




Elasticsearch在转转公司的千亿规模深度应用实践

张化仁
转转



战略级赞助商  HUAWEI

钻石级赞助商  普翔

白金级赞助商  华夏博格

 神州数码
Digital China

金牌级赞助商  iDataAPI

合作伙伴  开源中国
oschina.net

 掘金

 知乎

 IT大数说

 otpub

 Broadview
www.broadview.com.cn

 百格活动
bagevent.com

 MAXHUB
高效会议平台



Elasticsearch 在转转的深度应用实践

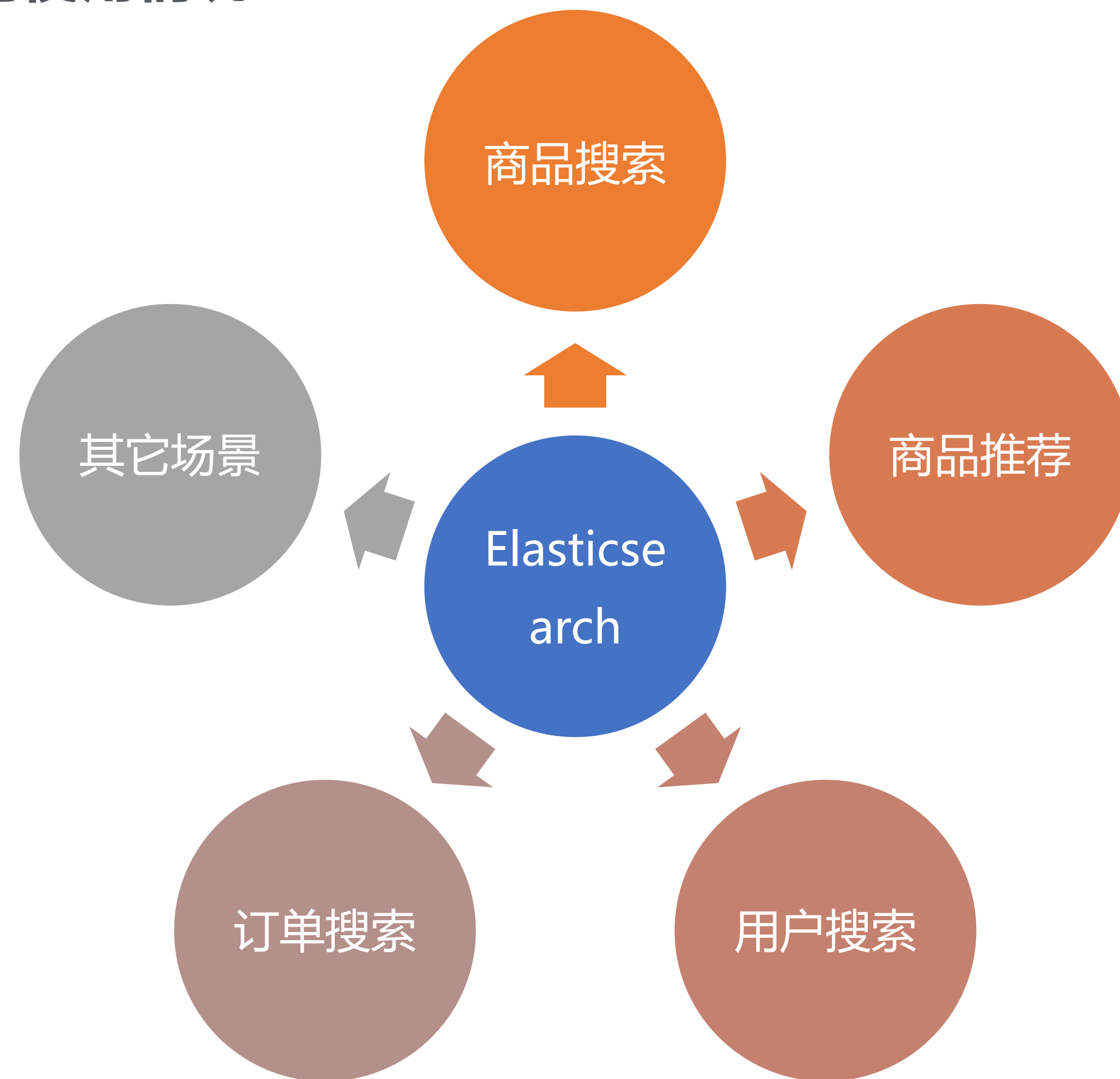
张化仁 @ 转转

目录

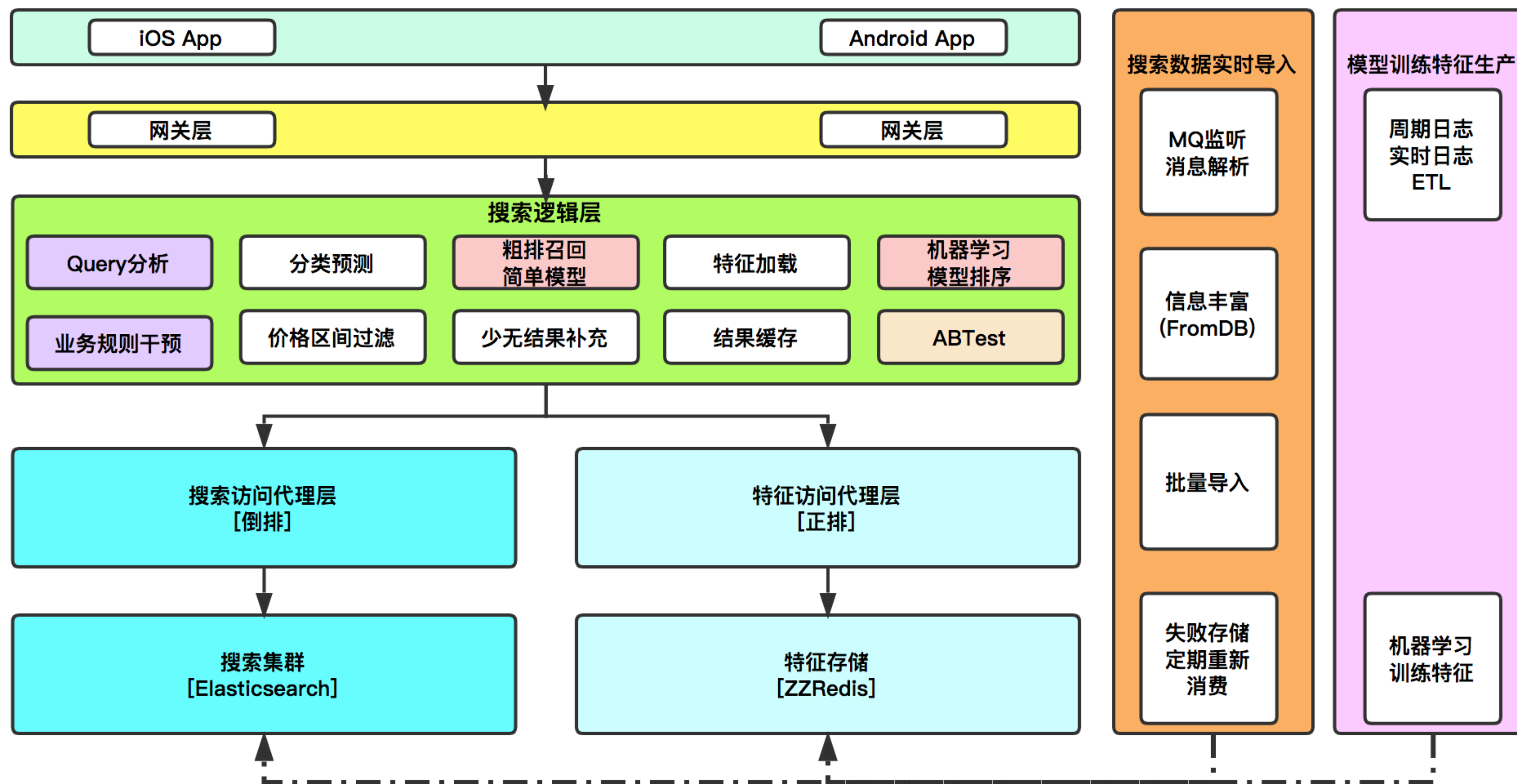
1. 转转简介
2. Elasticsearch在转转的使用情况
3. Elasticsearch在转转的架构实践
4. Elasticsearch在转转的应用实践
5. 转转Elasticsearch体系未来规划



