



# ELASTIC 深圳 MEETUP

🕒 2019年4月20日

📍 广东深圳市南山区高新科技园中区一路腾讯大厦2F多功能厅

<http://elasticsearch.cn>

# 阿里云Elasticsearch架构设计与性能优化实践

阿里巴巴搜索推荐事业部 欧阳楚才

# Agenda

1. 运维Elasticsearch的痛点
2. 阿里云Elasticsearch架构解析
3. Elasticsearch性能优化实践

# 阿里云Elasticsearch介绍



Elastic Stack

+



阿里云

商业插件	0部署成本	托管式智能运维
平滑扩缩容	跨机房高可用	100%兼容开源

# 阿里云Elasticsearch规模

3000+

集群数量

1w+

Node数

5PB+

数据量

# Elasticsearch运维的痛点

## 自建Elasticsearch服务成本



### 硬件成本

ECS、云盘、SLB、OSS...



### 运维成本

集群维护、配置维护、告警监控、日志查看...

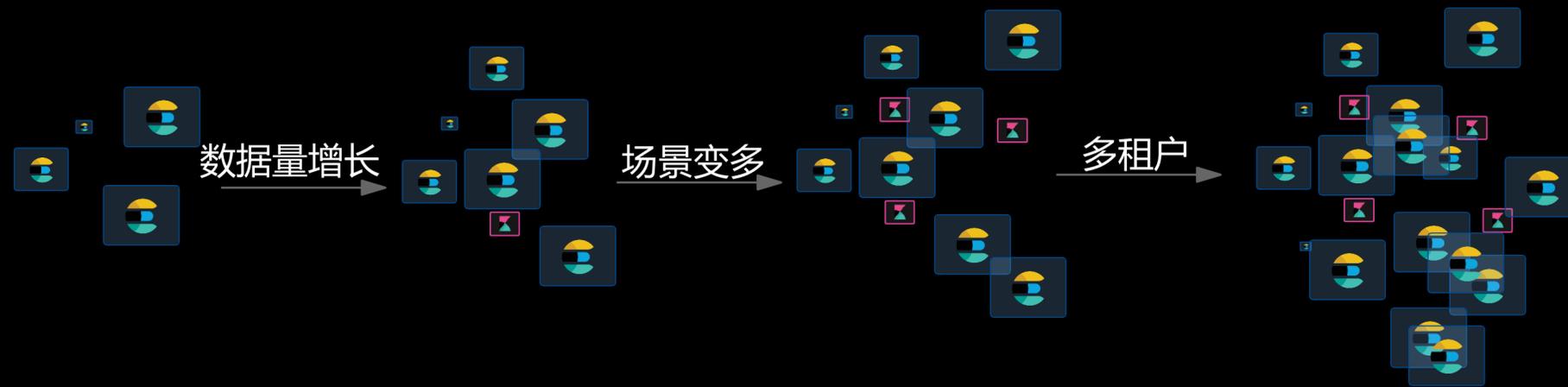


### 隐性成本

安全、备份、数据源、全增量、性能调优

# Elasticsearch运维的痛点

业务规模快速增长



# Elasticsearch运维的痛点

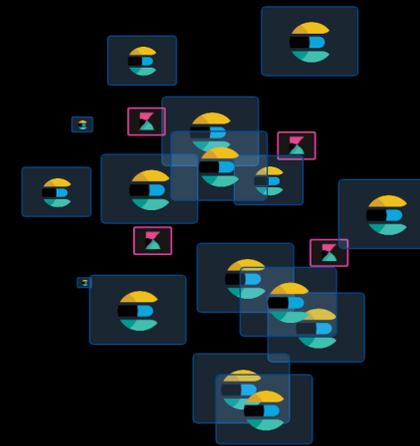
## 大集群需要解决的问题

### 多租户隔离

- Index or type
- 访问方式迥异
- 升级频率不同
- ...

### 运维需求

- 业务重要性不同，SLA不同
- 备份策略
- ES的版本升级需求
- 升级窗口不一致
- ...



