Elasticsearch 使用中的双刃剑特性

汇报人：金多安 @极限科技

时间：2025-03-26


双刃剑
吃资源
不支持事务
大聚合OOM
分词器达不到要求的效果
语法学习成本
可以做日志
文本搜索
商品标签
海量数据
restful
聚合丰富

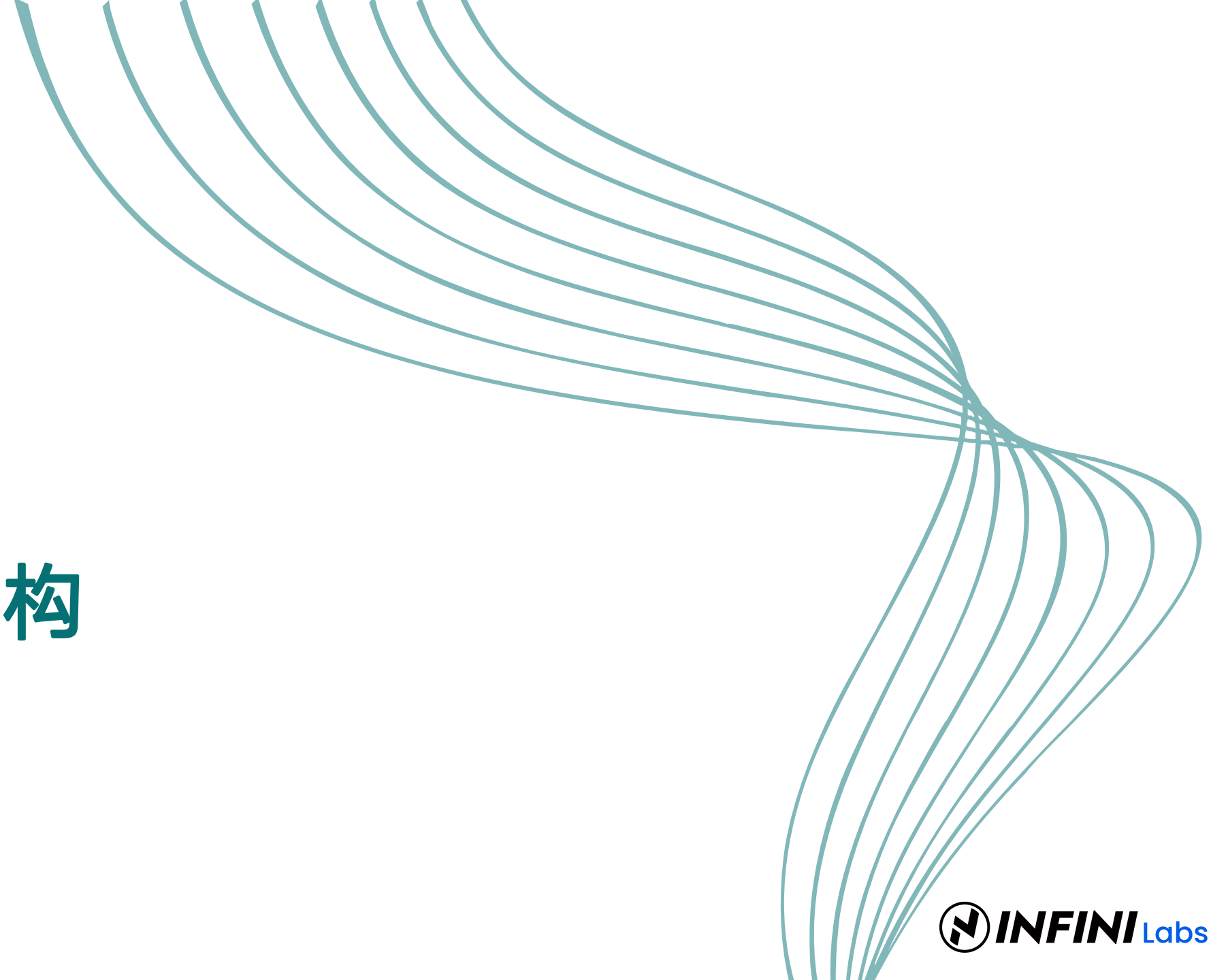
目录

Contents

INFINI Labs

01

分布式架构



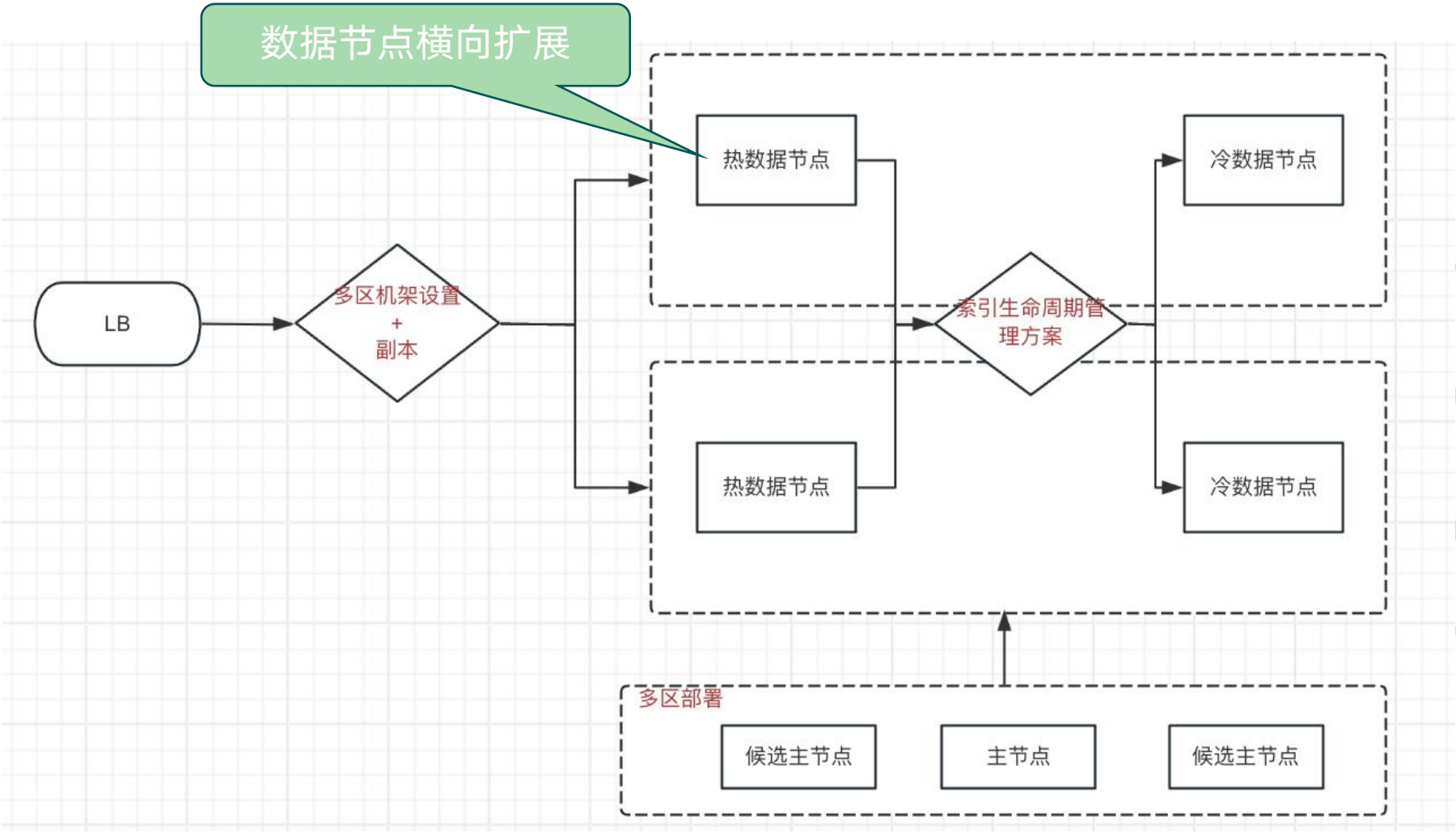
INFINI Labs

分布式架构

