

# 大数据入门基础系列之Hadoop1.X、Hadoop2.X和Hadoop3.X的多维度区别详解

zlsich 大数据躺过的坑 2017-09-27

在前面的博文里，我已经介绍了

[大数据入门基础系列之Linux操作系统简介与选择](#)

[大数据入门基础系列之虚拟机的下载、安装详解](#)

[大数据入门基础系列之Linux的安装详解](#)

[大数据入门基础系列之远程连接工具下载和安装详解](#)

[大数据入门基础系列之Apache版本的hadoop集群详细部署搭建（包括HA和非HA）（包括单节点、3节点、5节点）](#)

[大数据入门基础系列之CDH版本的hadoop集群详细部署搭建（3节点）](#)

[大数据入门基础系列之ClouderManager版本的hadoop集群详细部署搭建（1、3、4节点）（在线或离线）](#)

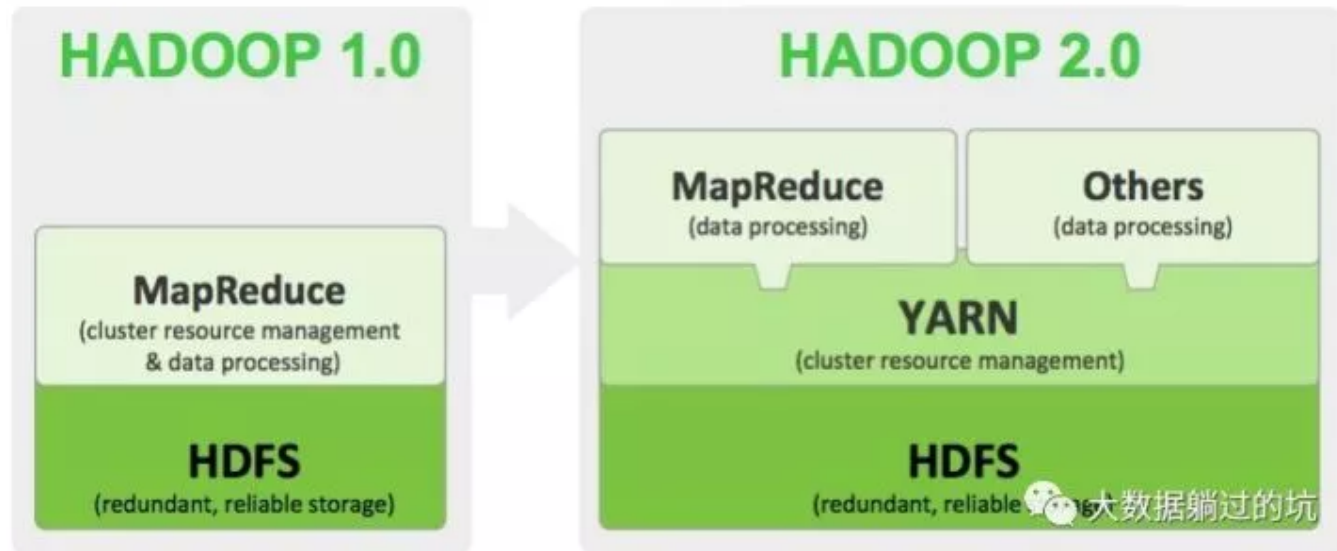
[大数据入门基础系列之Ambari版本的hadoop集群详细部署搭建（1、3、4节点）（在线或离线）](#)

[大数据入门基础系列之初步认识hadoop生态系统圈](#)

## 为什么要写这篇微信公众号博文，来特意谈Hadoop1.X、Hadoop2.X和Hadoop3.X的区别？

**答：最近校招，我在面试某一家知名互联网公司问及了此问题。**

现在，我们一般普遍用的比较多的一个稳定版本是hadoop2.6.0。当然，今年的6月份也已经诞生了hadoop3.X，那么肯定有一大批跟我一样喜欢敢于去稳中探索新版本的朋友，肯定已经对其研究一番了。



当我们安装完毕hadoop2.X的时候，我们看到为啥没有jobtracker，这是因为hadoop2中已经没有jobtracrer了，而是产生了yarn。

下面介绍了hadoop1.X与hadoop2.X的一些常用概念。

- 1.hadoop1.0的组成是什么？
  - 2.hadoop2.0为和变化如此大？
  - 3.什么是YARN？（Hadoop2.X开始）
  - 4.什么是HDFS Federation？（Hadoop2.X开始）
  - 5.什么是HDFS HA？（Hadoop2.X开始）
  - 6.什么是
- .... ..