## 免部署

基于容器编排和定制化OS

- DataNode、MasterNode...
- Os、jvm参数定制
- 15分钟拉起100个node

## 弹性扩容

规格、存储、节点数

- 基于rolling update、按需更新保证服务平滑

## 配置管理

Yml、词典、插件、黑白名单

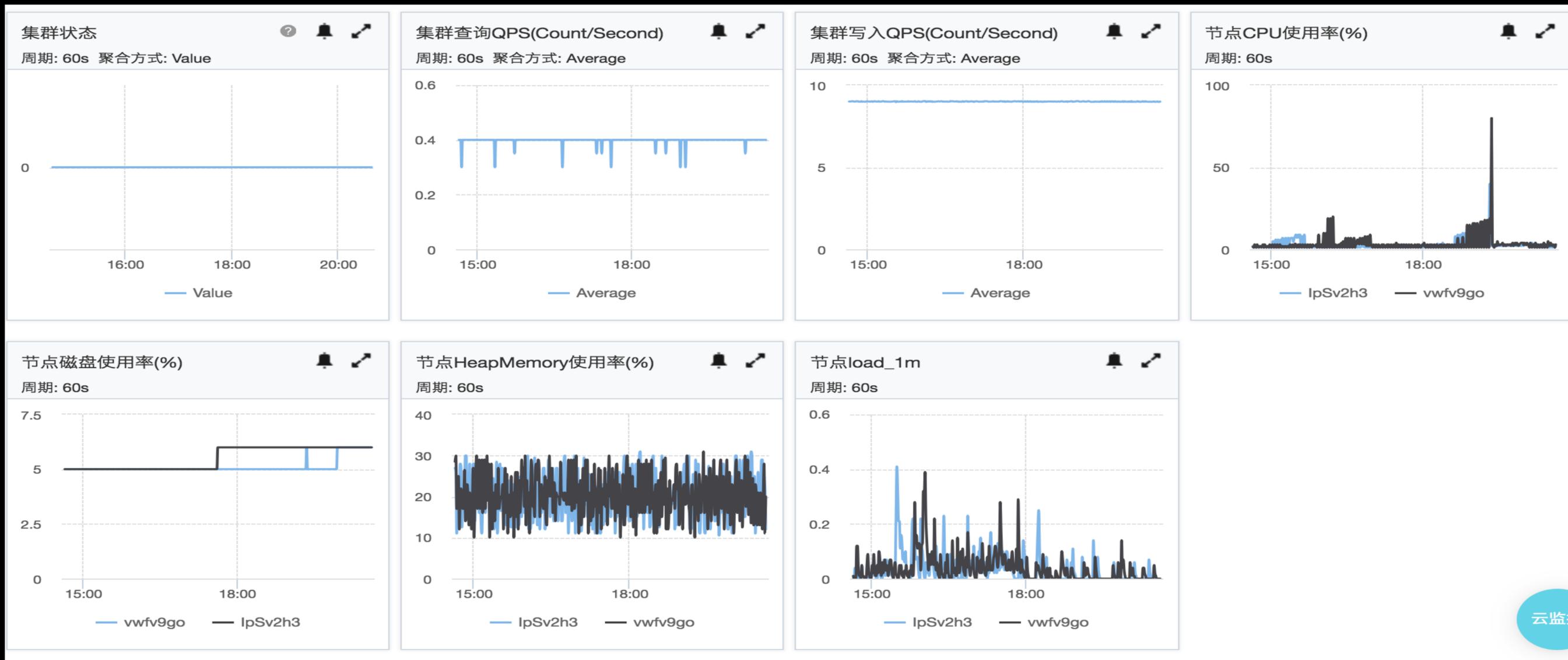
- 按需冷热更新

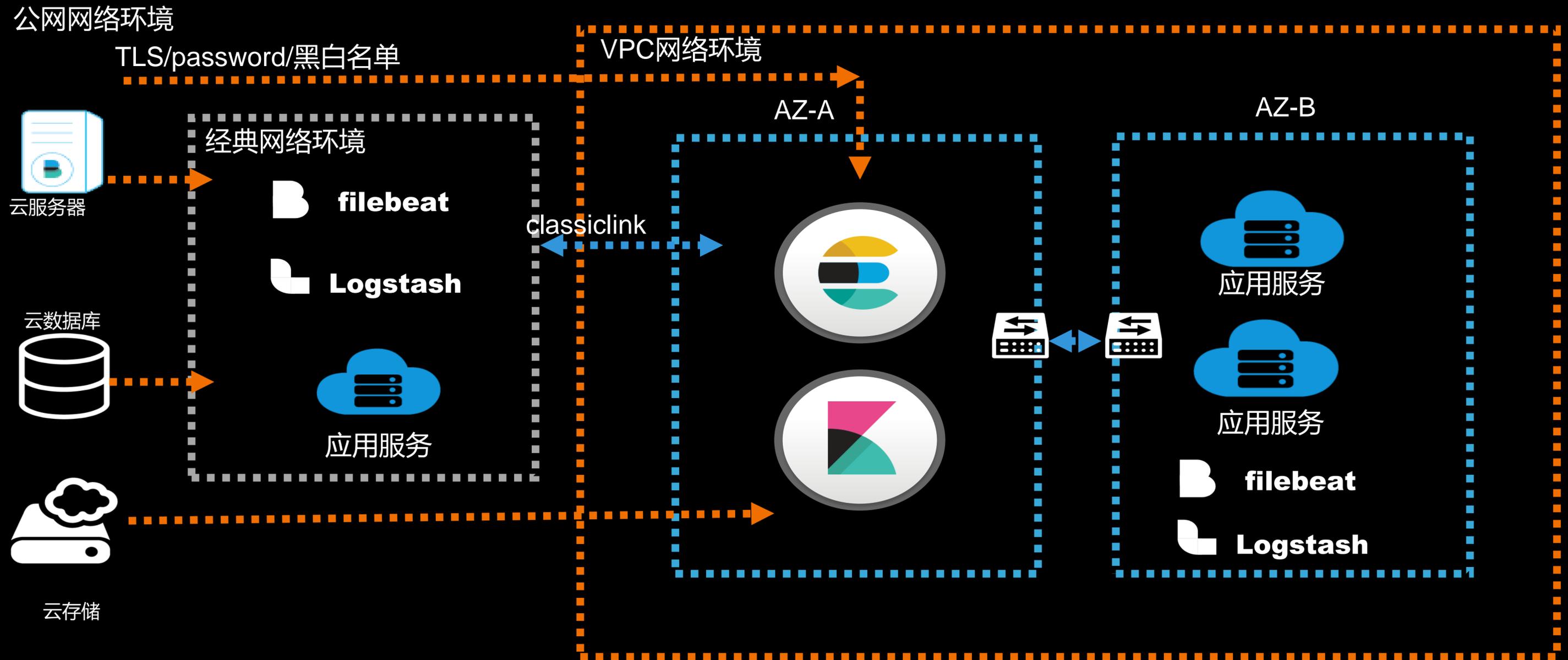
## 平滑升级

大小版本升级

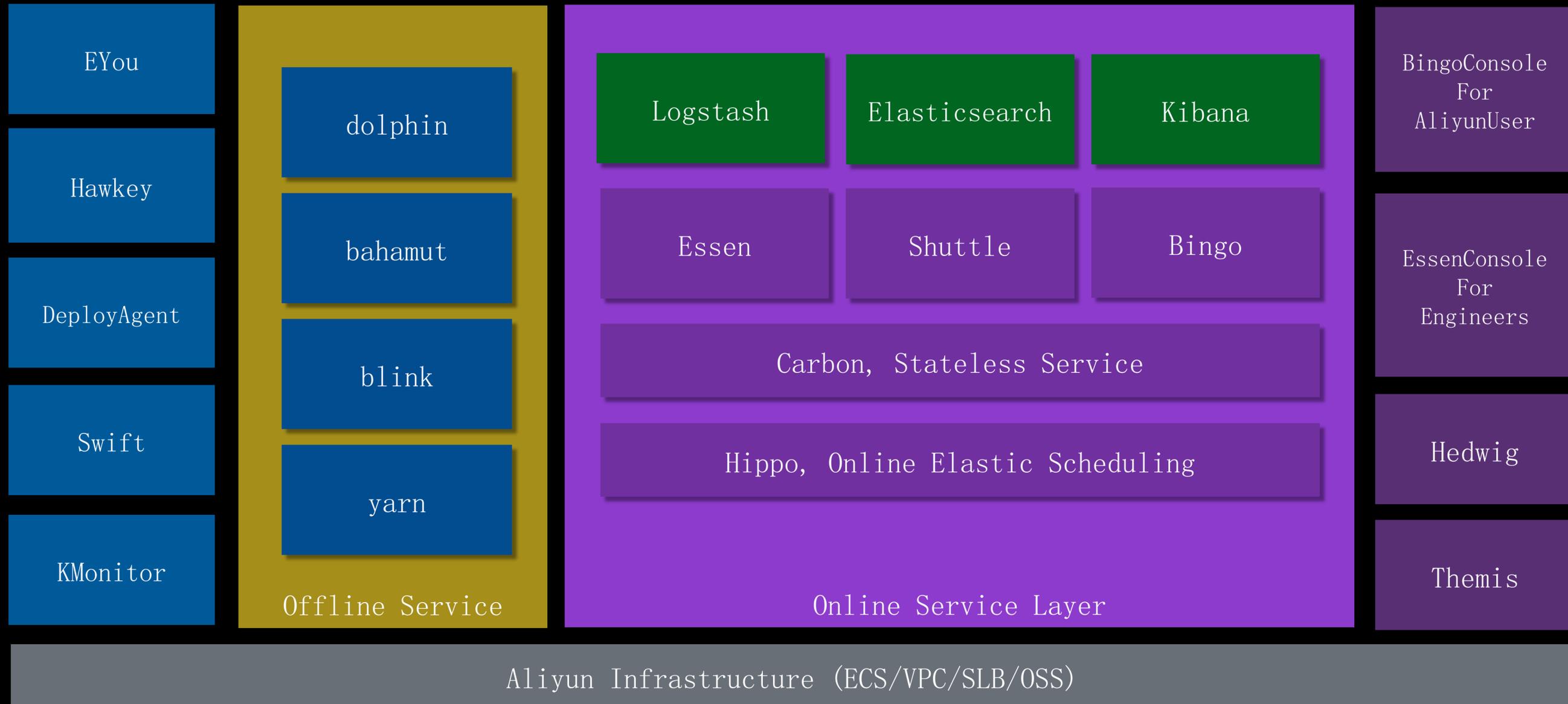
- 支持5.5/6.3/6.7。支持多版本升级

