**Filebeat性能测试文档**

V1.0.0

王智强 @ Enlink 数据中心

**目录**

[1. 概述 3](#_Toc27551)

[1.1. 说明 3](#_Toc3118)

[1.2. 测试环境 3](#_Toc31657)

[1.3. 测试方法 3](#_Toc10929)

[1.4. 测试数据 3](#_Toc4700)

[2. 配置文件参数测试 4](#_Toc30082)

[2.1. compression\_level 4](#_Toc21582)

[2.2. worker 4](#_Toc23328)

[2.3. bulk\_max\_size 5](#_Toc26939)

[2.4. 最优配置 5](#_Toc29354)

[3. 连接Elasticsearch节点数量测试 6](#_Toc10444)

[3.1. 节点测试 6](#_Toc1539)

[4. 硬件配置对Filebeat性能的影响 7](#_Toc19129)

[4.1. 测试环境 7](#_Toc5627)

[4.2. 测试方法 7](#_Toc7348)

[4.3. 内存大小对Filebeat的影响 7](#_Toc28922)

[4.4. CPU性能对Filebeat的影响 7](#_Toc5382)

[5. Elasticsearch对Filebeat的影响 8](#_Toc30650)

[5.1. 硬件配置 8](#_Toc3783)

[5.2. 节点数据转移 8](#_Toc18704)

[6. Filebeat稳定性 9](#_Toc9492)

[6.1. Filebeat长期运行稳定性测试 9](#_Toc5343)

[6.2. 容灾能力 9](#_Toc3246)

# 概述

## 说明

本文档描述Filbeat在不同配置下向Elasticsearch索引推送数据的速度，硬件配置和Elasticsearch对Filebeat索引性能的影响以及Filebeat的稳定性和容灾能力。

## 测试环境

本文中Filebeat的测试环境为VMware创建的虚拟服务器，虚拟服务器的操作系统为：CentOS Linux release 7.4.1708 (Core)。硬件配置信息为：CPU(英特尔 Xeon(至强) E5645 @ 2.40GHz 六核 (X2) )，内存（32G），硬盘（1.5T），带宽1000Mbps（多台服务器公用）。

## 测试方法

本文档关于性能测试采用的方法为单一变量法，及保持其他参数不变的情况下对某一参数进行改动分析其性能变化。

## 测试数据

本文中Filebeat扫描的数据源为res日志，res日志为CASBS后台EnWAS系统用户访问资源的操作记录，具体实例如表1-1所示：

表 1-1 res具体日志实例

|  |
| --- |
| 2018-12-04 13:40:49|WARNING|wangzq|zaf\_group|第三方用户|0254799d489846bc9a2674d8b8e03fac|192.168.32.181|www.baidu.com|1145|389|1534|DELETE|HTTP/1.1|www.baidu.com/images/index\_sy.js|404|128|text/xml|Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/67.0.3396.99 Safari/537.36|https://192.168.32.181:455/client/| |

res日志数据由Filbeat扫入Elasticsearch集群索引的每条数据格式实例为下图1-1：



图 1-1 res日志数据格式实例

# 配置文件参数测试

本章节描述Filebeat配置文件filebeat.yml中配置参数worker，compression\_level，

bulk\_max\_size，对Filebeat向Elasticsearch推送数据性能的影响。作为对比，首先测试worker,compression\_level,bulk\_max\_size为默认值1，0，50时Filebeat的推送能力如表2-1所示：

表 2-1 Filebeat性能测试数据

|  |  |  |  |
| --- | --- | --- | --- |
| **compression\_level** | **worker** | **bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| 0 | 1 | 50 | 1000 |

## compression\_level

compression\_level即Filebeat向Elasticsearch发送数据压缩级别，取值范围为0~9默认值为0，增加压缩级别将减少网络使用，但会增加CPU使用率。Filebeat性能随compression\_level配值的变化如表2-2所示：

表 2-2 Filebeat性能测试数据

|  |  |  |  |
| --- | --- | --- | --- |
| **compression\_level** | **worker** | **bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| 0 | 1 | 50 | 1000 |
| 3 | 1 | 50 | 920 |
| 5 | 1 | 50 | 880 |
| 7 | 1 | 50 | 920 |
| 9 | 1 | 50 | 800 |