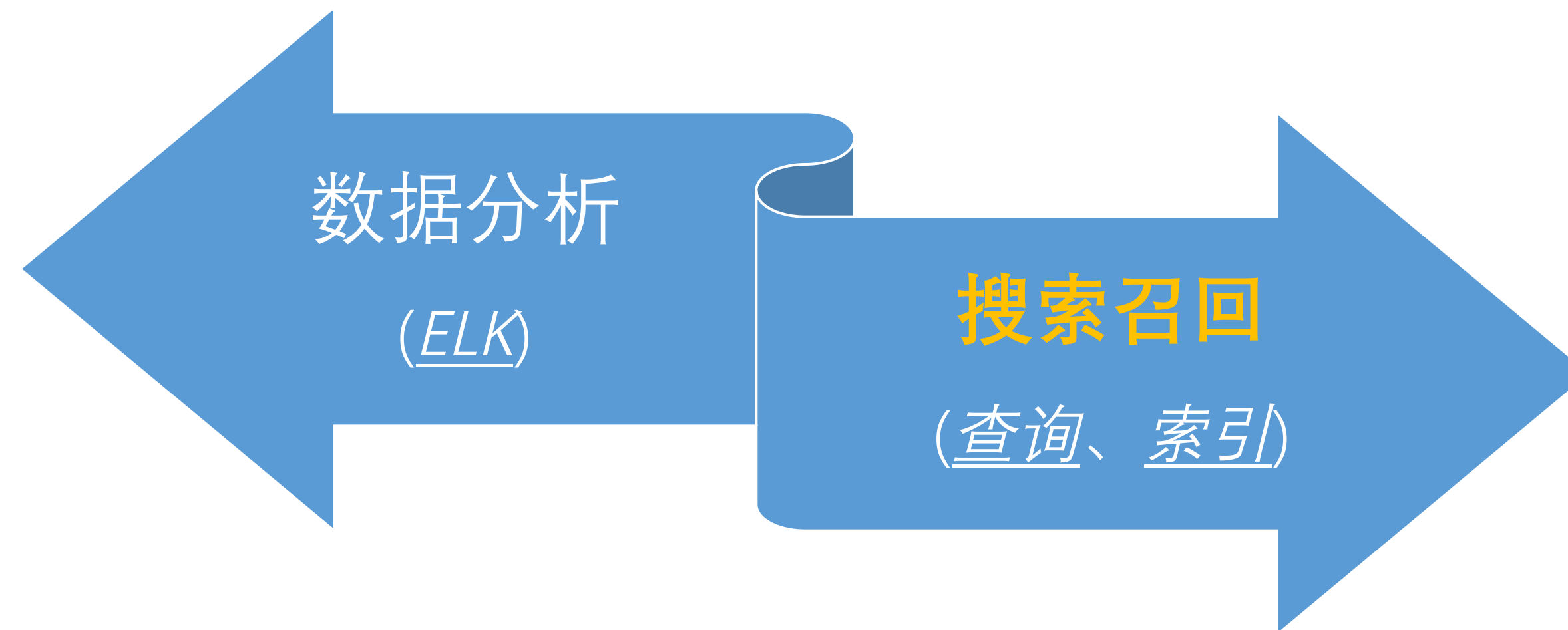
Elasticsearch在转转的使用情况



Elasticsearch在转转的架构实践



Elasticsearch的主要应用场景



转转搜索查询场景 & 如何借助Elasticsearch解决实现

Elasticsearch在转转的应用实践 - 查询 - 场景举例

搜索

用户输入关键词，选择分类、城市、价格等筛选条件，选择按默认、价格、时间、距离排序，返回匹配的商品

推荐

根据用户兴趣优先召回满足用户兴趣的商品

运营

自由组装搜索筛选条件(与或非)，过滤出目标商品集

附近

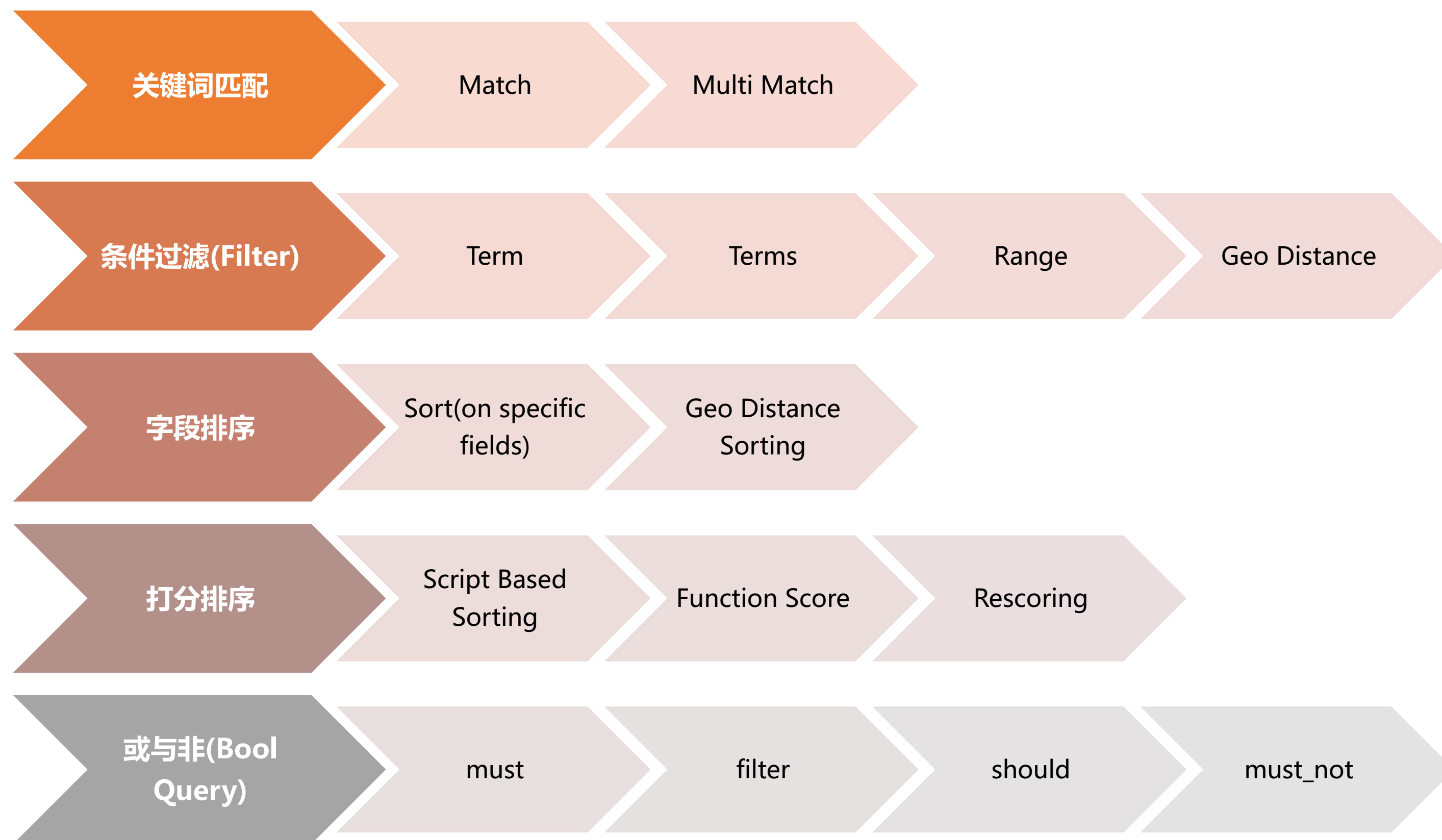
召回附近的商品

排序

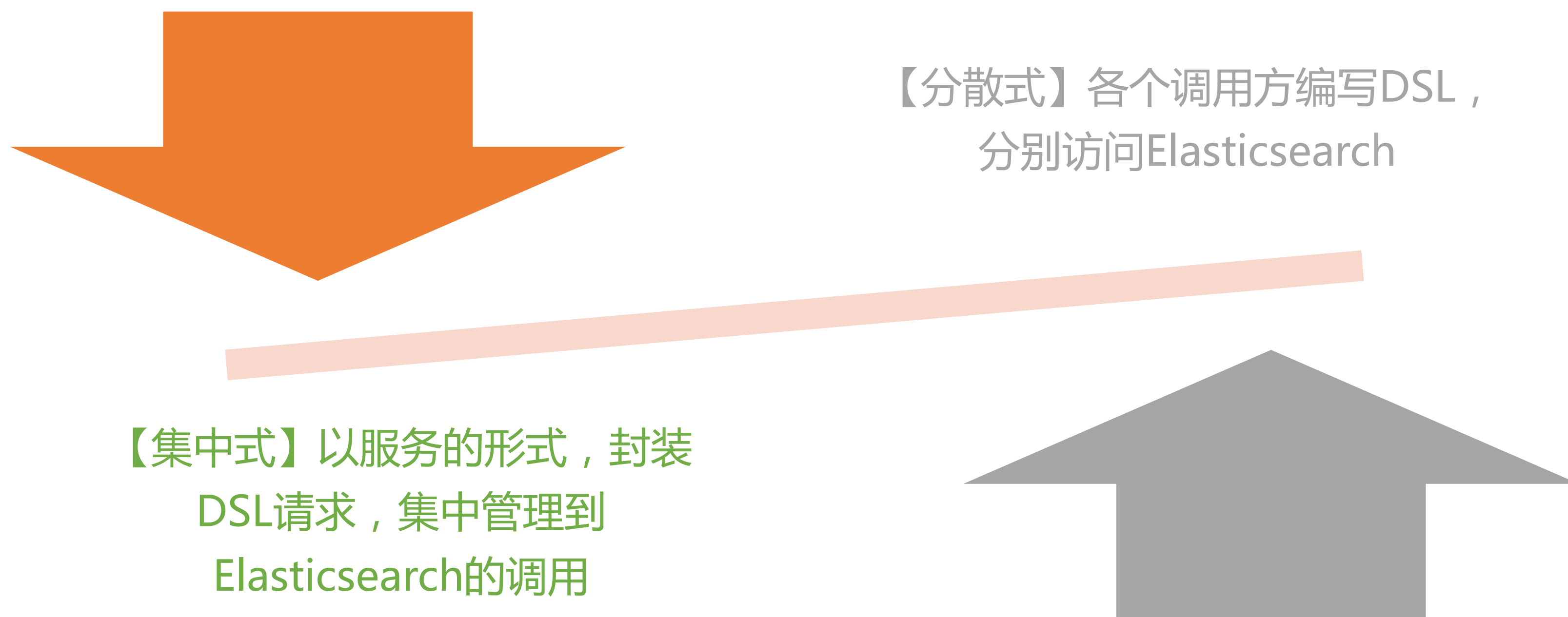
需要兼顾「商品新鲜度」、「商品质量」、「卖家信誉」、「匹配程度」等多种因素

关键词匹配、条件过滤、字段排序、打分排序、与或非

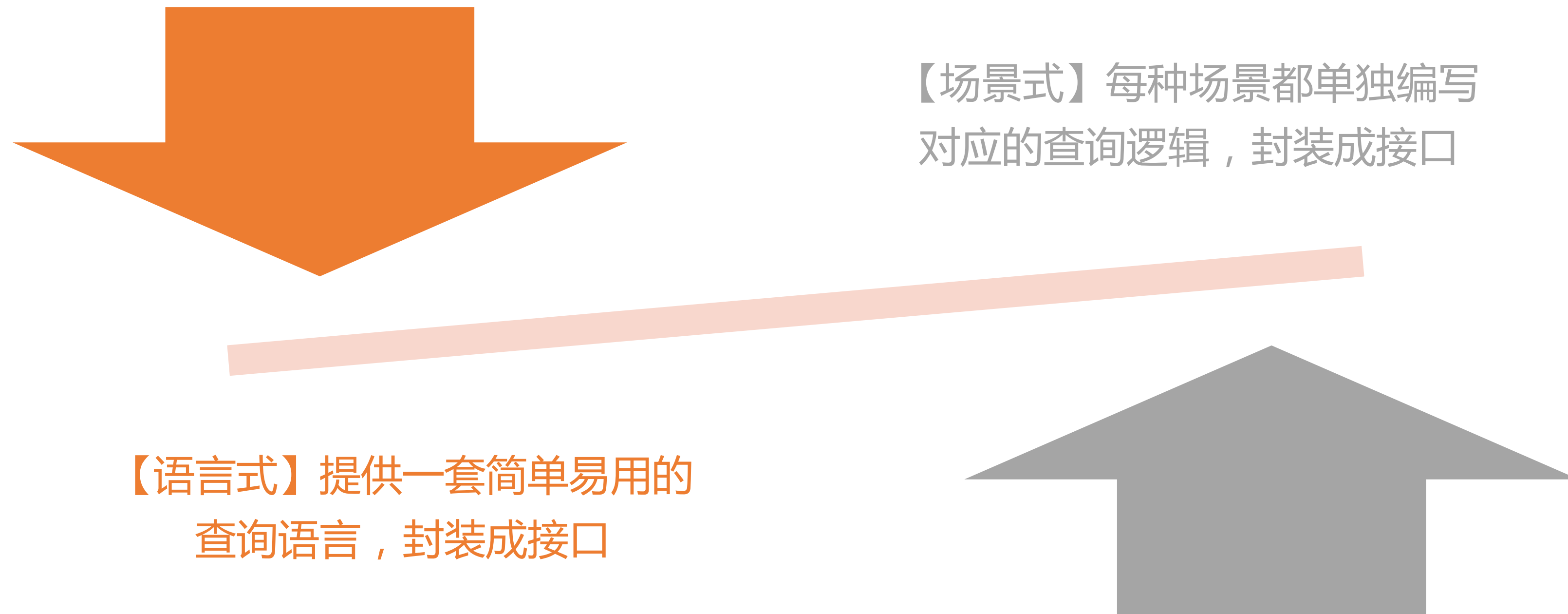
Elasticsearch在转转的应用实践 - 查询 - 场景分析



Elasticsearch在转转的应用实践 - 查询 - 微服务架构下以何种形式引入Elasticsearch ?



Elasticsearch在转转的应用实践 - 查询 - 如何封装到Elasticsearch的调用?



Elasticsearch在转转的应用实践 - 查询 - Query样例

keyword=elasticsearch&category!=2&city=1|2|3&price=2000_3000&sort=default

key: keyword、category、city、price、sort

op: =、!=

value: elasticsearch、2、1|2|3、2000_3000、default

link: &

fields: keyword->title:Text、category:Keyword、city:Keyword、price:Integer

