



基于ES构建贝壳找房搜索中台

高攀

贝壳找房搜索平台负责人

2019-12-07

提纲

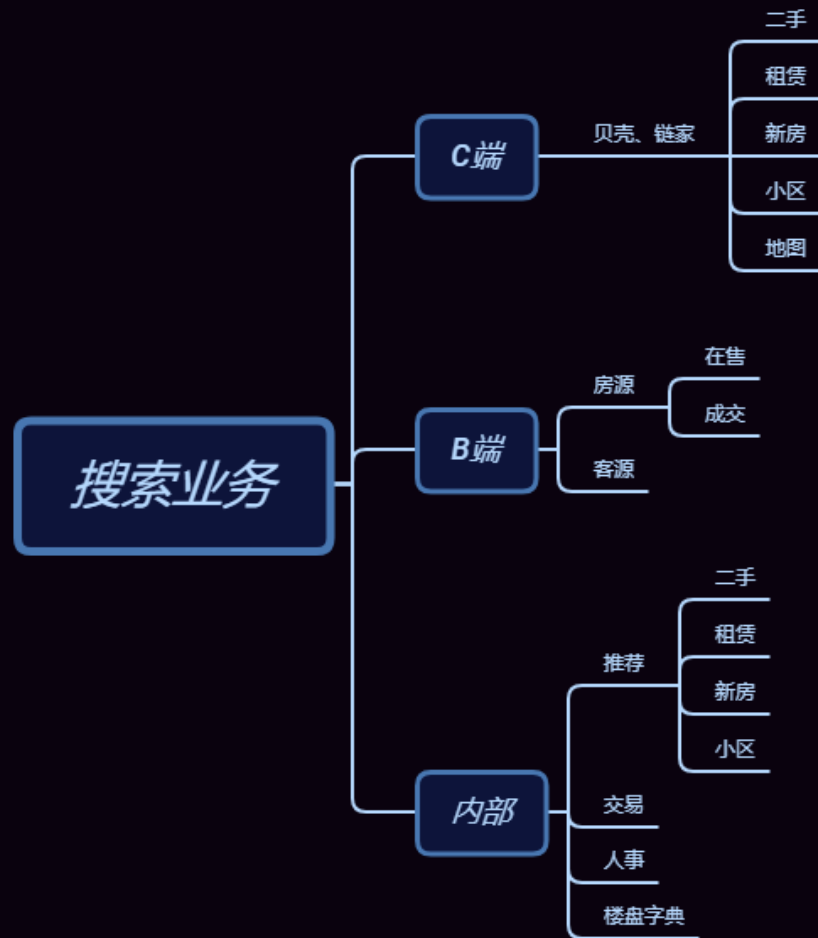
- 1 ES在贝壳的使用情况
- 2 贝壳搜索中台演化之路
- 3 稳定性99.99%的挑战
- 4 未来规划



ES在贝壳的使用情况

使用场景

- 贝壳、链家主搜
- C端、B端、内部
- 二手、新房、租赁
- 房源、客源、推荐
- 共支持350多个业务线
- **最核心服务之一**



ES在贝壳的使用情况

使用规模

21套集群



300ES节点



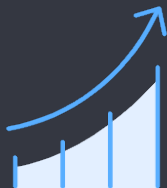
1100 index



20亿数据



350业务



10亿查询/天



1亿写入/天



峰值3W/QPS



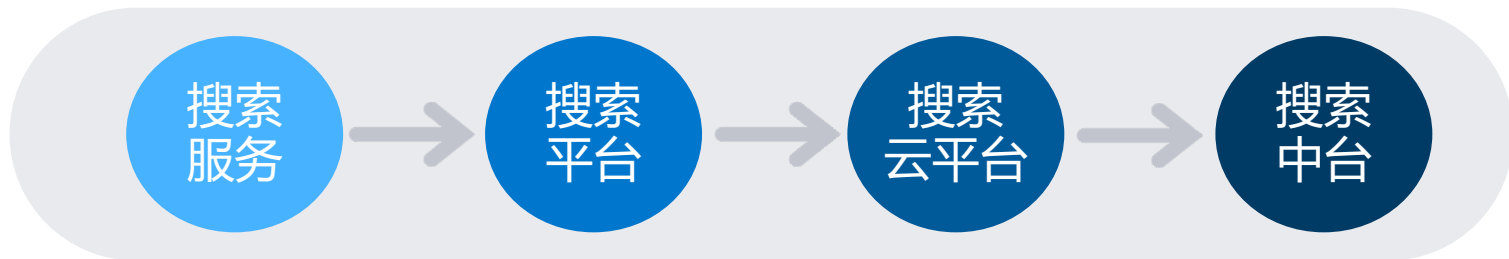
稳定性99.99%



提纲

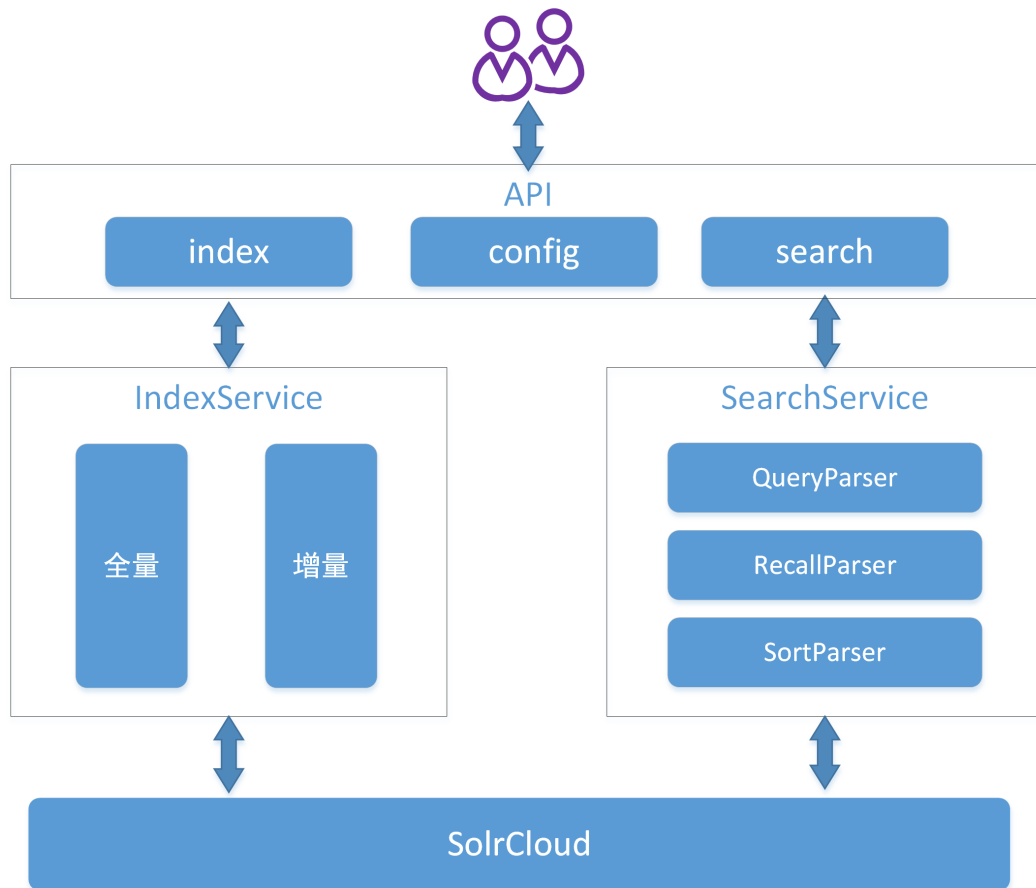
- 1 ES在贝壳的使用情况
- 2 贝壳搜索中台演化之路
- 3 稳定性99.99%的挑战
- 4 未来规划

