

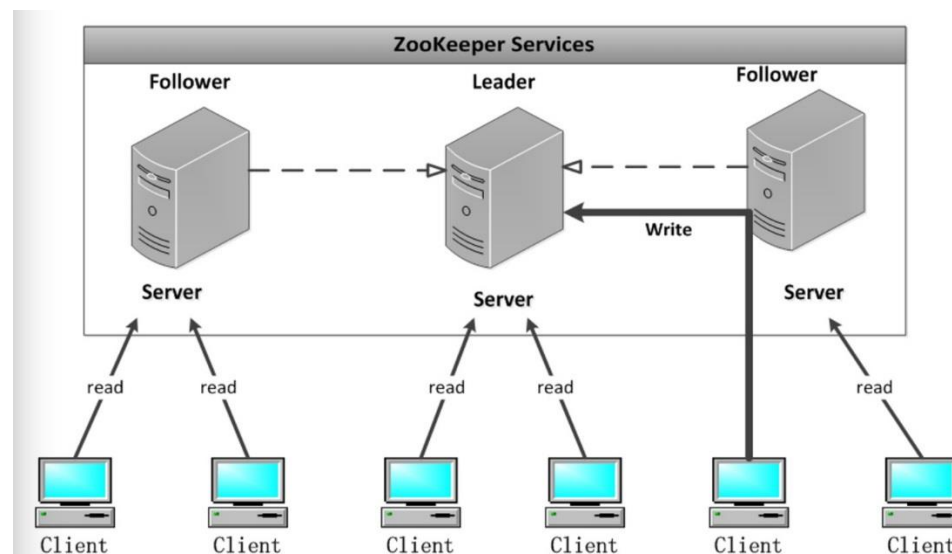
# zookeeper 的介绍以及集群环境搭建

## 1、ZooKeeper 概述

**Zookeeper** 是一个**分布式协调服务**的开源框架。 主要用来解决分布式集群中应用系统的**一致性**问题，例如怎样避免同时操作同一数据造成脏读的问题。

ZooKeeper 本质上是一个**分布式的小文件存储系统**。 提供基于类似于文件系统的目录树方式的数据存储，并且可以对树中的节点进行有效管理。从而用来维护和监控你存储的数据的状态变化。通过监控这些数据状态的变化，从而达到基于数据的集群管理。 诸如： 统一命名服务 (dubbo)、分布式配置管理 (solr 的配置集中管理)、分布式消息队列 (sub/pub)、分布式锁、分布式协调等功能。

## 2、zookeeper 的架构图



**Leader:**

Zookeeper 集群工作的核心

事务请求（写操作） 的唯一调度和处理者，保证集群事务处理的顺序性；

集群内部各个服务器的调度者。

对于 create, setData, delete 等有写操作的请求，则需要统一转发给 leader 处理， leader 需要决定编号、执行操作，这个过程称为一个

事务。

**Follower:**

处理客户端非事务（读操作）请求，

转发事务请求给 Leader；

参与集群 Leader 选举投票  $2n-1$  台可以做集群投票。

此外，针对访问量比较大的 zookeeper 集群，还可新增观察者角色。

**Observer:**

观察者角色，观察 Zookeeper 集群的最新状态变化并将这些状态同步过来，其对于非事务请求可以进行独立处理，对于事务请求，则会转发给 Leader 服务器进行处理。

不会参与任何形式的投票只提供非事务服务，通常用于在不影响集群事务处理能力的前提下提升集群的非事务处理能力。

扯淡：说白了就是增加并发的读请求

### 3、zookeeper 的特性

1. **全局数据一致**：每个 server 保存一份相同的数据副本，client 无论连接到哪个 server，展示的数据都是一致的，**这是最重要的特征**；
2. **可靠性**：如果消息被其中一台服务器接受，那么将被所有的服务器接受。
3. **顺序性**：包括全局有序和偏序两种：全局有序是指如果在一台服务器上消息 a 在消息 b 前发布，则在所有 Server 上消息 a 都将在消息 b 前被发布；偏序是指如果一个消息 b 在消息 a 后被同一个发送者发布，a 必将排在 b 前面。
4. **数据更新原子性**：一次数据更新要么成功（半数以上节点成功），要么失败，不存在中间状态；
5. **实时性**：Zookeeper 保证客户端将在一个时间间隔范围内获得服务器的更新信息，或者服务器失效的信息。

# 三台机器 zookeeper 的集群环境搭建

## Zookeeper 的简介

为什么搭建 zookeeper 集群?