（1）：特性

Hadoop1.x：批量应用

Hadoop2.x：批量应用，交互式，实时性，流式，内存计算。

（2）：HDFS与HDFS2

HDFS2多了高可用性。

（3）：MapReduce

Hadoop2.x中将原来1.x中的MapReduce的功能分开了，只保留了数据处理。将集群资源管理放到了Yarn中。从而让2.x的生态圈极大的扩展。

## （1）Hadoop 1.0

Hadoop 1.0即第一代Hadoop，由分布式存储系统HDFS和分布式计算框架MapReduce组成，其中，HDFS由一个NameNode和多个DataNode组成，**MapReduce由一个JobTracker和多个TaskTracker组成**。

Hadoop1.X对应Hadoop版本为Apache Hadoop 0.20.x、1.x、0.21.X、0.22.x和CDH3。

**hadoop1.0需要解决的几个问题：**

1：单点故障问题，HDFS和MapReduce都是Master-Slave的结构，但是Master节点都是单点，一旦出现故障，那么集群就不能正常运行。

2：HDFS存储海量数据是按照Block来存储的，整个存储只有一种格式，企业存储受限，企业的数据非常多样，存储起来容易造成资源的浪费。

当namenode所在机器的内存不够时，集群不能正常工作。

3：MapReduce进行离线的批处理，处理速度慢，以Map和Reduce进程的方式来运行，一般是晚间来进行计算，MapReduce集群资源利用率低。

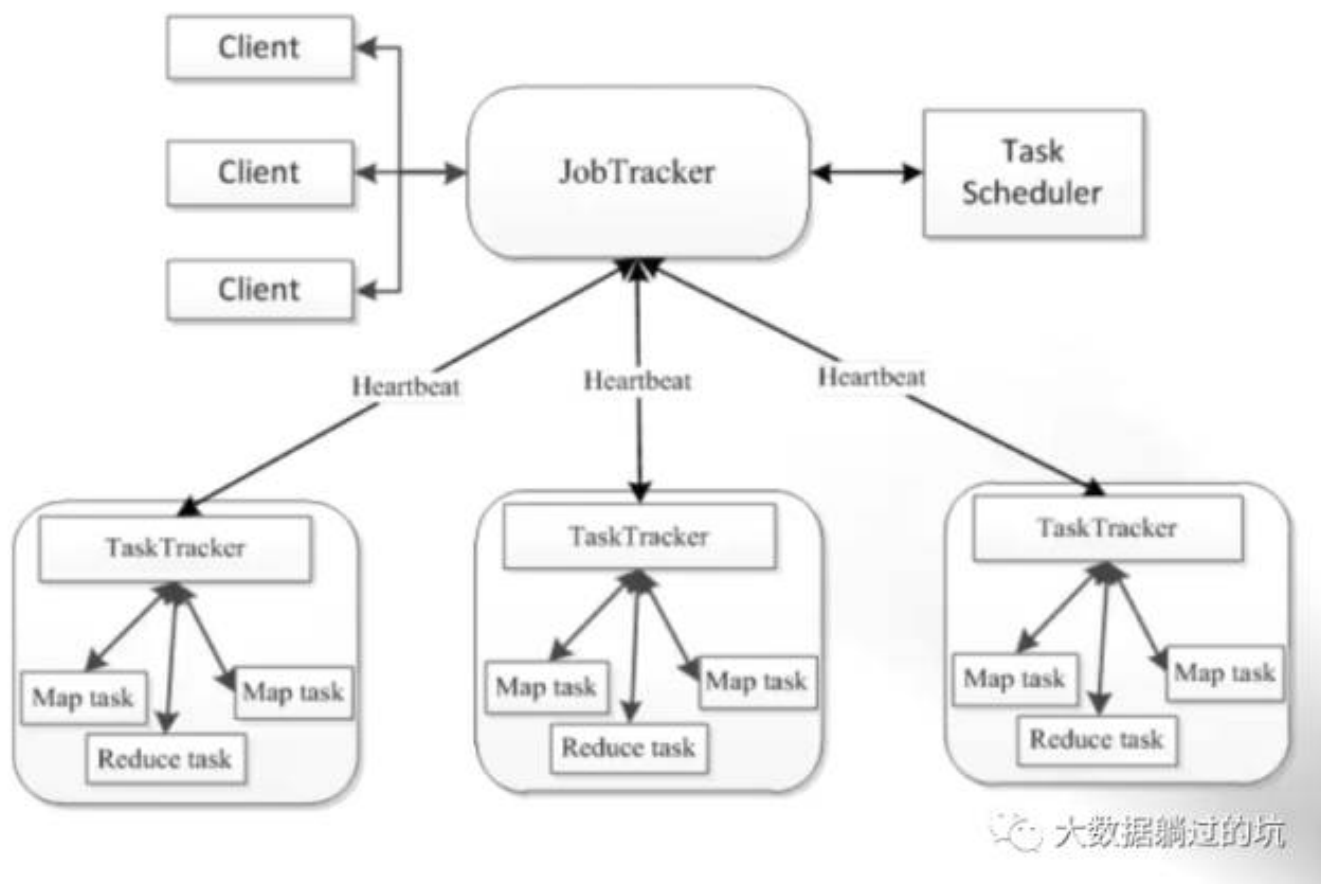
## ( 2 ) Hadoop 2.0

Hadoop 2.0即第二代Hadoop，**为克服Hadoop 1.0中HDFS和MapReduce存在的各种问题而提出的**。针对Hadoop 1.0中的单NameNode制约HDFS的扩展性问题，提出了HDFS Federation，它让多个NameNode分管不同的目录进而实现访问隔离和横向扩展，同时它彻底解决了NameNode 单点故障问题；针对Hadoop 1.0中的MapReduce在扩展性和多框架支持等方面的不足，它将JobTracker中的资源管理和作业控制功能分开，分别由组件ResourceManager和ApplicationMaster实现。