贝壳搜索中台演化之路



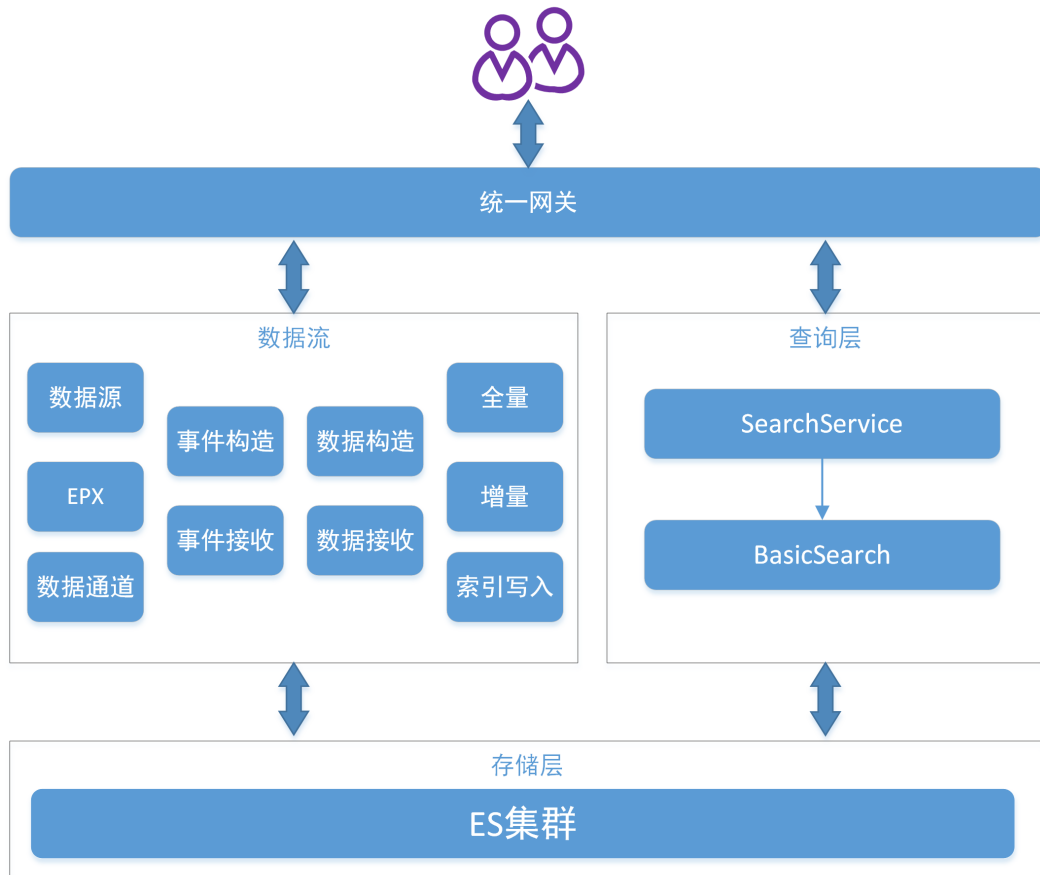
贝壳搜索中台演化之路

搜索服务



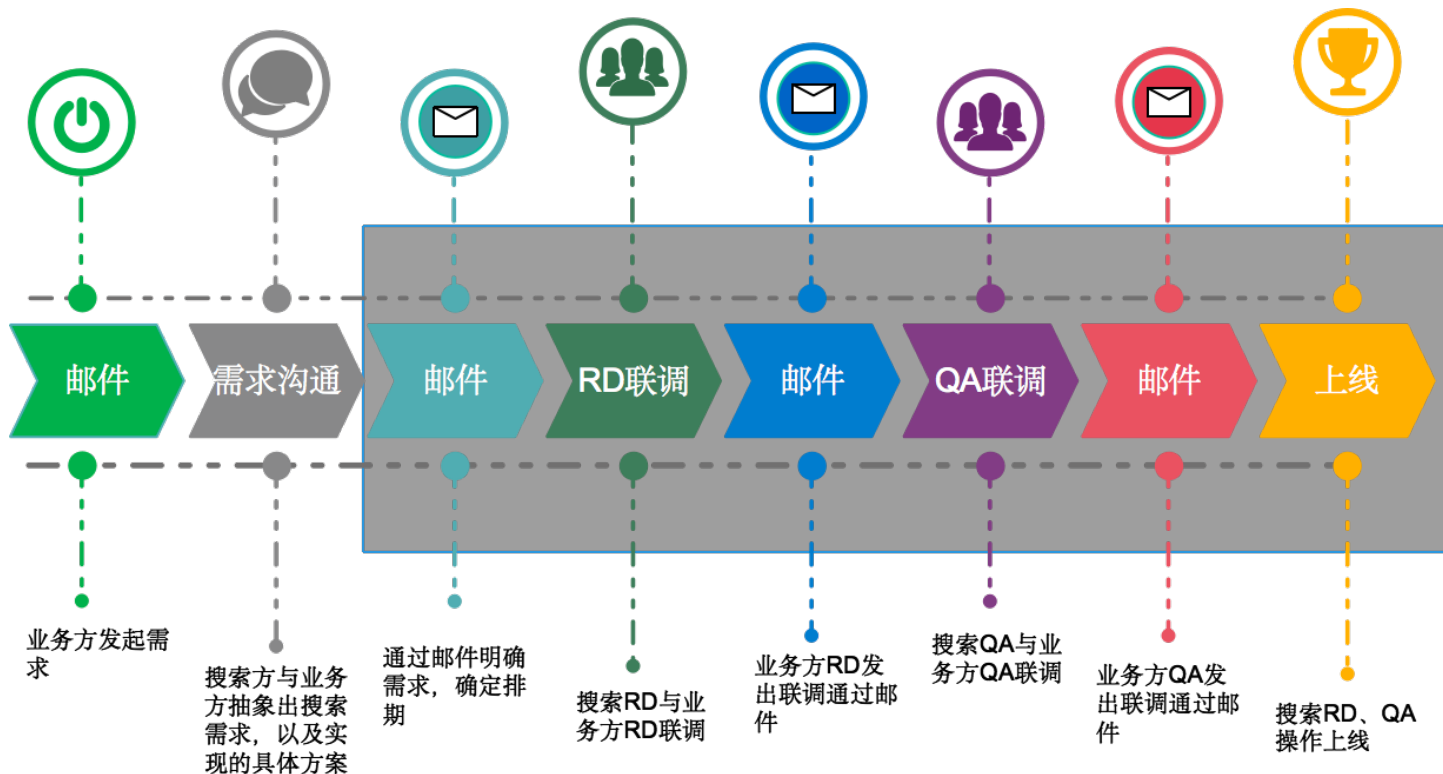
贝壳搜索中台演化之路

搜索平台



贝壳搜索中台演化之路

搜索云平台



贝壳搜索中台演化之路

搜索云平台



搜索云平台



应用列表 / 创建新应用

创建新应用

* 应用描述: 高攀测试

示例: 新房-内容-评论数据索引接入

* To B or to To B To C

* 预估数据总量 千万

* 每日新增量级 十万

* 搜索需求描述 高攀测试

* 数据接入方式 服务端拉取 客户端推送[点击了解“服务端拉取”与“客户端推送”](#)