# 阿里云监控报警与Elasticsearch整合



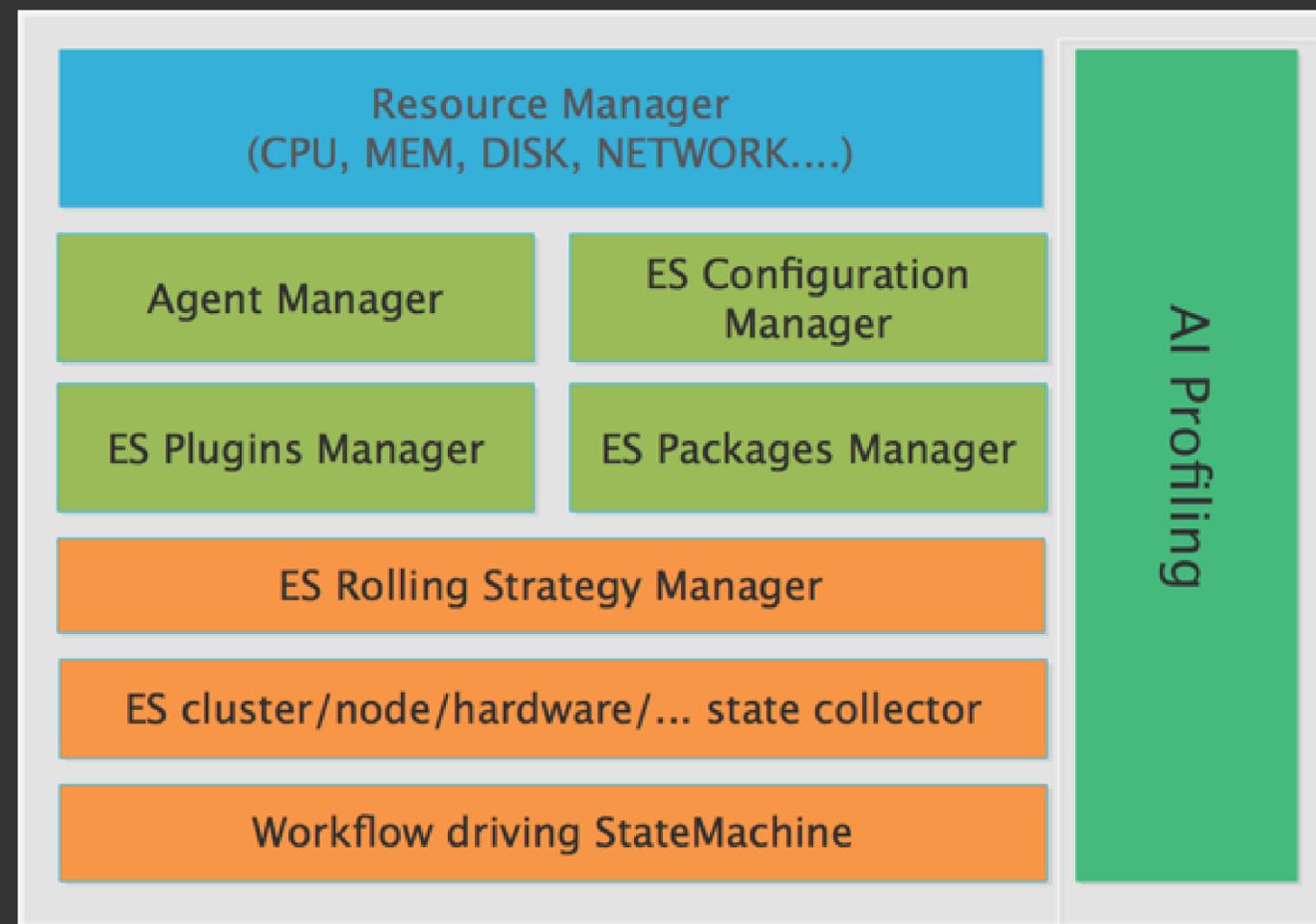


# 阿里云Elasticsearch产品架构



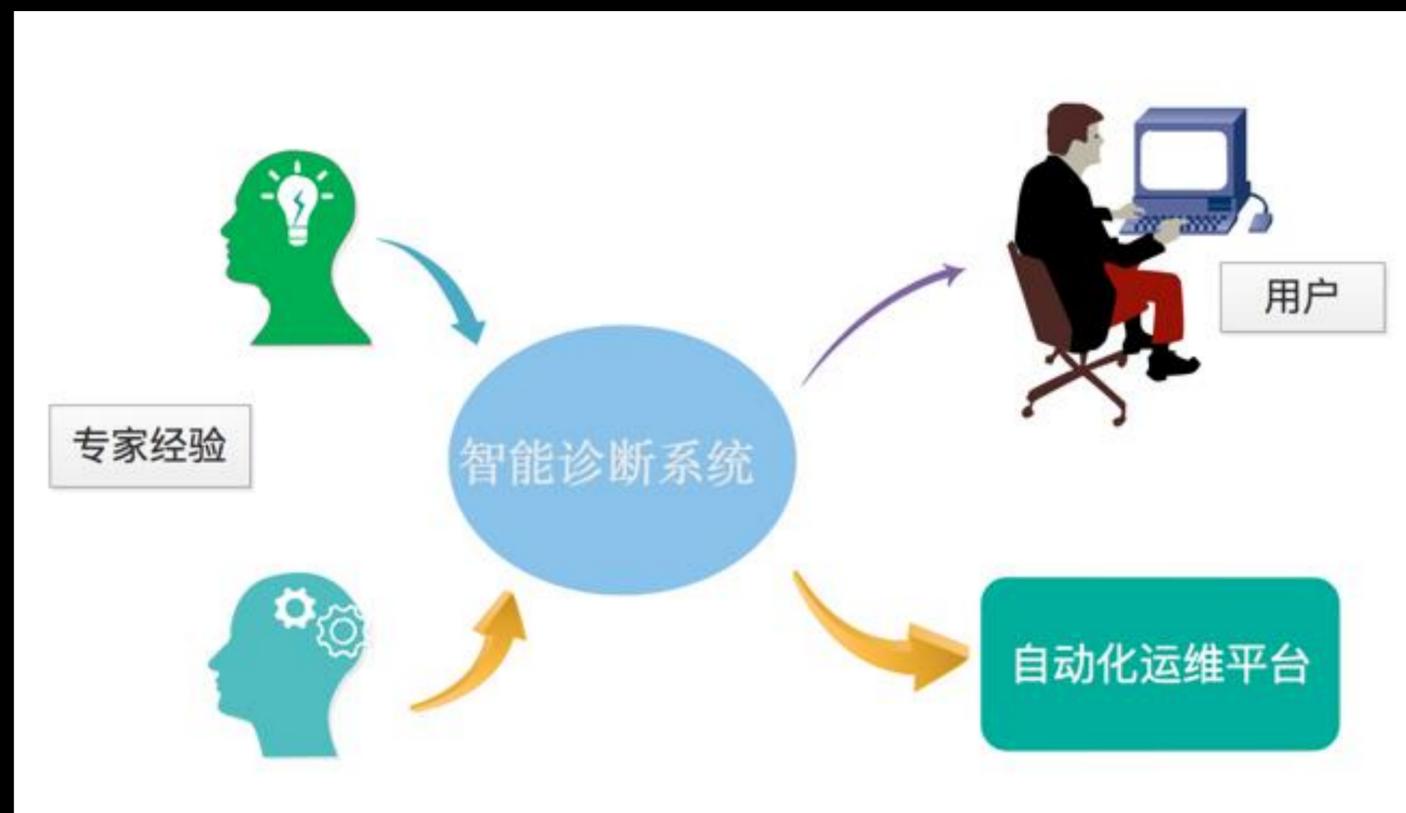
# Elasticsearch自动化运维核心架构

- 资源管理
  - 硬件管理, CPU, MEM, NET...
  - 软件管理(插件/YML/多版本5.5, 6.3和6.7)
- ES集群生命周期管理
  - Rolling策略决策管理
  - 集群, 节点, 硬件等状态收集管理
  - Rolling状态机
- Eyou: AI Profiling 智能分析



