

# ES性能优化在同程艺龙的实践

刘帅 ( Ricky )

# 个人简介

## 刘帅 ( Ricky Lau )

搜索工程师、60+ ES集群负责人，  
有着丰富的信息检索和分布式搜索架构经验，  
专注于海量数据的存储、检索和实时分析。  
阿里云Elasticsearch MVP，目前就职于同程艺龙。

个人微信



# 挑战

1

## 高吞吐

- a. 为离线数据提供检索功能
- b. 实时分析类业务 ( e.g. 监控、BI分析 )
- c. 日志收集

2

## 低延迟、高并发

- a. 1W+ QPS
- b. 99.9% 小于等于50毫秒
- c. 无拒绝请求

3

## 资源利用率

- a. 最大利用率及最大吞吐量
- b. 最大利用率与性能之争



# 目录

COMPANY

01 写入优化

02 查询优化

03 Q & A



01

写入优化

# 写入优化

## 1. 主分片优化

过大、过小的影响是什么？

- ◆ 过大
  - 线程资源争抢
  - 资源浪费
- ◆ 过小
  - 查询延迟高
  - 无扩展性

注意：一旦索引创建成功，主分片数无法变更



# 写入优化

## 合适的主分片数

### □ 主分片

- 主分片数 = 节点数
- 主分片数 = magic number (e.g. 5 or 10)
- 主分片数 = 数据总大小 / ( 30G ~ 50G )

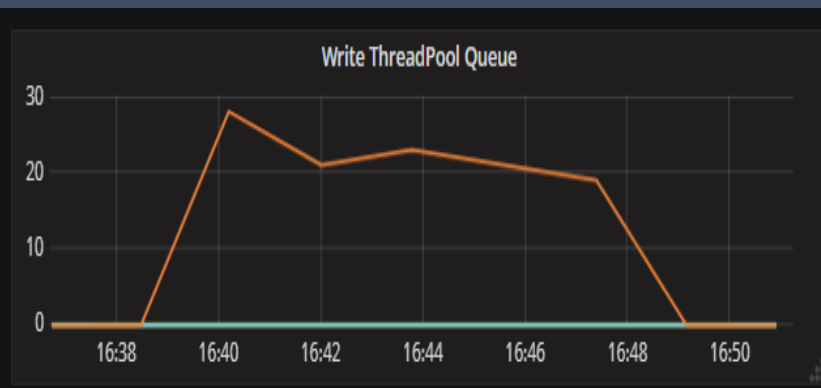
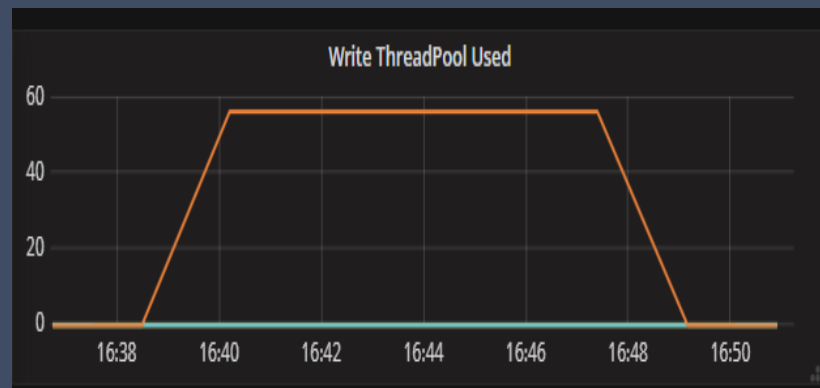
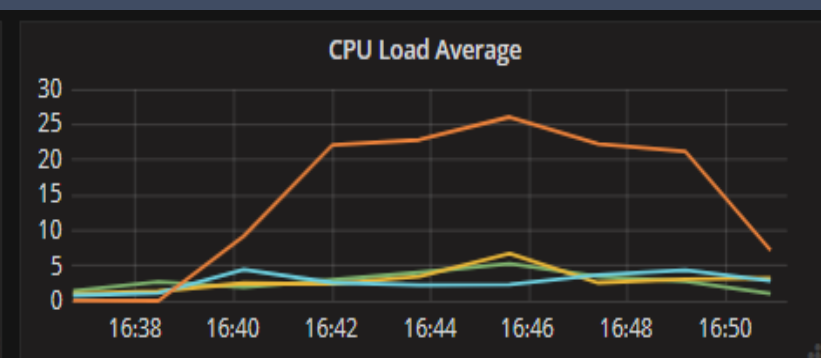
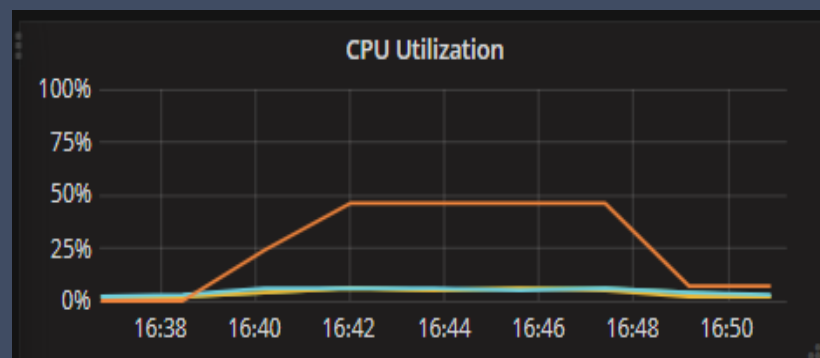


# 写入优化

## 1. 主分片优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 1
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1000
- ❑ 索引刷新频率 1s

Segment count		292	
Min Throughput	index-append	208766	docs/s
Median Throughput	index-append	209518	docs/s
Max Throughput	index-append	217042	docs/s
50th percentile latency	index-append	367.288	ms
90th percentile latency	index-append	557.191	ms
99th percentile latency	index-append	826.304	ms
99.9th percentile latency	index-append	1038.13	ms
99.99th percentile latency	index-append	1197.86	ms
100th percentile latency	index-append	1353.21	ms



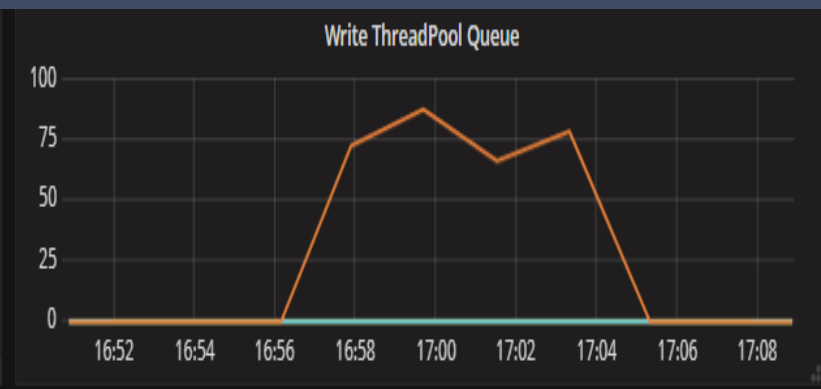
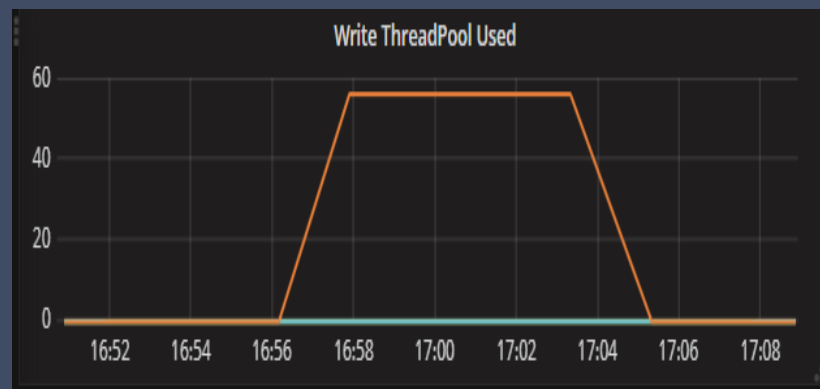
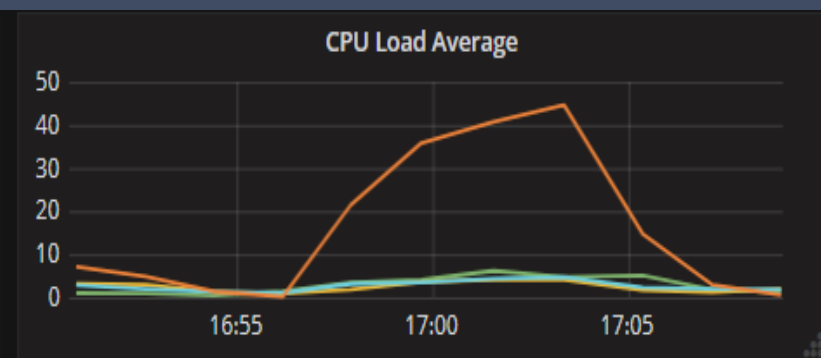
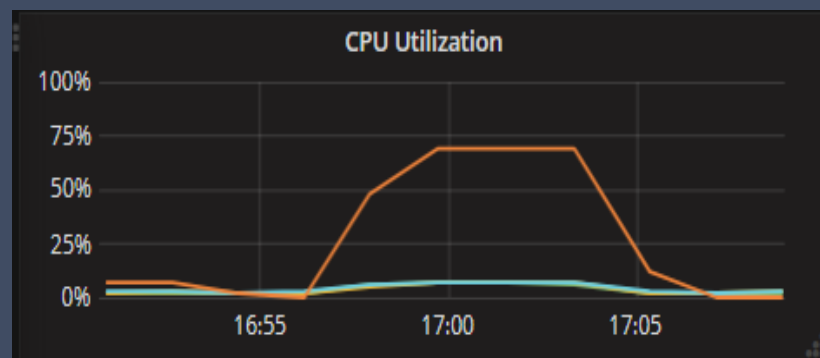


# 写入优化

## 1. 主分片优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 2
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1000
- ❑ 索引刷新频率 1s

Segment count		198	
Min Throughput	index-append	284766	docs/s
Median Throughput	index-append	288703	docs/s
Max Throughput	index-append	326142	docs/s
50th percentile latency	index-append	267.007	ms
90th percentile latency	index-append	408.624	ms
99th percentile latency	index-append	597.733	ms
99.9th percentile latency	index-append	969.17	ms
99.99th percentile latency	index-append	1500.58	ms
100th percentile latency	index-append	1826.63	ms

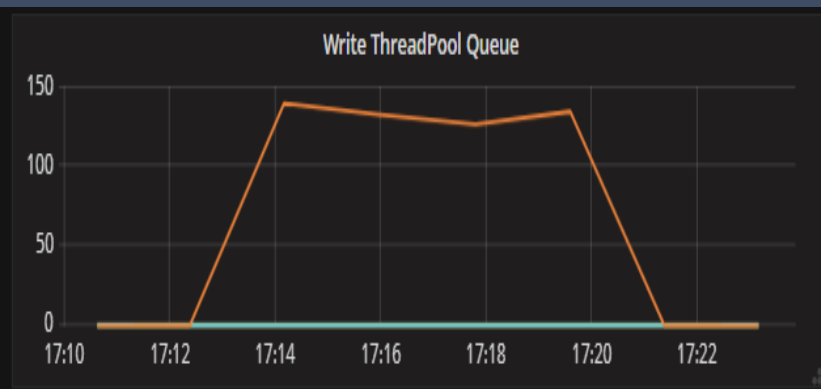
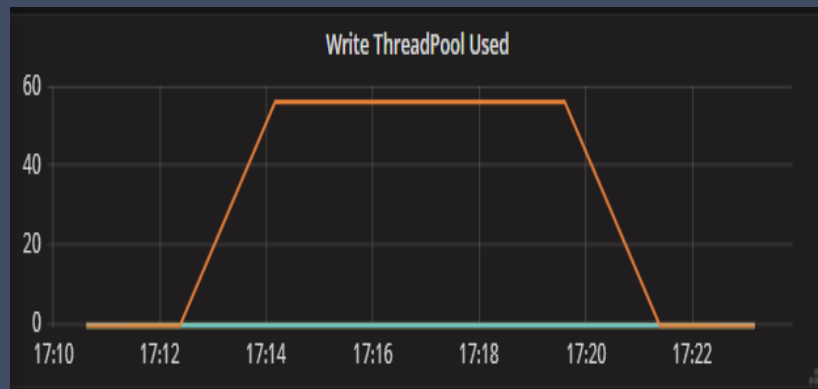
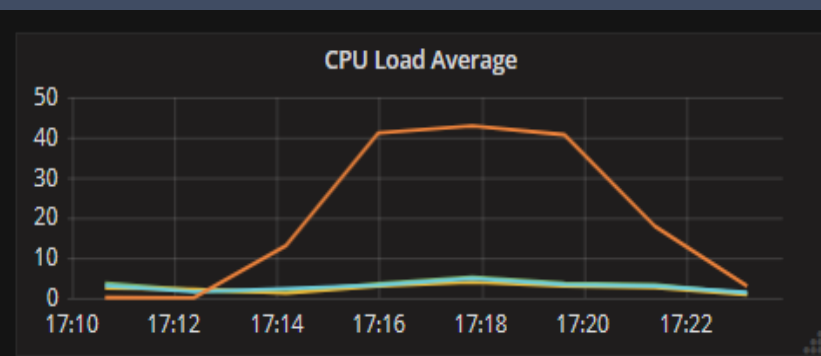
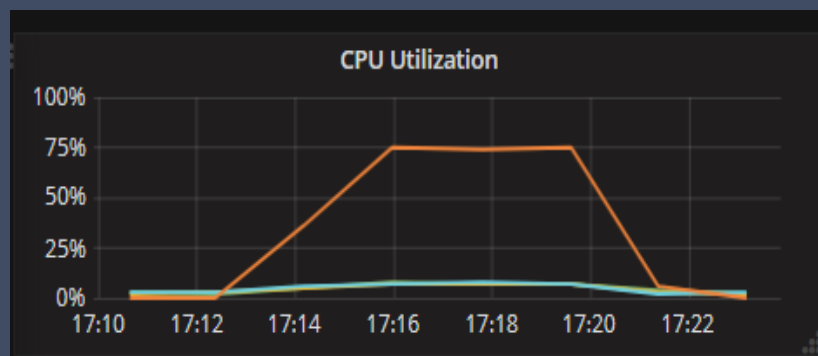


# 写入优化

## 1. 主分片优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1000
- ❑ 索引刷新频率 1s

Segment count		170	
Min Throughput	index-append	293438	docs/s
Median Throughput	index-append	305542	docs/s
Max Throughput	index-append	359086	docs/s
50th percentile latency	index-append	259.533	ms
90th percentile latency	index-append	392	ms
99th percentile latency	index-append	663.315	ms
99.9th percentile latency	index-append	1094.6	ms
99.99th percentile latency	index-append	1312.59	ms
100th percentile latency	index-append	1495.22	ms