其中，ResourceManager负责所有应用程序的资源分配，而ApplicationMaster仅负责管理一个应用程序，进而诞生了全新的通用资源管理框架YARN。基于YARN，用户可以运行各种类型的应用程序（不再像1.0那样仅局限于MapReduce一类应用），从离线计算的MapReduce到在线计算（流式处理）的Storm等。

Hadoop 2.0对应Hadoop版本为Apache Hadoop 0.23.x、2.x和CDH4。

### ( 3 ) MapReduce 1.0或MRv1

MapReduce 1.0计算框架主要由三部分组成，分别是编程模型、数据处理引擎和运行时环境。它的基本编程模型是将问题抽象成Map和Reduce两个阶段，其中Map阶段将输入数据解析成key/value，迭代调用map()函数处理后，再以key/value的形式输出到本地目录，而Reduce阶段则将key相同的value进行规约处理，并将最终结果写到HDFS上；它的数据处理引擎由MapTask和ReduceTask组成，分别负责Map阶段逻辑和Reduce阶段逻辑的处理；它的运行时环境由（一个）JobTracker和（若干个）TaskTracker两类服务组成，其中，JobTracker负责资源管理和所有作业的控制，而TaskTracker负责接收来自JobTracker的命令并执行它。该框架在扩展性、容错性和多框架支持等方面存在不足，这也促使了MRv2的产生。



### ( 4 ) MapReduce 2.0或MRv2

MRv2具有与MRv1相同的编程模型和数据处理引擎，唯一不同的是运行时环境。MRv2是在MRv1基础上经加工之后，运行于资源管理框架YARN之上的计算框架MapReduce。它的运行时环境不再由JobTracker和TaskTracker等服务组成，而是变为通用资源管理系统YARN和作业控制进程ApplicationMaster，其中，YARN负责资源管理和调度，而ApplicationMaster仅负责一个