* 申请方研发人 高攀 (1000078719) X

* 申请方测试人 赵宁 (26034659) 数据智能组 X

提交

取消

* searchService: search-common#platform-test#das-stream-achetype

线上环境配置

* 服务类型: 10, 搜索服务

* 数据类型: 11, 二手房

* 项目类型: 100, 大C

* 集群分组: idc

* searchService: search-content#prod#master

* 数据流kafka集群: platform-sit-kafka

* shard数: 10 注: 只能填数字

* ES副本数: 2 注: 只能填数字

* ES刷新频率(秒): 1 注: 只能填数字

* 最大返回数目: 1000 注: 只能填数字

* 尾号: 001 注: 只能填数字或字母

* 索引数: 1 注: 只能填数字

□ service_id: 1011100001

通过

拒绝

注意：一套配置可以包含多个数据库的监听，但要求这些库下监听的表信息一致，若不一致请拆分为多套配置。

数据库信息

操作	IP	端口	数据库名 (一行只能输入一个)
+ -	10.26.XX.XX	3306	rpms

数据表信息

操作	数据表	业务主键	其他字段 (多个字段以英文逗号分隔, 不填为全部监听)
+ -	rent_apartment_project	id	

监听校验

已有5XP监听, 集群名为: epk-kafka-test, Topic为: epk-10x26x21x27-3306-rpms

取消 保存

RD环境配置 / 编辑配置环境

编辑配置环境

- 配置监听数据
- 配置回调地址
- 配置索引结构**
- 数据合规性校验
- 配置生效

提示: 数据库字段类型为string时, 若无需分词, 选择keyword类型; 若需分词, 请选择text类型. 各分词器效果, [点击查看](#)

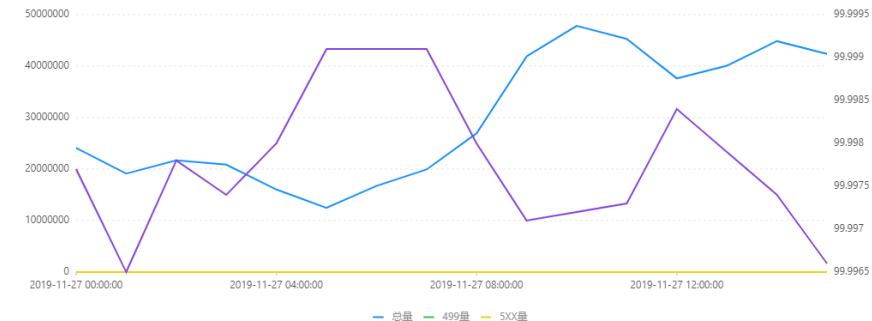
操作	索引字段名	字段类型
+ -	name	text 中文分词
+ -	id	keyword
+ -	age	integer

返回 上一步 下一步 保存到草稿

整体数据

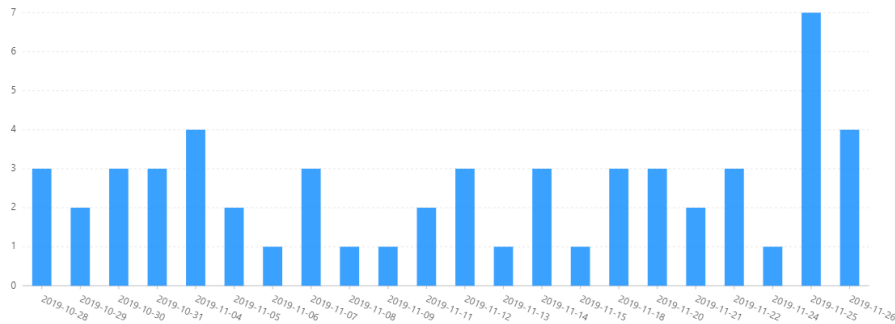
搜索总量	平均响应时间 (ms)	稳定性	5xx量	5xx占比	499量	499占比
477,818,227	18	99.9976%	1,417	0.0003%	9,971	0.0021%

数据趋势



总体 新接入 索引变更

应用数 共56个



[应用管理](#)
[搜索研发](#)
[搜索测试](#)
[平台管理](#)
[接入数据统计](#)
[报警配置](#)
[ES集群管理](#)
[平台值班](#)
[搜索工具](#)

+ 新建集群

集群信息

集群名称	节点数	master地址	kibanna地址	部署区域	操作
lianjia-search-tencent-content-c01	9	:20031 :20034 :20037	es.docker.ke.com/lianjia-search-tencent-content-c01	TENCENT	新增节点 查看节点信息
lianjia-search-tencent-fang-c02	12	:20071 :20074 :20077	es.docker.ke.com/lianjia-search-tencent-fang-c02	TENCENT	新增节点 查看节点信息
lianjia-search-tencent-fang-c01	19	:20111 :20114 :20117	es.docker.ke.com/lianjia-search-tencent-fang-c01	TENCENT	新增节点 查看节点信息

* 集群名称:

* 数据节点个数:

* 部署区域:

以下为高级配置(可不填):

容器镜像版本:

CPU限制:

内存限制:

开始机器名称:

恢复快照名称:

[取消新增](#) [确认新增](#)

节点信息

节点名称	角色	外部地址	容器镜像版本	jvmXmx	CPU限制	内存限制	是否有数据	操作
es-ten12-search.mars.ljnode.com_node0	data	:20056	1.1.0	16g	14	32Gi	是	修改 迁移 重启 删除 驱赶数据 取消驱赶数据
es-ten04-search.mars.ljnode.com	master	:20031	1.1.0	16g	14	32Gi	否	修改 迁移 重启