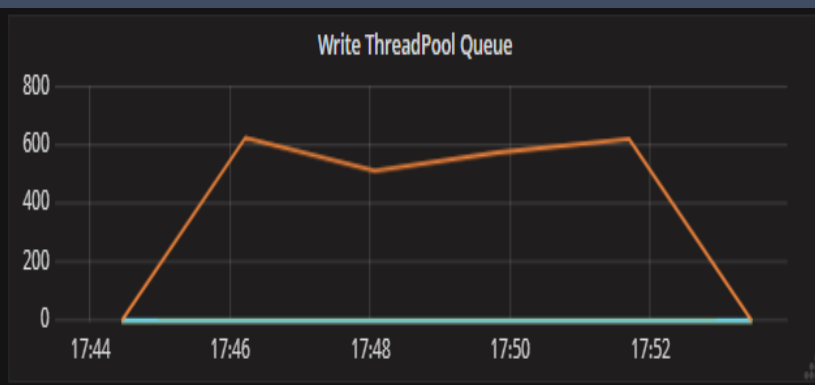
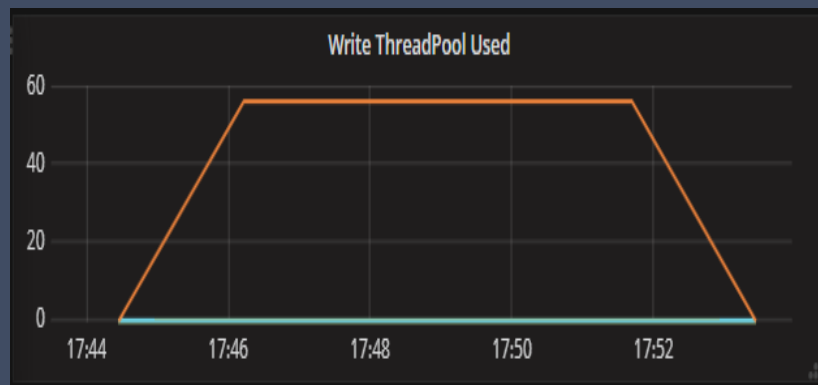
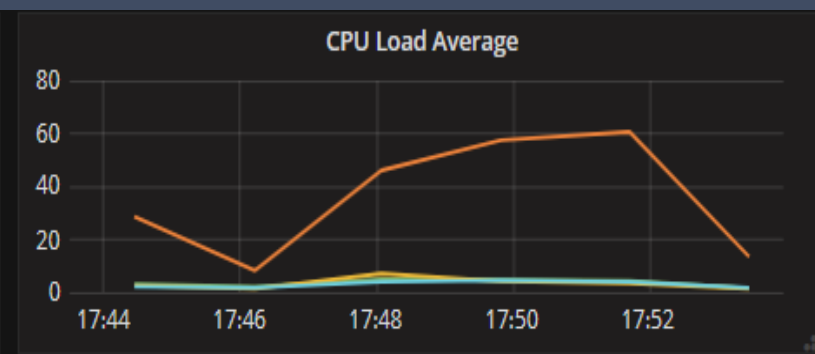
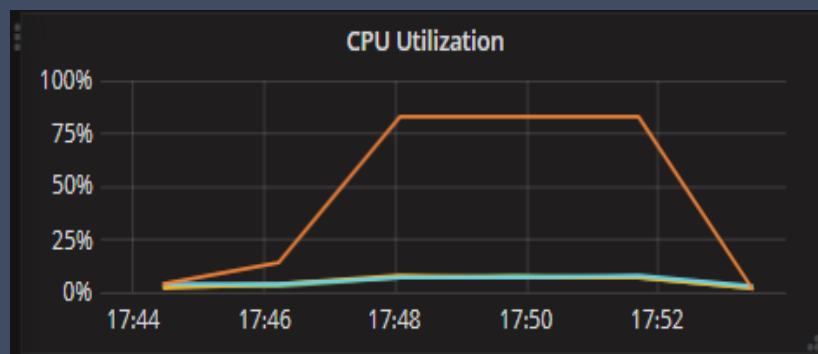


# 写入优化

## 1. 主分片优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 10
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1000
- ❑ 索引刷新频率 1s

Segment count		440	
Min Throughput	index-append	328129	docs/s
Median Throughput	index-append	350179	docs/s
Max Throughput	index-append	447702	docs/s
50th percentile latency	index-append	207.941	ms
90th percentile latency	index-append	342.846	ms
99th percentile latency	index-append	940.485	ms
99.9th percentile latency	index-append	1485.43	ms
99.99th percentile latency	index-append	1823.91	ms
100th percentile latency	index-append	2217.87	ms



# 写入优化

## 1. 主分片优化

- 主分片数等于节点，单分片吞吐量最高，  
无扩展能力
- 单节点分片数过多会，会导致写入线程资源争抢  
极限情况可能会导致写入请求拒绝

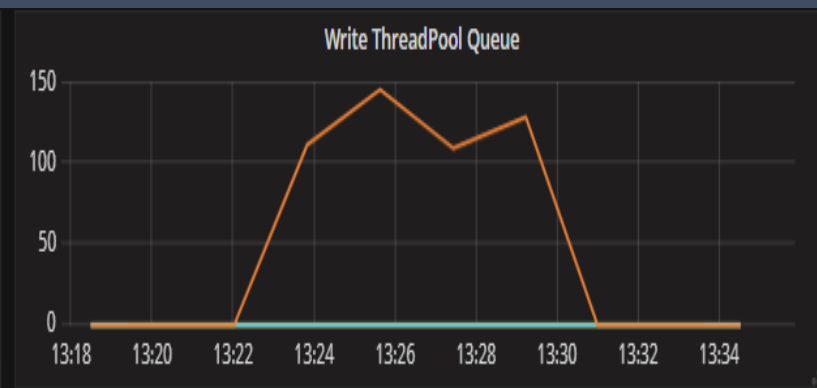
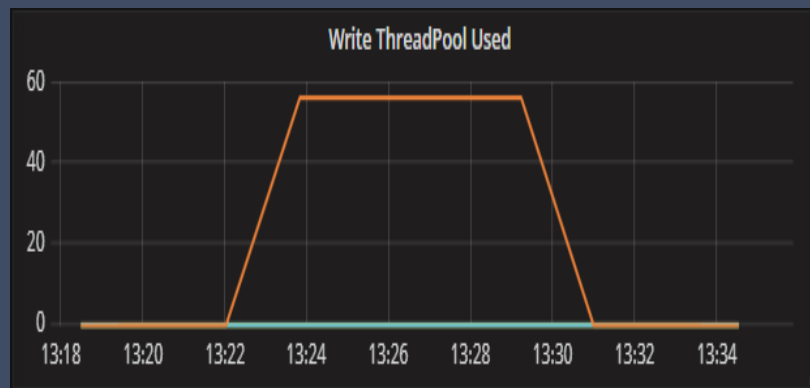
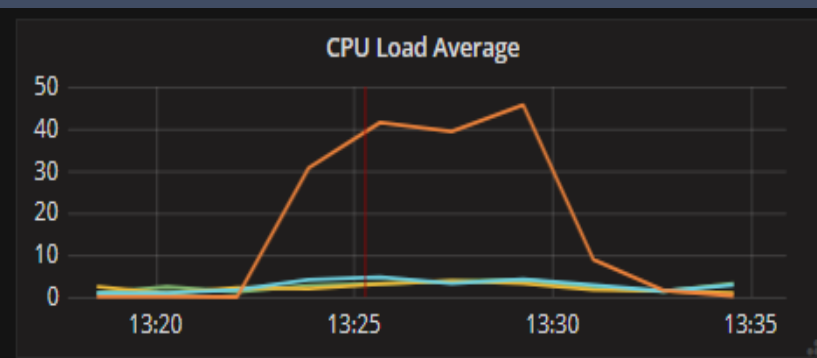
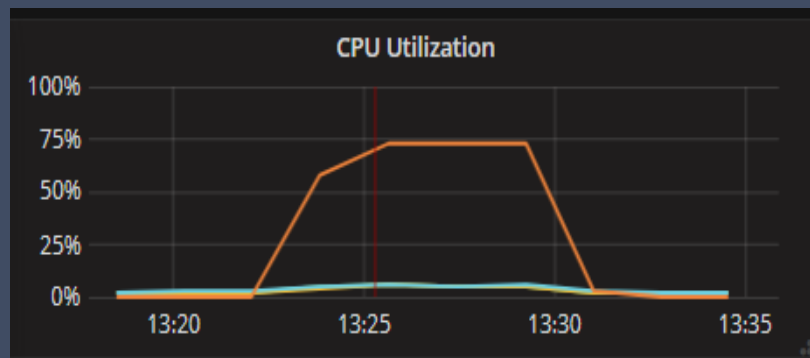


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 500
- ❑ 索引刷新频率 1s

Segment count		165	
Min Throughput	index-append	289074	docs/s
Median Throughput	index-append	301336	docs/s
Max Throughput	index-append	357213	docs/s
50th percentile latency	index-append	123.815	ms
90th percentile latency	index-append	202.592	ms
99th percentile latency	index-append	427.203	ms
99.9th percentile latency	index-append	853.101	ms
99.99th percentile latency	index-append	1029.44	ms
100th percentile latency	index-append	1639.78	ms

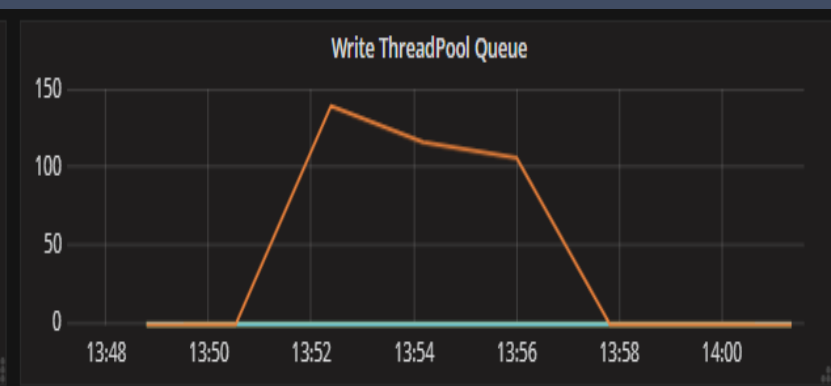
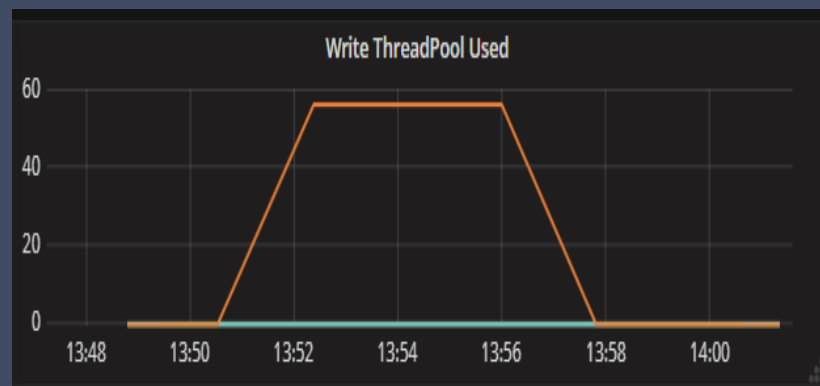
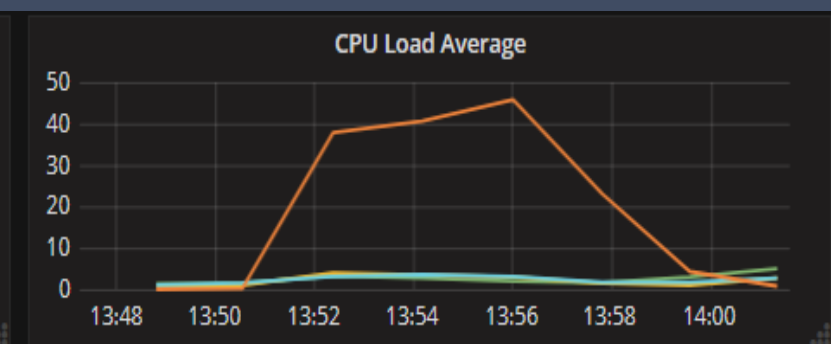
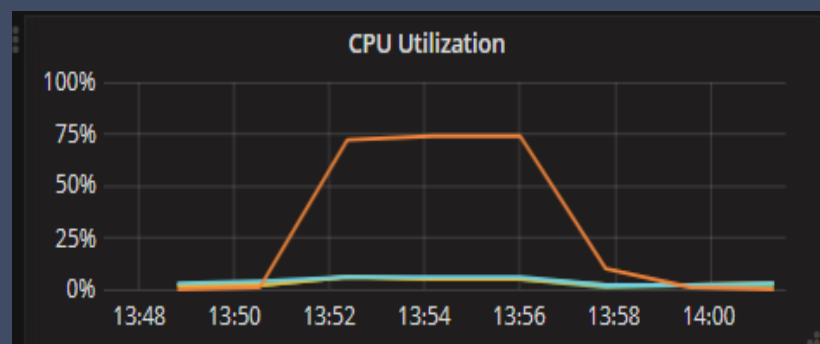


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1000
- ❑ 索引刷新频率 1s

Segment count		221	
Min Throughput	index-append	299594	docs/s
Median Throughput	index-append	312142	docs/s
Max Throughput	index-append	363890	docs/s
50th percentile latency	index-append	254.893	ms
90th percentile latency	index-append	385.335	ms
99th percentile latency	index-append	622.733	ms
99.9th percentile latency	index-append	1115.43	ms
99.99th percentile latency	index-append	1408.71	ms
100th percentile latency	index-append	1523.54	ms

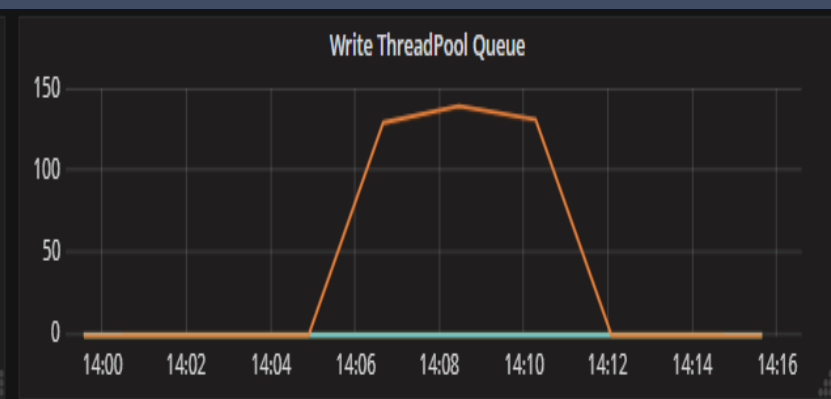
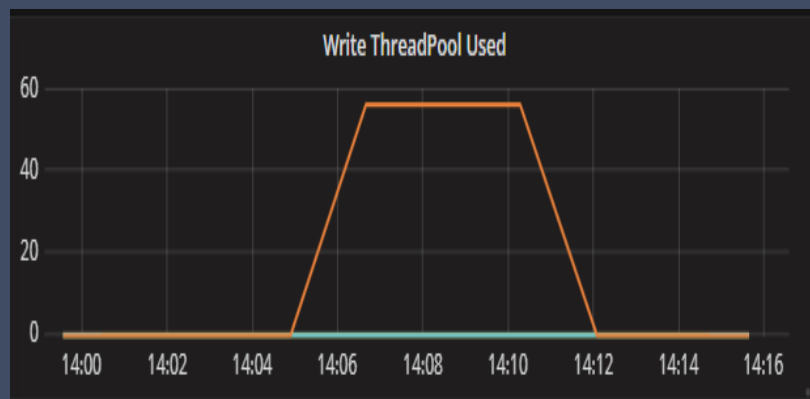
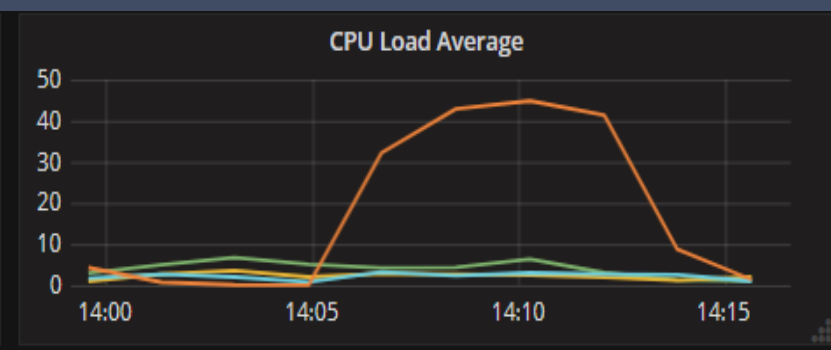
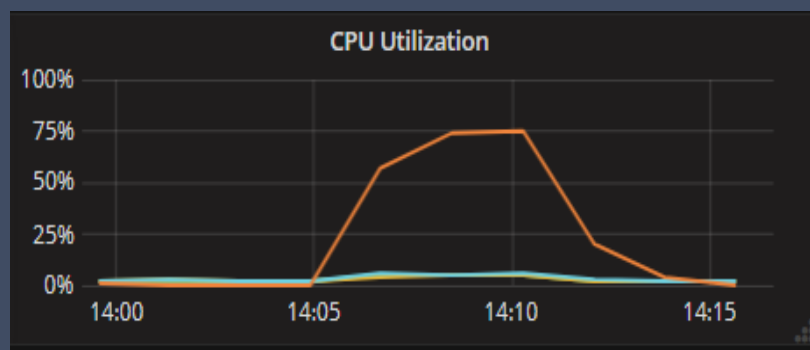


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 1500
- ❑ 索引刷新频率 1s

Segment count		304	
Min Throughput	index-append	301654	docs/s
Median Throughput	index-append	313163	docs/s
Max Throughput	index-append	364824	docs/s
50th percentile latency	index-append	389.628	ms
90th percentile latency	index-append	564.295	ms
99th percentile latency	index-append	821.281	ms
99.9th percentile latency	index-append	1440.25	ms
99.99th percentile latency	index-append	1742.44	ms
100th percentile latency	index-append	1821.15	ms



