

ES在AdMaster社会化 数据分析系统中的大规模实践

AdMaster(精硕科技)
宋兵强

自我介绍

- 百度基础架构部分分布式存储组
 - 贴吧、文库，移动云 —> 网盘
 - 技术、用户 —> 成功（数百PB，过万节点，RS coding, ARM, 大容量、高速、稳定）
 - 商业 —> it's hard to say...
- AdMaster 社会化数据系统研发

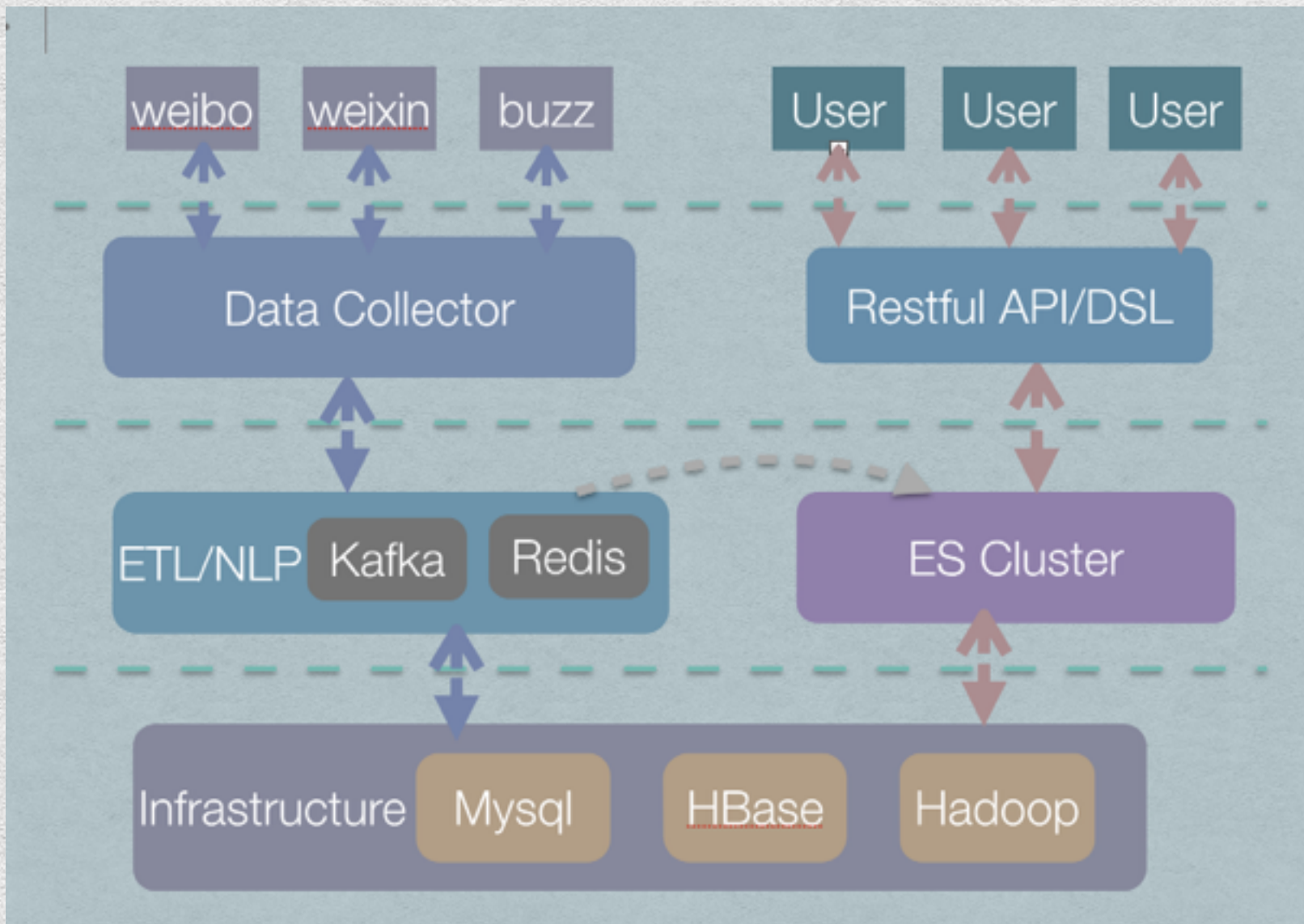
目录

- 数据
- 系统
- 索引
- 分词
- 查询
- 运维
- 思考

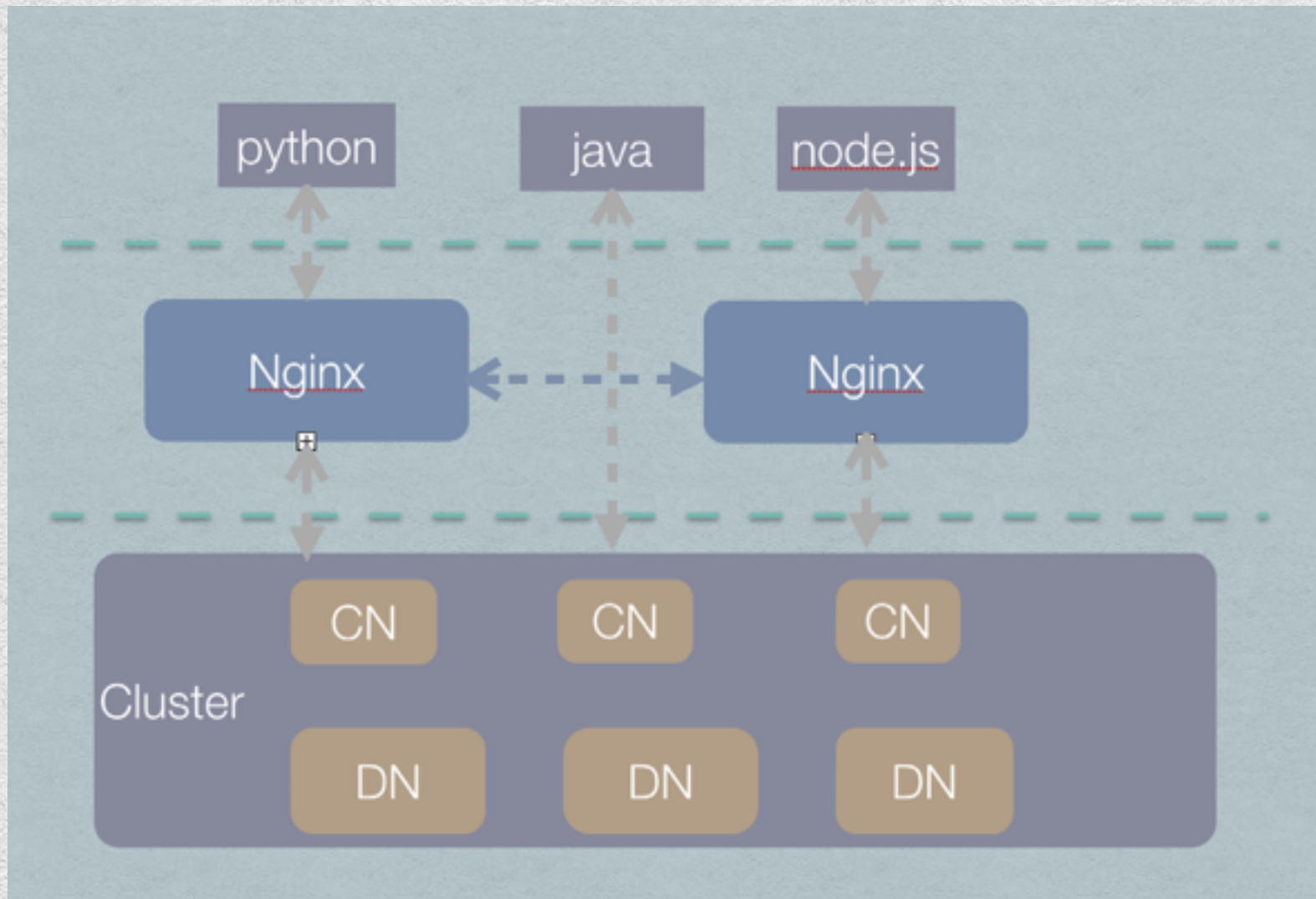
数据

- 2014.9 上线，20+节点
- 2015.7 跨地域迁移
- 接近100节点，千亿级 doc（微博，微信，电商，Buzz，广告DMP）
- 复杂查询过百条件
- 典型用户 —>解决问题、支撑业务、不断演化

系统



集群



索引

- 实时, python —> http
- 历史, hadoop —> transport
- 批量、并发写入, doc values, dynamic mapping
- 天然时间序列分库, alias
- groovy, upsert, 安全问题
- client library —> node.js, python, java, raw http
- 按需 open index
- 并发 Reindex Task System