# • 阿里云Elasticsearch智能诊断系统Eyou

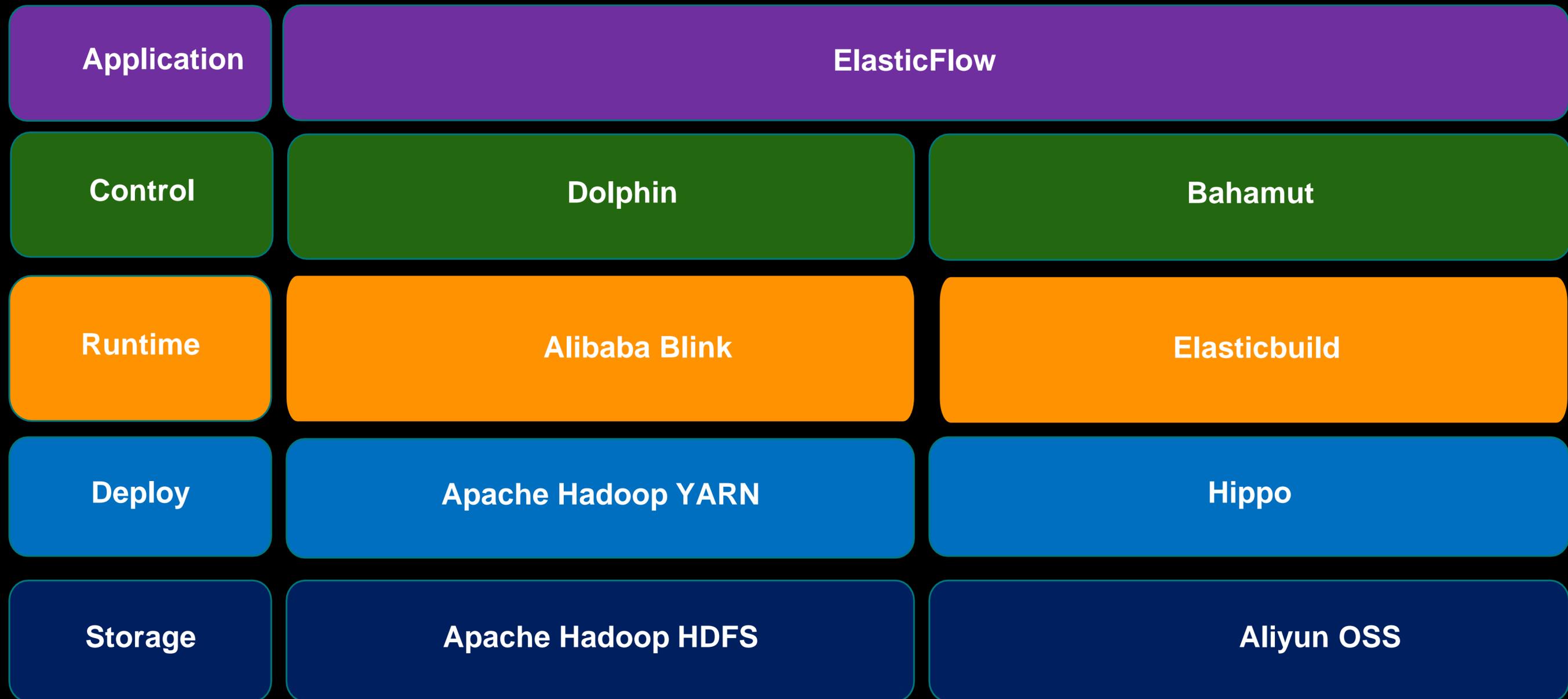
- 汇集专家经验沉淀
- 更全方面了解ES集群健康
- 寻找更佳的使用方式
- 智能化运维