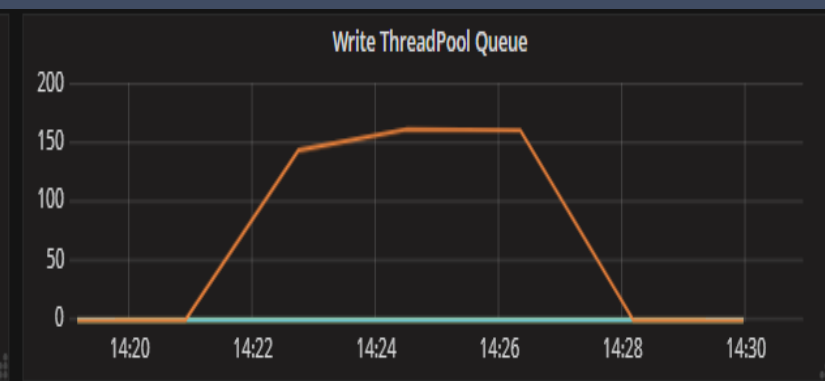
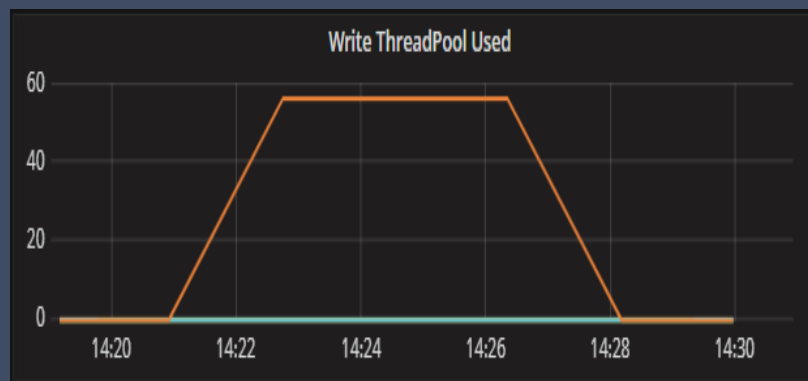
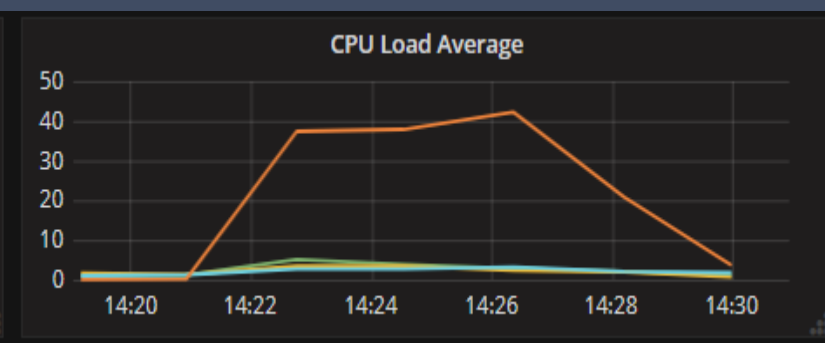
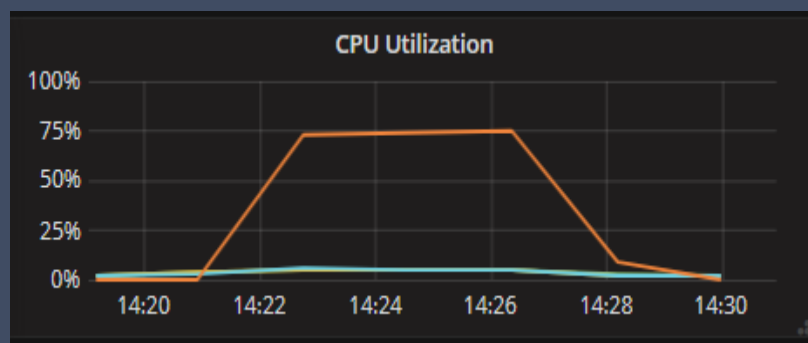


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 3000
- ❑ 索引刷新频率 1s

Segment count		261	
Min Throughput	index-append	307396	docs/s
Median Throughput	index-append	317716	docs/s
Max Throughput	index-append	359466	docs/s
50th percentile latency	index-append	795.554	ms
90th percentile latency	index-append	1023.63	ms
99th percentile latency	index-append	1393.64	ms
99.9th percentile latency	index-append	2171.94	ms
99.99th percentile latency	index-append	2392.3	ms
100th percentile latency	index-append	2630.19	ms



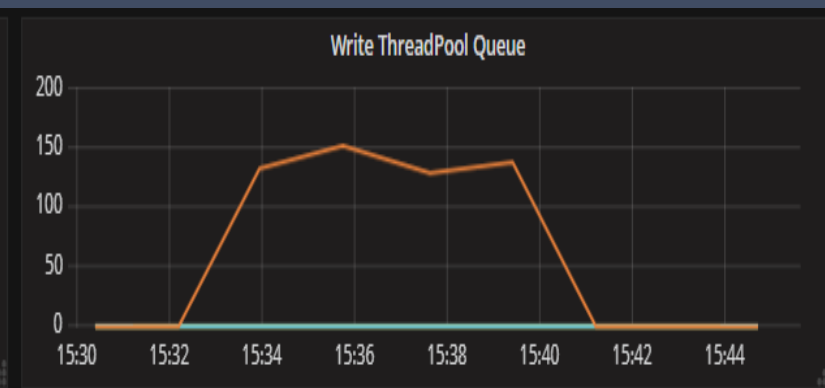
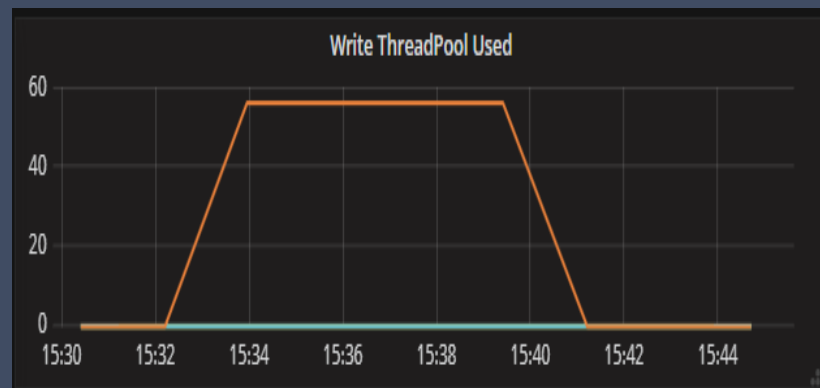
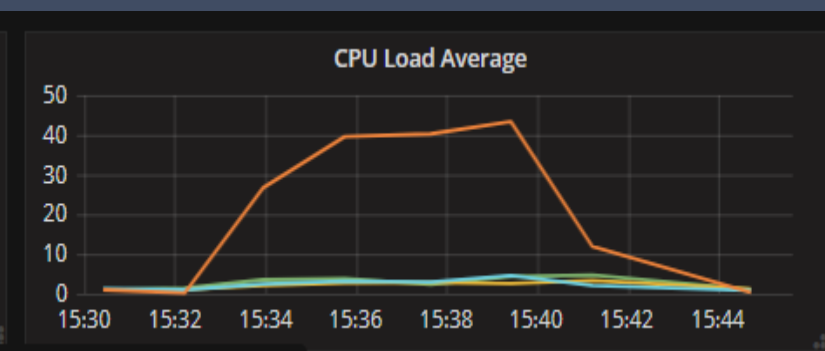
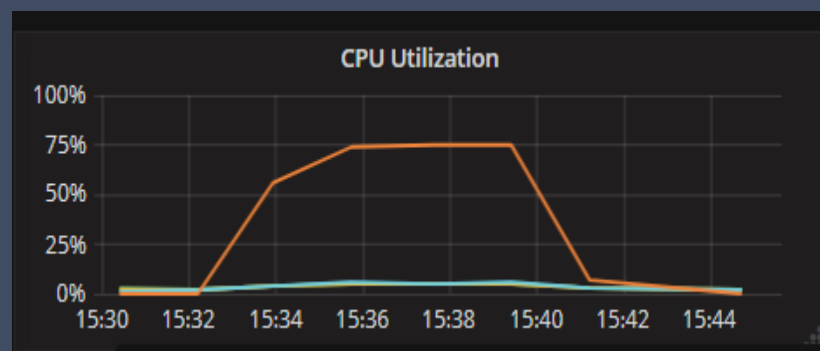


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 5000
- ❑ 索引刷新频率 1s

Segment count		212	
Min Throughput	index-append	305030	docs/s
Median Throughput	index-append	315299	docs/s
Max Throughput	index-append	360937	docs/s
50th percentile latency	index-append	1345.01	ms
90th percentile latency	index-append	1681.78	ms
99th percentile latency	index-append	2152.5	ms
99.9th percentile latency	index-append	2464.56	ms
99.99th percentile latency	index-append	2718.04	ms
100th percentile latency	index-append	3084.74	ms

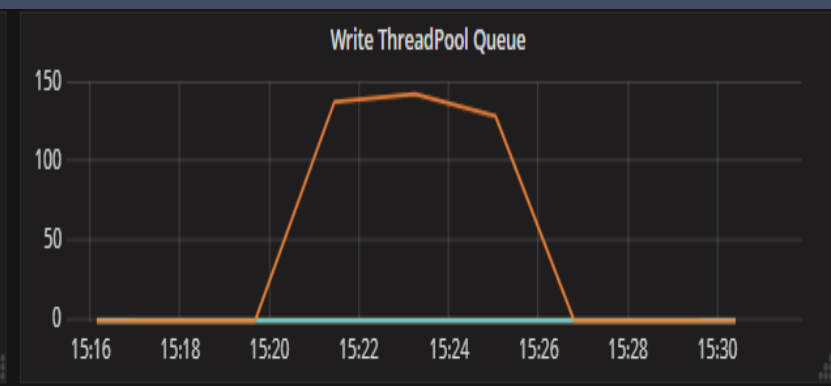
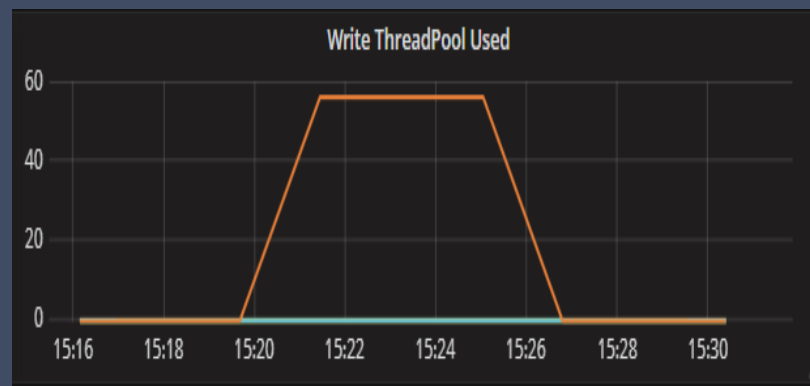
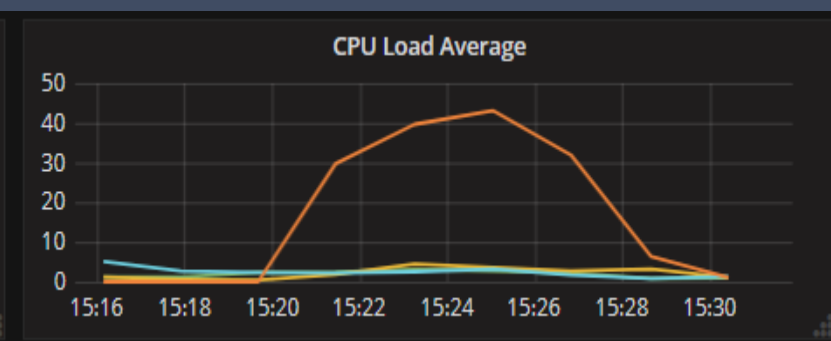
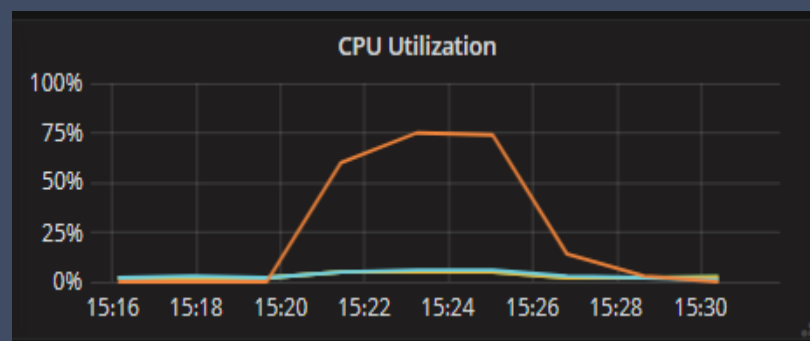


# 写入优化

## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 10000
- ❑ 索引刷新频率 1s

Segment count		299	
Min Throughput	index-append	278721	docs/s
Median Throughput	index-append	288335	docs/s
Max Throughput	index-append	320241	docs/s
50th percentile latency	index-append	2973.69	ms
90th percentile latency	index-append	3569.98	ms
99th percentile latency	index-append	4209.94	ms
99.9th percentile latency	index-append	4843.56	ms
99.99th percentile latency	index-append	5472.5	ms
100th percentile latency	index-append	5550.65	ms

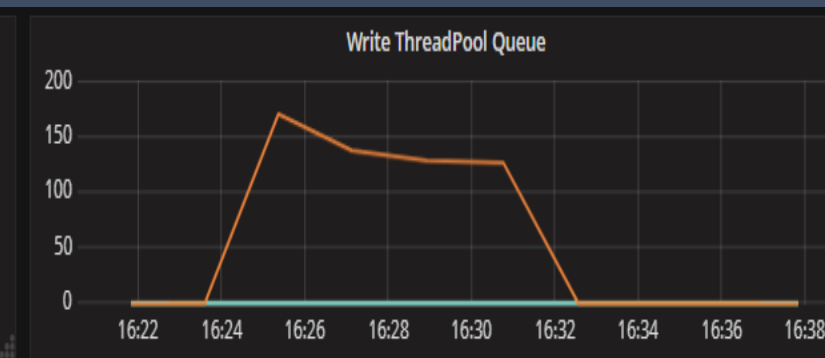
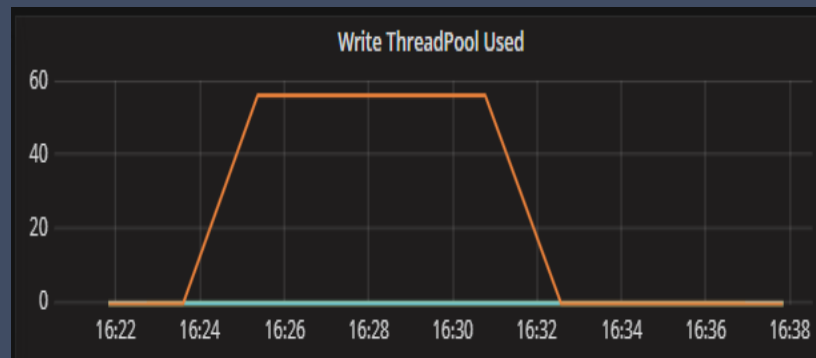
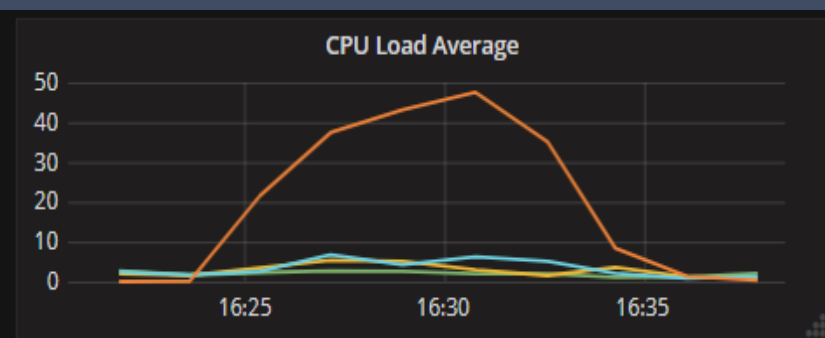
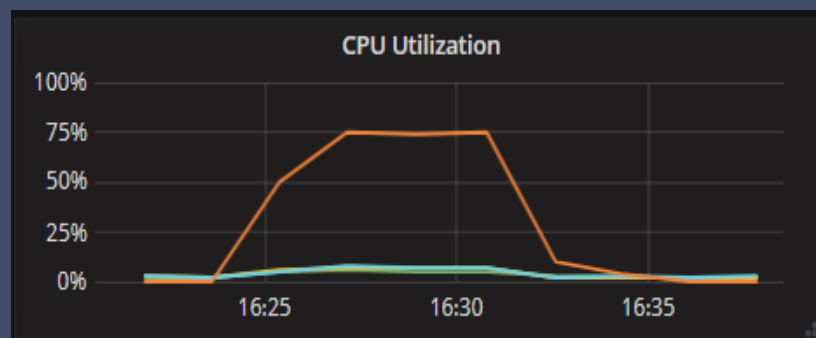


