

# 目 录

第 1 章 Elasticsearch 集群入门 .....	1	1.5.3 Lucene 查询语法 .....	26
1.1 全文检索 .....	1	1.6 小结 .....	27
1.1.1 Lucene 词汇表和架构 .....	1	第 2 章 索引 .....	28
1.1.2 输入数据分析 .....	3	2.1 Elasticsearch 索引 .....	28
1.1.3 评分和查询相关性 .....	4	2.1.1 分片和副本 .....	28
1.2 Elasticsearch 基础 .....	4	2.1.2 创建索引 .....	29
1.2.1 数据架构的主要概念 .....	4	2.2 映射配置 .....	31
1.2.2 Elasticsearch 主要概念 .....	6	2.2.1 类型确定机制 .....	31
1.2.3 索引建立和搜索 .....	6	2.2.2 索引结构映射 .....	33
1.3 安装并配置集群 .....	8	2.2.3 不同的相似度模型 .....	43
1.3.1 安装 Java .....	8	2.2.4 信息格式 .....	45
1.3.2 安装 Elasticsearch .....	8	2.2.5 文档值 .....	47
1.3.3 在 Linux 上用二进制包安装 Elasticsearch .....	9	2.3 批量索引以提高索引速度 .....	48
1.3.4 目录布局 .....	9	2.3.1 为批量索引准备数据 .....	48
1.3.5 配置 Elasticsearch .....	10	2.3.2 索引数据 .....	48
1.3.6 运行 Elasticsearch .....	11	2.3.3 更快的批量请求 .....	50
1.3.7 关掉 Elasticsearch .....	12	2.4 用附加的内部信息扩展索引结构 .....	50
1.3.8 Elasticsearch 作为系统服务 运行 .....	13	2.4.1 标识符字段 .....	50
1.4 用 REST API 操作数据 .....	14	2.4.2 _type 字段 .....	51
1.4.1 理解 Elasticsearch 的 RESTful API .....	14	2.4.3 _all 字段 .....	52
1.4.2 在 Elasticsearch 中存储数据 .....	15	2.4.4 _source 字段 .....	52
1.4.3 新建文档 .....	15	2.4.5 _index 字段 .....	53
1.4.4 检索文档 .....	16	2.4.6 _size 字段 .....	54
1.4.5 更新文档 .....	17	2.4.7 _timestamp 字段 .....	54
1.4.6 删除文档 .....	18	2.4.8 _ttl 字段 .....	55
1.4.7 版本控制 .....	18	2.5 段合并介绍 .....	56
1.5 使用 URI 请求查询来搜索 .....	20	2.5.1 段合并 .....	56
1.5.1 示例数据 .....	20	2.5.2 段合并的必要性 .....	56
1.5.2 URI 请求 .....	20	2.5.3 合并策略 .....	57
		2.5.4 合并调度器 .....	57
		2.5.5 合并因子 .....	57