由上表所述，compression\_level的改动对Filebeat向Elasticsearch推送数据的性能影响并不大，考虑到compression\_level值增高对cpu的消耗增高则设置compression\_level的值为0。

## worker

worker即当前主机向每个Elasticsearch节点发布事件的线程数量，取值范围不小于1，默认值为1。worker不同值时候Filebeat的推送性能如表2-3所示：

表 2-3 Filebeat性能测试数据

| **compression\_level** | **worker** | **bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| --- | --- | --- | --- |
| 0 | 1 | 50 | 1000 |
| 0 | 2 | 50 | 1200 |
| 0 | 4 | 50 | 1500 |
| 0 | 6 | 50 | 1800 |
| 0 | 8 | 50 | 1900 |
| 0 | 10 | 50 | 2050 |
| 0 | 12 | 50 | 2200 |
| 0 | 14 | 50 | 2300 |
| 0 | 16 | 50 | 1900 |
| 0 | 20 | 50 | 1900 |
| 0 | 30 | 50 | 2000 |
| 0 | 40 | 50 | 2100 |
| 0 | 50 | 50 | 2100 |
| 0 | 100 | 50 | 2100 |
| 0 | 150 | 50 | 2000 |
| 0 | 200 | 50 | 2000 |

综上所述，Filebeat向Elasticsearch推送数据的性能随着worker增加而增加，但是当worker大于14以后则性能变化不大反而有下降趋势，为避免不必要的消耗则设置worker值为14。

## bulk\_max\_size

bulk\_max\_size即Elasticsearch bulk api批次处理数据的数量，默认值为50，worker不同值时候Filebeat的推送性能如表2-4所示：

表 2-4 Filebeat性能测试数据

| **compression\_level** | **worker** | **bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| --- | --- | --- | --- |
| 0 | 1 | 25 | 850 |
| 0 | 1 | 50 | 1000 |
| 0 | 1 | 100 | 1250 |
| 0 | 1 | 200 | 1400 |
| 0 | 1 | 400 | 1550 |
| 0 | 1 | 500 | 1650 |
| 0 | 1 | 600 | 1680 |
| 0 | 1 | 800 | 1680 |
| 0 | 1 | 1500 | 1700 |
| 0 | 1 | 2000 | 1700 |
| 0 | 1 | 5000 | 1700 |

综上所述，bulk\_max\_size值小于500时候，Filebeat推送数据的性能跟bulk\_max\_size值有明显的正向关系，值大于500时候性能趋于稳定跟bulk\_max\_size值无明显关系。

## 最优配置

根据上述三小节测试结果表得知，Filebeat推送数据的性能随compression\_level值增大而衰减，随worker以及bulk\_max\_size值的增大先是增加然后趋于稳定，则可以先确定compression\_level的值为0后再根据之前所得结果测试Filebeat配置文件的最优配置，具体结果如表2-5所示：

表 2-5 Filebeat性能测试数据

| **compression\_level** | **worker** | **bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| --- | --- | --- | --- |
| 0 | 14 | 500 | 4000 |
| 0 | 14 | 600 | 4300 |
| 0 | 14 | 700 | 4000 |
| 0 | 14 | 1000 | 3800 |
| 0 | 14 | 3000 | 2800 |
| 0 | 10 | 600 | 3800 |
| 0 | 14 | 600 | 4300 |
| 0 | 20 | 600 | 4000 |
| 0 | 30 | 600 | 4000 |
| 0 | 40 | 600 | 4000 |

通过以上测试得知，Filebeat配置文件filebeat.yml中compression\_level，worker，bulk\_max\_size，分为配置为0，14，600时，Filebeat向Elasticsearch推送数据的性能最优为4300条/s。

# 连接Elasticsearch节点数量测试

## 节点测试

通过改变Filebeat连接Elasticsearch集群服务器的节点数量来观察Filebeat的变化，测试结果如表3-1所示：

|  |  |  |
| --- | --- | --- |
| **节点数量** | **compression\_level，worker，bulk\_max\_size** | **Filebeat最大推送(条/s)** |
| 1 | 0，14，600 | 4200 |
| 2 | 0，14，600 | 4200 |
| 3 | 0，14，600 | 4100 |
| 4 | 0，14，600 | 4200 |