sorts: default->samecity, quality_score

Elasticsearch在转转的应用实践 - 查询 - 翻译过程

fields(Match&Filter)

- text->match
- text->multi-match
- keyword->term|terms
- boolean->term
- number->term|terms|range
- date->range
- geo-point->geo distance
- ...

field sorts

- score sort
- number field sort
- date field sort
- geo distance sort
- painless script sort

score sorts

- field value factor function
- guass decay function
- liner decay function
- exp decay function
- natic script function
- painless script function
- filter function
- ...

Elasticsearch在转转的应用实践 - 查询 - 翻译结果

keyword=elasticsearch&category!=2&city=1|2|3&price=2000_3000&sort=default

```
1 {
2   "from": 0,
3   "size": 20,
4   "query": {
5     "function_score": {
6       "query": {
7         "bool": {
8           "must": [
9             { "match": { "title": { "query": "elasticsearch", "minimum_should_match": "100%" } } }
10          ],
11          "filter": [
12            { "terms": { "city": [ 1, 2, 3 ] } },
13            { "range": { "price": { "from": 2000, "to": 3000, "include_lower": true, "include_upper": true } } }
14          ],
15          "must_not": [
16            { "term": { "category": { "value": 2 } } }
17          ]
18        }
19      },
20      "functions": [
21        { "weight": 10, "script_score": { "script": { "source": "samecity", "lang": "native", "params": { "city": 1 } } } },
22        { "field_value_factor": { "field": "quality_score" } }
23      ],
24      "score_mode": "sum",
25      "boost_mode": "sum"
26    }
27  },
28  "_source": false
29 }
```

Elasticsearch在转转的应用实践 - 查询 - 如何解决特殊排序需求？

e.g.: 特殊类别优先展示同城商品

- 这部分类别数量可能较多，不适合直接放到DSL中
- 自定义排序插件(Advanced scripts using script engines)

e.g.: 希望ES召回后排序使用更加的复杂排序模型

- Function Score Query只能提供基本的线性模型，而且无法做一些特征离散化之类的处理
- 开源排序插件(elasticsearch-learning-to-rank)

