芒果TV ELK日志系统实践

关于我

刘波涛

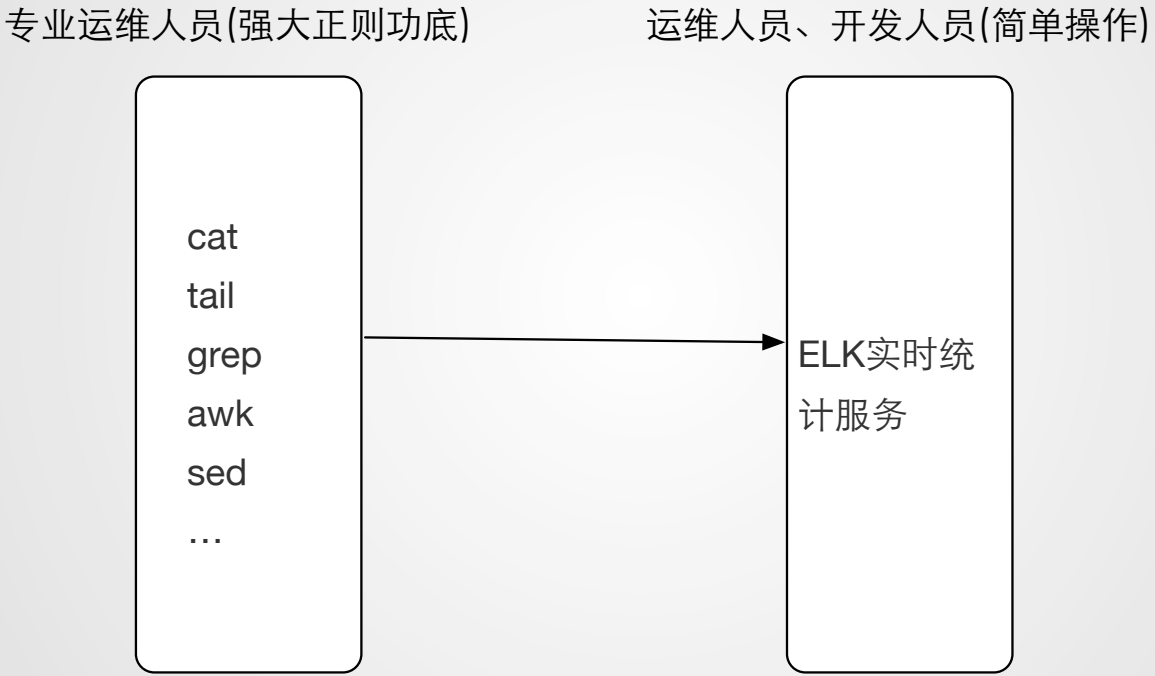
芒果TV研发工程师

日志文件重要性

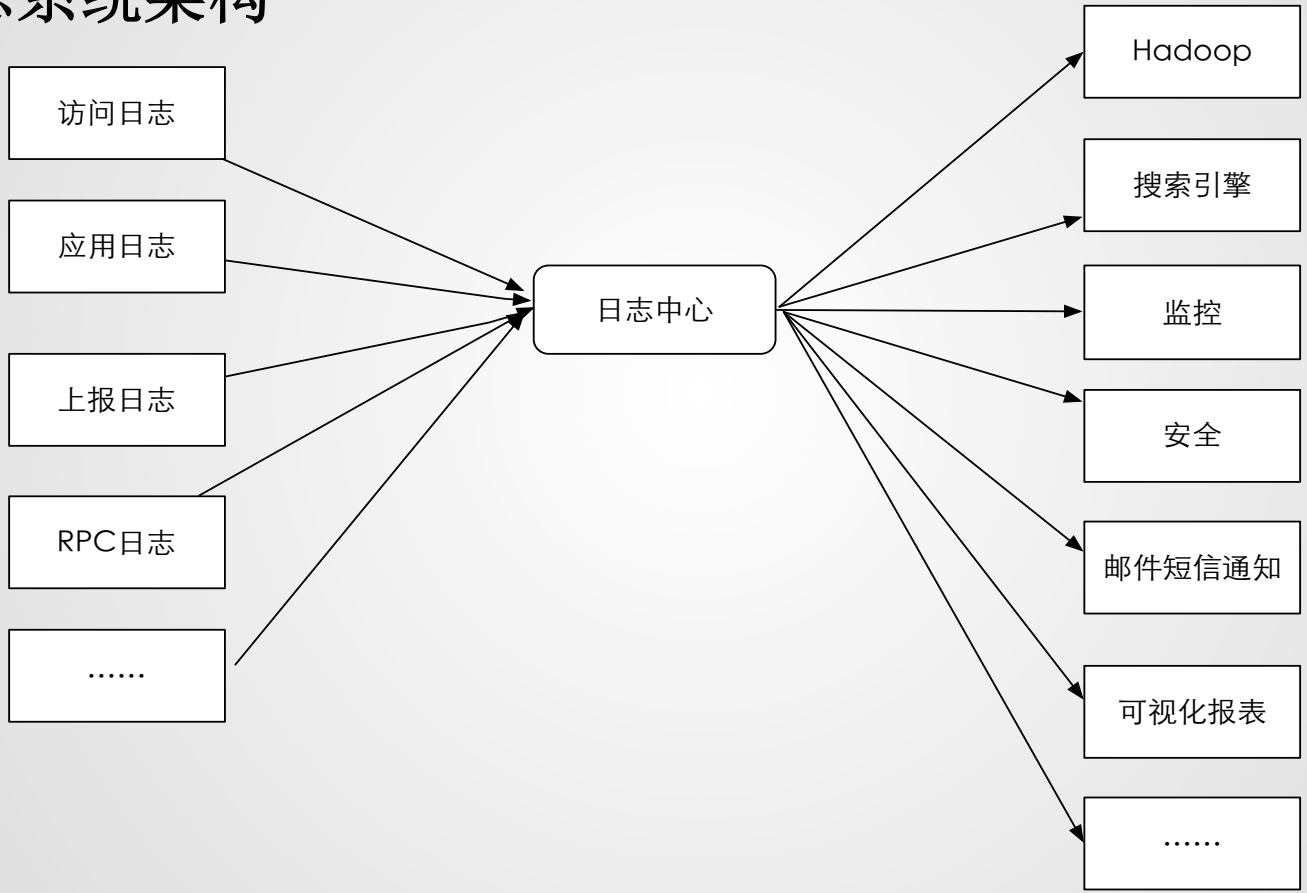
Linux哲学：万物皆文件

运维哲学：日志管理是保障高质量服务的基础

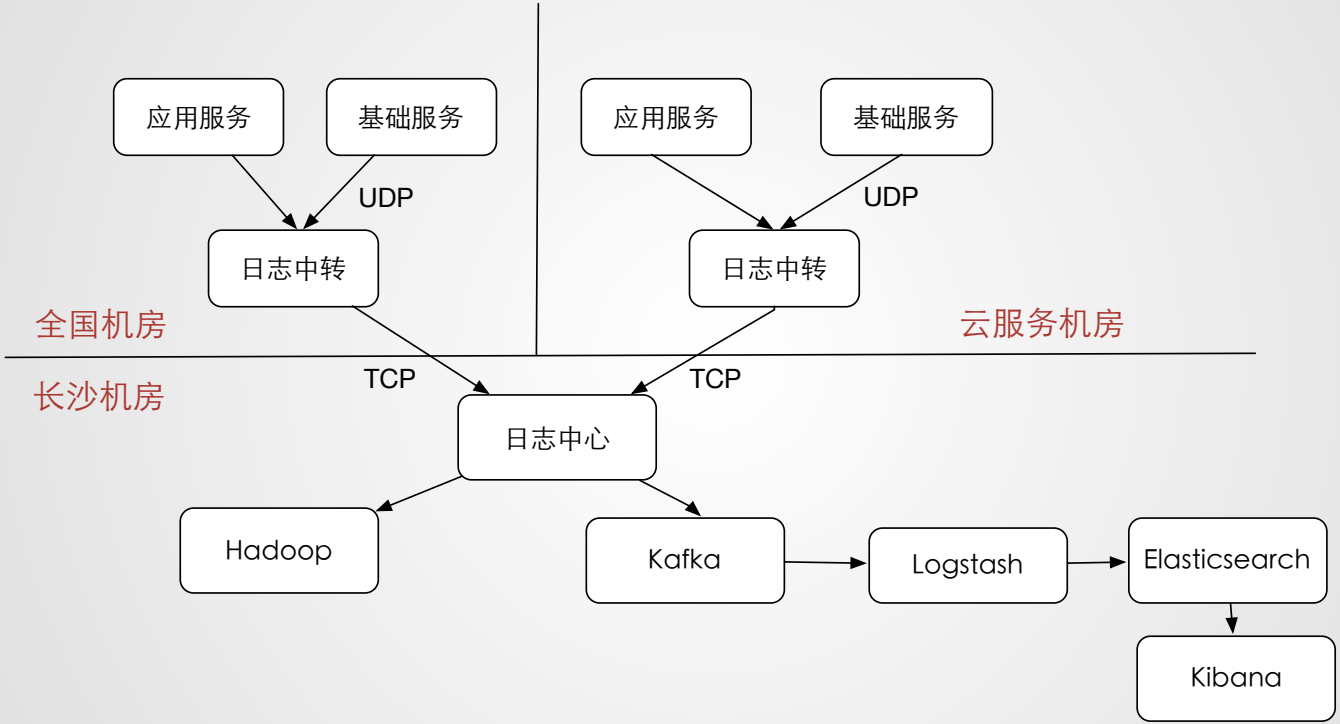
日志架构演变


专业运维人员(强大正则功底)
运维人员、开发人员(简单操作)
cat
tail
grep
awk
sed
...
ELK实时统计服务

日志系统架构



ELK系统架构



Rsyslog

Nginx： 通过syslog模块转发

应用程序: 通过调用syslog函数写入

规范统一格式

v8版本以上(原生支持Kafka)