综上所述，连接节点的改变对Filebeat性能的并无影响。

# 硬件配置对Filebeat性能的影响

## 测试环境

考虑到硬件配置可能也会对Filebeat的性能产生影响以及硬件资源有限，所以在主机搭建虚拟服务器通过改变CPU核数以及内存大小来测试Filebeat的性能，虚拟服务器的其他配置如下：CPU：Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz（可变），硬盘：40G，内存：可变（0.5G~3G）,带宽：100MPbs,操作系统：CentOS Linux release 7.5.1804 (Core) 。Filebeat配置文件filebeat.yml的主要参数配置为：compression\_level: 0，worker: 14， bulk\_max\_size: 512。

## 测试方法

本文档关于性能测试采用的方法为单一变量法，及保持其他参数不变的情况下对某一参数进行改动分析其性能变化。

## 内存大小对Filebeat的影响

其他配置不变，通过改变虚拟机的内存大小来观察Filebeat向Elasticsearch的最大推送能力，具体数据变化如表4-1所示：

表 5-1 Filebeat性能数据

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **内存 (G)** | **CPU核数 (个)** | **最大推送速度（条/s）** | **CPU使用率** | **内存占用（G）** | **磁盘读取 (kb/s)** | **磁盘写入 (kb/s)** |
| 0.5 | 1 | 4000 | 9%~11% | 0.075 | 850~3300 | 10~20 |
| 1 | 1 | 4000 | 10%~12% | 0.071 | 850~3300 | 10~20 |
| 2 | 1 | 4000 | 6%~9% | 0.065 | 850~2200 | 10~20 |
| 3 | 1 | 3800 | 6%~10% | 0.065 | 850~1700 | 10~20 |

由上表所得，虚拟机内存大小的变化对Filebeat的推送性能并未实质性变化。

## CPU性能对Filebeat的影响

因主机CPU型号无法改变，所以只能通过改变虚拟机CPU核数来间接改变虚拟机的CPU性能，具体数据变化如表4-2所示：

表 4-2 Filebeat性能数据

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **内存 (G)** | **CPU核数 (个)** | **最大推送速度（条/s）** | **CPU使用率** | **内存占用（G）** | **磁盘读取 (kb/s)** | **磁盘写入 (kb/s)** |
| 3 | 1 | 3800 | 6%~10% | 0.065 | 850~1800 | 10~20 |
| 3 | 2 | 4200 | 7%~10% | 0.080 | 820~2400 | 15~20 |
| 3 | 4 | 4400 | 6%~9% | 0.075 | 820~2400 | 10~20 |
| 3 | 8 | 4300 | 6%~9% | 0.079 | 820~2400 | 15~20 |

由上表所得，虚拟机CPU核数的提高一定程度上会提高Filebeat的推送性能，但并无实质性的提高。

# Elasticsearch对Filebeat的影响

## 硬件配置

虚拟机硬件的变化虽然对Filebeat有影响但其影响远小于预想，假设Filebeat的性能瓶颈在与其联系最密切的Elasticsearch上，那么Elasticsearch所在虚拟服务器的硬件配置的改变是否对Filebeat推送数据产生影响。配置4台虚拟机（除内存大小CPU核数其它和上述虚拟服务器配置相同），搭建和测试服务器参数配置相同的Elasticsearch集群。不同Elasticsearch集群配置的测试结果如表5-1所示：

表 5-1 集群配置测试结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **CPU （型号）** | **CPU核数 (个)** | **最大推送速度（条/s）** | **CPU使用率** | **内存占用（G）** | **磁盘读取 (kb/s)** | **磁盘写入 (kb/s)** |
| E5645 | 6 | 4000 | 20%~40% | 0.075 | 850~3300 | 10~20 |
| i7-8700 | 6 | 10000 | 70%~80% | 0.077 | 2700~4000 | 150~200 |