贝壳搜索中台演化之路

搜索云平台

3倍

业务接入效率提升
3倍

之前：平均9天

之后：平均3天

6倍

搜索RD人效提升6倍

之前：3人/日

之后：0.5人/日

60%

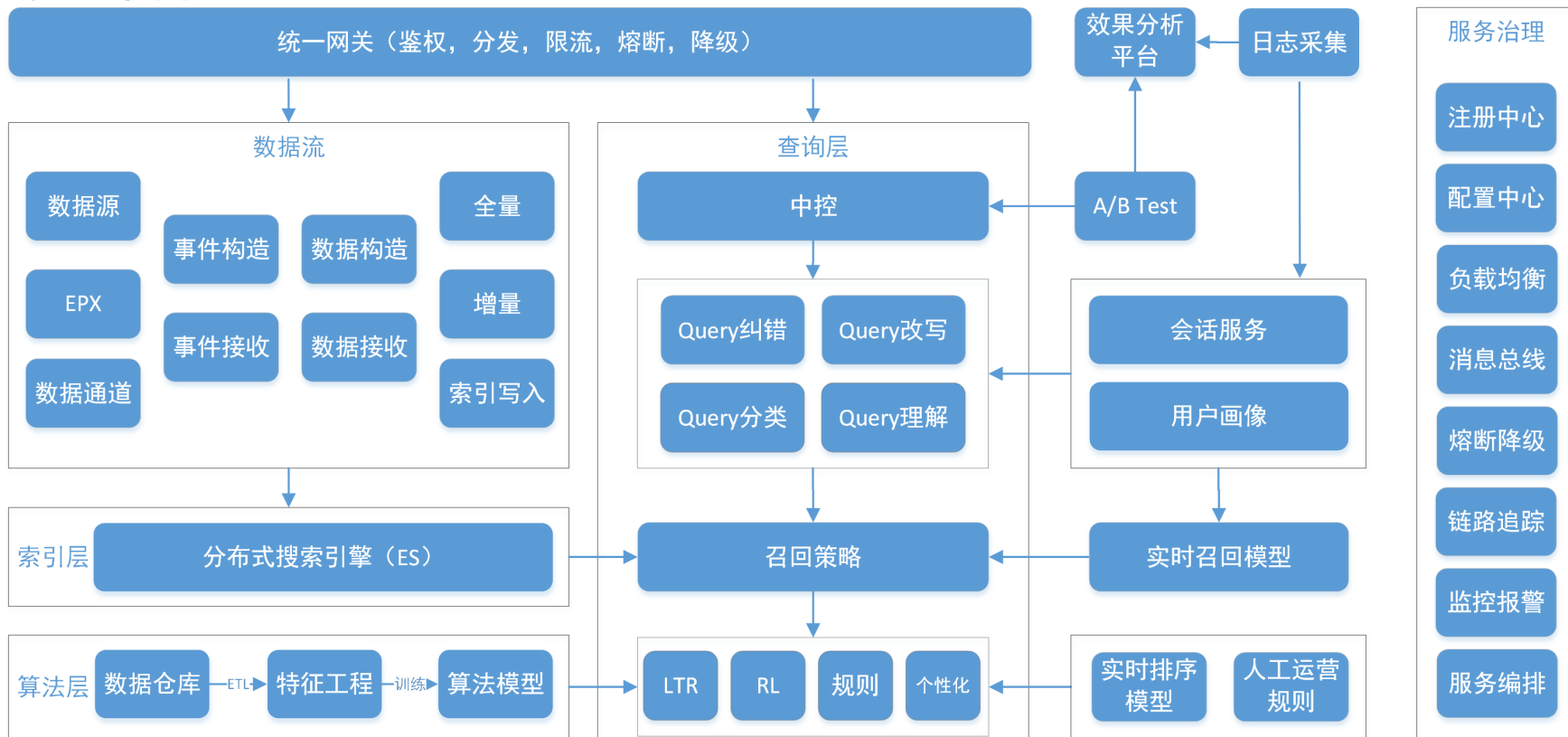
故障率降低60%

之前：人工改配置误
操作造成6起线上故障

之后：0

贝壳搜索中台演化之路

搜索中台



模型管理

任务调度

配置管理

二手

新房

租赁

策略配置

召回流

配置管理-租赁 / 排序配置详情

实验配置

service_id: 1013901005

步骤: 排序

实验: 3

生效时间: 2019-11-11 00:00:00-2020-12-31 23:59:59

场景: all

端: all

用户: all

城市: all

修改流量或有效期

生成新节点

分组配置

节点修改记录

实验组	流量	版本	说明	更新时间	更新人	操作
gbdv v2(0)	25%	1573469259251	gbdv v2	2019-11-11 18:47:39	刘妍(26020419)	查看详情 修改 版本记录
lr pctcvr(1)	25%	1573469326991	lr pctcvr	2019-11-11 18:48:47	刘妍(26020419)	查看详情 修改 版本记录
gbdv no encode 0(2)	25%	1573469147379	无	2019-11-11 18:45:47	刘妍(26020419)	查看详情 修改 版本记录
gbdv no encode 1(3)	25%	1573469147379	无	2019-11-11 18:45:47	刘妍(26020419)	查看详情 修改 版本记录

