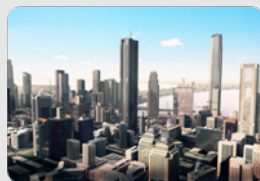


Elasticsearch结合Hbase的应用

中控 汪兴



主要内容

- 1、 Hbase简单说明
- 2、 Hbase不足
- 3、 Elasticsearch优势
- 4、 Elasticsearch与Hbase的结合使用
- 5、 ES的一些坑（经验之谈，虚拟机文件损坏）

Hbase

Hbase是运行在Hadoop上的NoSQL数据库，它是一个分布式的和可扩展的大数据仓库，也就是说HBase能够利用HDFS的分布式处理模式，并从Hadoop的MapReduce程序模型中获益。这意味着在一组商业硬件上存储许多具有数十亿行和上百万列的大表。除去Hadoop的优势，HBase本身就是十分强大的数据库，它能够融合key/value存储模式带来实时查询的能力，以及通过MapReduce进行离线处理或者批处理的能力。总的来说，Hbase能够让你在大量的数据中查询记录，也可以从中获得综合分析报告。

Hbase不足

- 多条件查询不方便（例如：表里包含字段车牌、颜色、车速、部门、道路）。
- count问题，无法高效获取数据总量。
- 没有高效的实时灵活的统计方法

ES优势

- ① es支持非常灵活的多条件查询，不仅仅可以设置多个查询条件，还能进行多组条件逻辑查询，还可以通过分值进行匹配度查询，除此之外还支持全文检索。
- ② 查询速度很快，分布式，动态可扩展。
- ③ 可做一些分析统计，提供分组和统计文档的能力，这个统计还是比较强大的，可以同时做分组统计。
- ④ 支持Restful Api，这是非常方便的，不仅仅在实际生产环境中还是调试阶段。

ES与Hbase的结合使用

数据生成方式:

- 1.定时同步
- 2.提前生成关联，同时向两边写数据
- 3.使用Coprocessors

数据存储方式:

- 1.Hbase和ES各存一份
- 2.ES只存索引

@Override

dependencies

APACHE HBASE

Home

Table Details

Local Logs

Log Level

Debug Dump

Metrics Dump

HBase Configuration

Online Regions	Offline Regions	Failed Regions	Split Regions	Other Regions	Description
192	0	0	17	0	'AfterVehicle', {TABLE_ATTRIBUTES => {coprocessor\$1 => ' org.apache.hadoop.hbase.coprocessor.AggregateImplementation 1073741823 ', METADATA => {'COMPACTION_ENABLED' => 'true'}}}, {NAME => 'search'}
390	0	0	50	0	'ElectronicLicensePlate', {TABLE_ATTRIBUTES => {coprocessor\$1 => ' org.apache.hadoop.hbase.coprocessor.AggregateImplementation 1073741823 ', METADATA => {'COMPACTION_ENABLED' => 'true'}}}, {NAME => 'search'}

```
int value = Integer.parseInt(cValue);
contentBuilder.field(cFamily, value);
}
```


ES的一些坑

```
#
#<domain>
#
##
#root
##
#@student
#@faculty
#@faculty
#ftp
#ftp
#@student

* hard nofile 655
* soft nofile 655
root hard nofile
root soft nofile
# End of file
```

<type> <item> <value>

```
{
  "mappings": {
    "AfterVehicle": {
      "_source": {
        "excludes": [],
        "compress": true,
      },
      "_id": {
        "path": "id"
      },
      "properties": {
        "eid": {
          "index": "not_analyzed",
          "store": true,
          "type": "string"
        },
        "recordIp": {
          "index": "not_analyzed",
          "store": true,
          "type": "string"
        }
      }
    }
  }
}
```




← 返回博客列表

原

摘要

本文主要介绍了...
帮助, 对于已经

分布式搜索 介绍

Elasticsearch是...
发, 提供Restful...
重启服务, 基本...
化。

概念和设计

分布式搜索引擎Elasticsearch

介绍

概念和设计

索引

文档

文档类型

节点与集群

分片

副本

安装和配置

Elasticsearch集群

tar.gz安装及配置

rpm安装及配置

运行elasticsearch

关闭elasticsearch

插件

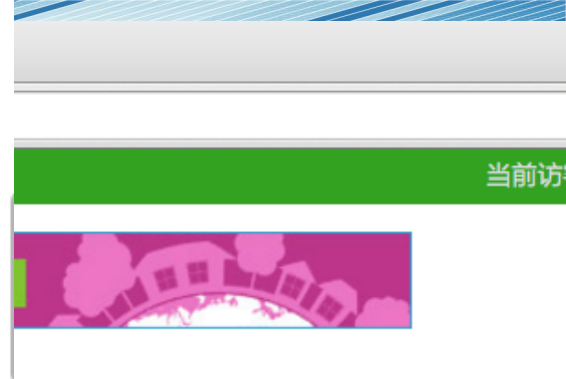
站点插件 (以网页形式展现)

插件安装

分词插件IK

插件安装

测试



安装配置

评论: 0

用。对于新入门的同学应该很有

力。Elasticsearch是用java开发, 扩展能力非常强, 不需要...
决, bug存在, API繁多并且变

谢谢大家！

Thank You