- 1. 集群有高可用的能力。
- 2. 高并发的情况下，单机版性能低下

## Zookeeper 选举策略:

Zookeeper 集群搭建指的是 ZooKeeper 分布式模式安装。通常由  $2n+1$  台 servers 组成。这是因为为了保证 Leader 选举（基于 Paxos 算法的实现）能过得到多数的支持，所以 ZooKeeper 集群的数量一般为奇数。  
Zookeeper 运行需要 java 环境，所以需要提前安装 jdk。对于安装 leader+follower 模式的集群，大致过程如下：

- 配置主机名称到 IP 地址映射配置
- 修改 ZooKeeper 配置文件
- 远程复制分发安装文件
- 设置 myid
- 启动 ZooKeeper 集群

如果要想使用 Observer 模式，可在对应节点的配置文件添加如下配置：

peerType=observer

其次，必须在配置文件指定哪些节点被指定为 Observer，如：

server.1:localhost:2181:3181:observer

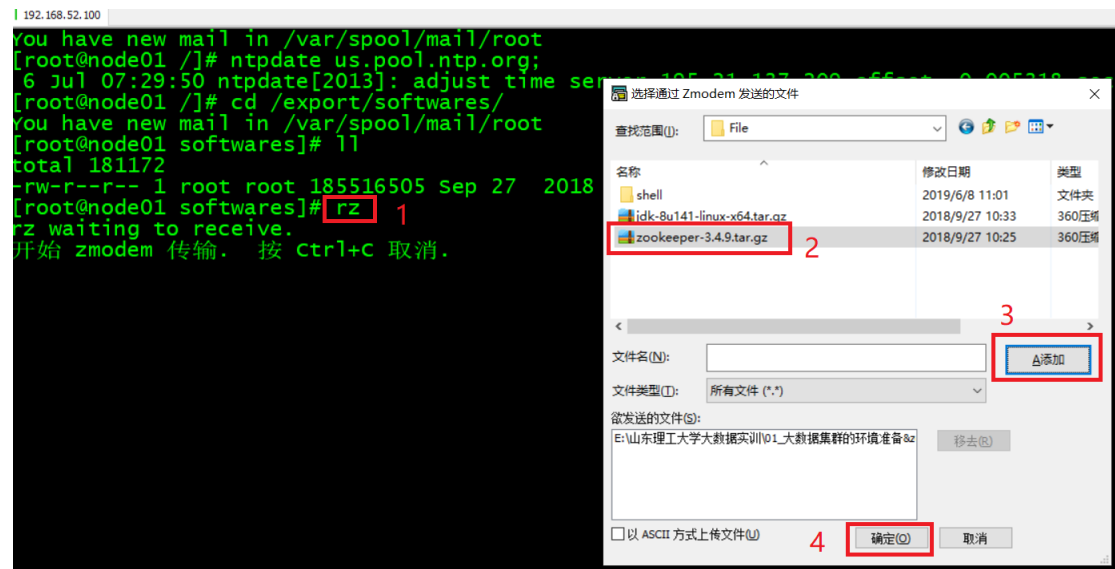
服务器 IP	主机名	myid 的值
192.168.52.100	node01	1
192.168.52.110	node02	2
192.168.52.120	node03	3