# 写入优化

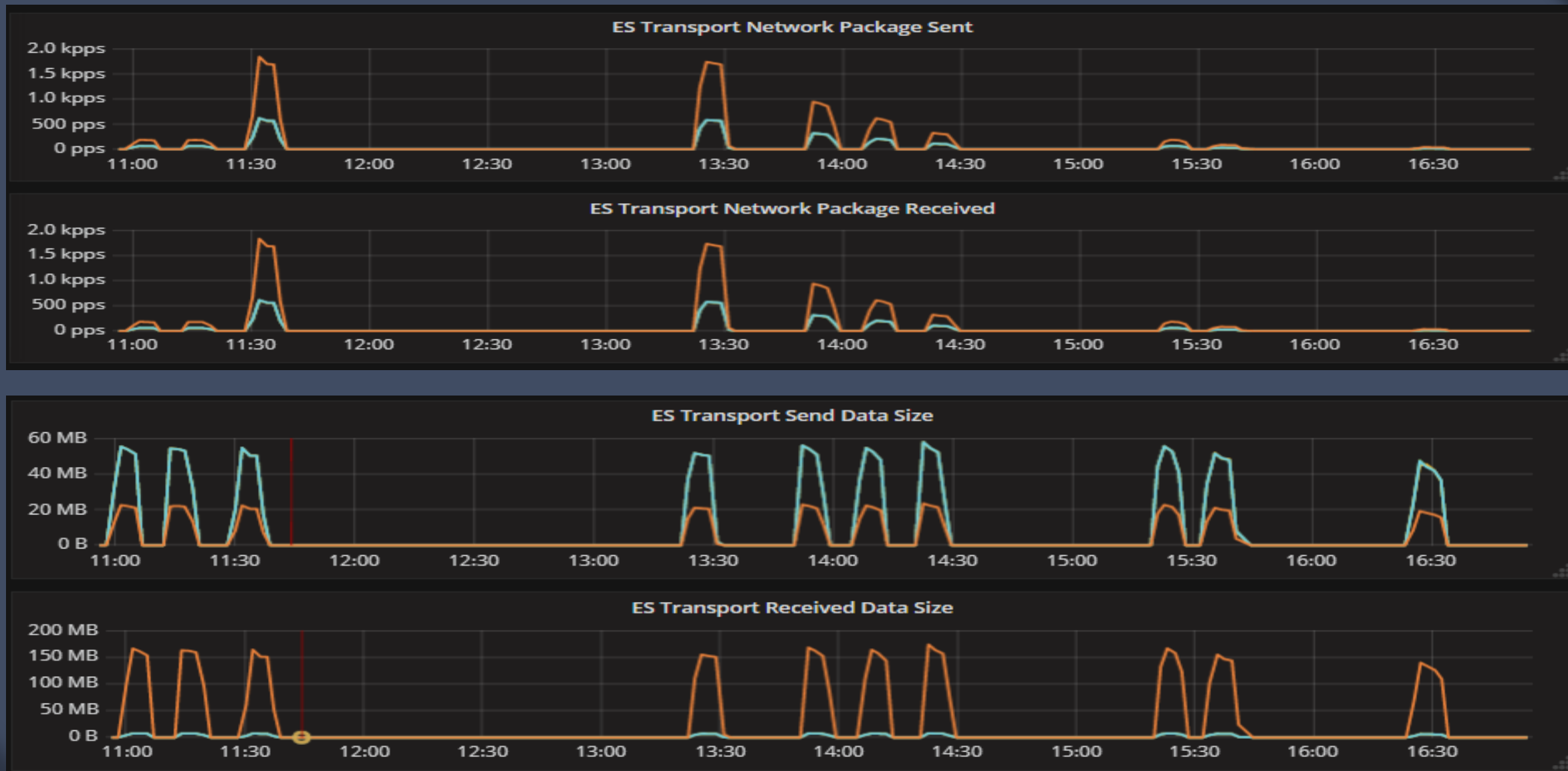
## 2. 批量提交优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 20000
- ❑ 索引刷新频率 1s

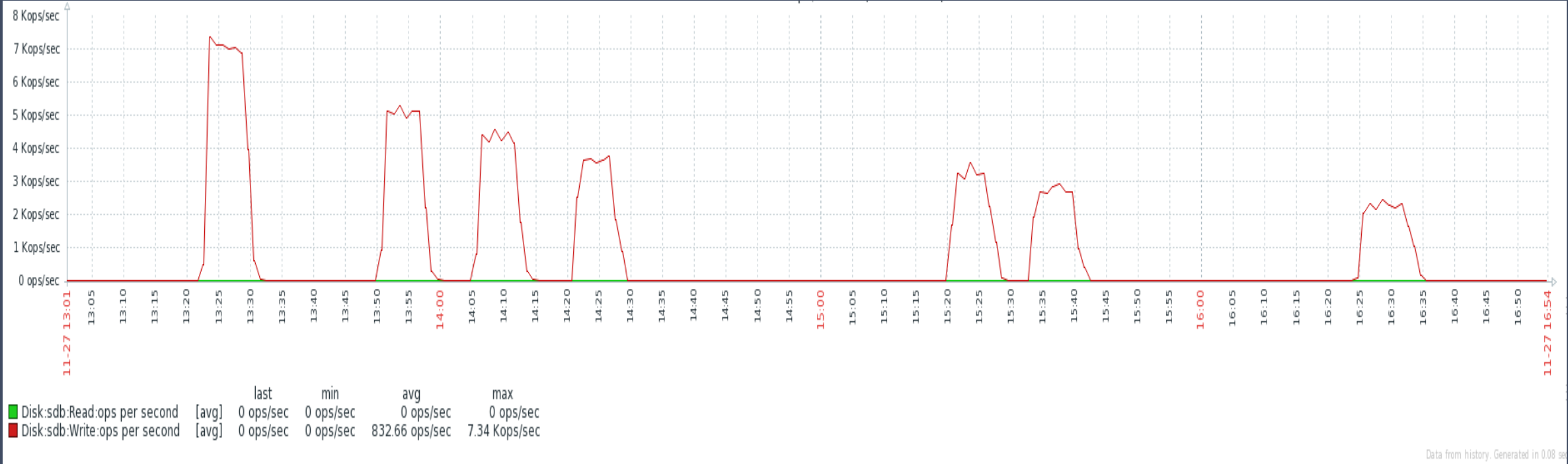
Segment count		182	
Min Throughput	index-append	247515	docs/s
Median Throughput	index-append	257909	docs/s
Max Throughput	index-append	279505	docs/s
50th percentile latency	index-append	6762.89	ms
90th percentile latency	index-append	7659.79	ms
99th percentile latency	index-append	8341.83	ms
99.9th percentile latency	index-append	8846.83	ms
100th percentile latency	index-append	9501.2	ms



# 写入优化



# 写入优化



# 写入优化

## 2. 批量提交优化

测试验证，适当的提升Bulk Size  
可以减少网络包的传输和降低磁盘IOPS  
提升写入吞吐量

注意：

- 过大的Bulk Size 会占用过多的内存
- 过大的Bulk Size 会降低吞吐量
- 过小的Bulk Size 会增加网络压力
- 过小的Bulk Size 会增加磁盘压力



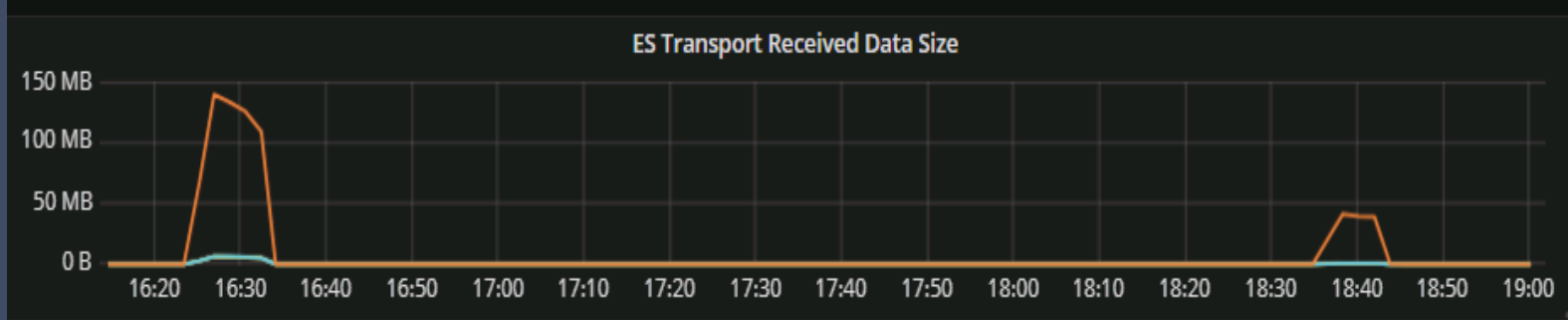
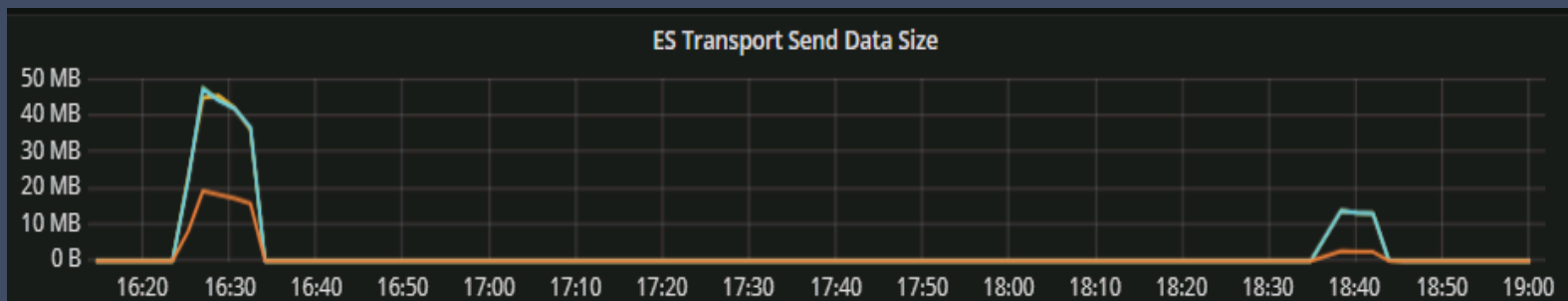
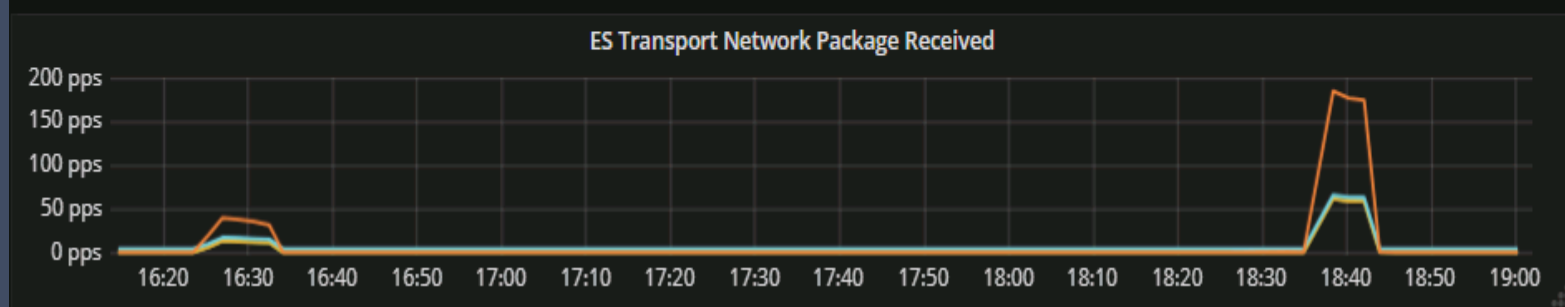
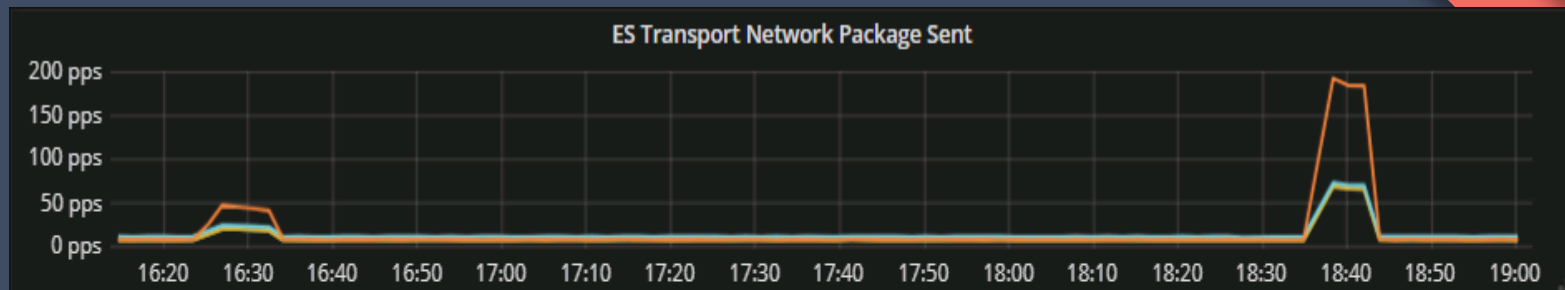
# 写入优化

## 3. 网络流量优化

测试验证，采用 DEFLATE 算法  
压缩流量，可节省3倍以上的带宽

# transport 模块主要用于集群内部节点间的通信  
# 默认网络流量是未压缩的  
# 支持使用 DEFLATE 算法无损压力数据，  
# 降低网络传输流量

`transport.tcp.compress: true`



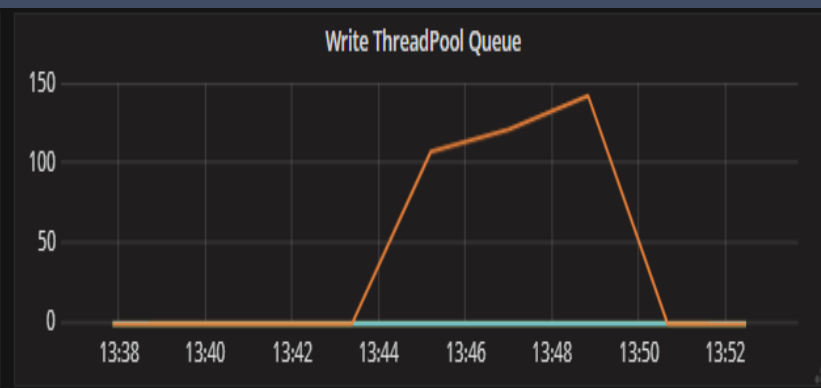
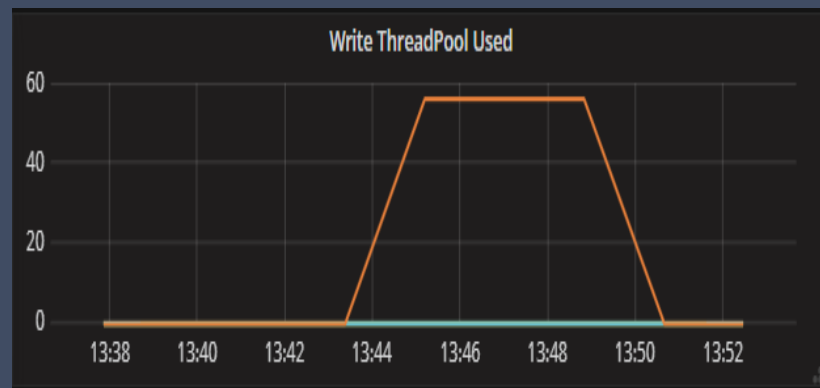
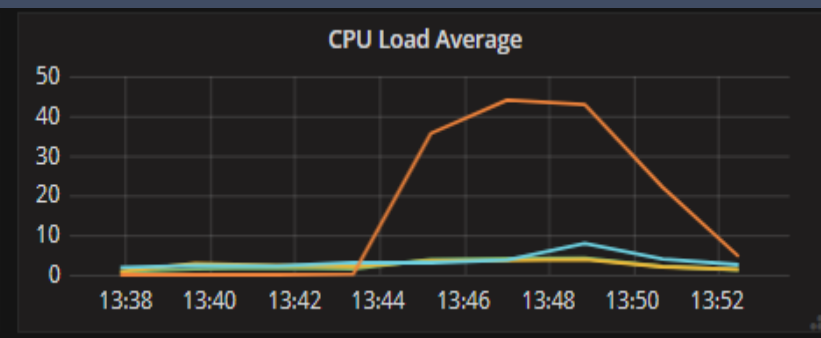
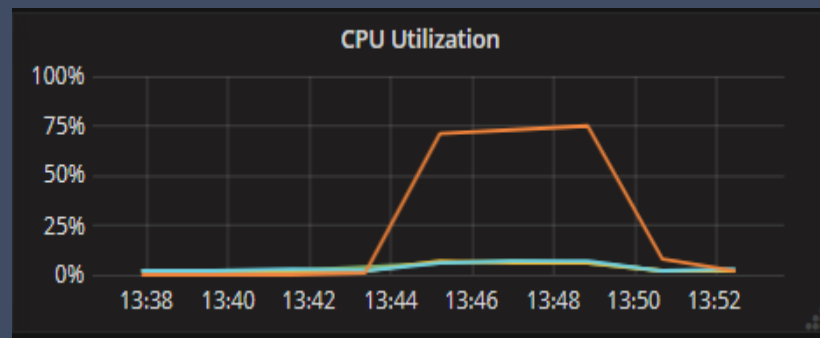