分词

- ik, jieba
- 识别特殊字符
- 自定义词库
- 分词服务
- NLP, 情感判断

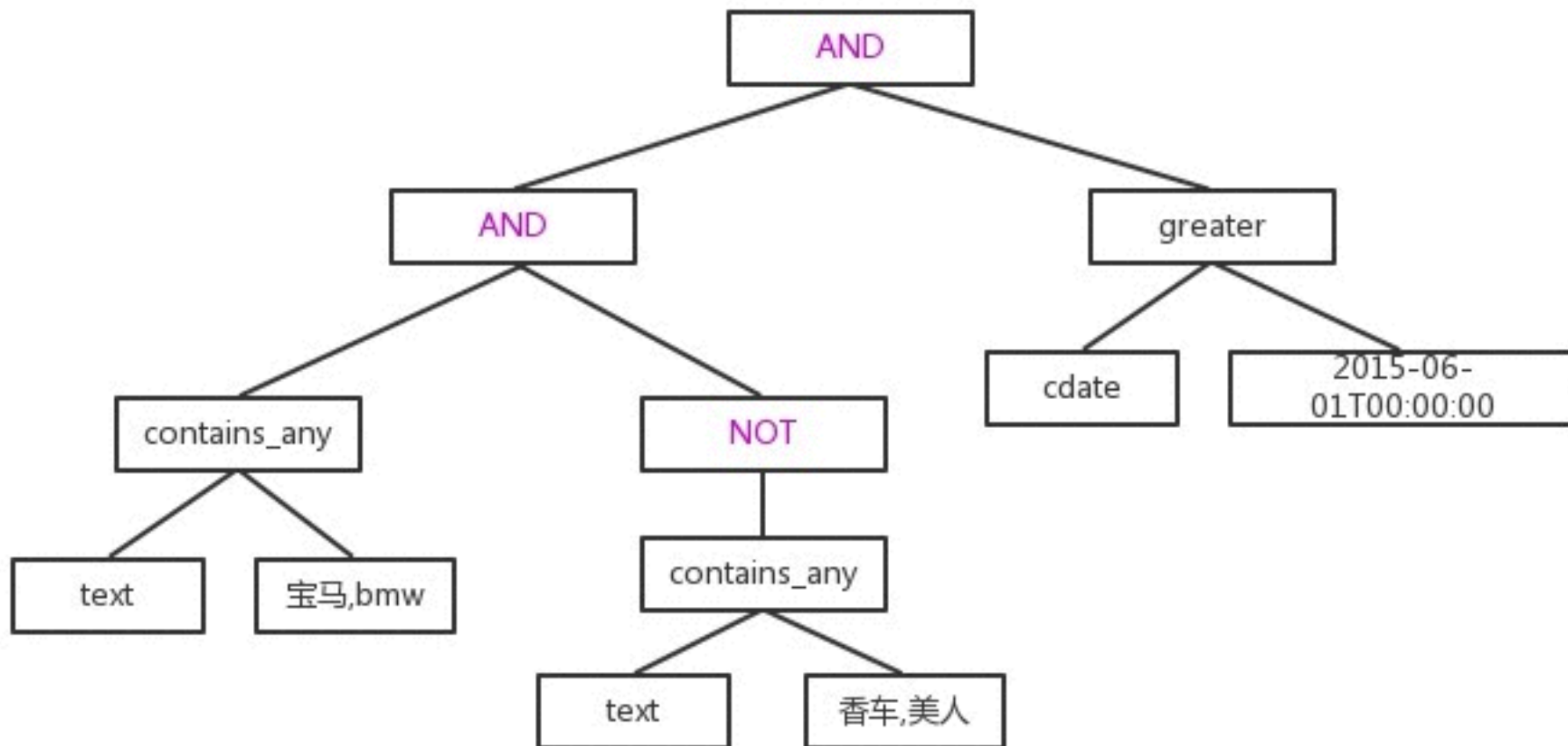
查询

- 查询缓存
- filter vs query string
- scan, Scroll
- **Restful API , DSL Parser**
- 语义 —> **contains, contains_all, contains_any, AND, NOT, LESS**

查询 — DSL 转换

- `weibo.text contains_any "宝马,bmw" AND not weibo.text contains_any "香车,美人" and weibo.cdate > "2015-06-01T00:00:00"`

查询 — DSL转换



查询 — DSL 转换

```
• {  
•   "query": {  
•     "filtered": {  
•       "filter": {  
•         "bool": {  
•           "should": [],  
•           "must_not": [  
•             {  
•               "query": {  
•                 "query_string": {  
•                   "query": "\"香车\" OR \"美人\"",  
•                   "default_field": "text"  
•                 }  
•             }  
•           ]  
•         }  
•       }  
•     ],  
•   }
```


运维 — 1

- heap size, gc, 32G 限制
- max open files
- memlock
- **plugin —> bigdesk/kopf/head**
- **Docker、Ansible**

运维 — 2

- type problem
- max condition num
- **multicast problem**
- **shard allocate failed**
- **elk —> logstash/kibana4**

Thinking —— 规模1

- 数百，上千节点扩展
- leadership election
- P2P → Master ?

Thinking —— 规模2

- Auto Recovery —> speed, discovery
- Balance —> node/disk/index/mapping —> 异构集群
- HA —> Primary/Multi Secondary —> Rolling Update
- Monitor —> 更全, 更细
- 资源隔离 —> 多租户
- Lucence、Ssd

Thinking —— 模型

- OLAP
- 一致性
- 事务
- 部分取代传统db —> Mysql计划迁移
- 分布式数据库?
- 多引擎

Thinking —— 生态

- 非常易用
- 易于推广
- 更适合中小企业
- 打造 生态系统
- vs Hadoop —> 计算服务迁移
- Pinot, Drill, etc... —> 选型
- 发掘数据价值！

The Mists of Time

- Many years ago, a newly married unemployed developer called **Shay Banon** followed his wife to London, where she was studying to be a chef. While looking for gainful employment, he started playing with an early version of Lucene, with the intent of building his wife a recipe search engine.
-
- Shay's wife is still waiting for the recipe search...

Q & A

- 谢谢!