

一页纸精华 | ZooKeeper

原创 2016-04-07 牛家浩 中兴大数据

»» 这是中兴大数据第219篇原创文章

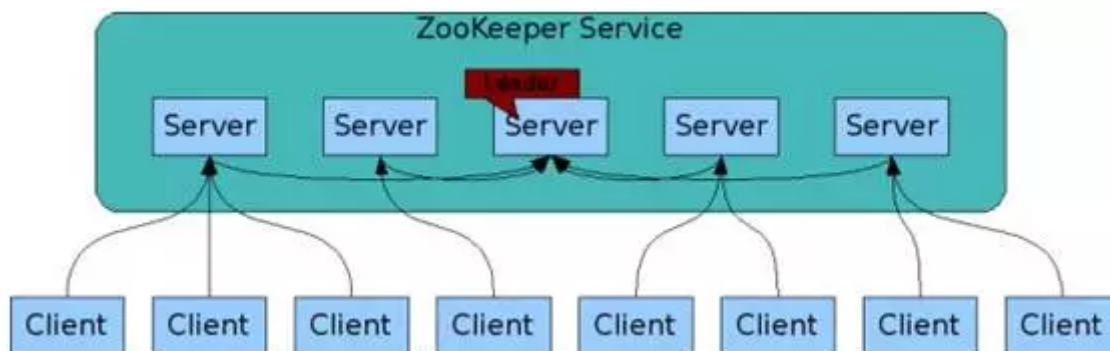
要入门大数据，最好的办法就是理清Hadoop的生态系统。中兴大数据公众号将推出“一页纸精华”栏目，将用最精炼的语言，陆续为你介绍Hadoop生态系统的各个组件。本期为你介绍分布式集群管理系统ZooKeeper。



ZooKeeper是一个针对大型分布式系统的可靠协调系统。ZooKeeper是分布式系统中的一个重要组件，它能为HDFS、HBase、MapReduce、YARN、Hive等组件提供重要的功能支撑。在分布式应用中，通常需要ZooKeeper来提供可靠的、可扩展的、分布式的、可配置的协调机制来统一各系统的状态。

1 ZooKeeper体系架构

ZooKeeper的体系架构如下图所示：



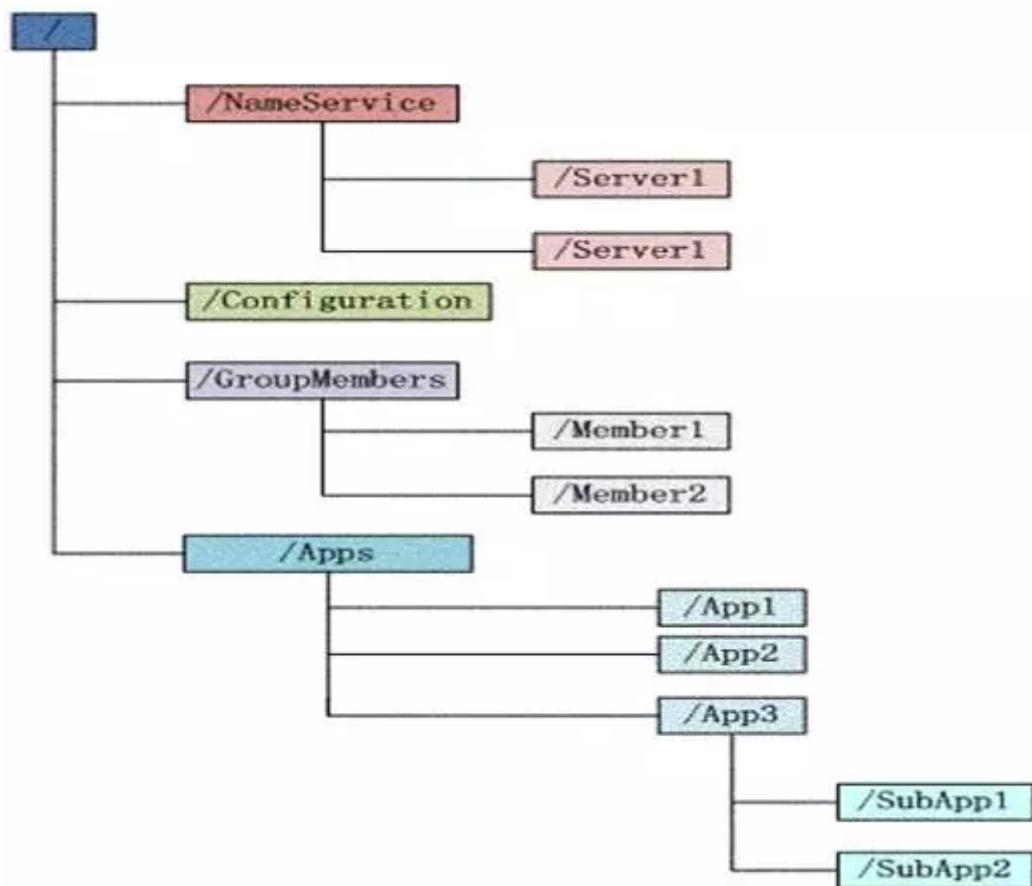
客户端可以连接到每个Server，每个Server的数据完全相同，每个Follower都和Leader有连接，接受Leader的数据更新操作，Server记录事务日志和快照到持久存储；过半数Server可用，整体服务就可用。Leader只有一个，宕机后会重新选出一个Leader。