# 写入优化

## 4. 索引刷新频率优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 5000
- ❑ 索引刷新频率 1s

Segment count		266	
Min Throughput	index-append	294811	docs/s
Median Throughput	index-append	302150	docs/s
Max Throughput	index-append	344716	docs/s
50th percentile latency	index-append	1391.45	ms
90th percentile latency	index-append	1692.99	ms
99th percentile latency	index-append	2293.46	ms
99.9th percentile latency	index-append	4294.62	ms
99.99th percentile latency	index-append	4510.63	ms
100th percentile latency	index-append	4592.79	ms



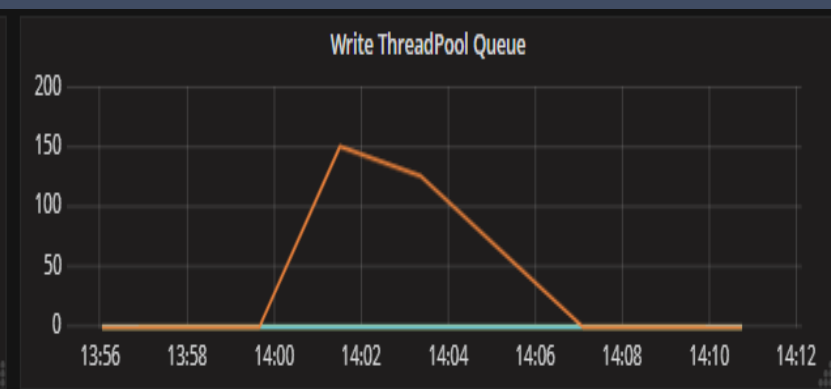
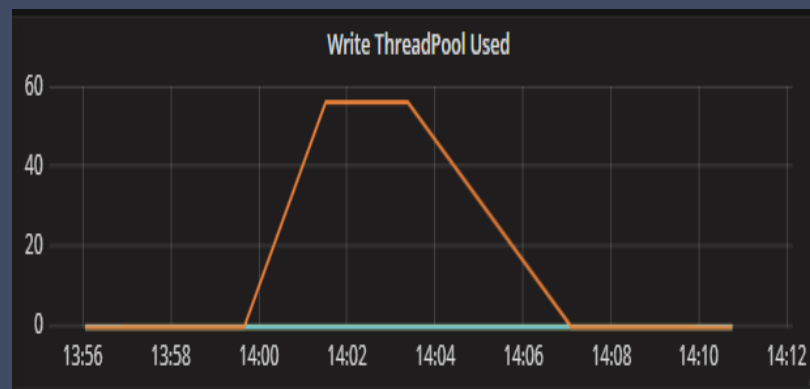
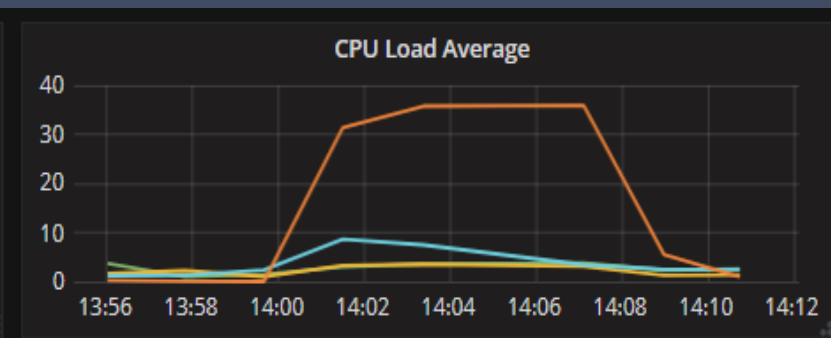
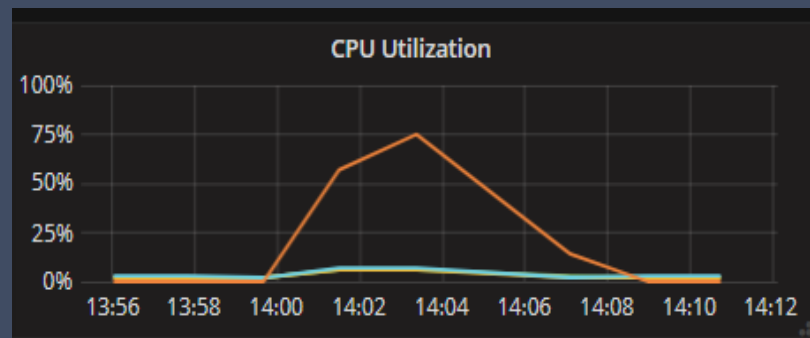


# 写入优化

## 4. 索引刷新频率优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 5000
- ❑ 索引刷新频率 3s

Segment count		272	
Min Throughput	index-append	293946	docs/s
Median Throughput	index-append	303302	docs/s
Max Throughput	index-append	366849	docs/s
50th percentile latency	index-append	1392.87	ms
90th percentile latency	index-append	1816.18	ms
99th percentile latency	index-append	2213.44	ms
99.9th percentile latency	index-append	2734.22	ms
99.99th percentile latency	index-append	3241.33	ms
100th percentile latency	index-append	3471.01	ms

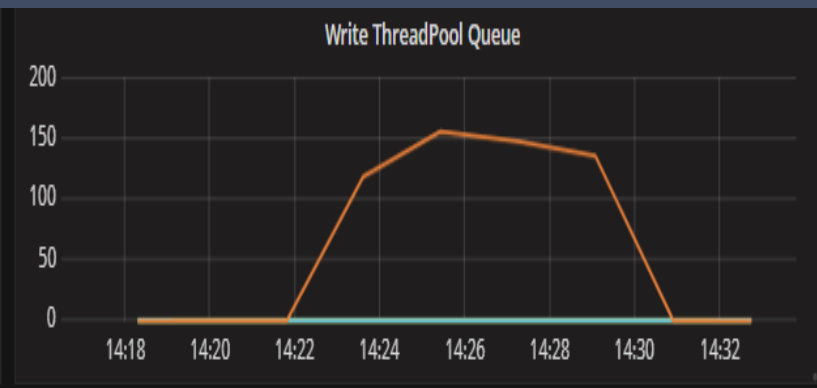
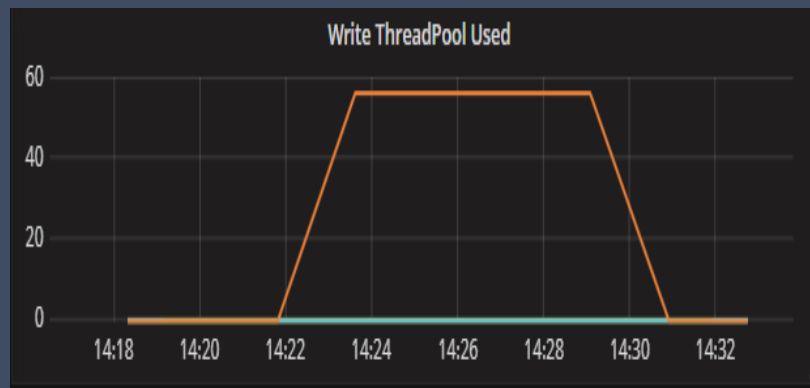
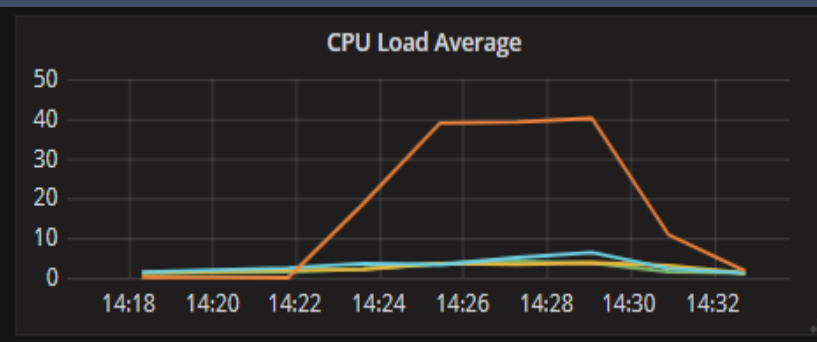
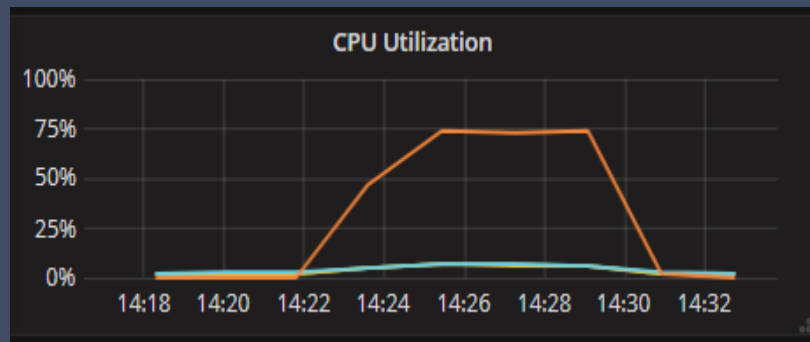


# 写入优化

## 4. 索引刷新频率优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 5000
- ❑ 索引刷新频率 10s

Segment count		322	
Min Throughput	index-append	295641	docs/s
Median Throughput	index-append	306557	docs/s
Max Throughput	index-append	380160	docs/s
50th percentile latency	index-append	1386.63	ms
90th percentile latency	index-append	1798.86	ms
99th percentile latency	index-append	2302.17	ms
99.9th percentile latency	index-append	2685.13	ms
99.99th percentile latency	index-append	3044.71	ms
100th percentile latency	index-append	3411.25	ms

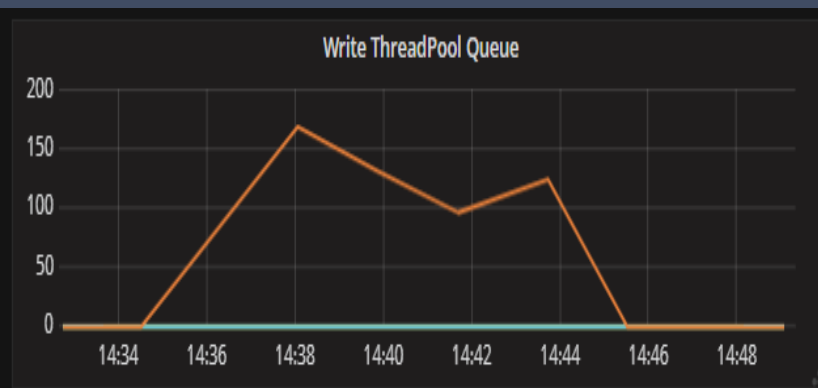
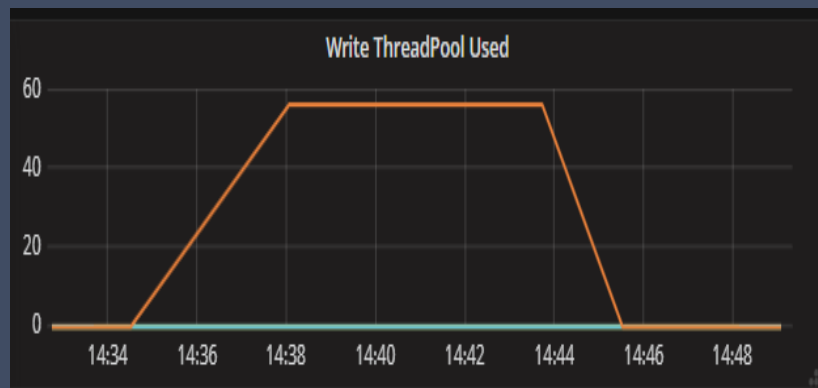
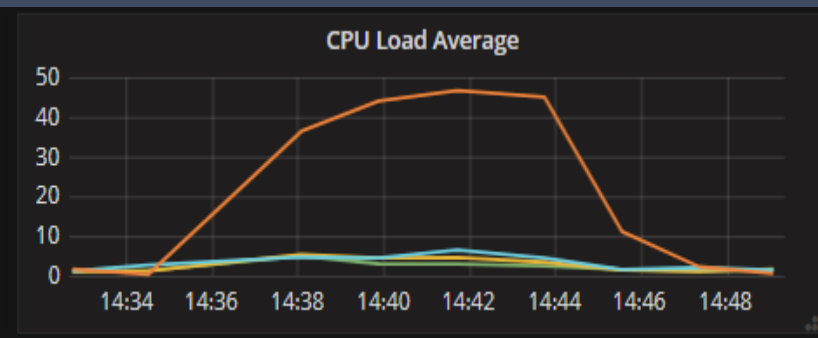
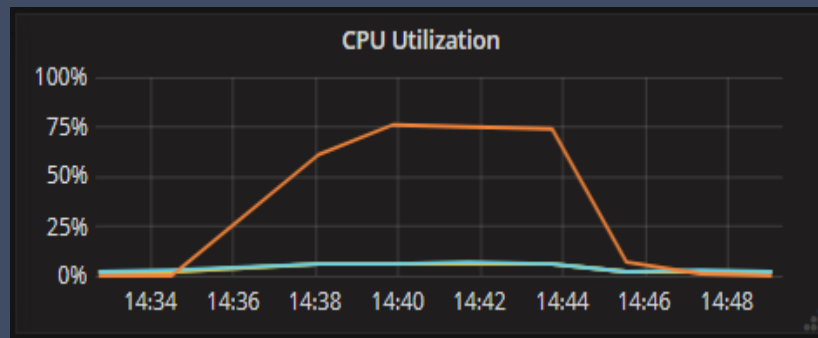


# 写入优化

## 4. 索引刷新频率优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ bulk-size : 5000
- ❑ 索引刷新频率 100ms

Segment count		148	
Min Throughput	index-append	268801	docs/s
Median Throughput	index-append	279442	docs/s
Max Throughput	index-append	302481	docs/s
50th percentile latency	index-append	1528.45	ms
90th percentile latency	index-append	1870.08	ms
99th percentile latency	index-append	2396.68	ms
99.9th percentile latency	index-append	2965.87	ms
99.99th percentile latency	index-append	3869.03	ms
100th percentile latency	index-append	4112.3	ms