2.5.6 调节 .....	58	3.3.16 more_like_this_field 查询 .....	89
2.6 路由介绍 .....	58	3.3.17 范围查询 .....	90
2.6.1 默认索引过程 .....	59	3.3.18 最大分查询 .....	90
2.6.2 默认搜索过程 .....	59	3.3.19 正则表达式查询 .....	91
2.6.3 路由 .....	61	3.4 复合查询 .....	91
2.6.4 路由参数 .....	62	3.4.1 布尔查询 .....	92
2.6.5 路由字段 .....	62	3.4.2 加权查询 .....	93
2.7 小结 .....	63	3.4.3 constant_score 查询 .....	94
第 3 章 搜索 .....	64	3.4.4 索引查询 .....	94
3.1 查询 Elasticsearch .....	64	3.5 查询结果的过滤 .....	95
3.1.1 示例数据 .....	65	3.5.1 使用过滤器 .....	95
3.1.2 简单查询 .....	66	3.5.2 过滤器类型 .....	96
3.1.3 分页和结果集大小 .....	67	3.5.3 过滤器的缓存 .....	104
3.1.4 返回版本值 .....	68	3.6 高亮显示 .....	105
3.1.5 限制得分 .....	69	3.6.1 高亮显示入门 .....	105
3.1.6 选择需要返回的字段 .....	69	3.6.2 字段配置 .....	106
3.1.7 使用脚本字段 .....	71	3.6.3 深入底层 .....	107
3.2 理解查询过程 .....	72	3.6.4 配置 HTML 标签 .....	107
3.2.1 查询逻辑 .....	72	3.6.5 控制高亮片段 .....	108
3.2.2 搜索类型 .....	73	3.6.6 全局设置与局部设置 .....	108
3.2.3 搜索执行偏好 .....	74	3.6.7 需要匹配 .....	109
3.2.4 搜索分片 API .....	75	3.6.8 信息高亮器 .....	111
3.3 基本查询 .....	76	3.7 验证查询 .....	113
3.3.1 词条查询 .....	76	3.8 数据排序 .....	115
3.3.2 多词条查询 .....	77	3.8.1 默认排序 .....	115
3.3.3 match_all 查询 .....	77	3.8.2 选择用于排序的字段 .....	116
3.3.4 常用词查询 .....	78	3.8.3 指定缺少字段的行为 .....	118
3.3.5 match 查询 .....	79	3.8.4 动态条件 .....	118
3.3.6 multi_match 查询 .....	81	3.8.5 排序规则和国家特有字符 .....	119
3.3.7 query_string 查询 .....	82	3.9 查询重写 .....	119
3.3.8 simple_query_string 查询 .....	84	3.9.1 重写过程示例 .....	119
3.3.9 标识符查询 .....	84	3.9.2 查询重写的属性 .....	120
3.3.10 前缀查询 .....	84	3.10 小结 .....	121
3.3.11 fuzzy_like_this 查询 .....	85	第 4 章 扩展索引结构 .....	122
3.3.12 fuzzy_like_this_field 查询 .....	86	4.1 索引树形结构 .....	122
3.3.13 fuzzy 查询 .....	86	4.1.1 数据结构 .....	122
3.3.14 通配符查询 .....	88	4.1.2 分析 .....	123
3.3.15 more_like_this 查询 .....	88	4.2 索引非扁平数据 .....	124
		4.2.1 数据 .....	124

4.2.2 对象.....	125	5.6 同义词.....	161
4.2.3 数组.....	125	5.6.1 同义词过滤器.....	161
4.2.4 映射.....	125	5.6.2 定义同义词规则.....	162
4.2.5 向 Elasticsearch 发送映射.....	127	5.6.3 查询时或索引时的同义词扩展.....	164
4.2.6 动态还是非动态.....	127	5.7 理解解释信息.....	164
4.3 使用嵌套对象.....	128	5.7.1 理解字段分析.....	164
4.4 使用父子关系.....	131	5.7.2 解释查询.....	165
4.4.1 索引结构和数据索引.....	131	5.8 小结.....	167
4.4.2 查询.....	132		
4.4.3 父子关系和过滤.....	134	第 6 章 超越全文检索.....	168
4.4.4 性能考虑.....	134	6.1 聚合.....	168
4.5 使用更新 API 修改索引结构.....	135	6.1.1 一般查询结构.....	168
4.5.1 映射.....	135	6.1.2 可用的聚合.....	170
4.5.2 添加一个新字段.....	135	6.1.3 聚合的嵌套.....	185
4.5.3 修改字段.....	136	6.1.4 桶排序和嵌套聚合.....	187
4.6 小结.....	137	6.1.5 全局和子集.....	187
第 5 章 更好的搜索.....	138	6.2 切面.....	190
5.1 Apache Lucene 评分简介.....	138	6.2.1 文档结构.....	190
5.1.1 当文档被匹配时.....	138	6.2.2 返回的结果.....	190
5.1.2 默认评分公式.....	139	6.2.3 使用查询进行切面计算.....	191
5.1.3 相关性的意义.....	140	6.2.4 使用过滤器进行切面计算.....	192
5.2 Elasticsearch 的脚本功能.....	140	6.2.5 terms 切面.....	193
5.2.1 脚本执行过程中可用的对象.....	140	6.2.6 基于范围的切面.....	194
5.2.2 MVEL.....	141	6.2.7 数值和日期直方图切面.....	196
5.2.3 使用其他语言.....	141	6.2.8 数值型字段统计数据的计算.....	197
5.2.4 使用自定义脚本库.....	142	6.2.9 词条统计数据的计算.....	198
5.3 搜索不同语言的内容.....	145	6.2.10 地理切面.....	199
5.3.1 区分处理不同语言.....	145	6.2.11 切面结果的过滤.....	200
5.3.2 多语言处理.....	145	6.2.12 内存考虑.....	201
5.3.3 检测文档的语言.....	146	6.3 使用建议器.....	201
5.3.4 示例文档.....	146	6.3.1 可用的建议器类型.....	201
5.3.5 映射文件.....	147	6.3.2 包含建议器.....	201
5.3.6 查询.....	148	6.3.3 term 建议器.....	203
5.4 使用查询加权影响得分.....	150	6.3.4 phrase 建议器.....	204
5.4.1 加权.....	150	6.3.5 completion 建议器.....	205
5.4.2 为查询添加加权.....	150	6.4 预匹配器.....	209
5.4.3 修改得分.....	153	6.4.1 示例索引.....	209
5.5 索引时加权何时有意义.....	160	6.4.2 预匹配器的准备.....	209
5.5.1 在输入数据中定义字段加权.....	160	6.4.3 深入.....	211
5.5.2 在映射中定义加权.....	161	6.5 文件的处理.....	214
		6.6 地理.....	217