```
module(load="omkafka")
local5.*        action(type="omkafka" topic="mpp-elk" broker="1[redacted]
[redacted]
[redacted]
[redacted]" partitions.number="14
" confParam=["queue.buffering.max.messages=2000000"] queue.size="360000000" queue.ma
xdiskspace="99G" queue.highwatermark="216000000" queue.discardmark="288000000" queue
.type="LinkedList" queue.dequeuebatchsize="4096" queue.timeoutenqueue="0" queue.maxf
ilesize="4G" queue.saveonshutdown="on" queue.workerThreads="14" template="access")
```

Rsyslog-Avoid-Block

关闭HUPIsRestart配置选项(低版本)

监控rsyslog服务,一旦crashes能够马上重启

传输方式由TCP改为UDP(恶性循环)

Kafka

Kafka vs Redis

强大消息堆积能力

日志领域高度成熟

支持Hadoop数据并行加载

高性能(顺序写单机写入TPS约在百万秒/s)

Kafka-Options

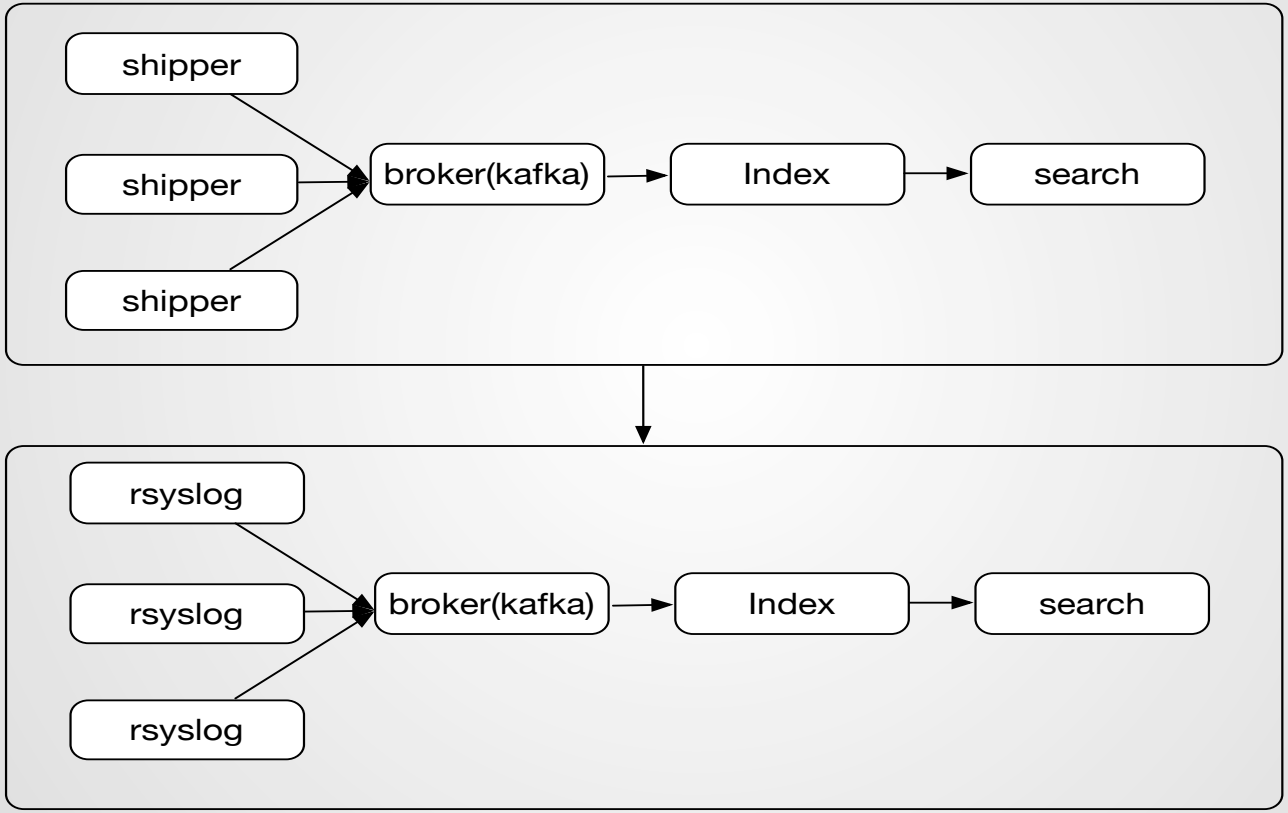
Partition Number(数量必须大于消费者数量)