Elasticsearch在转转的应用实践 - 查询 - 如何解决索引升级、集群迁移之类的问题？

需要持续对外服务，如何做到对业务方无感？

✓ 小索引升级：*alias*

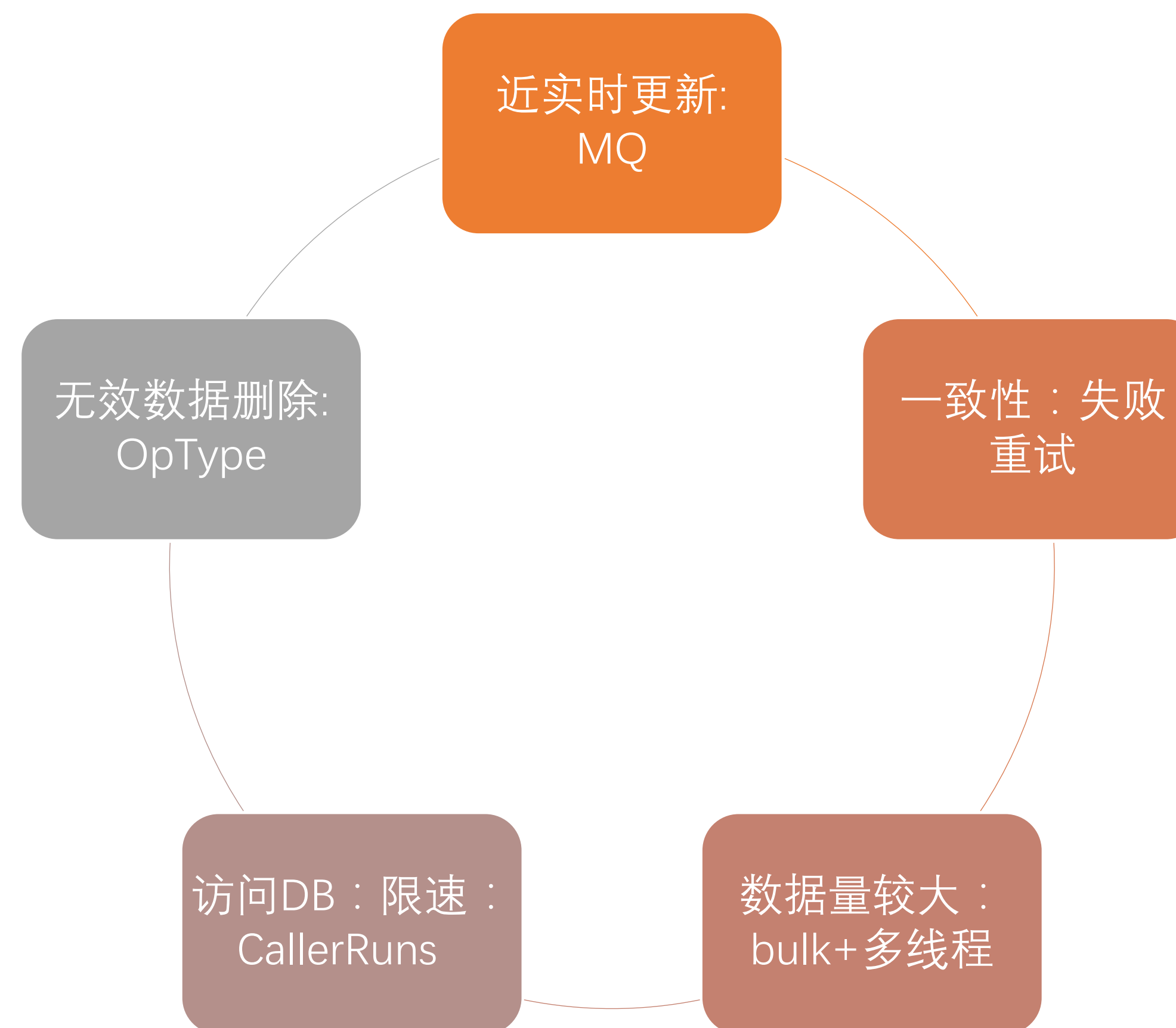
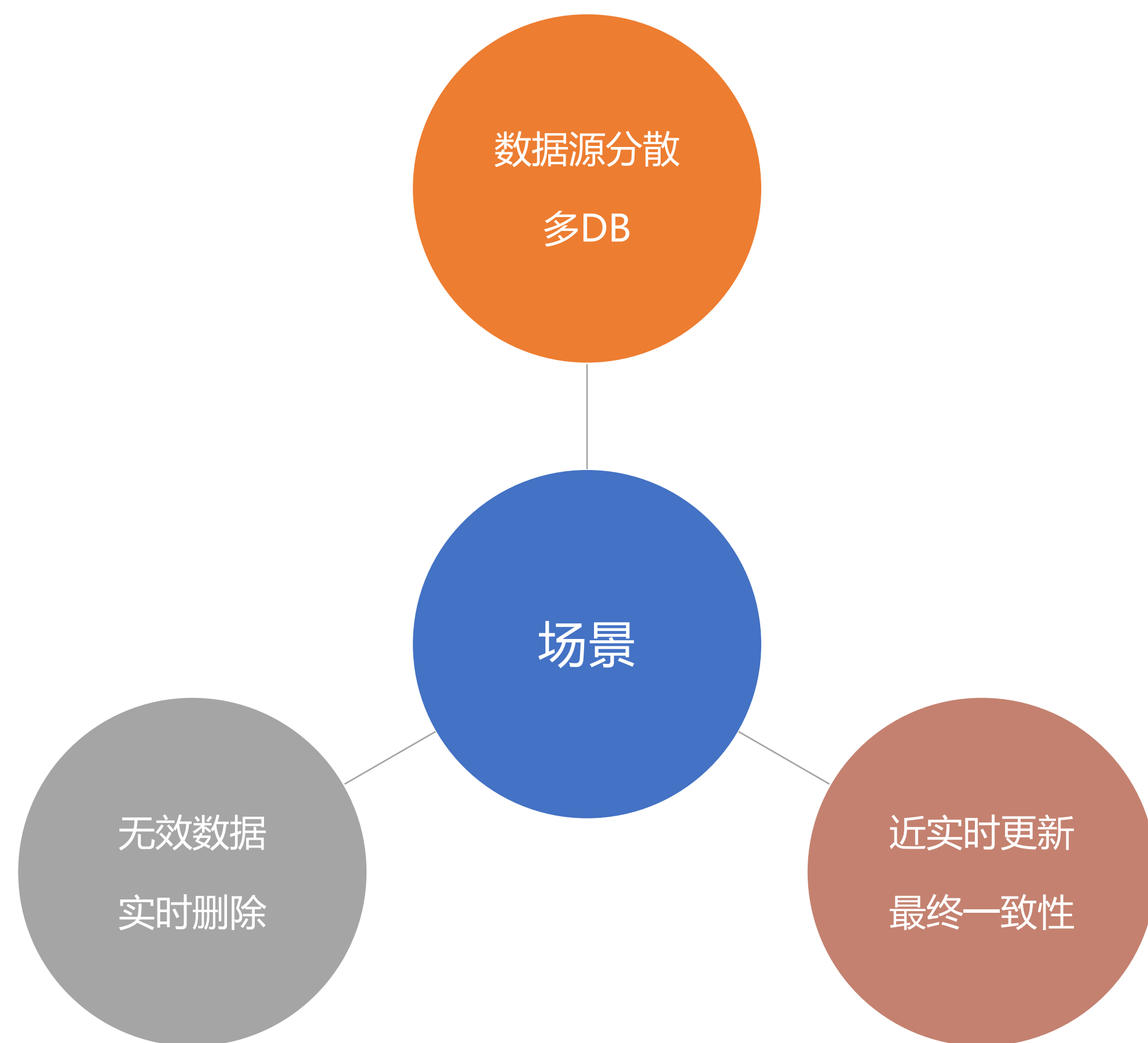
- 当前集群上创建新版本索引 -> 导入数据 -> 切换别名