6.6.1 为空间搜索准备映射	217	8.2.2 索引统计 API	253
6.6.2 示例数据	218	8.2.3 状态 API	256
6.6.3 示例查询	218	8.2.4 节点信息 API	256
6.6.4 任意地理形状	222	8.2.5 节点统计 API	257
6.7 卷动 API	226	8.2.6 集群状态 API	257
6.7.1 问题定义	226	8.2.7 挂起任务 API	258
6.7.2 作为解决方案的卷动	226	8.2.8 索引段 API	258
6.8 多词条过滤器	228	8.2.9 cat API	258
6.9 小结	232	8.3 控制集群的再平衡	260
第 7 章 深入 Elasticsearch 集群	233	8.3.1 再平衡	260
7.1 节点发现	233	8.3.2 集群的就绪	260
7.1.1 发现的类型	233	8.3.3 集群再平衡设置	260
7.1.2 主节点	234	8.4 控制分片和副本的分配	261
7.1.3 设置集群名	235	8.4.1 显式控制分配	262
7.1.4 节点的 ping 设置	236	8.4.2 集群范围的分配	264
7.2 时光之门与恢复模块	236	8.4.3 每个节点上的分片和副本数量	265
7.2.1 时光之门	236	8.4.4 手动移动分片和副本	265
7.2.2 恢复控制	237	8.5 预热	267
7.3 为高查询和高索引吞吐量准备 Elasticsearch 集群	238	8.5.1 定义一个新的预热查询	267
7.3.1 过滤器缓存	238	8.5.2 获取定义的预热查询	268
7.3.2 字段数据缓存和断路器	238	8.5.3 删除一个预热查询	269
7.3.3 存储模块	239	8.5.4 禁用预热功能	269
7.3.4 索引缓冲和刷新率	240	8.5.5 查询的选择	270
7.3.5 线程池的配置	240	8.6 使用索引别名来简化你的日常工作	270
7.3.6 结合起来, 一些通用建议	241	8.6.1 别名	271
7.4 模板和动态模板	244	8.6.2 创建别名	271
7.4.1 模板	244	8.6.3 修改别名	271
7.4.2 动态模板	245	8.6.4 合并命令	272
7.5 小结	246	8.6.5 获取所有别名	272
第 8 章 集群管理	248	8.6.6 移除别名	273
8.1 Elasticsearch 时光机	248	8.6.7 别名中的过滤	273
8.1.1 创建快照存储库	248	8.6.8 别名和路由	273
8.1.2 创建快照	249	8.7 Elasticsearch 插件	274
8.1.3 还原快照	251	8.7.1 基础知识	274
8.1.4 清理: 删除旧的快照	252	8.7.2 安装插件	274
8.2 监控集群的状态和健康度	252	8.7.3 移除插件	275
8.2.1 集群健康度 API	252	8.8 更新设置 API	275
		8.9 小结	276