第一步：下载 zookeeper 的压缩包，下载网址如下

<http://archive.apache.org/dist/zookeeper/>

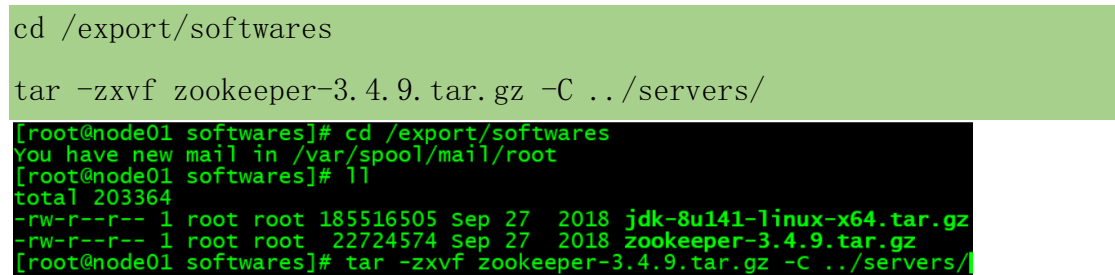
我们在这个网址下载我们使用的 zk 版本为 3.4.9

下载完成之后，上传到我们的 linux 的 /export/softwares 路径下准备进行安装



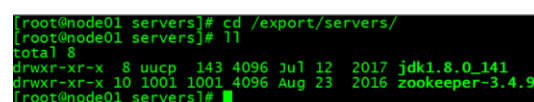
第二步：解压

解压 zookeeper 的压缩包到 /export/servers 路径下去，然后准备进行安装



第三步：修改配置文件

第一台机器修改配置文件



```
cd /export/servers/zookeeper-3.4.9/conf/
```

```
[root@node01 servers]# cd /export/servers/
[root@node01 servers]# ll
total 8
drwxr-xr-x 8 uucp 143 4096 Jul 12 2017 jdk1.8.0_141
drwxr-xr-x 10 1001 1001 4096 Aug 23 2016 zookeeper-3.4.9
[root@node01 servers]# cd /export/servers/zookeeper-3.4.9/conf/
You have new mail in /var/spool/mail/root
[root@node01 conf]# ll
total 12
-rw-rw-r-- 1 1001 1001 535 Aug 23 2016 configuration.xml
-rw-rw-r-- 1 1001 1001 2161 Aug 23 2016 log4j.properties
-rw-rw-r-- 1 1001 1001 922 Aug 23 2016 zoo_sample.cfg
[root@node01 conf]#
```

```
cp zoo_sample.cfg zoo.cfg
```

```
[root@node01 conf]# cp zoo_sample.cfg zoo.cfg
You have new mail in /var/spool/mail/root
[root@node01 conf]# ll
total 16
-rw-rw-r-- 1 1001 1001 535 Aug 23 2016 configuration.xml
-rw-rw-r-- 1 1001 1001 2161 Aug 23 2016 log4j.properties
-rw-rw-r-- 1 root root 922 Jul 6 07:39 zoo.cfg
-rw-rw-r-- 1 1001 1001 922 Aug 23 2016 zoo_sample.cfg
[root@node01 conf]#
```

```
mkdir -p /export/servers/zookeeper-3.4.9/zkdatas/
```

```
[root@node01 conf]# mkdir -p /export/servers/zookeeper-3.4.9/zkdatas/
You have new mail in /var/spool/mail/root
[root@node01 conf]# cd /export/servers/zookeeper-3.4.9/
[root@node01 zookeeper-3.4.9]# ll
total 1596
drwxr-xr-x 2 1001 1001 4096 Aug 23 2016 bin
-rw-rw-r-- 1 1001 1001 83539 Aug 23 2016 build.xml
-rw-rw-r-- 1 1001 1001 90824 Aug 23 2016 CHANGES.txt
drwxr-xr-x 2 1001 1001 4096 Jul 6 07:39 conf
drwxr-xr-x 10 1001 1001 4096 Aug 23 2016 contrib
drwxr-xr-x 2 1001 1001 4096 Aug 23 2016 dist-maven
drwxr-xr-x 6 1001 1001 4096 Aug 23 2016 docs
-rw-rw-r-- 1 1001 1001 1953 Aug 23 2016 ivysettings.xml
-rw-rw-r-- 1 1001 1001 3499 Aug 23 2016 ivy.xml
drwxr-xr-x 4 1001 1001 4096 Aug 23 2016 lib
-rw-rw-r-- 1 1001 1001 11938 Aug 23 2016 LICENSE.txt
-rw-rw-r-- 1 1001 1001 171 Aug 23 2016 NOTICE.txt
-rw-rw-r-- 1 1001 1001 1770 Aug 23 2016 README_packaging.txt
-rw-rw-r-- 1 1001 1001 1585 Aug 23 2016 README.txt
drwxr-xr-x 5 1001 1001 4096 Aug 23 2016 recipes
drwxr-xr-x 8 1001 1001 4096 Aug 23 2016 src
drwxr-xr-x 2 root root 4096 Jul 6 07:40 zkdatas
-rw-rw-r-- 1 1001 1001 1368238 Aug 23 2016 zookeeper-3.4.9.jar
-rw-rw-r-- 1 1001 1001 819 Aug 23 2016 zookeeper-3.4.9.jar.asc
-rw-rw-r-- 1 1001 1001 33 Aug 23 2016 zookeeper-3.4.9.jar.md5
-rw-rw-r-- 1 1001 1001 41 Aug 23 2016 zookeeper-3.4.9.jar.sha1
[root@node01 zookeeper-3.4.9]#
```

```
[root@node01 conf]# vim zoo.cfg
```

```
vim zoo.cfg
```

```
dataDir=/export/servers/zookeeper-3.4.9/zkdatas
autopurge.snapRetainCount=3
autopurge.purgeInterval=1
server.1=node01:2889:3889
server.2=node02:2889:3889
server.3=node03:2889:3889
```

```

# synchronization phase can take
initLimit=10
# The number of ticks that can pass between
# sending a request and getting an acknowledgement
syncLimit=5
# the directory where the snapshot is stored.
# do not use /tmp for storage, /tmp here is just
# example sake.
dataDir=/export/servers/zookeeper-3.4.9/zkdatas 更改 指定zk文件存储目录
# the port at which the clients will connect
clientPort=2181
# the maximum number of client connections.
# increase this if you need to handle more clients
#maxClientCnxns=60
#
# Be sure to read the maintenance section of the
# administrator guide before turning on autopurge.
#
# http://zookeeper.apache.org/doc/current/zookeeperAdmin.html#sc_maintenance
#
# The number of snapshots to retain in dataDir
autopurge.snapRetainCount=3 打开
# Purge task interval in hours
# Set to "0" to disable auto purge feature
autopurge.purgeInterval=1 这两行注释
server.1=node01:2888:3888
server.2=node02:2888:3888 添加这三行
server.3=node03:2888:3888
-- INSERT --