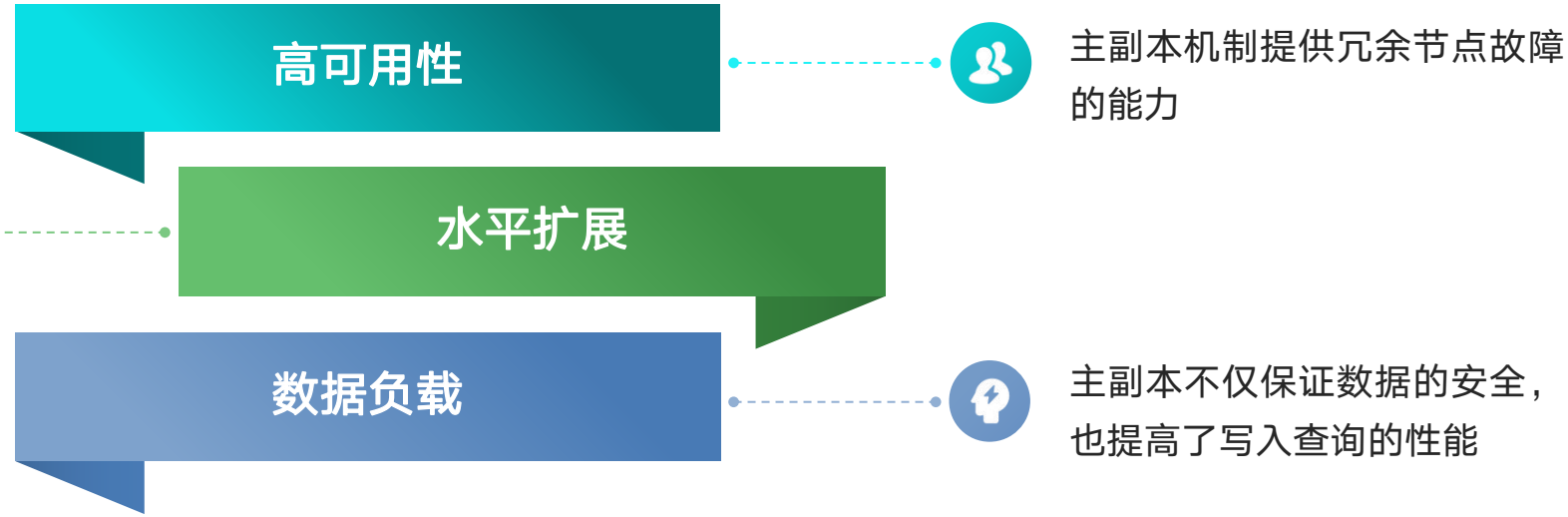
专用主

双活

冷热架构



优势



劣势

索引管理成本

- 索引数量过多
- 分片分配不均
- 索引元信息管理

资源消耗的要求

- 硬件要求较高
- 副本对写入资源的消耗
- 查询并发的扩散
- 甚至有资源挤占

运维复杂度

- 大规模集群的难度
- 众多的监控指标
- 业务场景的高要求

考虑因素

资源规划

主分片和数据节点数设计，**做好卡位**

基础的安全性要确保同角色一致

01

明确的要求

混布会带来资源挤压

清晰基本的性能要求

理想的使用是均衡的

02

业务特征

日志场景/搜索业务

业务规模与成本

业务要求越高，使用场景越简单

03

使用红线

集群分片上限

索引字段数量上限

JVM与分片数量的设计

04

可参考指标

1. JVM 分片配比 20分片：1GB
2. 分片大小 20–50GB 2亿lucene文档数
3. 主分片越多，分片越小，但长尾越明显。
4. 线程 / 分片配比：一个分片同时执行一个任务
5. 主节点规格与分片数上限，max 10w分片
6. 副本数越多，可以一定提高查询吞吐，写入资源消耗越大
7. 堆外内存越大，查询稳定性越高