✓ 大索引升级|集群迁移：*多集群*

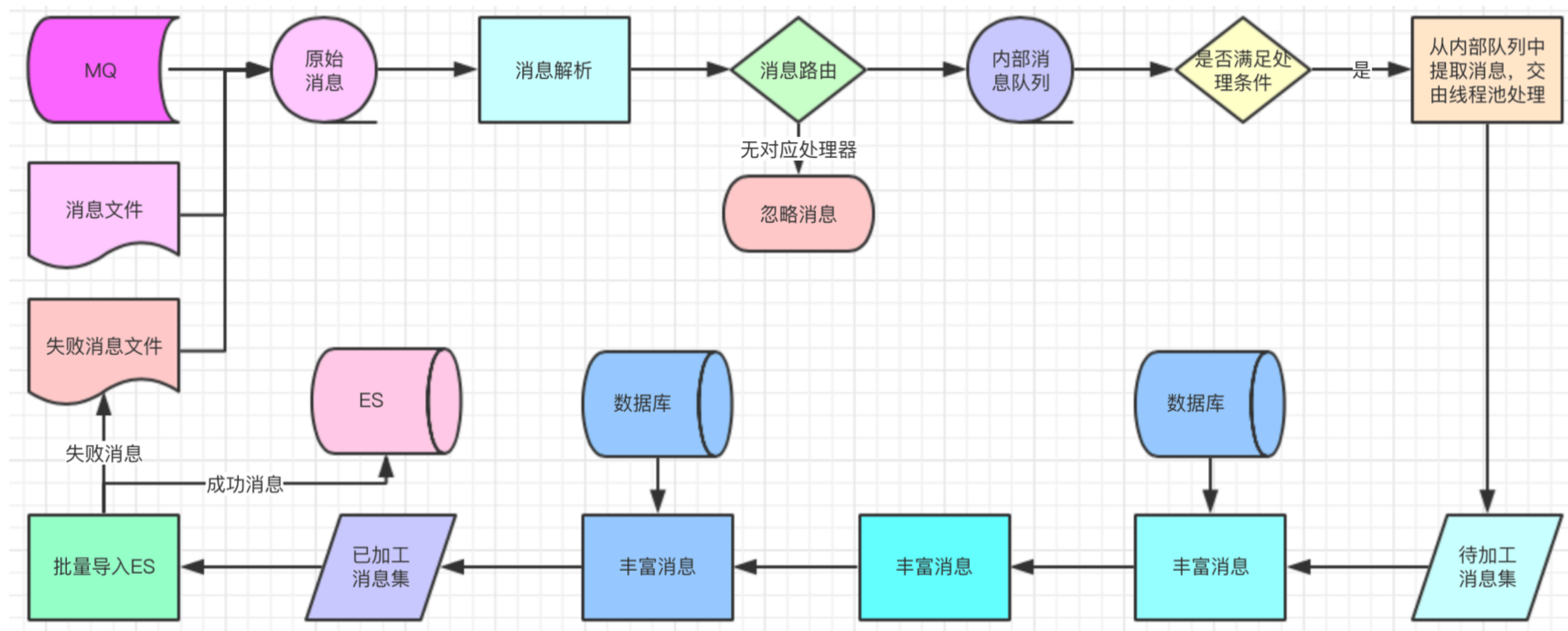
- 新集群上创建索引 -> 导入数据 -> 预热 -> 流量切换

转转索引数据导入场景 & 解决实现

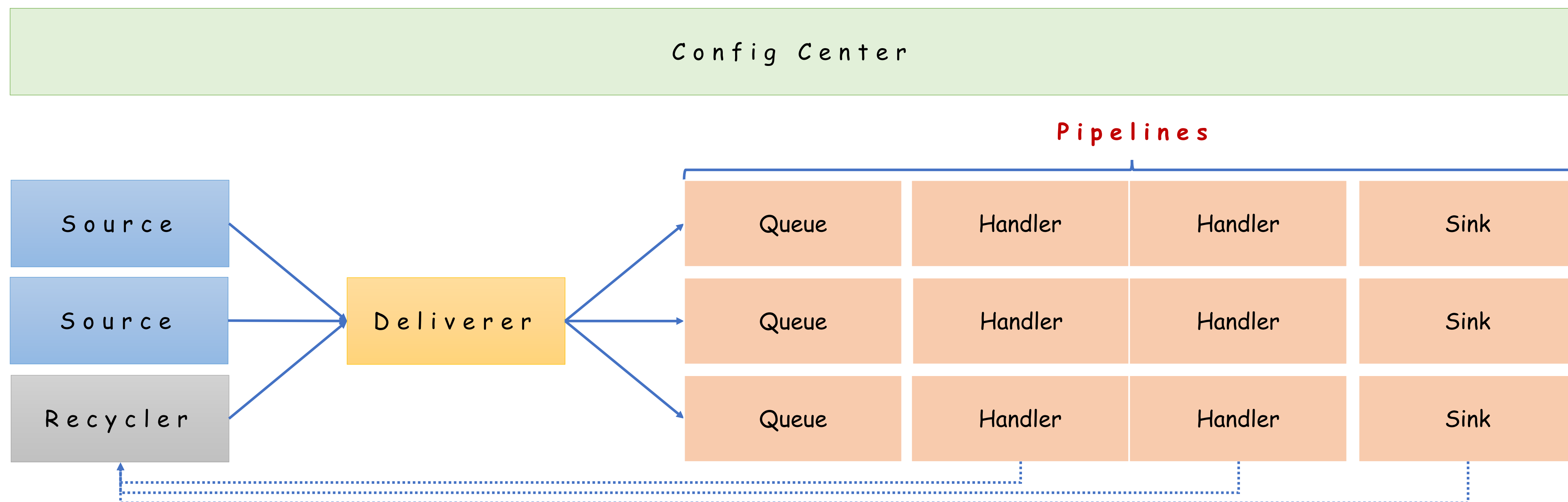
Elasticsearch在转转的应用实践 - 导入 - 场景&分析



