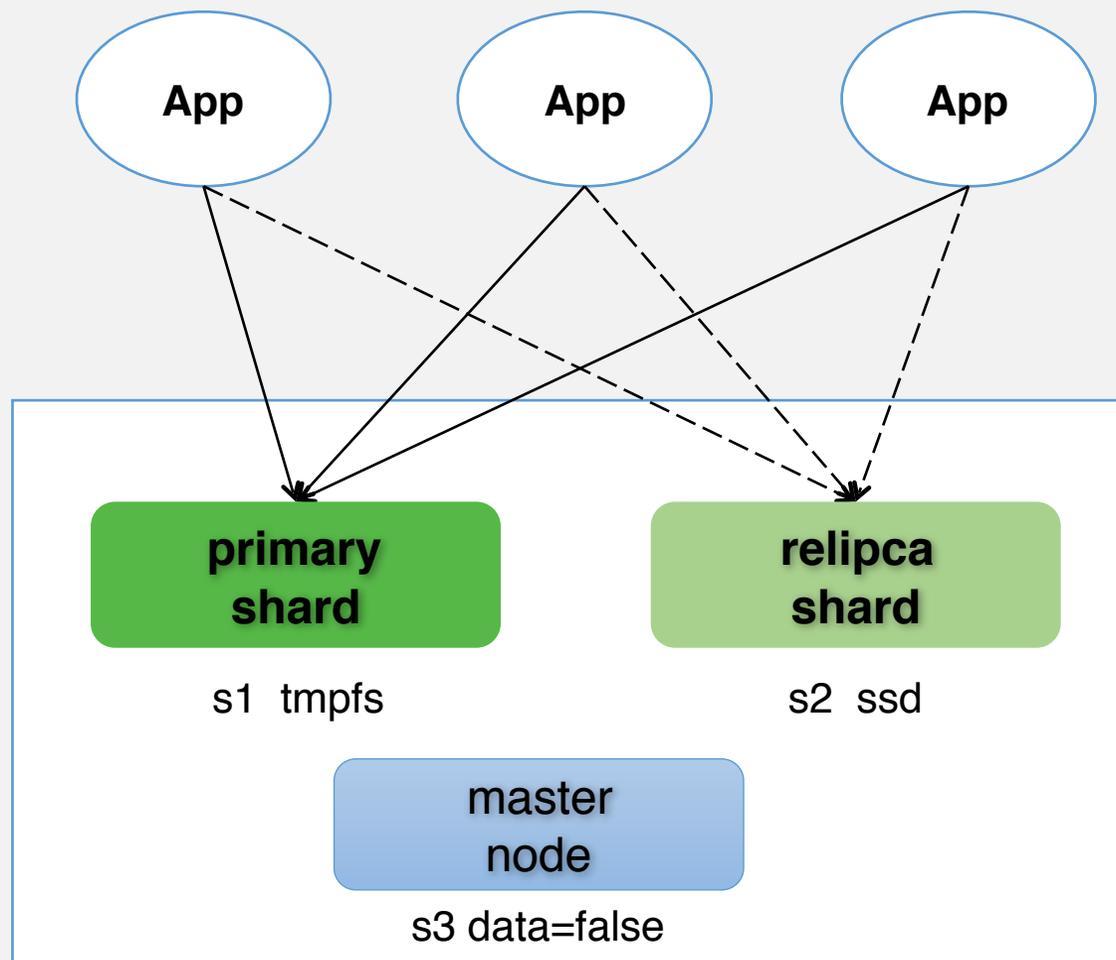
App分端：用户端，司机端

核心业务：

1. 用户查看司机
2. 用户下单
3. 司机抢单 + 用户返选
4. 记录里程



# 应用实践—位置搜索



## 应用实践—位置搜索

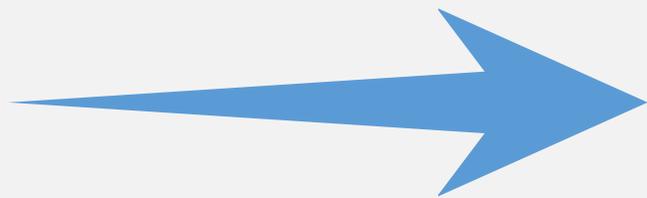
- 使用其它条件尽量减少数据集大小
- 使用memory来存储倒排索引
- 查询优先使用主分片
- 使用geo\_bounding\_box优化过滤模型
- 使用plane计算方式计算距离
- 压缩geohash长度，减少内存