Broker Number(配置和内核数相同)

num.network.threads

num.io.threads

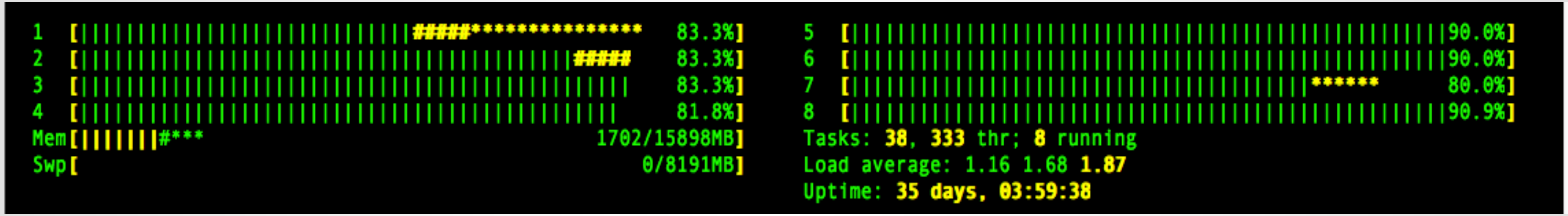
Logstash



Logstash

配置

性能问题,大量消耗CPU和内存



容易僵死

自定义JAVA程序替代Logstash(支持kafka,syslog输入,ES输出)

启动多个进程进行消费

Elasticsearch

减少副本数量(副本数为0)

以写为主,读为辅助(随机写 磁盘瓶颈 使用SSD替代传统硬盘)

设置filedldata: format :doc_value 避免Heap crash

增加Index.refresh_interval 时间(默认为一秒),降低压力

合理使用TCP,UDP索引模式(我们使用Http模式)

关闭系统swap

内核配置修改

对数据聚合进行处理string2int

定时删除旧索引(保存2个星期)

服务器参数调整

TCP参数

```
net.ipv4.tcp_fin_timeout = 30
net.ipv4.tcp_keepalive_time = 1200
net.ipv4.tcp_syncookies = 1
net.ipv4.tcp_tw_reuse = 1
net.ipv4.tcp_tw_recycle = 1
net.ipv4.ip_local_port_range = 1024 65000
net.ipv4.tcp_max_syn_baklog = 8192
net.ipv4.tcp_max_tw_bukets = 5000
```

调整Linux的最大文件数

ulimit -SHn 65535

磁盘

iostat

```
[root@localhost ~]# iostat -x 1 5
Linux 2.6.32-358.el6.x86_64 (localhost.localdomain)     2015年10月15日   _x86_64_        (24 CPU)

avg-cpu:  %user   %nice %system %iowait  %steal   %idle
          9.45    0.00    0.54    0.05    0.00   89.95

Device:         rrqm/s   wrqm/s     r/s     w/s    rsec/s    wsec/s avgrq-sz avgqu-sz   await  svctm  %util
sda               4.90     5.52    1.35    0.77     52.28     50.31    48.29     0.00    2.07   0.78   0.17
sdb               0.02  2224.07   20.03  122.72   1703.81  18730.76   143.15     0.79    5.52   0.16   2.25
```