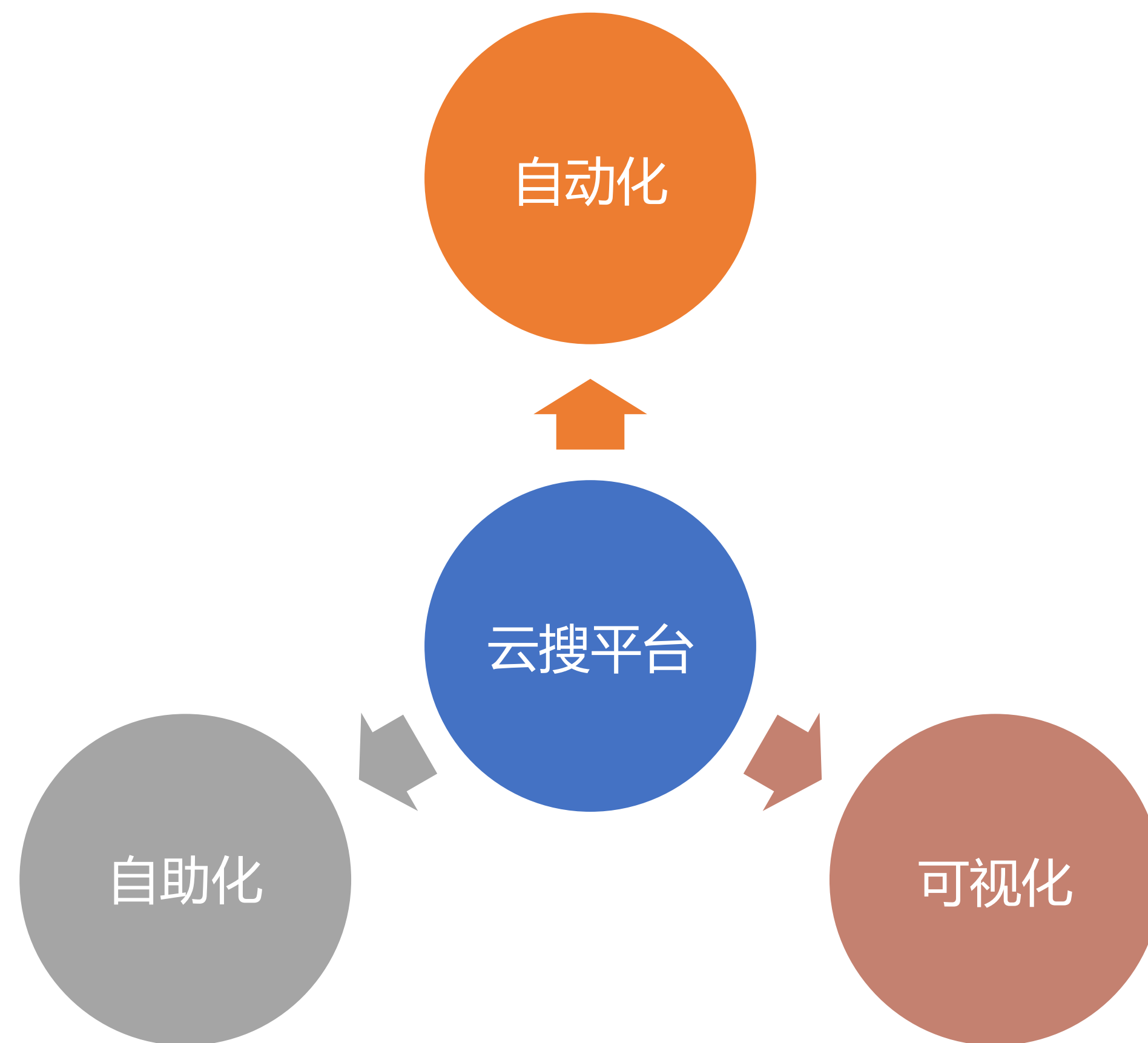
Elasticsearch在转转的应用实践 - 导入 - 数据流程



Elasticsearch在转转的应用实践 - 导入 - 系统设计



Elasticsearch在转转的应用实践 - 未来规划



THANKS !





elastic 中文社区

专业、垂直、纯粹的 Elastic 开源技术交流社区

<https://elasticsearch.cn/>