理想化的集群

集群架构

专主
冷热分离
双活机架
越清晰越好

01

02

索引主分片与数据节点数量

核心业务索引：主分片数=数据节点数量*（1-> cpu数量/n）

内存

jvm：不超过31GB
堆外：禁用swap，越多越好

03

04

分片状态

20GB-50GB
delete文档占比越低越好（可以尝试调整到20%）

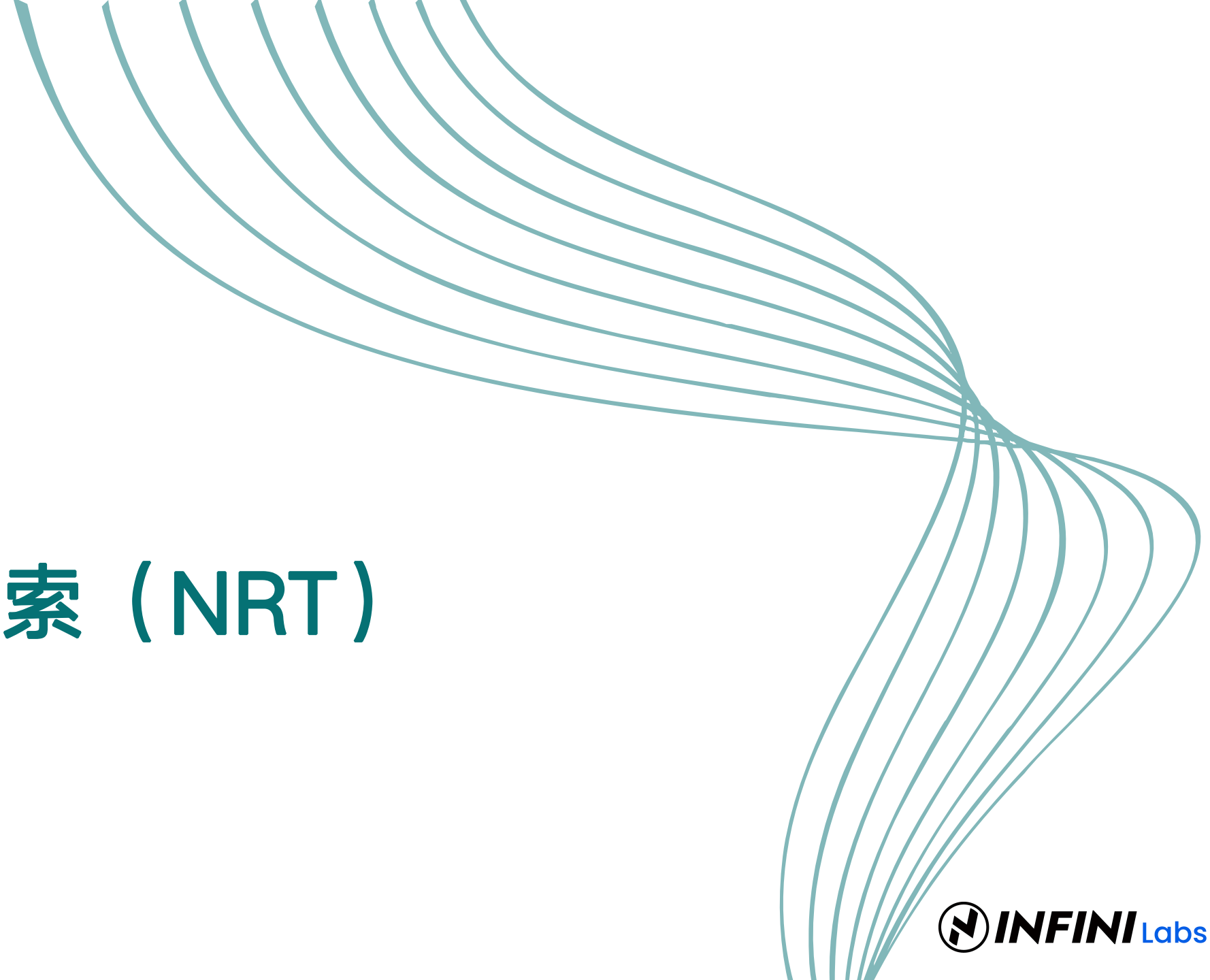
监控与测试

完善的监控数据
测试环境与数据

05

02

近实时搜索（NRT）



INFINI Labs

强大之处

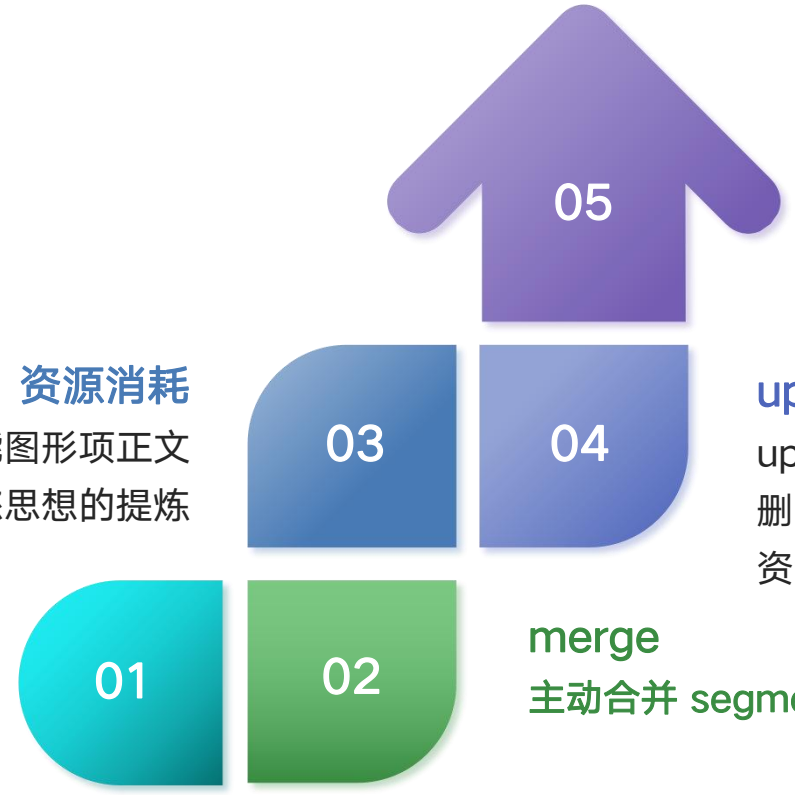
通过内存缓冲和定期刷新机制，平衡写入性能和搜索实时性。

高性能

快速可见性

数据写入后几乎立即可查，适合实时搜索和分析场景。

瓶颈



routing

集中几个分片更新
造成性能热点

资源消耗

单击此处输入你的智能图形项正文
文字是您思想的提炼

update

update 占比越高，
删除文档比越高，
资源消耗越大

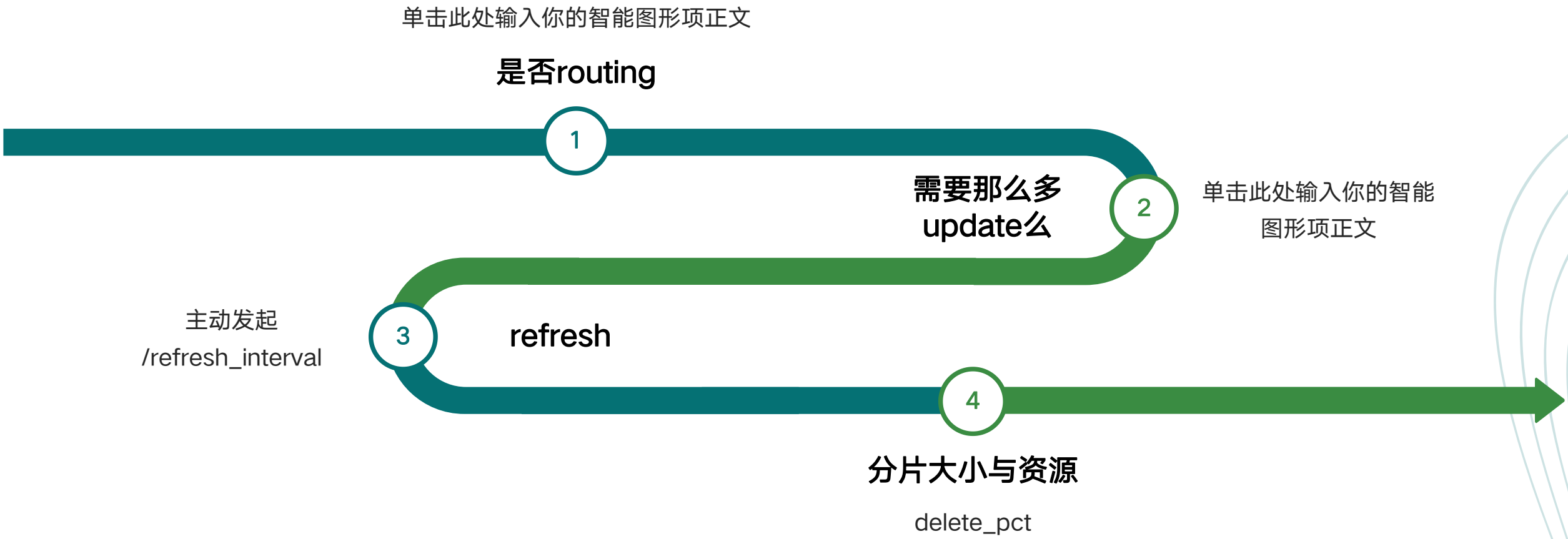
refresh 操作

主动发起
or 将 refresh_interval 设置较高频率

merge

主动合并 segment 文件

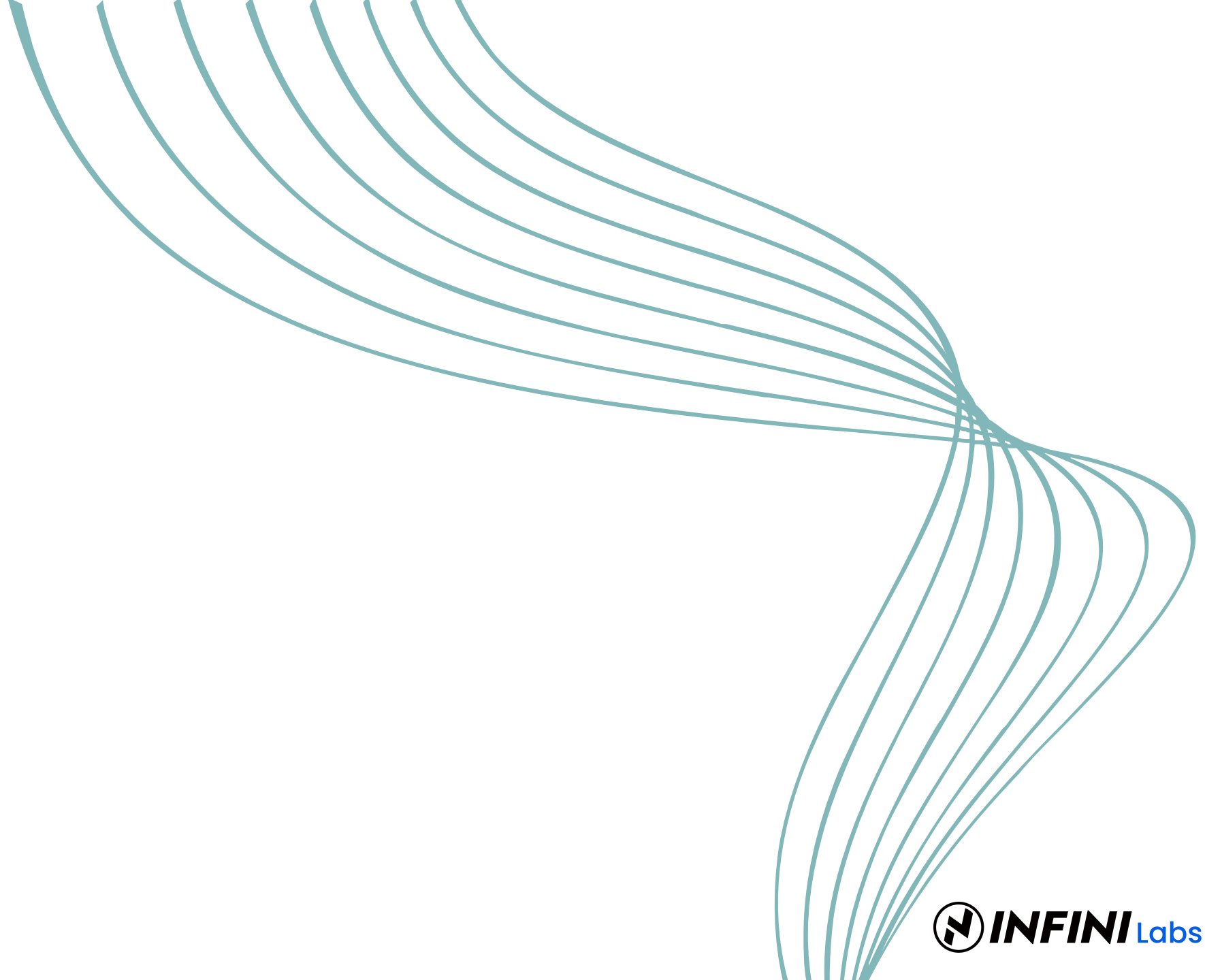
考虑因素



INFINI Labs

03

倒排索引



INFINI Labs

优点

词项查询效率高

全文匹配/精确查询/权重分配等等复杂场景

天然解决“哪些文档包含某词项”的核心问题

缺点

存储占用高

倒排索引会占用大量磁盘空间，尤其是文本字段较多的场景。

运行时资源消耗

