

ZooKeeper管理平台

工具

被[管理员](#)添加，被[章耿](#)最后更新于一月 19, 2015

项目简介

目前ZooKeeper管理平台与SAF管理端部署在一起，具有如下zk管理功能：ZooKeeper存活监控、请求积压状态监控、zookeeper日志删除worker，并提供四字命令执行接口。

ZooKeeper使用现状

ZooKeeper平台做为开源的工业级别的分布式协同器已在公司内部各项目中得到广泛的应用，但是目前在使用中存在如下的问题：

- 1. 各个项目往往自己维护所架设的ZooKeeper集群，而ZooKeeper所消耗的资源往往偏低从而造成服务器资源的浪费；
- 2. ZooKeeper平台缺乏运行监控，如节点数、watch数量、链接数等运行状态值，集群使用者无法得知；
- 3. ZooKeeper平台缺乏统一管理以及权限分配，目前许多项目是有限定数据访问权限，存在安全隐患；

为了解决以上问题，我们希望开发统一的ZooKeeper管理平台以加强对ZooKeeper平台的监控、管理；ZooKeeper管理平台的功能分为监控、管理两大模块。

监控模块

1.提供页面执行四字命令：

输入 ZooKeeper服务ip地址：ip；端口：port；所执行的四字命令；

输出：四字命令的执行结果；

ZooKeeper 四字命令	功能描述
conf	输出相关服务配置的详细信息。
cons	列出所有连接到服务器的客户端的完全的连接 / 会话的详细信息。包括"接受 / 发送"的包数量、会话 id 、操作延迟、最后的操作执行等等信息。
dump	列出未经处理的会话和临时节点。
envi	输出关于服务环境的详细信息（区别于 conf 命令）。
reqs	列出未经处理的请求
ruok	测试服务是否处于正确状态。如果确实如此，那么服务返回"imok "，否则不做任何相应。
stat	输出关于性能和连接的客户端的列表。
wchs	列出服务器 watch 的详细信息。
wchc	通过 session 列出服务器 watch 的详细信息，它的输出是一个与watch 相关的会话的列表。
wchp	通过路径列出服务器 watch 的详细信息。它输出一个与 session相关的路径

2.worker线程 定时对集群列表中的ZooKeeper节点，执行ruok命令，执行的结果显示在页面中，执行频率，以及执行的命令要求做到可配置；

管理模块

1.保存展示集群信息（有几个集群，集群有哪几个节点，ip是多少，服务端口是什么；）

1.树形展示ZooKeeper上的数据（区分临时节点与实际节点，展示节点上数据）；

2.通过页面执行ZooKeeper 命令，实际通过zkclient去执行，常见的命令有：

- 浏览某路径下的所有子节点；

动态

- 2012-08-30 ZooKeeper管理功能开发结束

项目成员

肖飞 [xiaofei@jd.com](#)
章耿 [zhanggeng@jd.com](#)
鲍宁天 [baoningtian@jd.com](#)
张海 [zhanghai@jd.com](#)

常用链接

[接入Zookeeper管理平台](#) [jsf@jd.com]

[测试管理端](#) [test.saf.jd.com]

[线上管理端](#) [saf.jd.com]