诊断项是可以直接反馈ES  
集群某一个状态或行为是  
否合理的指标。



# 阿里云Elasticsearch实时计算平台Eflow



# Eflow的优势

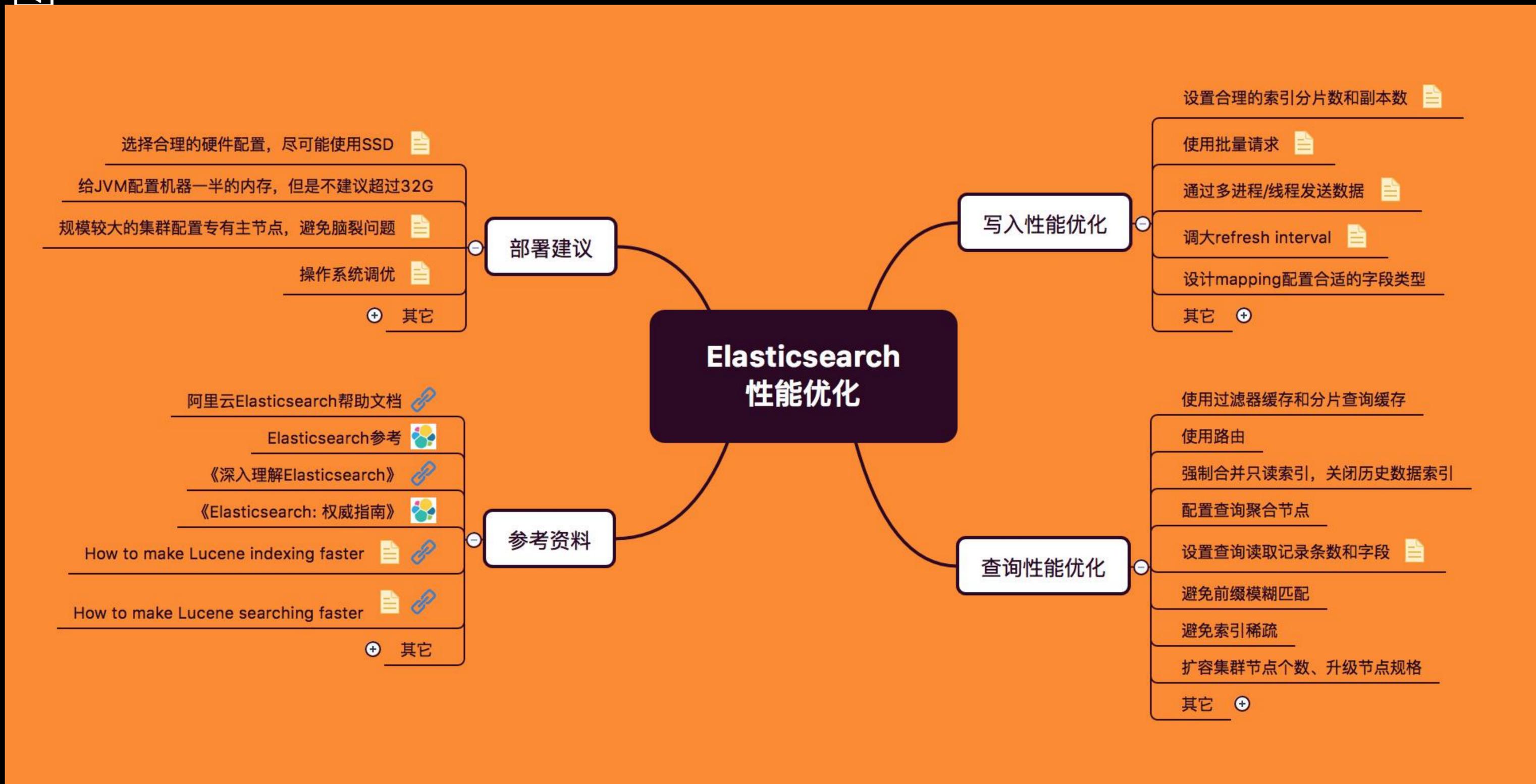
- 基于强大的阿里流计算Blink框架.
- 高效的离线数据处理, 流计算解决多表join.
- 支持多数据源导入至云ES, 比如自建mysql, RDS, ODPS...
- 离线索引构建性能提升4倍.
- 已内建丰富的处理数据算子. 推荐算法算子, 类目预测算子...

# Elasticsearch内核优化

- 硬件方向
  - Elasticsearch支持读写HDFS，分布式存储，高可用
- 软件方向
  - 系统架构
    - 规模弹性：当Shard数固定后，缺乏了弹性扩缩容的能力
    - 优化：将Shard数据进一步拆分，达到并发构建，随时扩缩集群规模的能力
  - Elasticsearch内核级别
    - Translog：Failover需要通过Translog回追数据，增加了大量IO消耗
    - 优化：利用Blink checkpoints机制替代Translog
    - 索引合并：索引合并的IO消耗巨大
    - 优化：通过内存Merge可以减少Merge索引的IO读写

# Elasticsearch性能调优思维导图

图



# Elasticsearch写入性能测试

环境:

Elasticsearch版本: 5.5.3, 3节点

规格: 2核4G、4核16G、16核64G

磁盘: 1T SSD云盘

压测工具:

