自定义敏感词性能测试报告

# 测试需求

PRD-5 自定义敏感词精确匹配功能接口

# 测试环境

**硬件环境：**

CPU：32核 Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz

内存：64G

硬盘：480G\*2 SSD INTEL SSDSC2KG480G7 + 3\*4THDD

**软件环境：**

Kubernetes

Helm

Pandora镜像

测试环境单节点容器化环境

**容器配置：**

|  |  |  |  |
| --- | --- | --- | --- |
| **副本数** | **Cpu限制（核数）** | **内存制限（MB）** | **Tornado服务参数** |
| **1** | **[0.5-1]** | **[500-20000]** | **max\_buffer\_size=500MB**  **max\_body\_sizy=500MB** |

**网络环境：**

本地环回网络

# 测试场景

1. Content长度：10k，并发数：100
2. Content长度：10k，并发数：200
3. Content长度：10k，并发数：400
4. Content长度：10k，并发数：800
5. Content长度：10k，并发数：1600
6. Content长度：20k，并发数：100
7. Content长度：20k，并发数：200
8. Content长度：20k，并发数：400
9. Content长度：20k，并发数：800
10. Content长度：20k，并发数：1600
11. Content长度：40k，并发数：100
12. Content长度：40k，并发数：200
13. Content长度：40k，并发数：400
14. Content长度：40k，并发数：800
15. Content长度：40k，并发数：1600
16. Content长度：80k，并发数：100
17. Content长度：80k，并发数：200
18. Content长度：80k，并发数：400
19. Content长度：80k，并发数：800
20. Content长度：80k，并发数：1600
21. Content长度：160k，并发数：100
22. Content长度：160k，并发数：200
23. Content长度：160k，并发数：400
24. Content长度：160k，并发数：800
25. Content长度：160k，并发数：1600
26. Content长度：320k，并发数：100
27. Content长度：320k，并发数：200
28. Content长度：320k，并发数：400
29. Content长度：320k，并发数：800
30. Content长度：320k，并发数：1600
31. Content长度：640k，并发数：50
32. Content长度：640k，并发数：100
33. Content长度：640k，并发数：200
34. Content长度：640k，并发数：400
35. Content长度：640k，并发数：800
36. Content长度：1M，并发数：50
37. Content长度：1M，并发数：100
38. Content长度：1M，并发数：200
39. Content长度：1M，并发数：400
40. Content长度：1M，并发数：800
41. Content长度：2M，并发数：25
42. Content长度：2M，并发数：50
43. Content长度：2M，并发数：100
44. Content长度：2M，并发数：200
45. Content长度：2M，并发数：400
46. Content长度：4M，并发数：12
47. Content长度：4M，并发数：25
48. Content长度：4M，并发数：50
49. Content长度：4M，并发数：100
50. Content长度：4M，并发数：200
51. Content长度：8M，并发数：12
52. Content长度：8M，并发数：25
53. Content长度：8M，并发数：50
54. Content长度：8M，并发数：100
55. Content长度：8M，并发数：200
56. Content长度：16M，并发数：5
57. Content长度：16M，并发数：10
58. Content长度：16M，并发数：20
59. Content长度：16M，并发数：40
60. Content长度：16M，并发数：80
61. Content长度：32M，并发数：1
62. Content长度：32M，并发数：2
63. Content长度：32M，并发数：4
64. Content长度：32M，并发数：8
65. Content长度：32M，并发数：16
66. Content长度：64M，并发数：1
67. Content长度：128M，并发数：1

# 测试结果









测试结果分析：

1. 服务的BPS与文本大小无关，约为1.3M（字/s）
2. 测试场景覆盖了从10K小文本到128M大文本递增大小数据的文本；
3. 每一个测试文本大小也涵盖了不同的并发请求数量；
4. 所有的测试场景的错误率都为0；
5. 随着文本大小的递增，内存使用是逐渐增多的；
6. 随着文本大小的递增，TPS（每秒处理的文件个数）是逐渐减小的；
7. 随着文本大小的递增，平均相应时间逐渐递增；
8. 根据测试结果，得出需要内存上限经验公式：

M=L\*100+C\*L\*6+1G

M:需要内存使用大小；

C：并发数

L:文本的长度；

L\*100：处理单个文本的过程中所需空间，100是经验系数。

C\*L\*6：后台需要的最大缓存来存储请求数据。

1G：pandora无负载运行时所需的最大内存。

1. 最大响应时间理论公式：

W = C\*L/V

W:最大相应时间

C:并发数量

L:文本字数

V:每秒处理文本的字数：约未1.3M（字/s）