2 ZooKeeper基本特性

- **强一致性**：Client不论连接到哪个Server，展示给它的都是同一个视图，这是ZooKeeper最重要的功能。
- **可靠性**：具有简单、健壮、良好的性能，如果消息Message被一台服务器接受，那么它将被所有的服务器接受。
- **实时性**：ZooKeeper保证客户端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。考虑到网络延时等原因，在需要最新数据时，应该在读数据之前调用sync实现。
- **等待无关 (wait-free)**：慢的或者失效的Client不得干预快速的Client的请求，使得每个Client都能有效的等待。
- **原子性**：更新只能成功或者失败，没有中间状态。
- **顺序性**：包括全局有序和偏序两种。

3 ZooKeeper数据模型

ZooKeeper的数据模型基于树型结构的命名空间，与文件系统类似，如下所示：

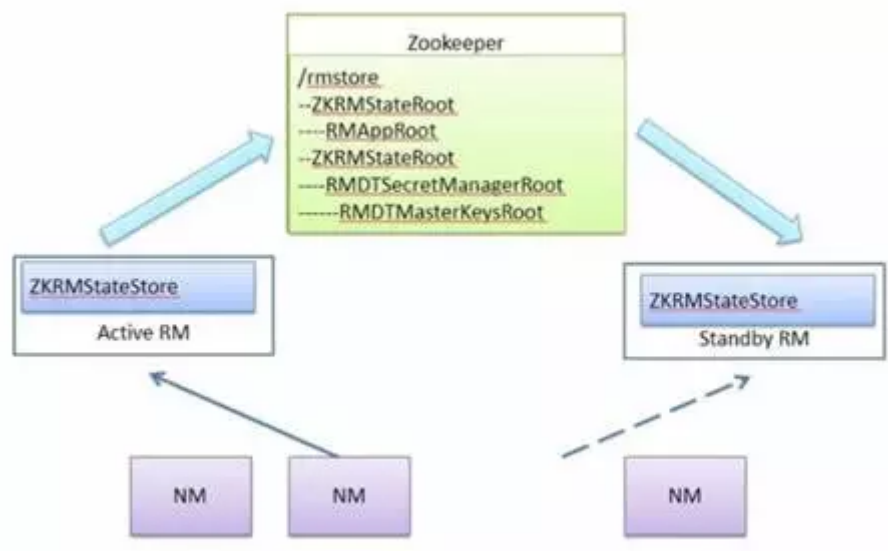


该数据模型有如下几个特点：

- 该数据模型是分布式的，数据节点被称为znode，客户端可以连接到每个Server，每个Server的数据完全相同；
- znode可以是临时节点也可以是持久性的。对应临时节点，一旦创建znode的客户端与服务器失去联系，这个znode也将自动删除，ZooKeeper的客户端和服务端通信采用长连接方式，每个客户端和服务端通过心跳来保持连接，这个连接状态称为session，如果znode是临时节点，这个session失效，znode也就删除了。
- znode可以被监控，包括这个目录节点中存储的数据的修改，子节点目录的变化等，一旦变化可以通知设置监控的客户端，这个是ZooKeeper的核心特性，ZooKeeper的很多功能都是基于这个特性实现的，通过这个特性可以实现的功能包括配置的集中管理，集群管理，分布式锁等。

4 ZooKeeper在Hadoop中的应用举例

ZooKeeper在YARN的应用如下图所示：



RM 的作业信息存储ZooKeeper的/rmstore下，Active RM向这个目录写App信息。RM启动的时候会通过向ZK的/rmstore目录下写一个Lock文件，写成功则成为Active，否则为Standby，Standby RM会一直监控Lock文件是否存在，如果不存在则会试图去创建，即争取成为Active RM。

当Active RM挂掉，另外一个Standby RM成功转换为Active RM后，会从/rmstore读取相应的作业信息，重新构建作业的内存信息。然后启动内部服务，开始接收NM的心跳，构建集群资源信息，并接收客户端提交作业的请求等。

如何阅读往期“一页纸精华”？

1 进入公众号对话框界面，选择“干货专区” — “基础课堂”子菜单



2 在弹出页面选择“一页纸”栏目

零基础可阅读 | 大数据初级班教程简介

零基础

一页纸



一页纸精华 | Hadoop生态圈的浓缩介绍

要入门大数据，最好的办法就是理清Hadoop的生态系统。本栏目用最精炼的语言，陆续为你介绍Hadoop生态系统的各个组件。



一页纸精华 | Hadoop概览

Hadoop基础知识栏目“一页纸精华”第1期。



一页纸精华 | YARN

本期为你介绍Hadoop统一资源管理框架YARN。



一页纸精华 | MapReduce

本期为你介绍Hadoop分布式计算框架MapReduce。



一页纸精华 | Hive

本期为你介绍Hadoop数据仓库工具Hive。



一页纸精华 | HBase

本期为你介绍Hadoop分布式数据库HBase。



再来一篇？

长按二维码关注