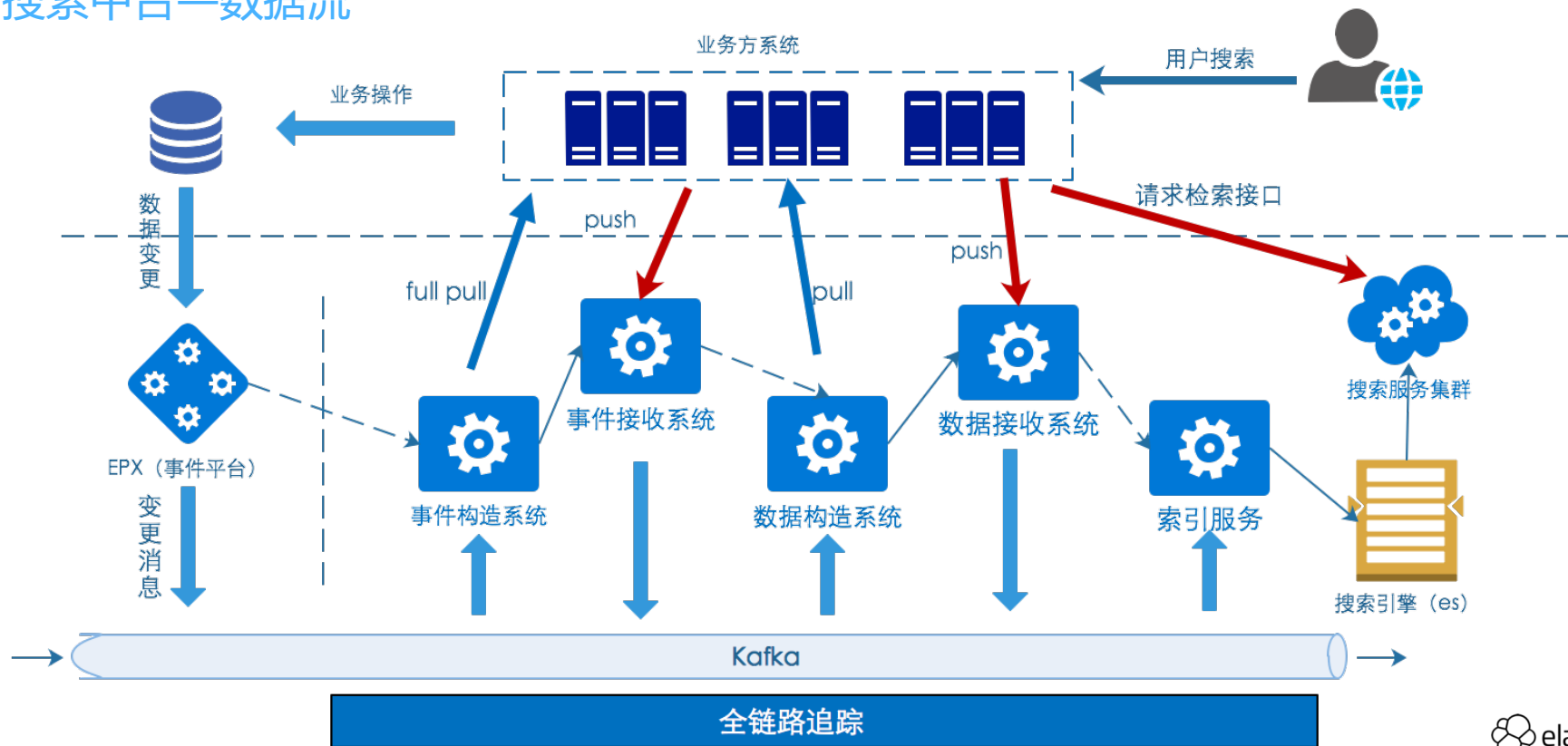
运行中模型

待审核模型

项目名称	业务线	模型名称	环境	描述	平台名称	部署路径	版本策略	部署状态	创建人	操作
ershov-sug-rank	搜索二手	tf-gbdv-v1	prod	二手app sug 排序	tensorflow	/ershov_sug_ran...	指定版本生效:2019...	● 未部署	李瑞东(23101559)	接口校验
cute-shell-v3	小贝	search-house-gb...	prod	小贝助手搜索找房模...	tensorflow	/cute_shell/cut...	指定版本生效:2019...	● 已部署	王文彬(23025260)	接口校验
er-xcv-filter-v...	搜索二手	gbdv	prod	增加selected_fil...	tensorflow	/ershov_xcv_fil...	指定版本生效:2019...	● 已部署	赵群(26041077)	接口校验