# 应用实践—位置搜索



## 应用实践—商品类搜索

- 服务项搜索
- SEO聚合
- 帖子查询
- 评价信息
- 配置管理后台

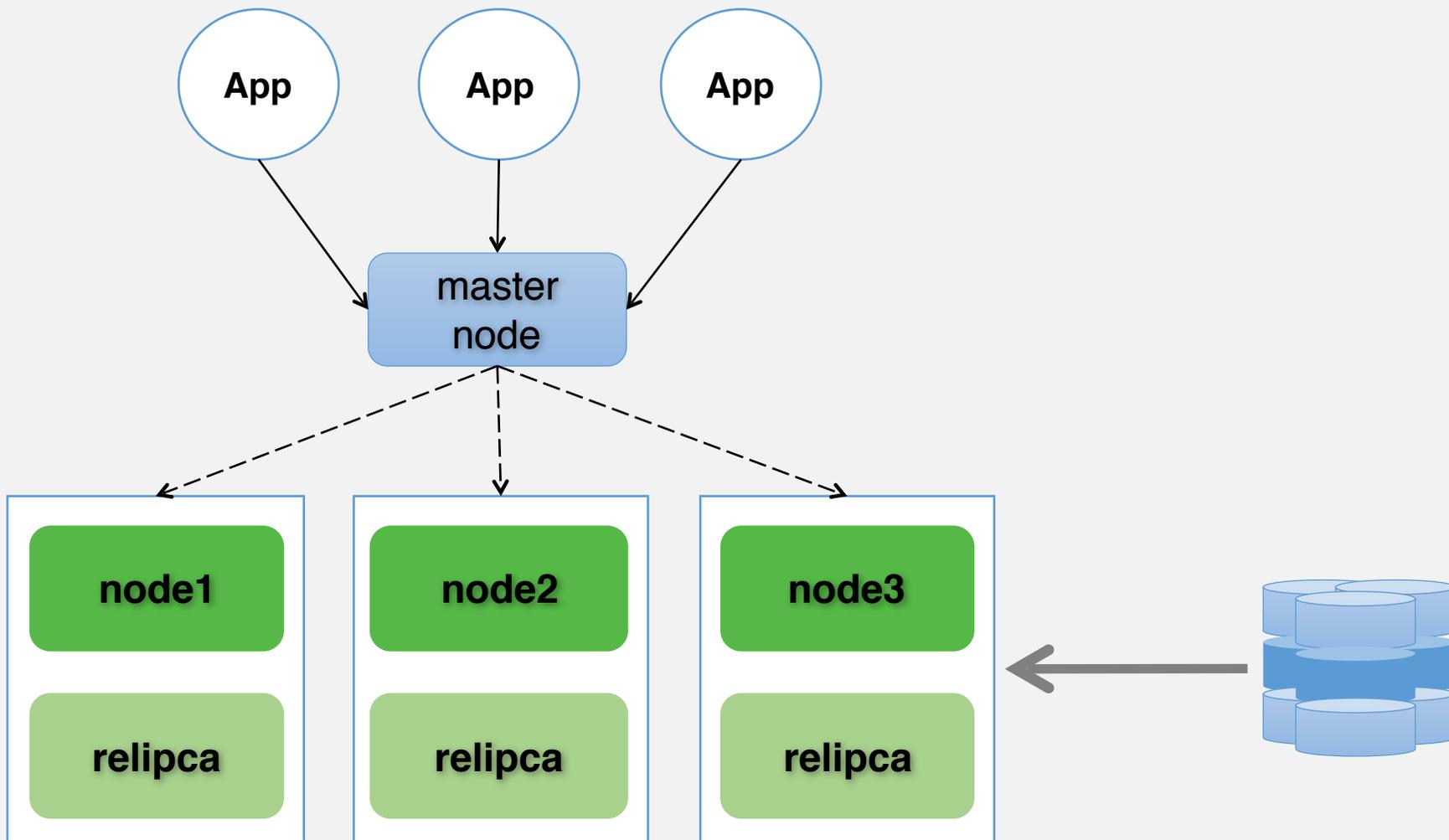


数据量适中

读写频率适中

数据有效期长

# 应用实践—商品类搜索



## 应用实践—商品类搜索

### 数据同步校验

- 业务控制
- 速度控制

### 集群重建

- token映射别名
- 服务可终止?
- 追数据+双写

## 应用实践—商品类搜索

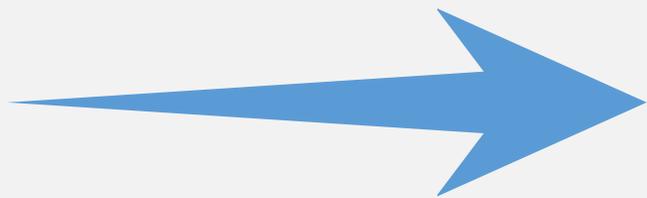
- 业务索引隔离
- 业务路由
- 减少非必要字段
- 合理分配节点角色
- 段合并
- 元数据备份