```

autopurge.snapRetainCount 这个参数指定了需要保留的文件数目。默认是保留 3 个。

autopurge.purgeInterval ZK 提供了自动清理事务日志和快照文件的功能，这个参数指定了清理频率，单位是小时，需要配置一个 1 或更大的整数，默认是 0，表示不开启自动清理功能。

server 后的数字是选举 id，在选举过程中会用到。注意:数字一定要能比较出大小。

2888 端口 原子广播端口，可以自定义 3888 端口 选举端口，

leader 会通过原子广播端口广播给其他节点，并收集每台服务器反馈信息。

## 第四步：添加 myid 配置

在第一台机器的

/export/servers/zookeeper-3.4.9/zkdatas/这个路径下创建一个文件，文件名为 myid ,文件内容为 1

```
echo 1 > /export/servers/zookeeper-3.4.9/zkdatas/myid
```

```
[root@node01 zkdatas]# more myid
1
[root@node01 zkdatas]#
```

## 第五步：安装包分发并修改 myid 的值

安装包分发到其他机器

第一台机器上面执行以下两个命令

```
scp -r /export/servers/zookeeper-3.4.9/ node02:/export/servers/
scp -r /export/servers/zookeeper-3.4.9/ node03:/export/servers/
```

第二台机器上修改 myid 的值为 2

```
echo 2 > /export/servers/zookeeper-3.4.9/zkdatas/myid
```

```
[root@node02 zkdatas]# echo 2 > /export/servers/zookeeper-3.4.9/zkdatas/myid
[root@node02 zkdatas]# more myid
2
[root@node02 zkdatas]#
```

第三台机器上修改 myid 的值为 3

```
echo 3 > /export/servers/zookeeper-3.4.9/zkdatas/myid
```

```
myid: no such file or directory
[root@node03 ~]# more /export/servers/zookeeper-3.4.9/zkdatas/myid
3
[root@node03 ~]#
```

## 第六步：三台机器启动 zookeeper 服务

三台机器启动 zookeeper 服务

这个命令三台机器都要执行

```
/export/servers/zookeeper-3.4.9/bin/zkServer.sh start
```

查看启动状态

```
/export/servers/zookeeper-3.4.9/bin/zkServer.sh status
```

```
192.168.52.100
[root@node01 zookeeper-3.4.9]# bin/zkServer.sh status
ZooKeeper JMX enabled by default
Using config: /export/servers/zookeeper-3.4.9/bin/../conf/zoo.cfg
Mode: follower
[root@node01 zookeeper-3.4.9]#
```

192.168.52.110

```
You have new mail in /var/spool/mail/root
[root@node02 zookeeper-3.4.9]# bin/zkServer.sh status
ZooKeeper JMX enabled by default
Using config: /export/servers/zookeeper-3.4.9/bin/../conf/zoo.cfg
Mode: leader
[root@node02 zookeeper-3.4.9]#
```