- 倒排索引的构建和维护需要消耗大量 CPU 和内存资源。
  - 高并发写入场景下，索引性能可能成为瓶颈。

数据类型更新

不可更改数据类型

倒排存储

对比了
BKD和倒排的大小

```
{
  "unique_numbers": {
    "aliases": {},
    "mappings": {
      "properties": {
        "number_keyword": {
          "type": "keyword"
        },
        "test": {
          "type": "float"
        }
      }
    },
    "settings": {
      "index": {
        "creation_date": "1742888925252",
        "number_of_shards": "3",
        "number_of_replicas": "1",
        "uuid": "roj5yzTiTKS8T7evUEfUQg",
        "version": {
          "created": "1080399"
        },
        "provided_name": "unique_numbers"
      }
    }
  }
}
```

```
    "number_keyword": {
      "total": "333.7mb",
      "total_in_bytes": 349977178,
      "inverted_index": {
        "total": "151.5mb",
        "total_in_bytes": 158914351
      },
      "stored_fields": "0b",
      "stored_fields_in_bytes": 0,
      "doc_values": "182.2mb",
      "doc_values_in_bytes": 191062827,
      "points": "0b",
      "points_in_bytes": 0,
      "norms": "0b",
      "norms_in_bytes": 0,
      "term_vectors": "0b",
      "term_vectors_in_bytes": 0
    },
    "test": {
      "total": "77.2mb",
      "total_in_bytes": 80960485,
      "inverted_index": {
        "total": "0b",
        "total_in_bytes": 0
      },
      "stored_fields": "0b",
      "stored_fields_in_bytes": 0,
      "doc_values": "33.3mb",
      "doc_values_in_bytes": 35000017,
      "points": "43.8mb",
      "points_in_bytes": 45960468,
      "norms": "0b",
      "norms_in_bytes": 0,
      "term_vectors": "0b",
      "term_vectors_in_bytes": 0
    }
```