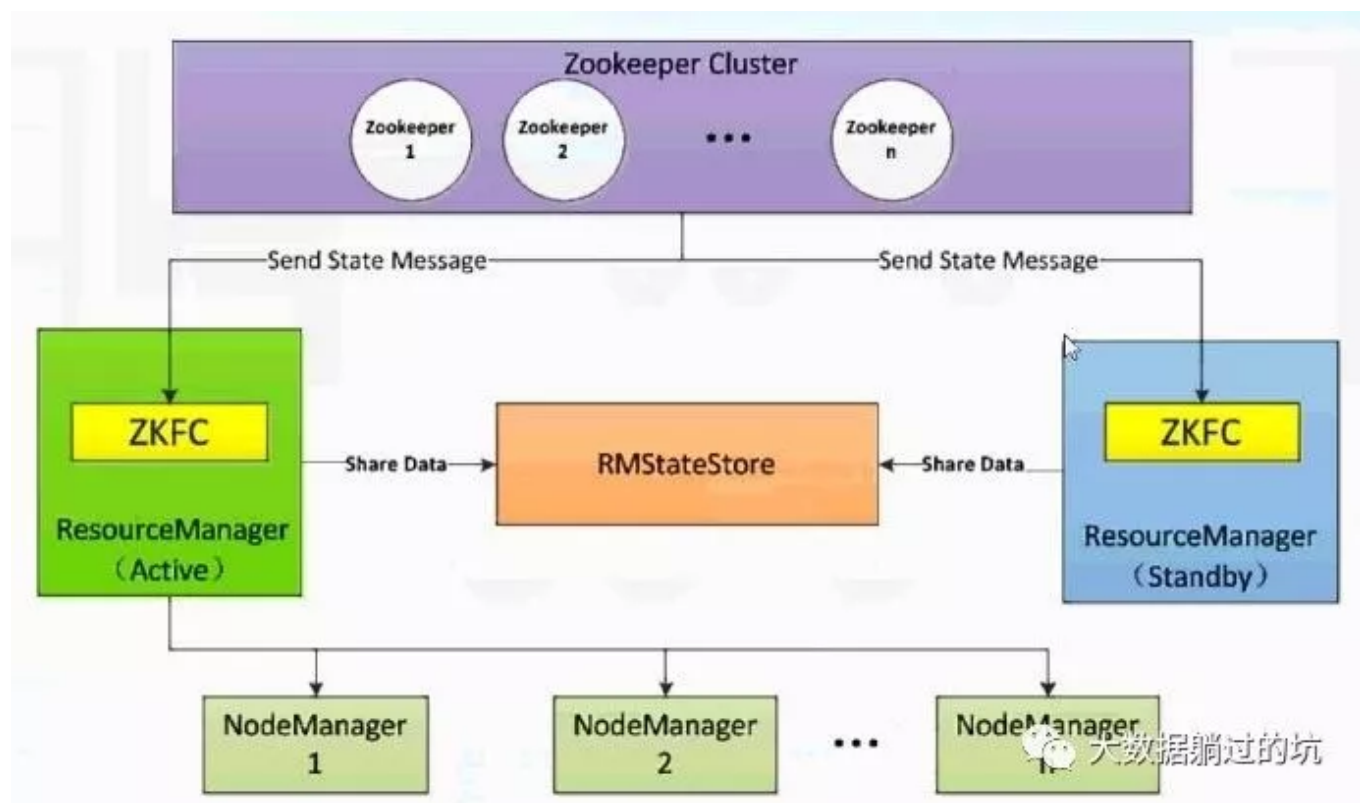
作业的管理。简言之，MRv1仅是一个独立的离线计算框架MapReduce，而MRv2则是运行于YARN之上的MapReduce。

## ( 5 ) YARN ( Hadoop2.X开始 )

YARN是Hadoop 2.0中的资源管理系统，它是一个通用的资源管理模块，可为各类应用程序进行资源管理和调度。YARN不仅限于MapReduce一种框架使用，也可以供其他框架使用，比如Tez、Spark、Storm等。YARN类似于几年前的资源管理系统Mesos和更早的Torque。由于YARN的通用性，下一代MapReduce的核心已经从简单的支持单一应用的计算框架MapReduce转移到通用的资源管理系统YARN。