esrally

rest api

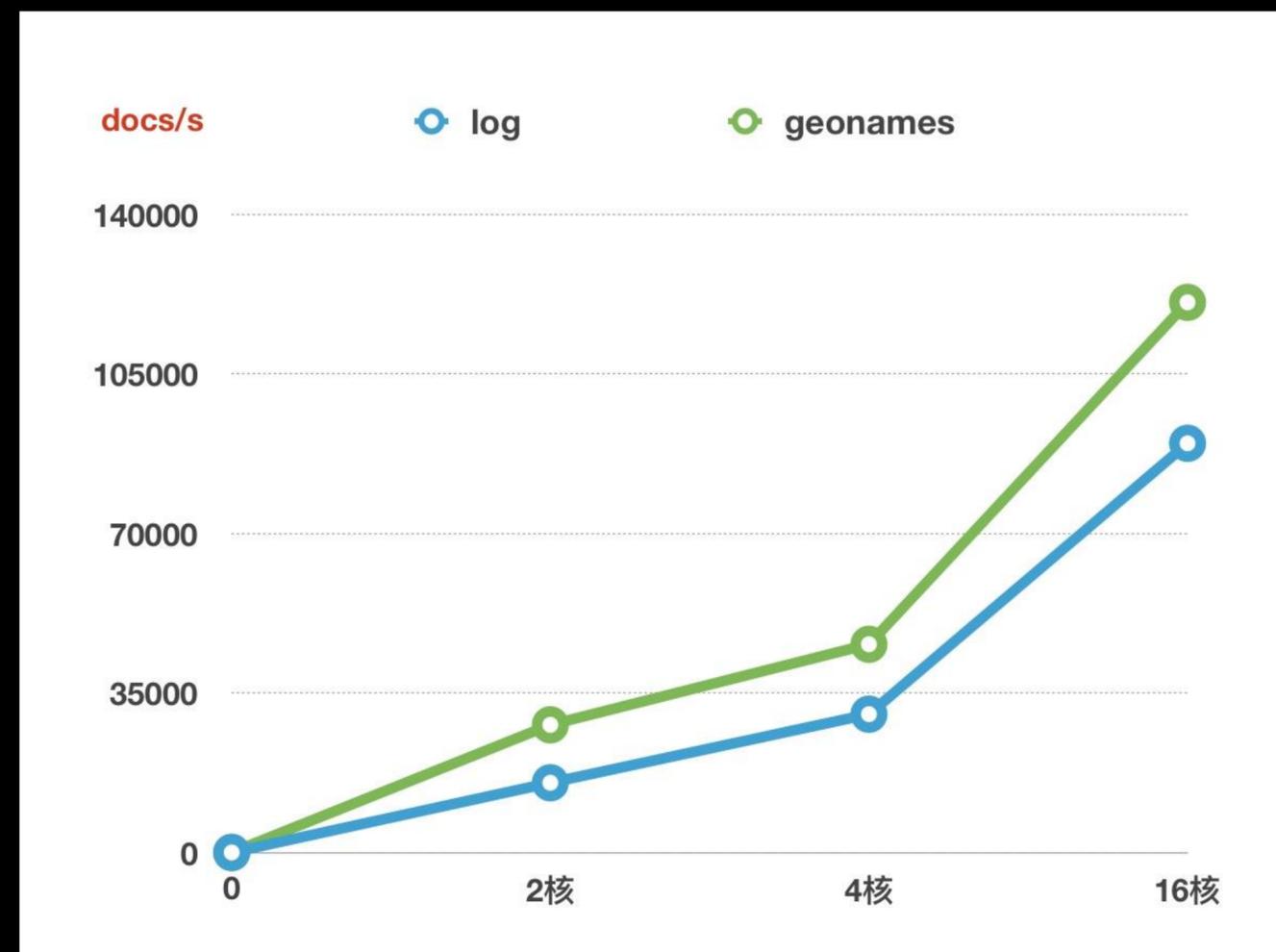
压测数据集:

esrally官方数据 geonames 3.3 GB, 单doc 311B

模拟某业务日志数据 80GB, 单doc 1432B

Log数据tps可以达到9W

Geonames数据tps达到12W



# Elasticsearch查询性能测试

在单机存储1TB数据场景下，SATA盘和SSD盘的全文检索性能对比（测试环境：Elasticsearch5.5.3，10亿条人口户籍登记信息，姓名作为查询关键词，单机16核CPU、64GB内存，12块6TB SATA盘，2块1.5 TB SSD盘）

磁盘类型	并发数	QPS	平均检索响应时间	50%请求响应时间	90%请求响应时间	IOPS
SATA盘	10并发	17	563ms	478ms	994ms	1200
SATA盘	50并发	64	773ms	711ms	1155ms	1800
SATA盘	100并发	110	902ms	841ms	1225ms	2040
SATA盘	200并发	84	2369ms	2335ms	2909ms	2400
SSD盘	10并发	94	105ms	90ms	200ms	25400
SSD盘	50并发	144	346ms	341ms	411ms	66000
SSD盘	100并发	152	654ms	689ms	791ms	60000
SSD盘	200并发	210	950ms	1179ms	1369ms	60000



欢迎交流(加微信好友建议备注姓名+单位)