# 写入优化

## 4. 索引刷新频率优化

测试验证，索引刷新频率越高，对性能的影响越大，实时性越好

如果对数据的可见性没有特殊要求  
`index.refresh_interval: 3s`

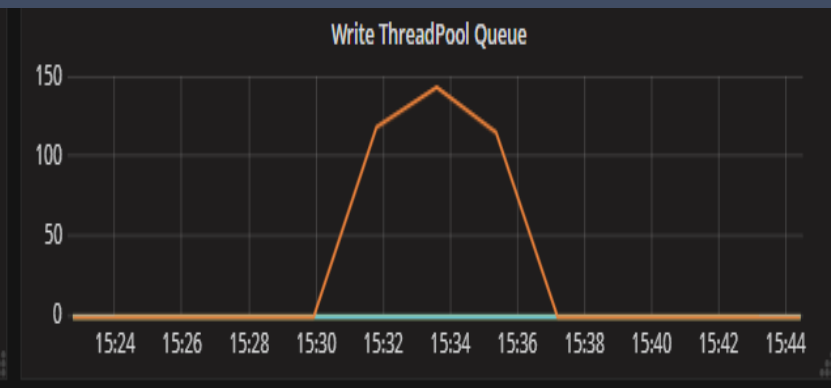
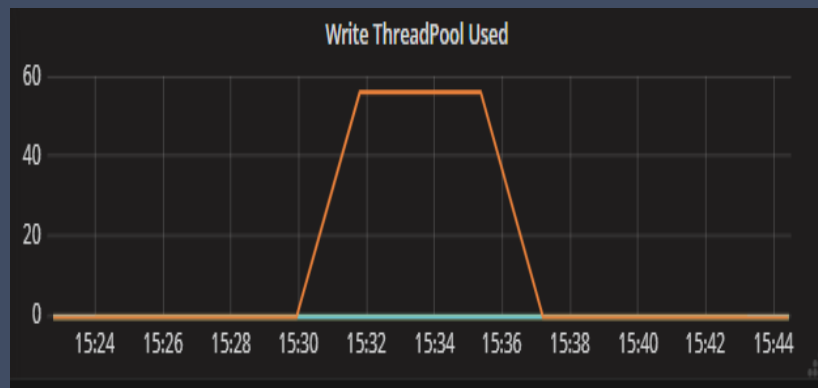
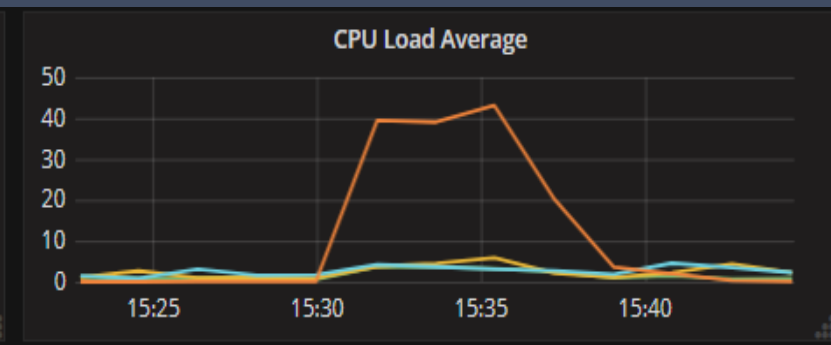
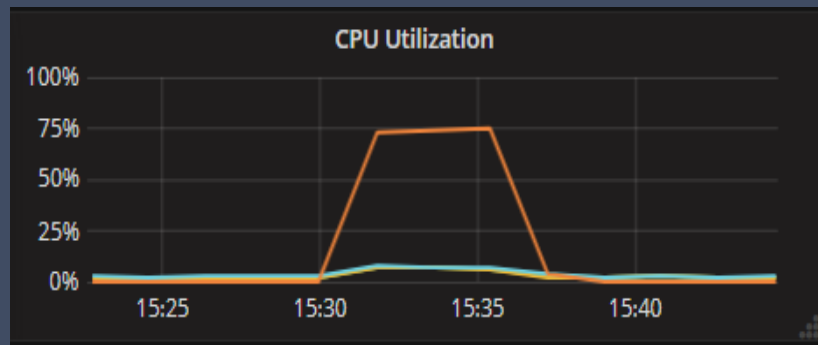


# 写入优化

## 5. 索引缓冲区优化

- CPU : 50+
- 内存分配 : 20G
- 节点个数: 1
- 分片数 : 3
- 副本数 : 0
- 并发度 : 340
- bulk-size : 5000
- 索引刷新频率 3s
- 索引缓冲区 48m

Segment count		331	
Min Throughput	index-append	248972	docs/s
Median Throughput	index-append	254597	docs/s
Max Throughput	index-append	267239	docs/s
50th percentile latency	index-append	1628	ms
90th percentile latency	index-append	2152.23	ms
99th percentile latency	index-append	2818.54	ms
99.9th percentile latency	index-append	3721.08	ms
99.99th percentile latency	index-append	4300.05	ms
100th percentile latency	index-append	4533.13	ms

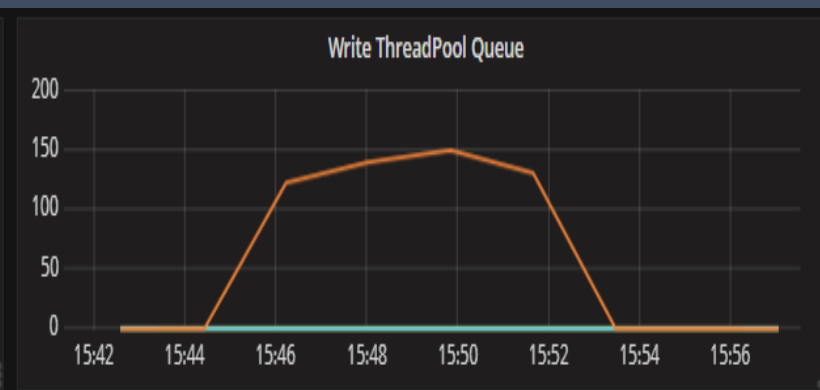
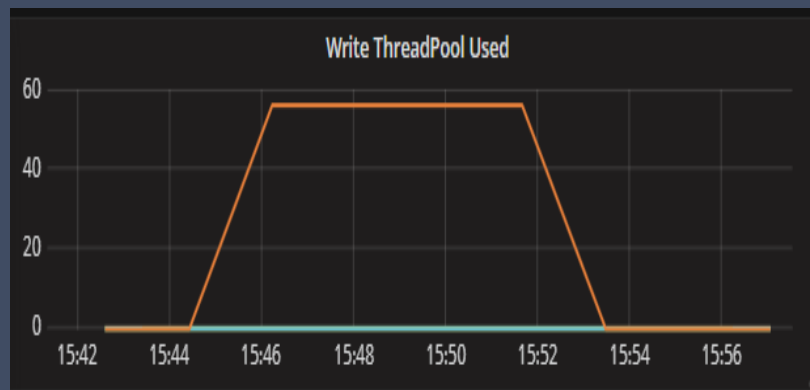
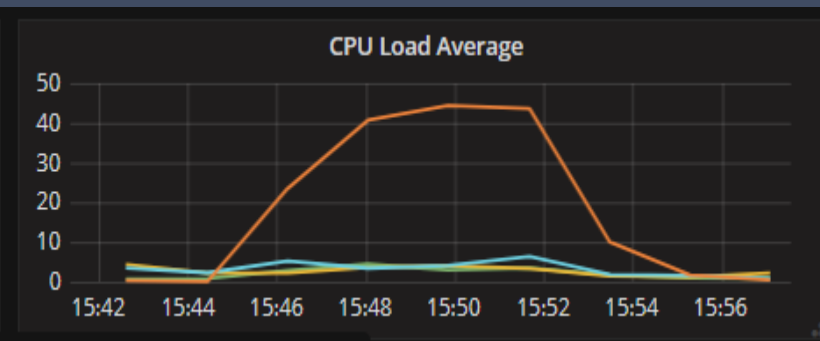
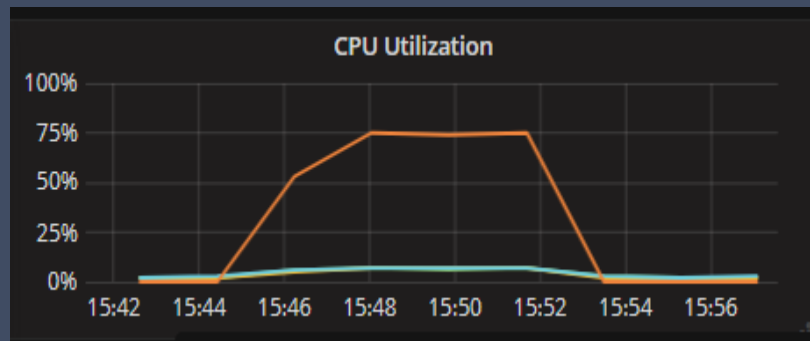


# 写入优化

## 5. 索引缓冲区优化

- CPU : 50+
- 内存分配 : 20G
- 节点个数: 1
- 分片数 : 3
- 副本数 : 0
- 并发度 : 340
- bulk-size : 5000
- 索引刷新频率 3s
- 索引缓冲区 512m

Segment count		285	
Min Throughput	index-append	298795	docs/s
Median Throughput	index-append	307513	docs/s
Max Throughput	index-append	357478	docs/s
50th percentile latency	index-append	1368.81	ms
90th percentile latency	index-append	1744.36	ms
99th percentile latency	index-append	2246.21	ms
99.9th percentile latency	index-append	2926.98	ms
99.99th percentile latency	index-append	3438.82	ms
100th percentile latency	index-append	3848.19	ms

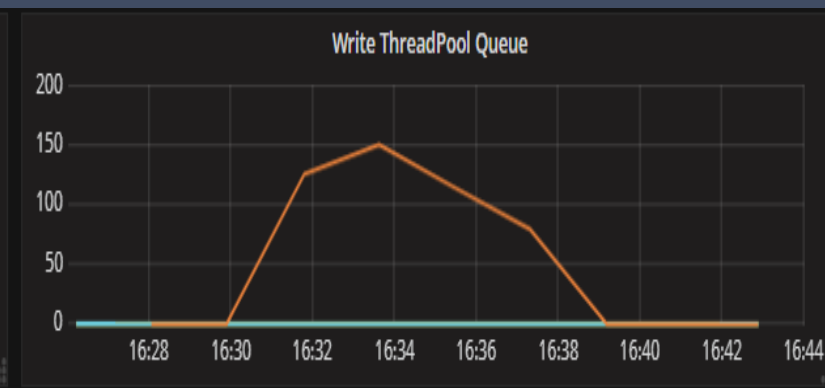
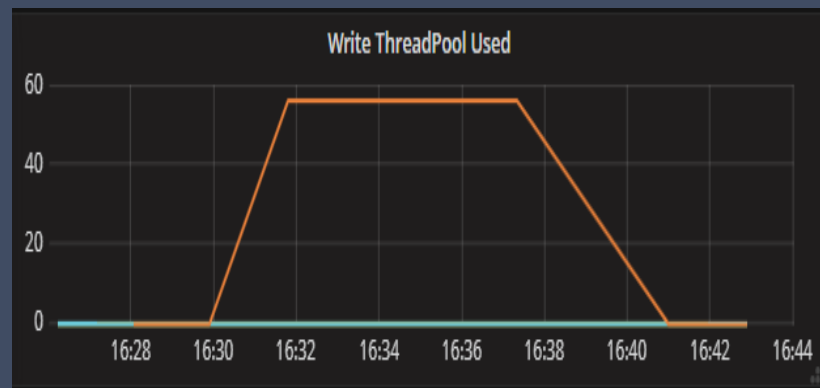
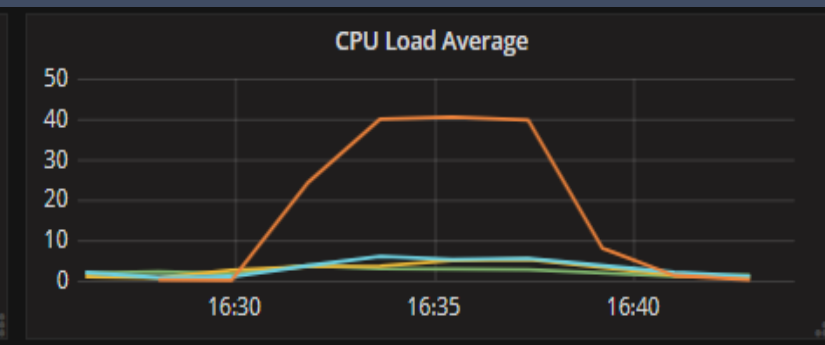
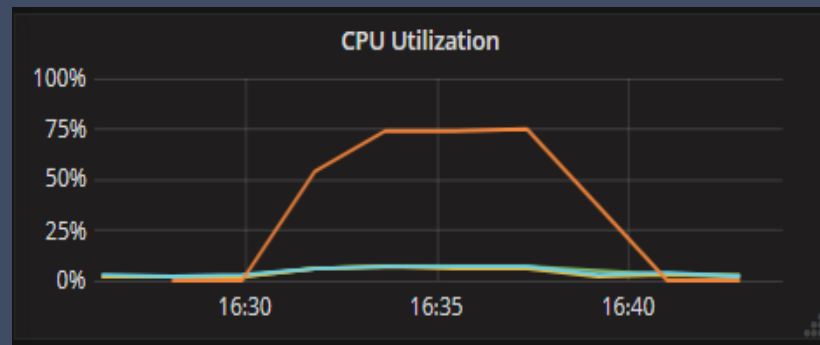


# 写入优化

## 5. 索引缓冲区优化

- CPU : 50+
- 内存分配 : 20G
- 节点个数: 1
- 分片数 : 3
- 副本数 : 0
- 并发度 : 340
- bulk-size : 5000
- 索引刷新频率 3s
- 索引缓冲区 1024m

Segment count		190	
Min Throughput	index-append	300265	docs/s
Median Throughput	index-append	309879	docs/s
Max Throughput	index-append	355984	docs/s
50th percentile latency	index-append	1354.73	ms
90th percentile latency	index-append	1805.48	ms
99th percentile latency	index-append	2293.45	ms
99.9th percentile latency	index-append	2771.52	ms
99.99th percentile latency	index-append	3251.81	ms
100th percentile latency	index-append	3470.64	ms



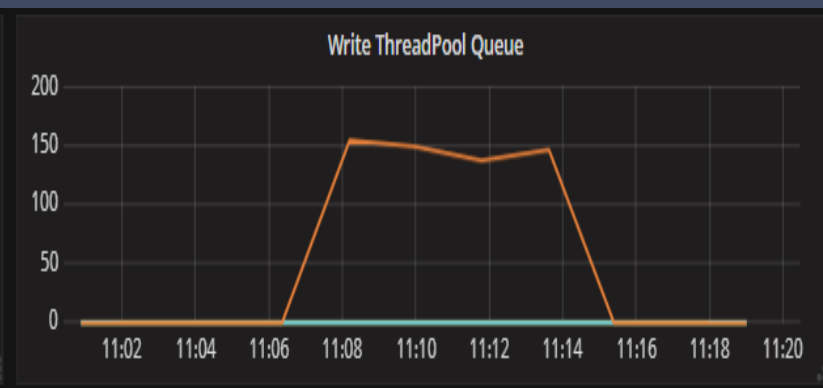
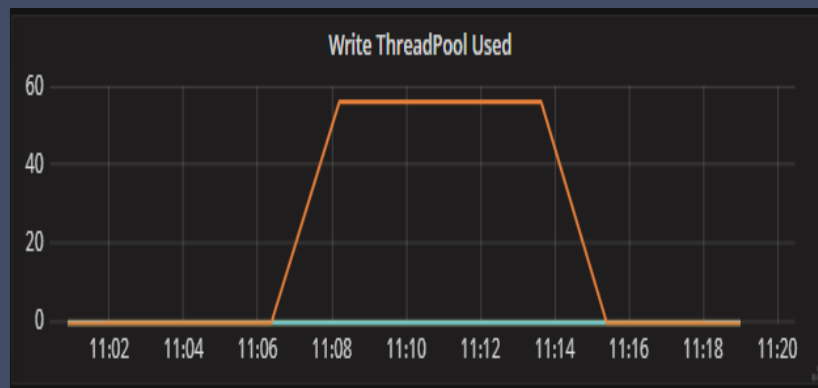
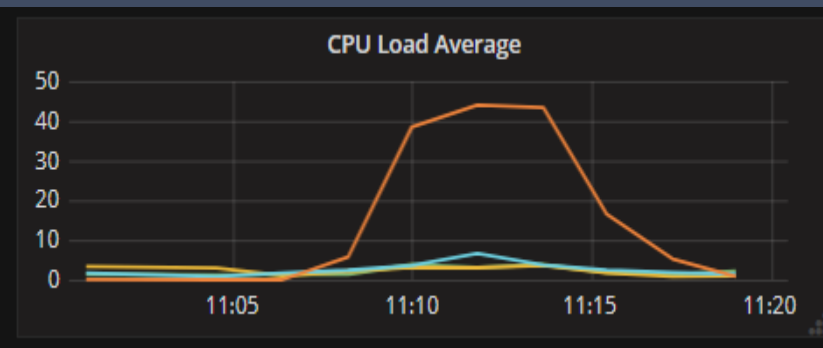
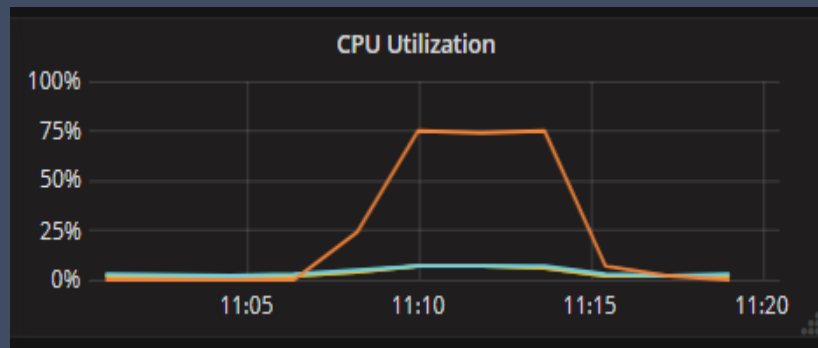