贝壳搜索中台演化之路

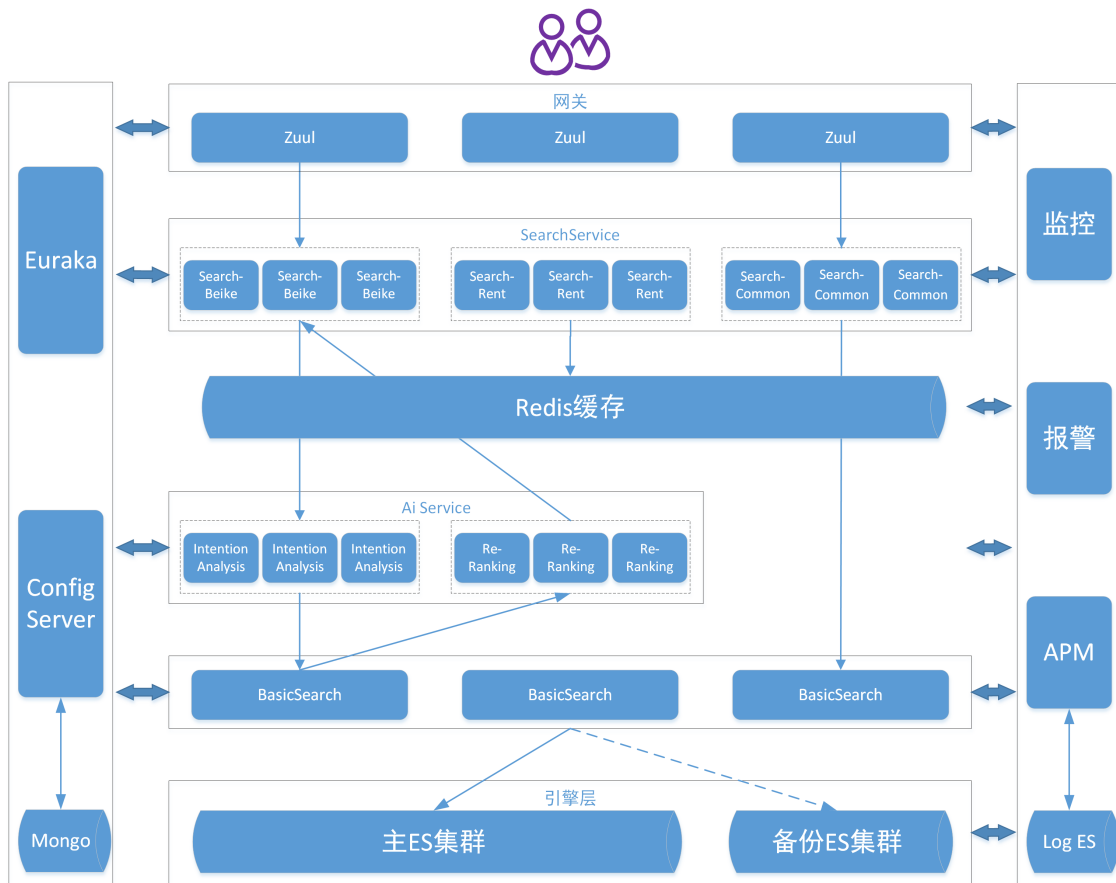
搜索中台—数据流



贝壳搜索中台演化之路

搜索中台—查询层

Spring Cloud



贝壳搜索中台演化之路

搜索中台—引擎层

SolrCloud -> ES

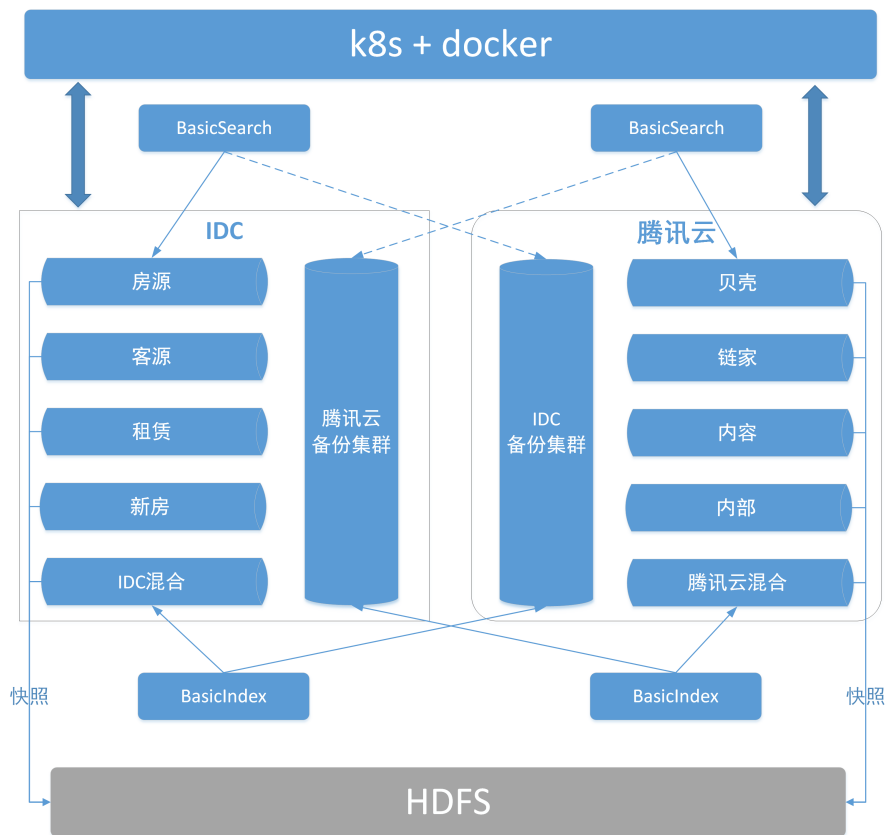
单集群 -> 多集群

单机房 -> 双机房

混布 -> 分组隔离

物理机部署 -> k8s + docker

部署效率提升10倍，资源使用率提升3倍



提纲

- 1 ES在贝壳的使用情况
- 2 贝壳搜索中台演化之路
- 3 稳定性99.99%的挑战
- 4 未来规划

稳定性99.99%的挑战

SLA	宕机时间/年	宕机时间/月	宕机时间/周	宕机时间/天
90%	36.5天	72小时	16.8小时	2.4小时
99%	3.65天	7.20小时	1.68小时	14.4分钟
99.9%	8.76小时	43.8分钟	10.1分钟	1.44分钟
99.99%	52.56分钟	4.38分钟	1.01分钟	8.66秒
99.999%	5.26分钟	25.9秒	6.05秒	0.87秒

稳定性99.99%的挑战

难点

业务敏感

- 100+ 一级核心业务
- 搜索故障超过**5分钟**即可能造成B级故障（严重）
- 超过20分钟升级为A级（重大）

流量变化

- 流量大小变化
- 业务查询模式变化
- 流量结构变化

服务过多

- 600+微服务
- 依赖复杂，牵一发而动全身
- 定位困难，运维成本高

稳定性99.99%的挑战

解决方案

举几个优化点：

1、平滑上线：

微服务下线先从Eureka摘除节点，上线等待预热成功后再注册到Eureka。优化后高峰上下线0报错。

2、配置刷新优化：

优化spring cloud配置刷新机制，预先初始化上下文后替换。防止重新加载配置时的短暂停顿。

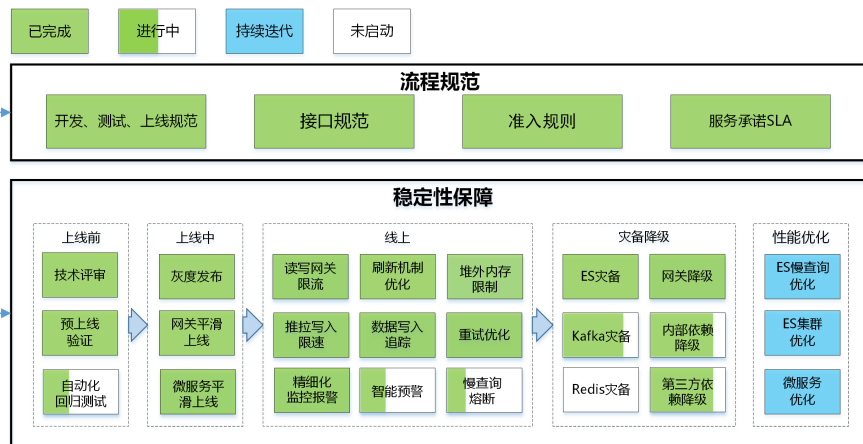
3、重试优化：

层层重试会造成流量翻倍。优化ribbon重试机制，过滤读超时重试，优化连接断开重试。优化后高峰任意kill进程0报错。

4、推拉写入限速：

防止业务高峰期刷数据影响查询。

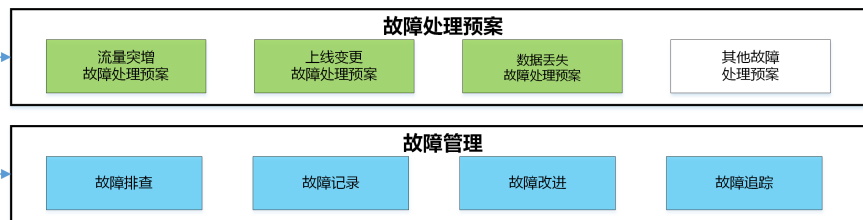
事前



事中



事后



稳定性99.99%的挑战

监控报警

fang-c03搜索监控 BOT

域名异常通知

时间: 2019-12-03 22:51 全天第1371分钟

域名: fang-c03-gw.search.ke.com

错误分钟级统计

22:514xx: 0.6029%5xx: 0.0032%总量: 30850 异常计数: {"499":186,"500":1}

22:504xx: 0.0000%5xx: 0.0000%总量: 32032 异常计数: null

22:494xx: 0.0000%5xx: 0.0000%总量: 34077 异常计数: null

出错明细:

22:51/api/1013901006/search/4xx: 0.2817%5xx: 0.0000%总量: 3550 异常计数: {"499":10}

22:51/api/1013901005-dup/search4xx: 0.5800%5xx: 0.0000%总量: 13277 异常计数: {"499":77}