INFINI Labs

字段类型选择

字段冗余

面向多场景

bkd tree

排序

范围查询

倒排

精确匹配

全文检索

需求场景

精确匹配

范围查询

聚合/排序

数据类型不可更改：reindex/update_by_query

reindex

适合大索引（上百GB）

变更方案较为复杂

update_by_query

适合小索引

变更逻辑简单，配合pipeline几乎无感

速度有限

04

聚合分析

INFINI Labs

支持复杂的聚合场景

指标聚合

max/min/avg

01

分桶聚合

bucket

02

嵌套聚合

聚合之间的嵌套

03

性能压力

高延迟，高压力

单击此处输入你的智能图形项正文

文字是您思想的提炼

数据的精度

min_doc_count:1

数据量大

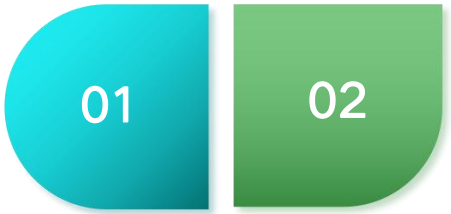
单击此处输入你的智能图形项正文

文字是您思想的提炼

聚合的复杂度

子聚合

嵌套聚合



数据的范围

size：10000

一些建议

01

查询参数

如果明确字段是低基数的，使用 keyword

配合 execution_hint:map

02

查询方式

异步查询

调整 max_concurrent_shard_requests

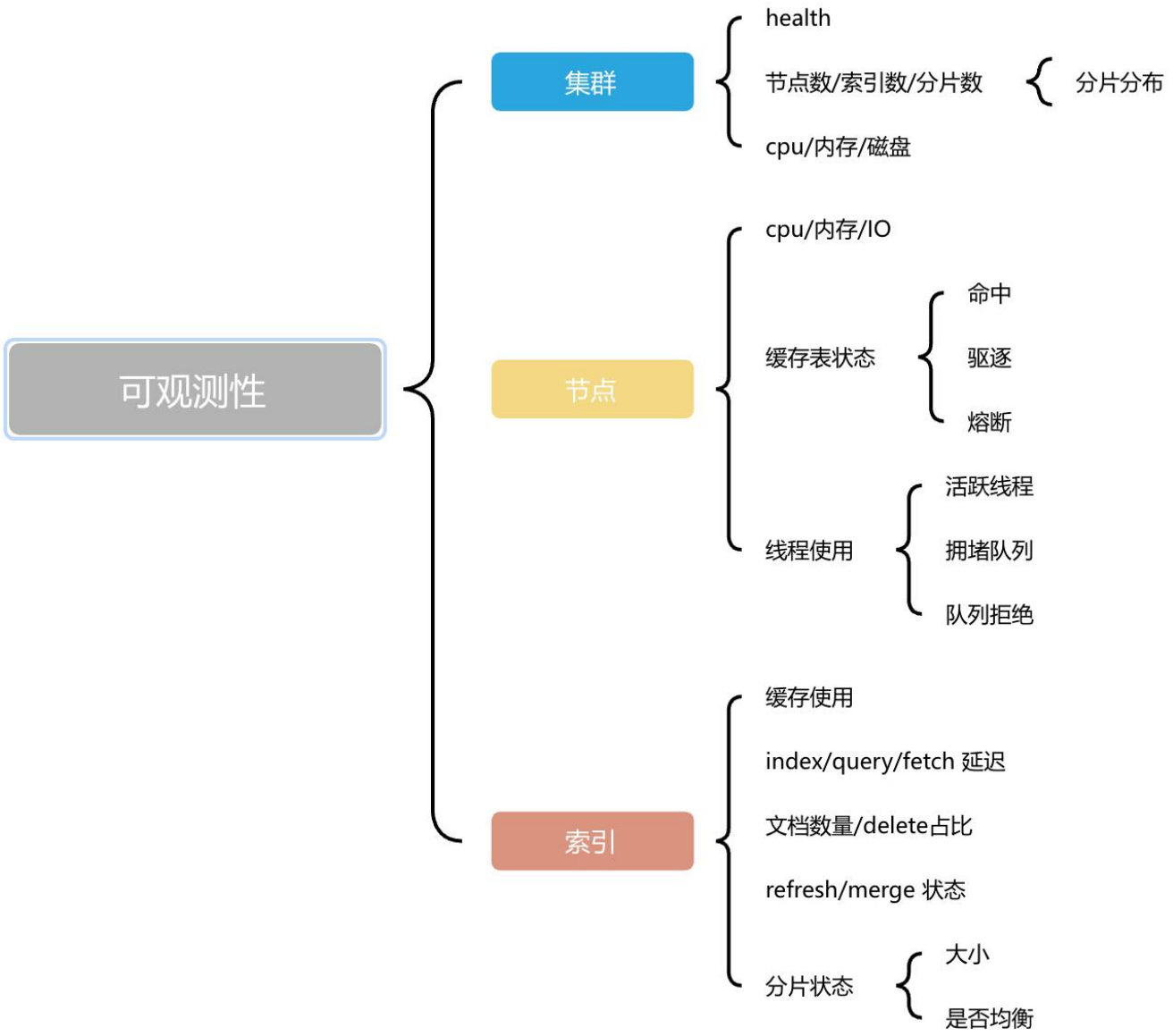
03

时序性数据

rollup精化指标后

进行进一步聚合

适合前四项：
提高可观测性
即完善的关键指标



05

生态系统与辅助组件

各类生态组件

带来的问题

使用复杂度

运维和开发的复杂度增加

兼容性

版本间的兼容性问题

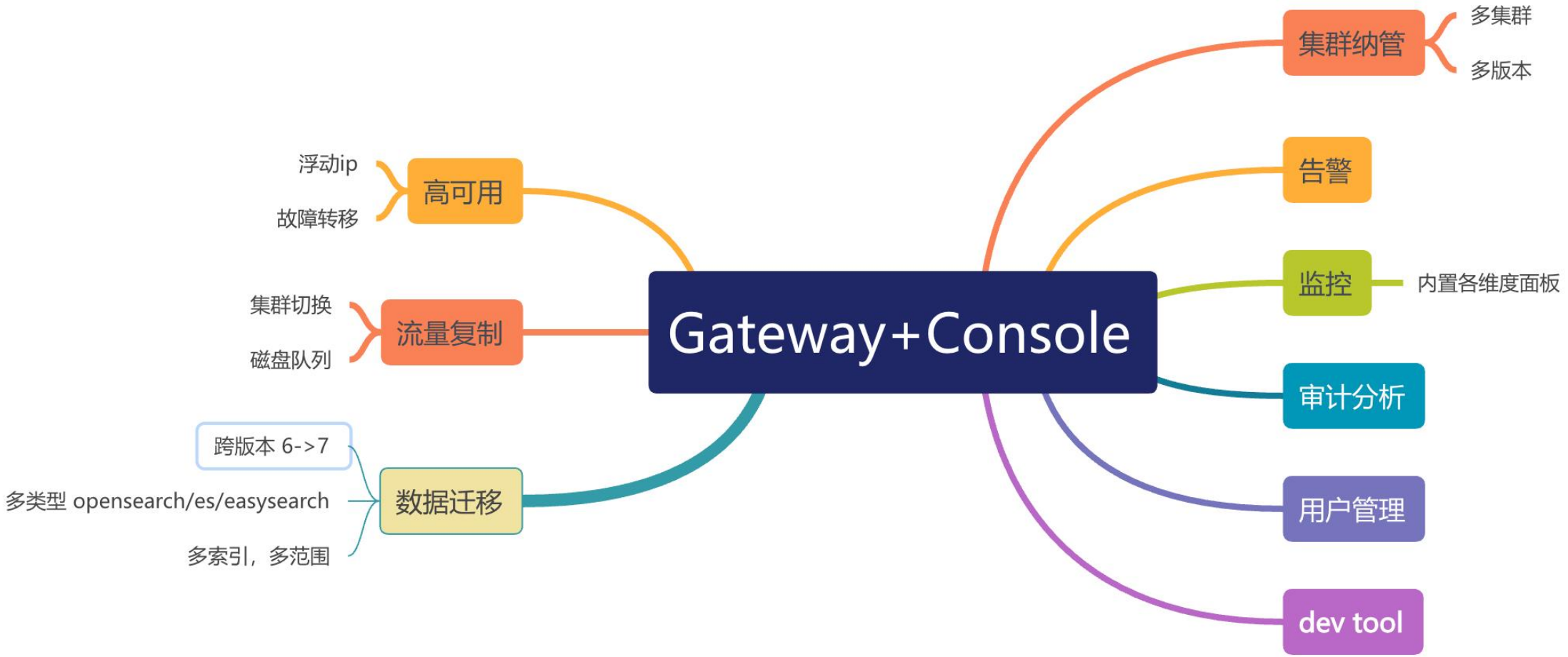
学习成本

同一场景下，不同组件的使用方式会有不同

众多的组件

多场景多类型的组件

Gateway + Console



INFINI Labs

感谢观看

汇报人：金多安

扫码加入Meetup用户交流群

搜索客 SearchKit

搜索人自己的社区

专业、垂直、纯粹的开源搜索技术交流社区

https://searchkit.cn