如果util接近100%则说明产生的I/O请求太多,I/O系统已经满负载
磁盘可能存在瓶颈如果idle小于70%,I/O的压力比较大,说明读取进程中有较多的wait

vmstat

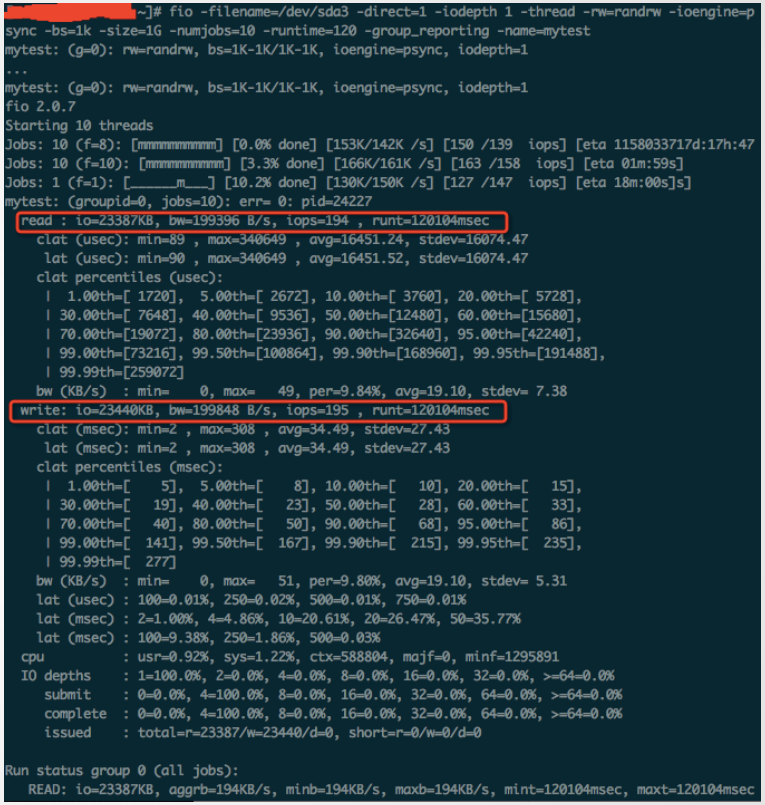
```
[root@localhost ~]# vmstat 2
procs -----------memory---------- ---swap-- -----io---- --system-- -----cpu-----
 r  b   swpd   free   buff  cache   si   so    bi    bo   in   cs us sy id wa st
12  0 4267028 558908 147512 32044600    1    1    37   392    1    1  9  1 90  0  0
 9  0 4267028 531364 147512 32073924    0    0     0     0 6888 5396 13  1 86  0  0
 1  0 4267028 529716 147512 32081780    0    0     0  9634 7645 5906 13  1 86  0  0
 0  0 4267028 511768 147528 32100348    0    0     0  7428 6867 5571 11  1 88  0  0
 5  0 4267028 489936 147528 32120188    0    0     0 11544 8170 5873 13  1 86  0  0
```

vmstat 2 查看b堵塞进程情况

磁盘

检查磁盘性能

fio -filename=/dev/sda3 -direct=1 -iodepth 1 -thread -rw=randrw -ioengine=psync -bs=1k –size=1G -numjobs=10 -runtime=120 -group_reporting -name=mytest



Elasticsearch

Mapping:压缩_source(compress:true)

Mapping:禁用all(include_in_all:false)

自定义分词:尽量不使用标准分词使用ik

ES_HEAP_SIZE: -Xms = -Xmx 不超过内存50%

index.cache.field.type = soft

index.cache.field.max_size:50000

index.cache.field.expire:10m

index.fielddata.cache: soft

Elasticsearch SSD优化参数

磁盘RAID0

mmap索引文件格式 (index.store.type: mmapfs)

indices.store.throttle.type:none

indices.memory.index_buffer_size: 30%

index.merge.scheduler.max_merge_count: 6

index.translog.flush_threshold_size:5gb

index.translog.flush_threshold_ops: 500000

index.gateway.local.sync:30s

index.merge.scheduler.max_thread_count: 3

关闭文件系统ATIME（atime off）

Kibana

K3 VS K4

原生 VS 自定义可视化

监控报警

Nginx 5xx/s

Mysql 慢日志、错误日志

Redis 慢日志、错误日志

程序错误日志

DNS劫持

统计报表

Nginx 响应时间

Nginx 正常响应占比

Nginx QPS统计

CDN视频流加载时间

搜索引擎

生成热门搜索

统计搜索转化率

调整搜索权重

感谢关注和支持

芒果tv