# 写入优化

## 5. 索引缓冲区优化

- CPU : 50+
- 内存分配 : 20G
- 节点个数: 1
- 分片数 : 3
- 副本数 : 0
- 并发度 : 340
- bulk-size : 5000
- 索引刷新频率 3s
- 索引缓冲区 20%

Segment count		339	
Min Throughput	index-append	302697	docs/s
Median Throughput	index-append	313416	docs/s
Max Throughput	index-append	384097	docs/s
50th percentile latency	index-append	1349.77	ms
90th percentile latency	index-append	1768.4	ms
99th percentile latency	index-append	2209.59	ms
99.9th percentile latency	index-append	3521.53	ms
99.99th percentile latency	index-append	4357.7	ms
100th percentile latency	index-append	4572.14	ms





# 写入优化

## 5. 索引缓冲区优化

测试验证，适当的提升Index Buffer可以降低磁盘刷写的频率以及使用更小的代价在内存中进行合并从而提升吞吐量

注意：

- 索引缓冲区内存使用过大会导致大量GC
- 极端情况会导致OOM



02

查询优化

# 查询优化

## 合适的主分片数

### □ 主分片

- 主分片数 = 节点数
- 主分片数 = magic number (e.g. 5 or 10)
- 主分片数 = 数据总大小 / ( 30G ~ 50G )

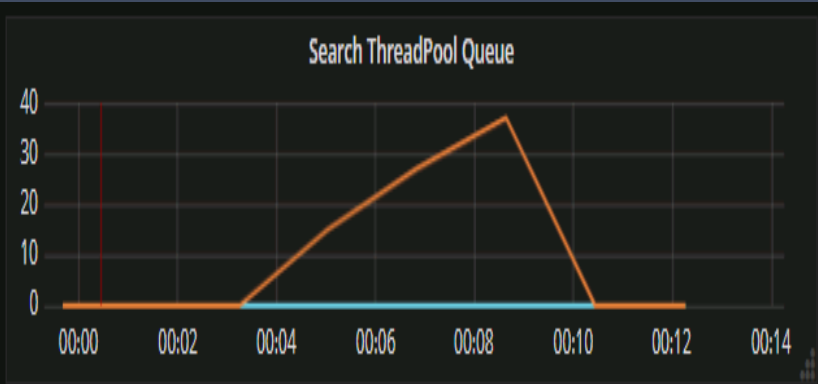
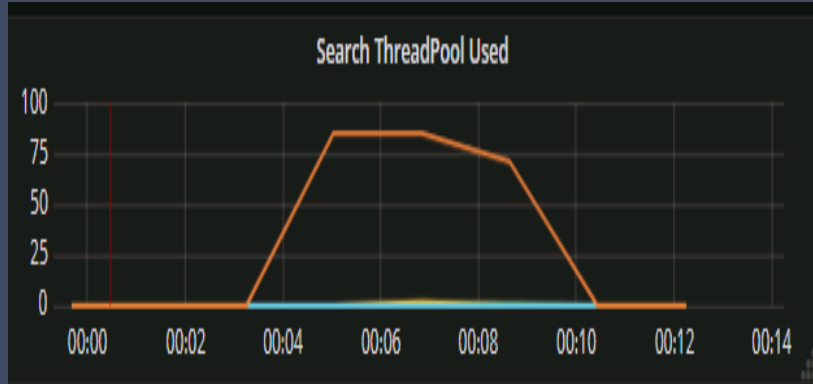
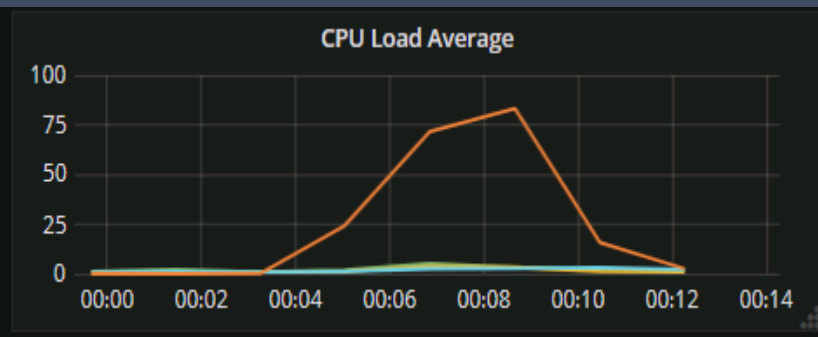
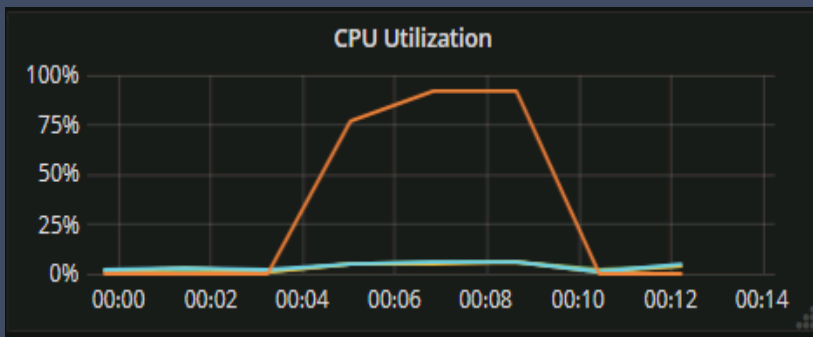


# 查询优化

## 1. 分片数优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 1
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ 数据大小 : 300G
- ❑ TPS : 5000

Segment count		141	
Min Throughput	random_term_time_range_query	1662.12	ops/s
Median Throughput	random_term_time_range_query	3447.8	ops/s
Max Throughput	random_term_time_range_query	3474.56	ops/s
50th percentile latency	random_term_time_range_query	38522.1	ms
90th percentile latency	random_term_time_range_query	67875	ms
99th percentile latency	random_term_time_range_query	74517.5	ms
99.9th percentile latency	random_term_time_range_query	75524.1	ms
99.99th percentile latency	random_term_time_range_query	75929.3	ms
100th percentile latency	random_term_time_range_query	76048.4	ms

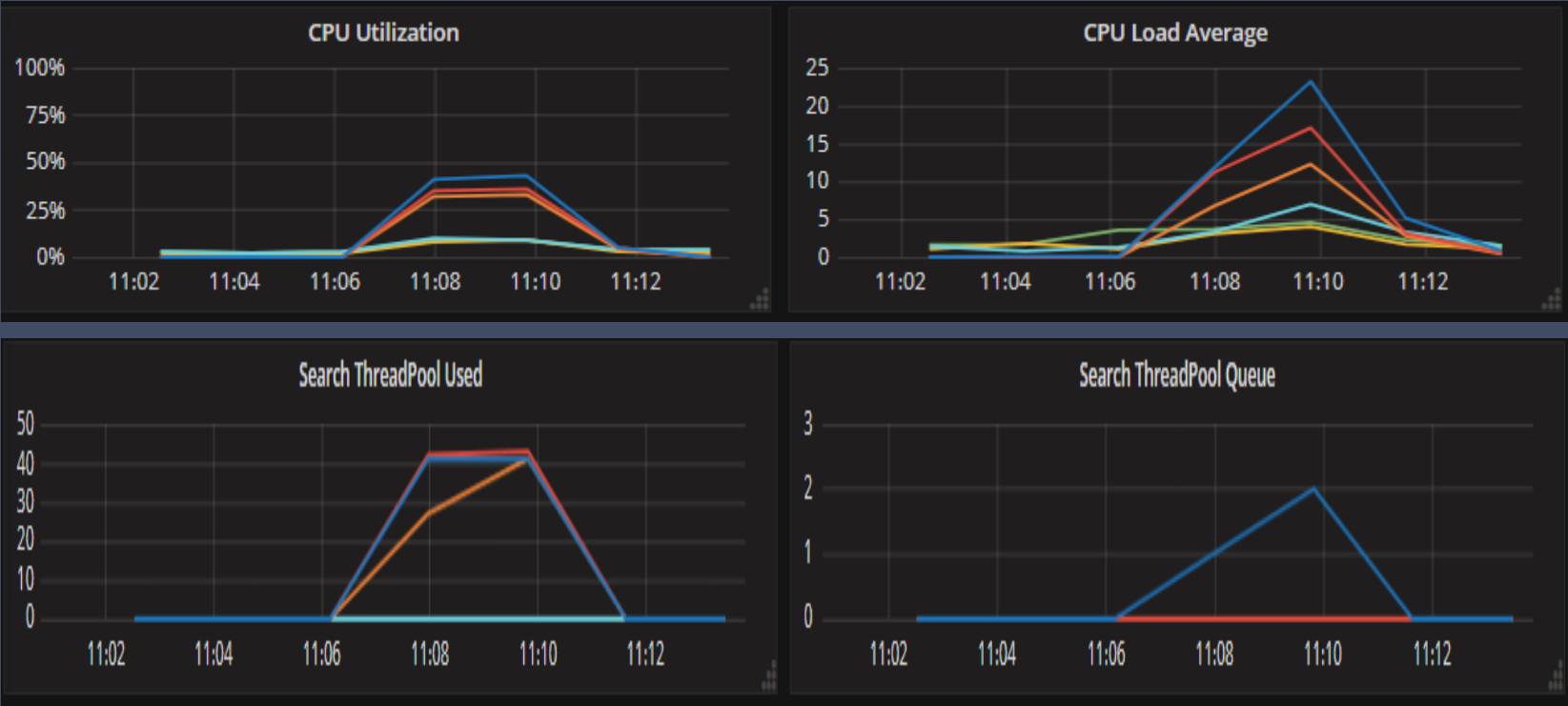


# 查询优化

## 1. 分片数优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 3
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ 数据大小 : 300G
- ❑ TPS : 5000

Segment count		239	
Min Throughput	random_term_time_range_query	4568.16	ops/s
Median Throughput	random_term_time_range_query	4987.81	ops/s
Max Throughput	random_term_time_range_query	4995.05	ops/s
50th percentile latency	random_term_time_range_query	9.79418	ms
90th percentile latency	random_term_time_range_query	309.782	ms
99th percentile latency	random_term_time_range_query	595.719	ms
99.9th percentile latency	random_term_time_range_query	731.309	ms
99.99th percentile latency	random_term_time_range_query	775.959	ms
100th percentile latency	random_term_time_range_query	821.991	ms

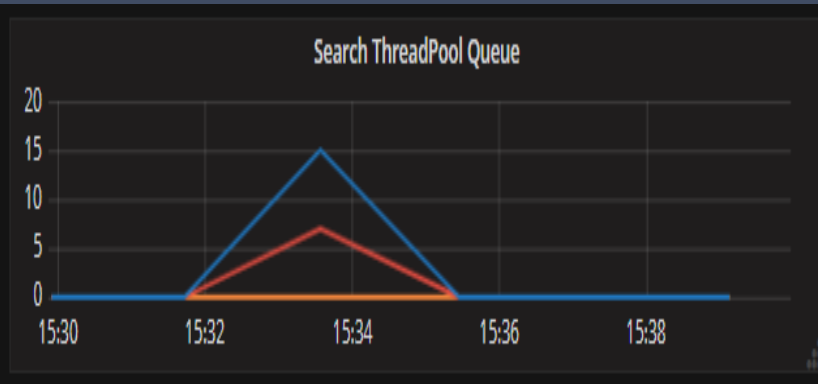
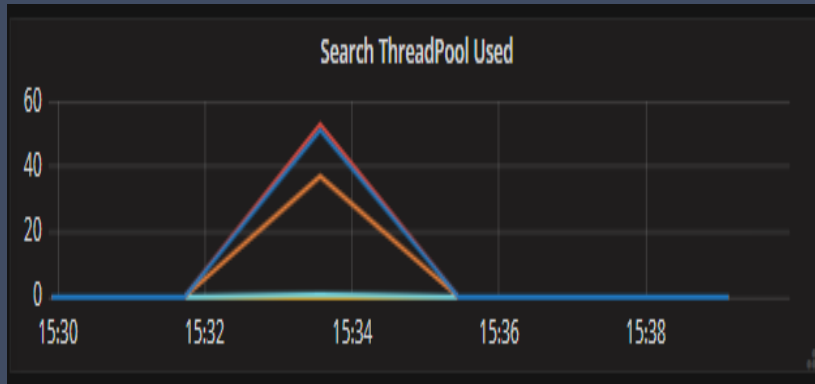
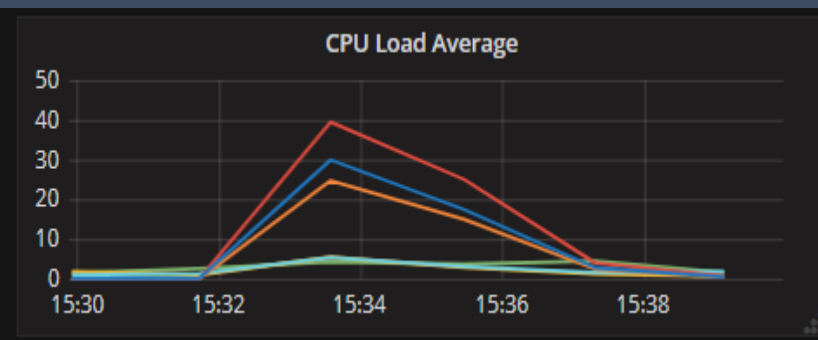
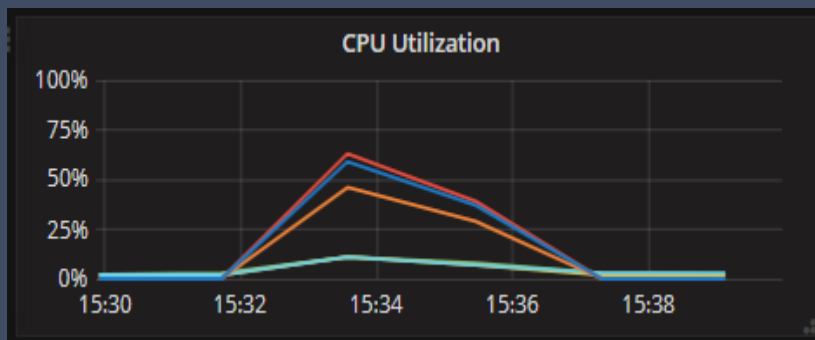


# 查询优化

## 1. 分片数优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 3
- ❑ 分片数 : 6
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ 数据大小 : 300G
- ❑ TPS : 2000

Segment count		217
Min Throughput	random_term_time_range_query	4149.2
Median Throughput	random_term_time_range_query	4973.61
Max Throughput	random_term_time_range_query	4990.9
50th percentile latency	random_term_time_range_query	128.244
90th percentile latency	random_term_time_range_query	438.452
99th percentile latency	random_term_time_range_query	635.002
99.9th percentile latency	random_term_time_range_query	715.671
99.99th percentile latency	random_term_time_range_query	760.229
100th percentile latency	random_term_time_range_query	801.351