最近的更新

[私有云平台 管理员](#)更新 (查看变动) 十月 13, 2014

[电商云方案 管理员](#)更新 (查看变动) 十月 31, 2013

[PAF 管理员](#)更新 (查看变动) 七月 11, 2013

[ELB 管理员](#)更新 (查看变动) 五月 10, 2013

[复制从 私有云平台 管理员](#)更新 (查看变动) 五月 10, 2013

[LOGBack 管理员](#)更新 (查看变动) 五月 06, 2013

[私有云平台 backup_2013-4-10 管理员](#)更新 (查看变动) 四月 16, 2013

[PAF中任务分配后及任务完成后的邮件提醒使用说明 管理员](#)更新 (查看变动) 四月 16, 2013

[PAF FAQ 管理员](#)更新 (查看变动) 四月 16, 2013

[PAF人工活动接口使用说明 管理员](#)更新 (查看变动) 四月 16, 2013

[AD LDAP 管理员](#)更新 (查看变动) 四月 16, 2013

[代码托管 管理员](#)创建 四月 08, 2013

- 在某路径下创建（临时）节点；
- 在某路径下删除节点；（需权限限制）

[JBOX 管理员更新](#) ([查看变动](#))
四月 03, 2013

[MSP 管理员更新](#) ([查看变动](#))
四月 02, 2013

[ZooKeeper管理平台 管理员更新](#)
([查看变动](#)) 四月 02, 2013

ZooKeeper最佳实践

在此纪录线上出现的一些问题以及解决方法：

1.系统服务时间不对，导致 ZooKeeper 集群不可用

发现环境：测试环境集群

现象：zookeeper集群不可用，系统重启和zookeeper重启都不管用。相当于没有集群启动起来。

原因分析：3台集群时间每台都不一样，相差10几分钟。

解决：修改3台集群机器的系统时间一致，重启zookeeper集群。

建议：检查系统的时间是否一致。

2.DiskIO过大或者IO不正常，导致 ZooKeeper 集群不稳定

发现环境：线上集群

现象：ZooKeeper集群突然不稳定，net IO较大，TCP连接数不稳定，使用mntr命令查看ZooKeeper的情况，可以看到排队请求（zk_outstanding_request）有250+

原因分析：ZooKeeper的历史日志较多，可能部分数据放到系统的交换分区里影响到磁盘IO性能。

解决：删除不稳定的机器的历史日志，只保留了当日的日志，ZooKeeper集群的IO瞬间飙升至一个高点，四五分钟后，趋于平稳，使用mntr命令查看ZooKeeper的情况，可以看到已基本无排队请求（zk_outstanding_request）。

PS:后来又删除两个和ZooKeeper无关的大日志文件（70G），可能看到ZooKeeper集群相应会有NetIO以及TCP链接数的变动，也是数分钟后又趋于平稳。

*建议：*通过ZooKeeper四字命令（如cons、mntr、ruok）监控ZooKeeper的当前运行状态；此外ZooKeeper日志配置要谨慎，日志目录与数据目录最好放在不同分区下，日志策略在配置时不必留太多日志，info级别、500M一个文件即可。

3.ZooKeeper客户端导致内存溢出

发现环境：线上集群

现象：客户端使用了netflix开源的curator客户端，在长时间运行、以及网络抖动情况下，很容易曝出connection loss错误并最终内存溢出

原因分析：curator对ZooKeeper的封装太复杂（PathCache），且其session过期处理有问题。

解决：自行基于官方客户端做简单封装，或者使用另一个非常简单的ZooKeeper—client。

建议：谨慎选择ZooKeeper客户端，上线前务必经过严格的稳定性测试；

4.ZooKeeper session过期时间以及watch的使用

发现环境：线上集群

现象：客户端总是抱connection loss错误，在服务端发现客户端重连时setwatch失败

原因分析：客户端 session过期时间设置过短，在网络抖动时很容易导致session过期，此时重连接需要重建watch，watch数量过多，或者watch的路径过长，而ZooKeeper目前版本在处理长数据时有限制会导致失败。

解决：加长ZooKeeper session过期时间（60S或者更长），分析数据结构，减少watch数量。

建议：session过期时间要设长，此外watch要谨慎使用，数量过多会严重影响ZooKeeper性能；

5. 大量的CLOSE_WAIT连接

发现环境:测试环境

现象：长连接数大于5w，实际只有200，大量的CLOSE_WAIT连接

原因分析：很多链接其实已经断开，只是没销毁。

解决：vi /etc/sysctl.conf 加入如下参数

```
net.core.somaxconn = 32768
net.core.wmem_default = 8388608
net.core.rmem_default = 8388608
net.core.rmem_max = 16777216
net.core.wmem_max = 16777216
net.ipv4.tcp_timestamps = 0
net.ipv4.tcp_synack_retries = 1
net.ipv4.tcp_syn_retries = 0
net.ipv4.tcp_tw_recycle = 1
net.ipv4.tcp_tw_reuse = 1
net.ipv4.tcp_mem = 94500000 915000000 927000000
net.ipv4.tcp_max_orphans = 3276800
net.ipv4.ip_local_port_range = 1024 65535
net.ipv4.tcp_fin_timeout = 10
net.ipv4.tcp_keepalive_time = 100
net.ipv4.tcp_syncookies = 1
net.ipv4.tcp_max_syn_backlog = 8192
net.ipv4.tcp_max_tw_buckets = 20000
net.ipv4.tcp_keepalive_probes = 3
net.ipv4.tcp_keepalive_intvl = 15
```

执行/sbin/sysctl -p 生效

重启网络 /etc/rc.d/init.d/network restart

建议重启zookeeper。

无