# 应用实践—商品类搜索



## 应用实践—日志类搜索

- nginx日志
- 业务日志
- 应用日志
- 调用链日志

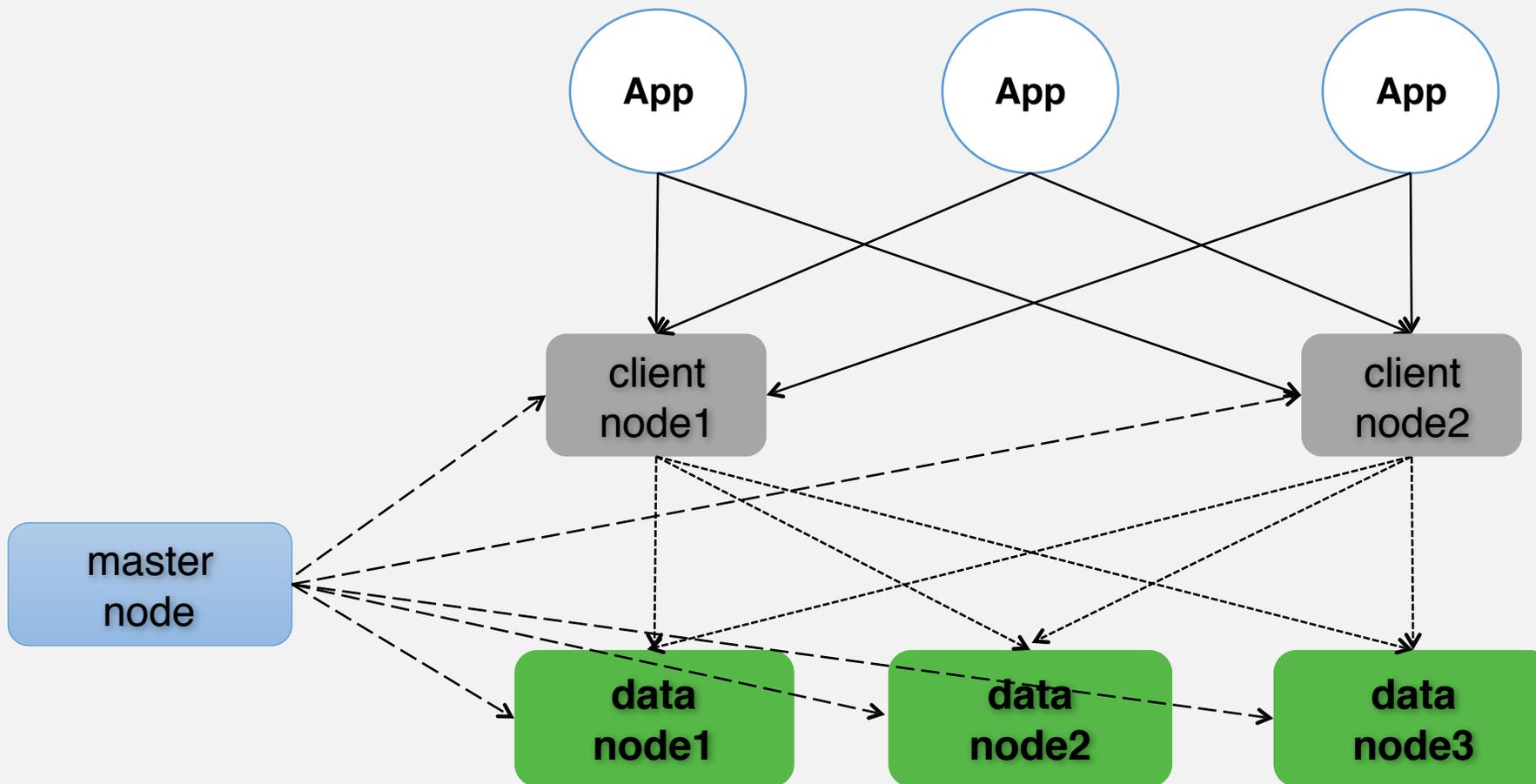


数据量大

写索引频繁

数据有效期短

# 应用实践—日志类搜索



## 应用实践—日志类搜索

- 批量写索引
- 使用SSD硬盘
- 减少副本冗余
- 增大刷新时间
- 按需创建切分索引
- 使用模板

# 应用实践—日志类搜索

2016-01-01 00:00:00 2016-11-20 17:19:02 URL 请输入查询内容 参数 起始IP 请输入查询内容 耗时: -  只显示异常

开始执行时间	耗时	起始IP	请求url	结构图
2016-11-20 17:17:59.900	38ms		http://.djco...	

集群名称	类型	状态	IP	接口	参数	时间轴
...arch_passportweb	dwf	OK		http://.web...	wp6SdRte...	20ms
...arch_passportweb	dsf(c)	OK		dus.AppUserTokenImpl.chec...	[79406125...	
dus	dsf(s)	OK		AppUserTokenImpl.checkUse...	[79406125...	

共 3 条数据

单位: ms <50 50~200 200~500 500~1000

wp6SdRtean7nz9B8StxIR6HRTm1924qCW3b60qt  
GdcY/MPAEnigYfi/FxEJRw2dRIILHhCLt  
CVcErjng9H2aen2Q0b15NzglRaro5L2Q04m84w4l  
szh+HZEPyC2hxfAuj5LaPvyVLgZDzH

# 未来规划

01

支持更灵活的打分策略  
让业务定制更加简单  
更强大的聚合功能

03

和DB绑定，做到CRUD  
的自动同步

满足高级用户的  
NativeAPI支持

02

进一步提高易用性，减  
少门槛，减少学习成本

04

# 总结

- 服务化构建
  - RPC, ORM, 约束
- 地理类集群
  - 读写密集, 减小数据集、内存、业务折中
- 商品类集群
  - 查询快, 隔离、路由、减少字段、定期优化
- 日志类集群
  - 写频繁, 批量、减少副本、角色分配、索引拆分

**Thanks !**

**Q&A?**