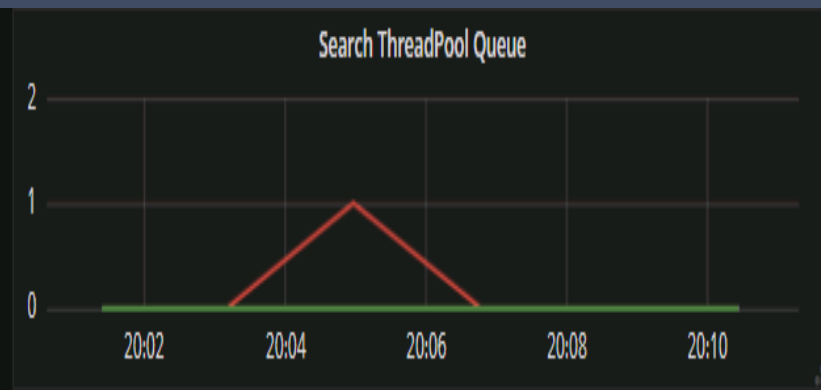
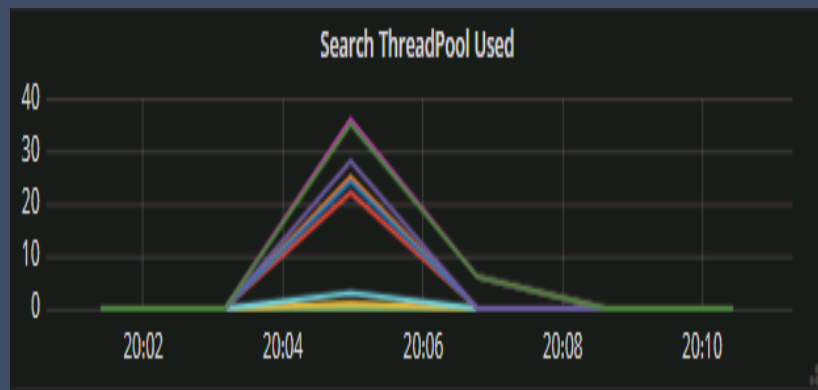
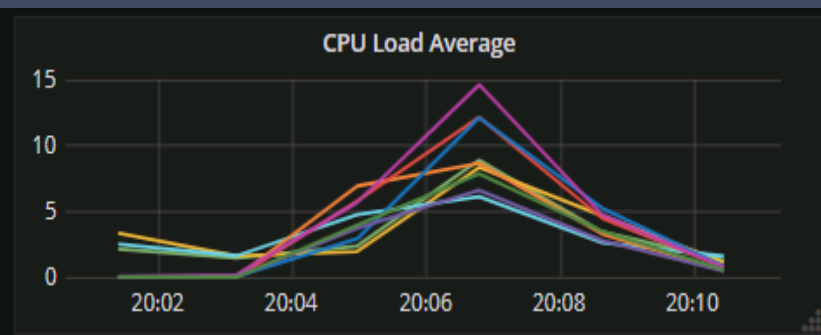
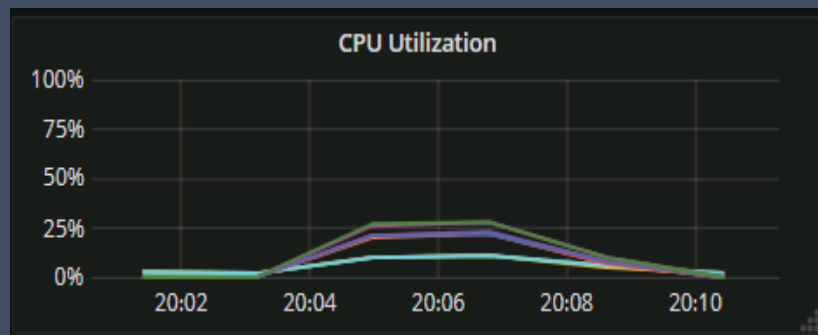


# 查询优化

## 1. 分片数优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 6
- ❑ 分片数 : 6
- ❑ 副本数 : 0
- ❑ 并发度 : 340
- ❑ 数据大小 : 300G
- ❑ TPS : 2000

Segment count		247	
Min Throughput	random_term_time_range_query	3405.6	ops/s
Median Throughput	random_term_time_range_query	4942.3	ops/s
Max Throughput	random_term_time_range_query	4972.26	ops/s
50th percentile latency	random_term_time_range_query	9.56611	ms
90th percentile latency	random_term_time_range_query	259.568	ms
99th percentile latency	random_term_time_range_query	418.099	ms
99.9th percentile latency	random_term_time_range_query	489.903	ms
99.99th percentile latency	random_term_time_range_query	521.388	ms
100th percentile latency	random_term_time_range_query	537.69	ms



# 查询优化

## 1. 分片数优化



## 吞吐量

TPS e.g. 10W / s



## 延迟

Latency e.g. 99% 50ms



## 扩展性

Saturability 60%



# 查询优化

## 2. 查询缓存优化

### □ 无缓存

- 范围 + 结构化限制条件

```
{
  "query": {
    "bool": {
      "must": [
        {
          "range": {
            "time": {
              "lte": 1512355285000,
              "gt": 1512355284000
            }
          }
        },
        {
          "term": {
            "key": "1506053021606006"
          }
        }
      ]
    }
  }
}
```

# 查询优化

## 2. 查询缓存优化

### □ 有缓存

- 范围 + 结构化限制条件

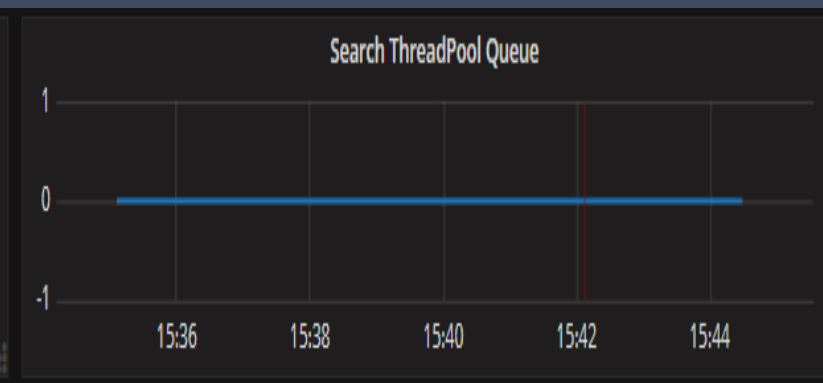
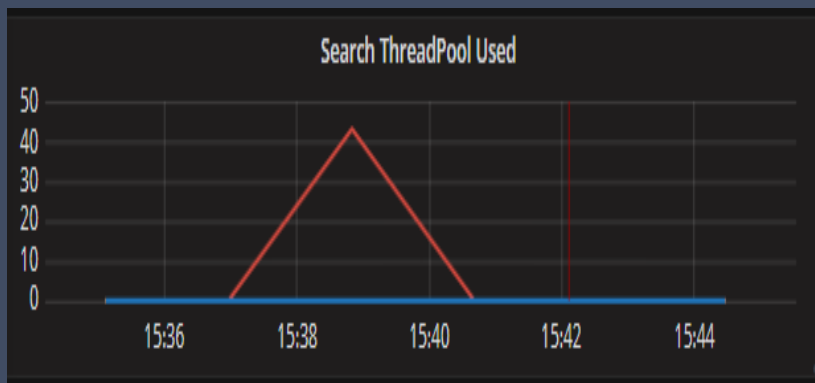
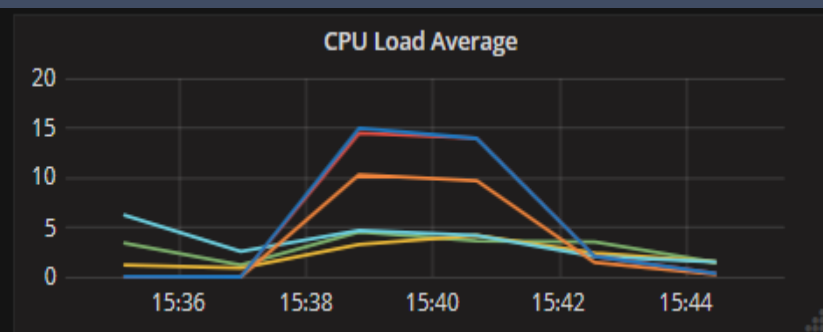
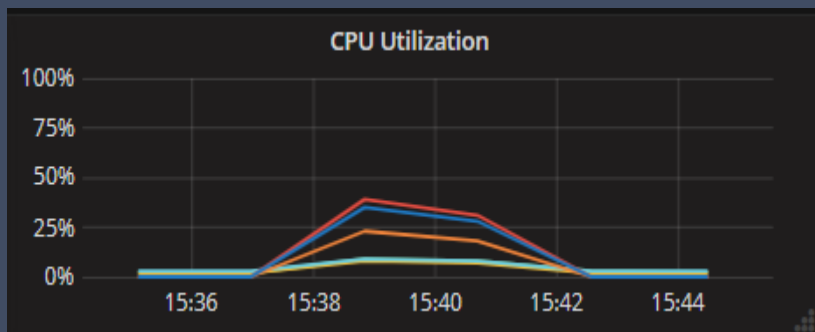
```
{
  "query": {
    "bool": {
      "filter": {
        "range": {
          "time": {
            "lte": 1512355285000,
            "gt": 1512355284000
          }
        }
      },
      "must": [
        {
          "term": {
            "key": "1506053021606006"
          }
        }
      ]
    }
  }
}
```

# 查询优化

## 2. 查询缓存优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 3
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 400
- ❑ 数据大小 : 300G
- ❑ TPS : 5000

Segment count		154	
Min Throughput	random_term_time_range_cache_query	2355.46	ops/s
Median Throughput	random_term_time_range_cache_query	4924.81	ops/s
Max Throughput	random_term_time_range_cache_query	4958.18	ops/s
50th percentile latency	random_term_time_range_cache_query	10.0121	ms
90th percentile latency	random_term_time_range_cache_query	163.977	ms
99th percentile latency	random_term_time_range_cache_query	1027.45	ms
99.9th percentile latency	random_term_time_range_cache_query	1541.91	ms
99.99th percentile latency	random_term_time_range_cache_query	1629.87	ms
100th percentile latency	random_term_time_range_cache_query	1688.32	ms

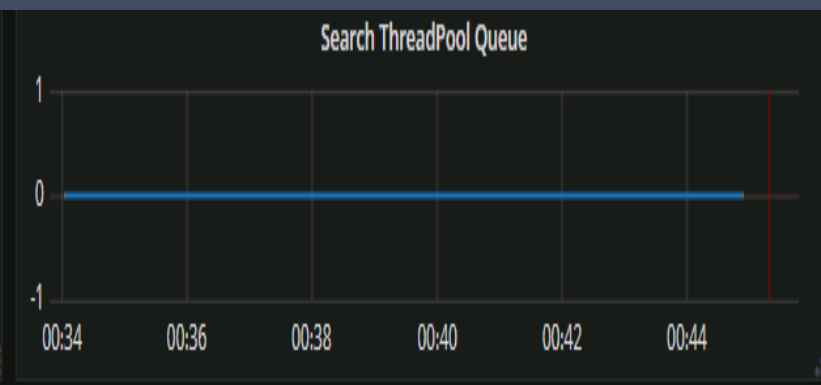
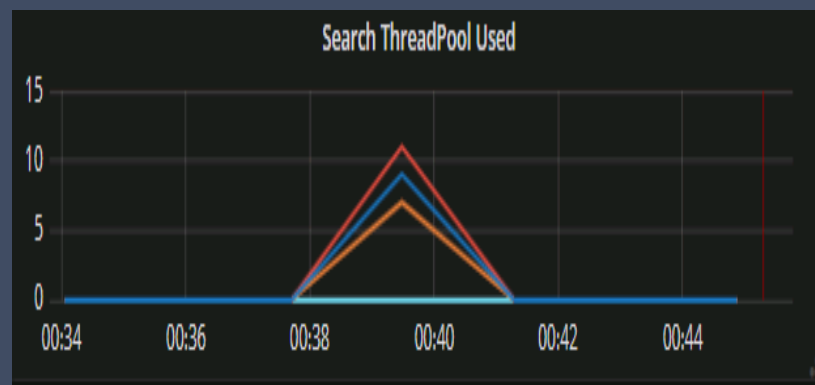
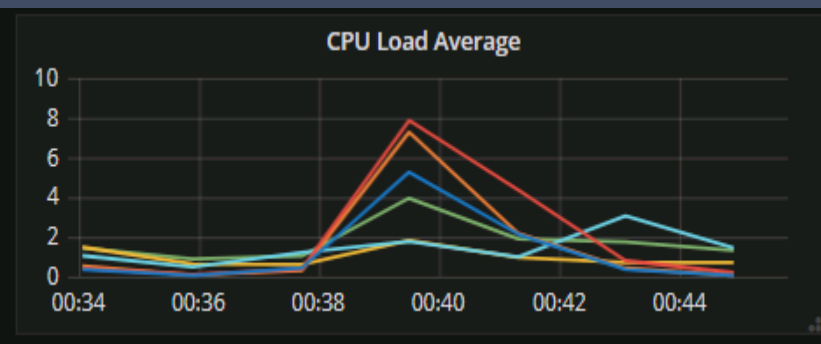
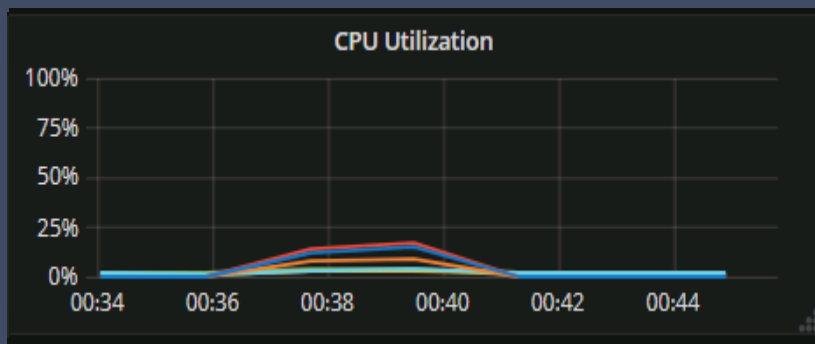


# 查询优化

## 2. 查询缓存优化

- ❑ 集群版本: 6.3.2
- ❑ CPU : 50+
- ❑ 内存分配 : 20G
- ❑ 节点个数: 3
- ❑ 分片数 : 3
- ❑ 副本数 : 0
- ❑ 并发度 : 400
- ❑ 数据大小 : 300G
- ❑ TPS : 5000

Segment count		239	
Min Throughput	random_term_time_range_query	3730.94	ops/s
Median Throughput	random_term_time_range_query	4935.03	ops/s
Max Throughput	random_term_time_range_query	4959.6	ops/s
50th percentile latency	random_term_time_range_query	9.08754	ms
90th percentile latency	random_term_time_range_query	301.419	ms
99th percentile latency	random_term_time_range_query	620.705	ms
99.9th percentile latency	random_term_time_range_query	848.906	ms
99.99th percentile latency	random_term_time_range_query	906.884	ms
100th percentile latency	random_term_time_range_query	929.795	ms



# 查询优化

## 3. 单索引与多索引优化



## 时间换空间

构建一个大而全的索引，通过查询获取结果

会员用户浏览轨迹

全国

# 查询优化

## 3. 单索引与多索引优化



## 空间换时间

多索引、预查询

会员用户浏览轨迹

上海

江苏

浙江

全国

# 查询优化

## 3. 单索引与多索引优化

### 单索引

```
a_es_nginx_access_log
{
  "query": {
    "bool": {
      "must": [
        {
          "range": {
            "st": {
              "gt": "2018-12-03 00:00:00",
              "lte": "2018-12-03 23:59:59"
            }
          }
        },
        {
          "term": {
            "indices_uri": "/test/index_2018-12-03"
          }
        }
      ]
    }
  }
}
```

### 多索引

```
es_nginx_access_log_2018-11-29
{
  "query": {
    "bool": {
      "must": [
        {
          "term": {
            "indices_uri": "/test/index_2018-12-03"
          }
        }
      ]
    }
  }
}
```



# 查询优化

## 3. 单索引与多索引优化

### 单索引

> [s98l0YHATOGFGoda8saWCA][7]				1.671s
Type		Self Time	Total Time	% Time
BooleanQuery +st:[1543420800001 TO 1543507199000] +in...		636.0ms	1.7s	100.00%
IndexOrDocValuesQuery st:[1543420800001 TO 1543507199000]		859.6ms	859.6ms	51.46%
TermQuery indices		175.0ms	175.0ms	10.47%

### 多索引

> [oY5u1tisS4qApecG3DOeKA][4]				378.183ms
Type		Self Time	Total Time	% Time
TermQuery indices_ur		378.2ms	378.2ms	100.00%



03

Q & A



The End

Thank You





专业、垂直、纯粹的 Elastic 开源技术交流社区  
<https://elasticsearch.cn/>