22:51/api/1013901005/search/4xx: 0.6409%5xx: 0.0000%总量: 5773 异常计数: {"499":37}

22:51/api/1013901007/search/4xx: 0.6590%5xx: 0.0000%总量: 1214 异常计数: {"499":8}

22:51/api/1013901003/search/4xx: 0.7746%5xx: 0.0000%总量: 6842 异常计数: {"499":53}

22:51/api/1013901009/search/4xx: 2.6316%5xx: 0.0000%总量: 38 异常计数: {"499":1}

22:51/api/1013902011/search4xx: 0.0000%5xx: 3.7037%总量: 27 异常计数: {"500":1}

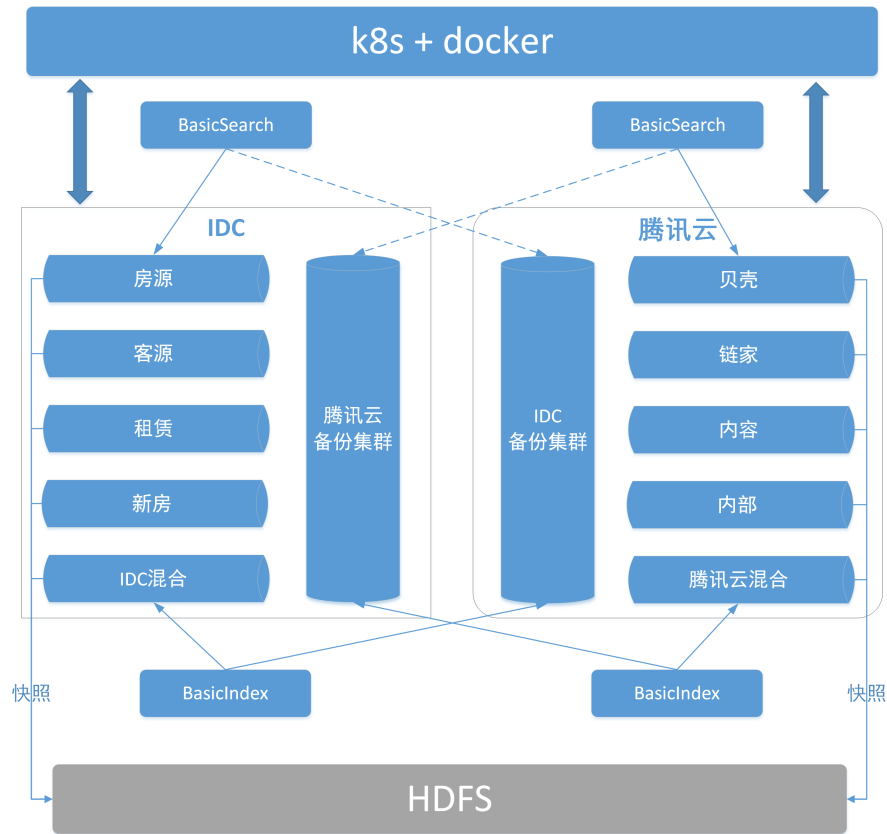
报警规则

状态规则: 连续3分钟, 4xx率超过0.5000% 4xx个数大于30, 或者5xx率超过0.0100% 5xx个数大于5, 触发电话报警

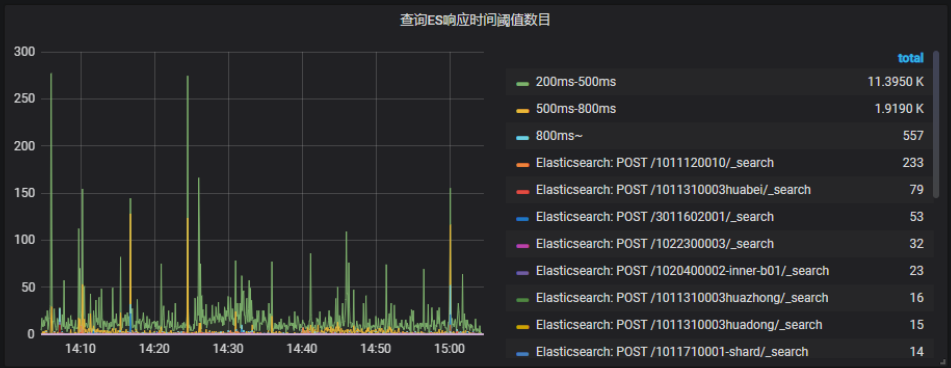
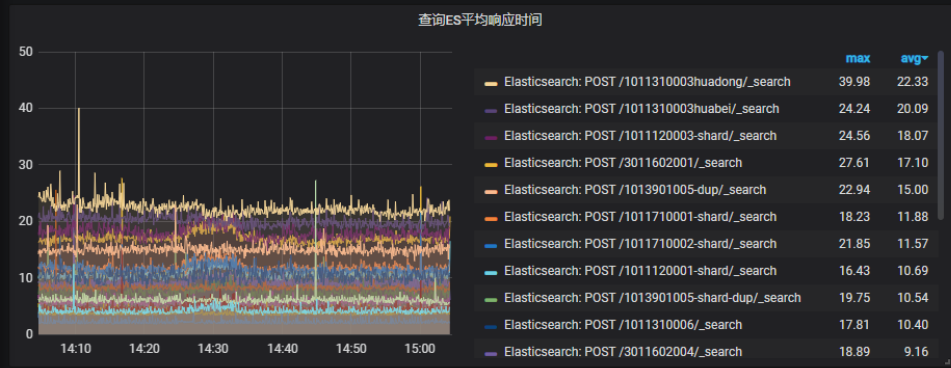
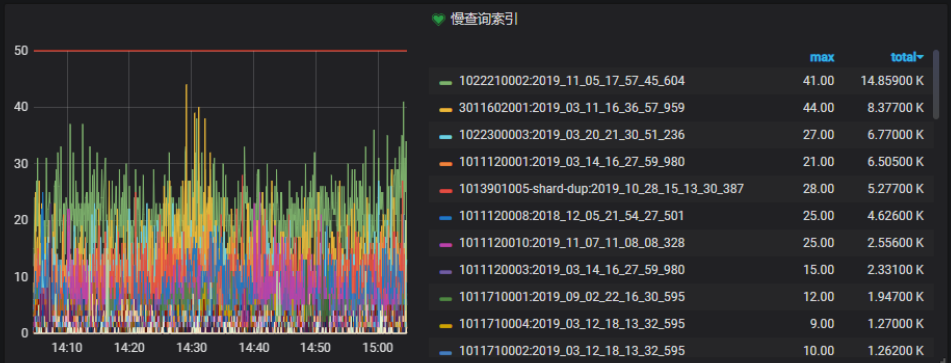
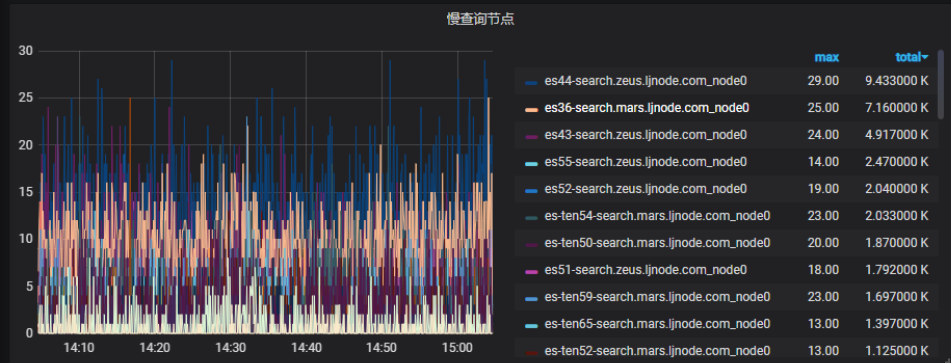
稳定性99.99%的挑战

ES引擎稳定性保障

- 1、双集群双机房备份
- 2、查询自动熔断降级
- 3、主备集群一键分流
- 4、故障节点自动迁移
- 5、每天定时快照备份
- 6、基础监控报警
- 7、慢查询监控报警
- 8、慢查询统计巡检
- 9、慢查询预测熔断