注：CPU型号以及核数为不同Elasticsearch集群所在服务器的配置，其它所得数据为Filebeat在测试服务器（192.168.100.49）向不同集群推送数据所得数据。

两种型号得CPU性能对比如图5-1所示：

图 5-1 CPU性能对比图

lALPDgQ9qZADN-wzzQOY_920_51

综上所述，同一服务器上Filebeat向不同配置下的Elasticsearch集群推送数据其性能会有差异，其配置越高Filebeat性能越好。

## 节点数据转移

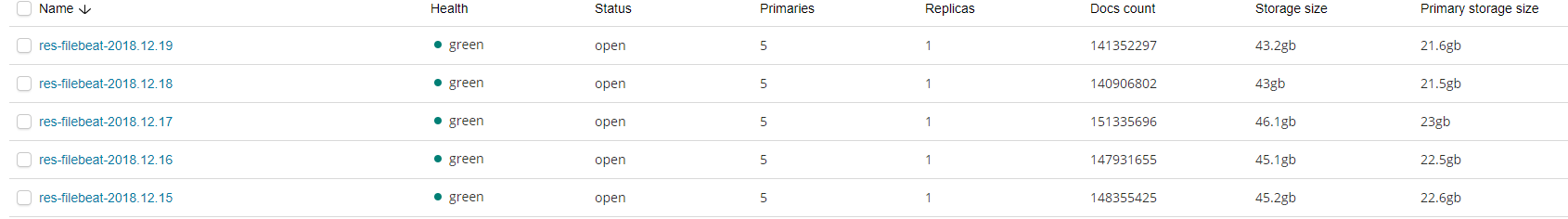
Elasticsearch集群自身会自动调整索引数据在各个节点的存储，当Elasticsearch进行节点间的数据转移时会有大量的磁盘IO操作，这时Filebeat向Elasticsearch集群推送数据的能力会大打折扣（性能下降到之前的二分之一），当Elasticsearch集群索引节点调整完成后，Filebeat则会很快恢复到正常工作的状态。

# Filebeat稳定性

## Filebeat长期运行稳定性测试

通过使用脚本模仿日志产生环境24小时为Filebeat提供伪日志（脚本日志生成速度：1800条/s），下面为Filebeat连续运行5天Elasticsearch集群对应得索引数据：

图 6-1 索引数据



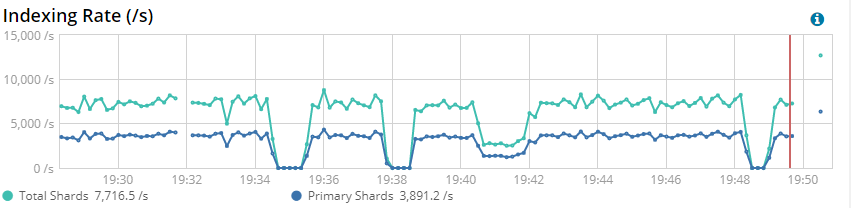
如图6-1所示，Filebeat每天向集群推送得索引数据大致相同且也符合Filebeat数据统计文档对日志数量及所占空间得统计。在对Filebeat长期的使用研究中也并没遇到Filebeat运行不稳定的情况。

## 容灾能力

通过对Filebeat所在虚拟机强制关机来模拟现实情况中服务器突然停电宕机的情况，目的是测试正在运行的Filebeat遇到宕机重新启动后其推送的数据日否丢失或增加。

通过脚本生成一份日志数据，在Filebeat对其扫描向集群推送数据时将虚拟机随机宕机3次，其数据推送速度如图6-2所示：

图 6-2 Filebeat数据推送



日志源文件日志条数为：7166056条，生成索引最终数据条数为：7172200，索引数据多出了6144条。为确认多出的数据是否因多次宕机造成，生成类似大小日志文件让Filebeat正常工作，最终索引所含数据条数跟日志文件条数相同，由此可判断索引多出的数据是多次宕机造成。

在多次宕机模拟测试后，均出现因Filebeat所在服务器宕机造成索引数据增多的现象，且数据增多的大小跟Filebeat扫描速度跟宕机次数有关，最后得出服务器突然宕机会使增在工作的Filebeat重复推送数据。对此问题的解决方法还在研究中。