ES-Slowlog

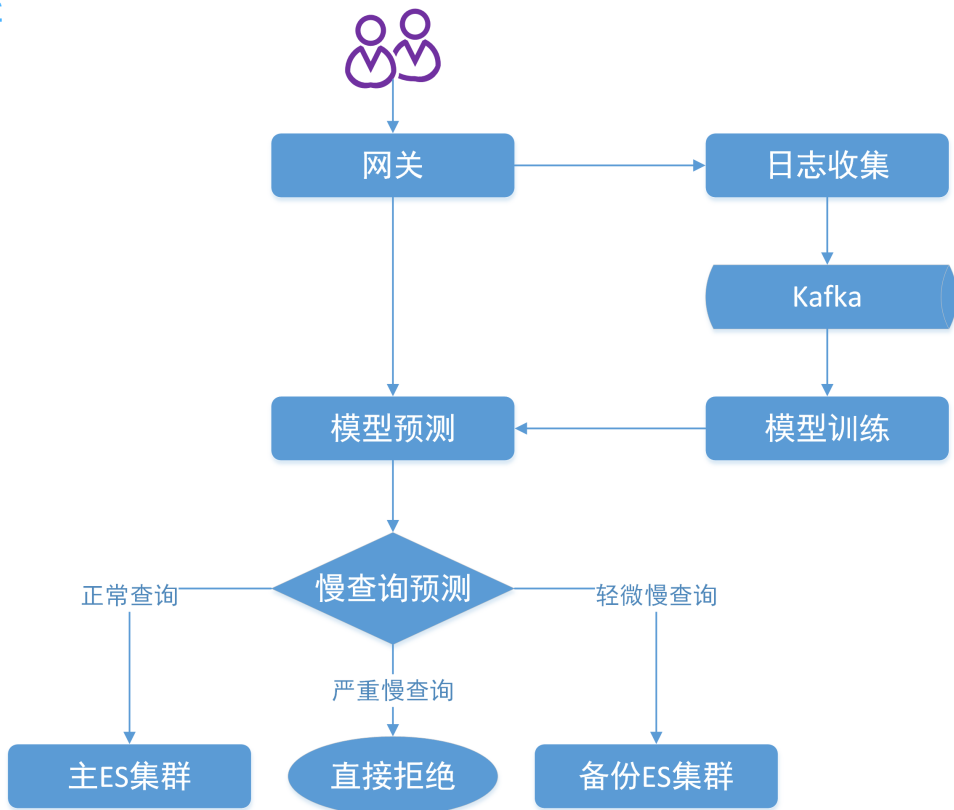


> APM流程时间 (6 panels)

> ES状态监控 (30 panels)

稳定性99.99%的挑战

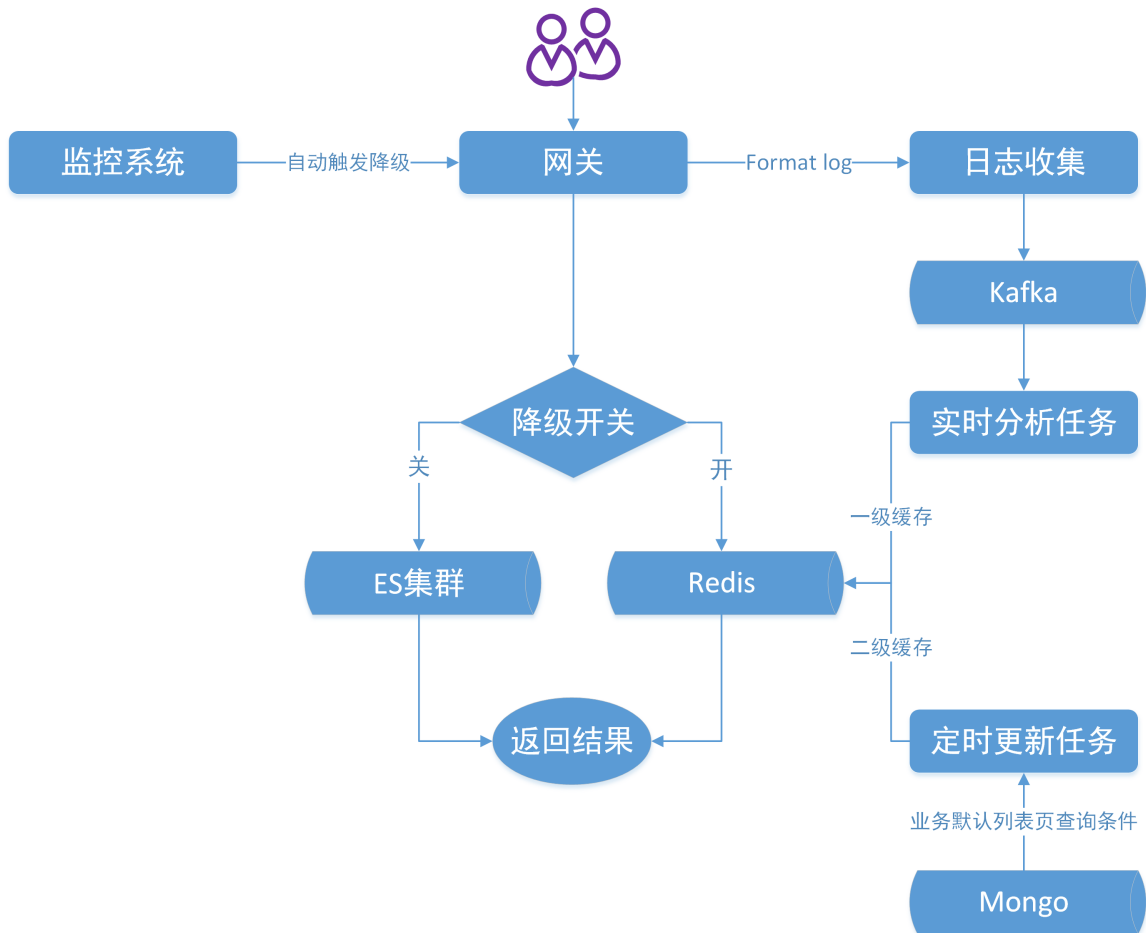
慢查询预测熔断



稳定性99.99%的挑战

网关自动降级

- 1、手动 & 自动
- 2、单业务 & 全分组
- 3、默认连续3分钟报错比例超过10%，自动打开网关降级
- 4、时间、阈值动态调整
- 5、一级缓存返回历史查询结果
- 6、二级缓存返回默认结果
- 7、C端用户基本无感知
- 8、争取故障修复时间，影响降到最低



稳定性99.99%的挑战

总结

监控报警

- 重视监控报警，专人专项
- 精准报警，提前预警

墨菲定律

- 任何可能出现的故障都必然出现

蝴蝶效应

- 配置审核，流程规范
- 重大变更，回滚策略

未雨绸缪

- 最坏打算，最好准备
- 风险预案，故障演练

包产到户

- 每个系统，每个子系统，每个模块，责任到人

技术修养

- 敬畏心，不犯低级错误
- 责任感，全面性

提纲

- 1 ES在贝壳的使用情况
- 2 贝壳搜索中台演化之路
- 3 稳定性99.99%的挑战
- 4 未来规划

未来规划

沉淀可定制的通用 搜索解决方案