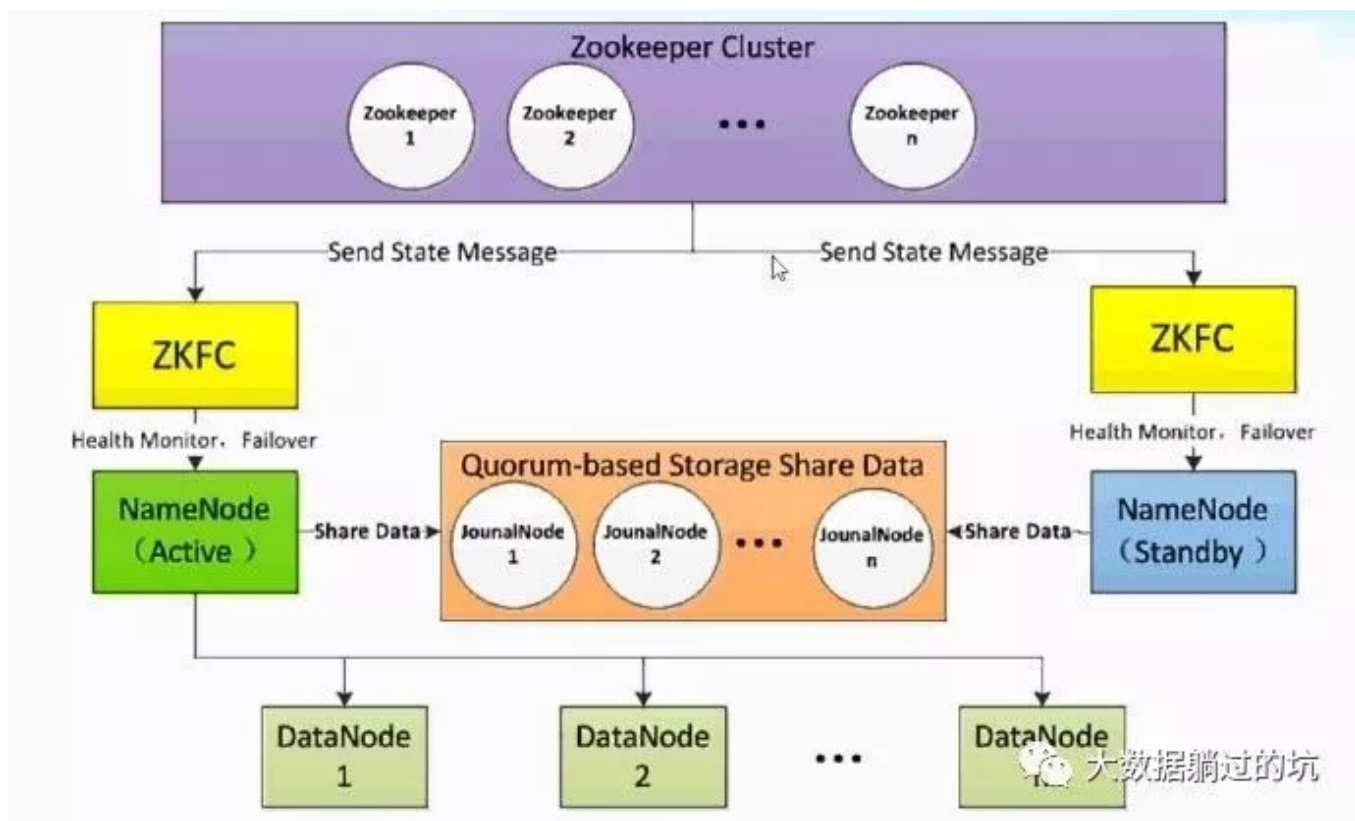
## ( 6 ) YARN HA ( Hadoop2.6.0开始 )

[Hadoop发展到了2.6.0](#)，[Yarn也有了HA](#)。Yarn把Datanode和应用的执行状态都存在在RMStateStore上面。从而具有long running services,Support for rolling upgrades。任务挂了就不用重新跑了。太牛的功能了。



## ( 7 ) Hadoop2.x中HDFS架构

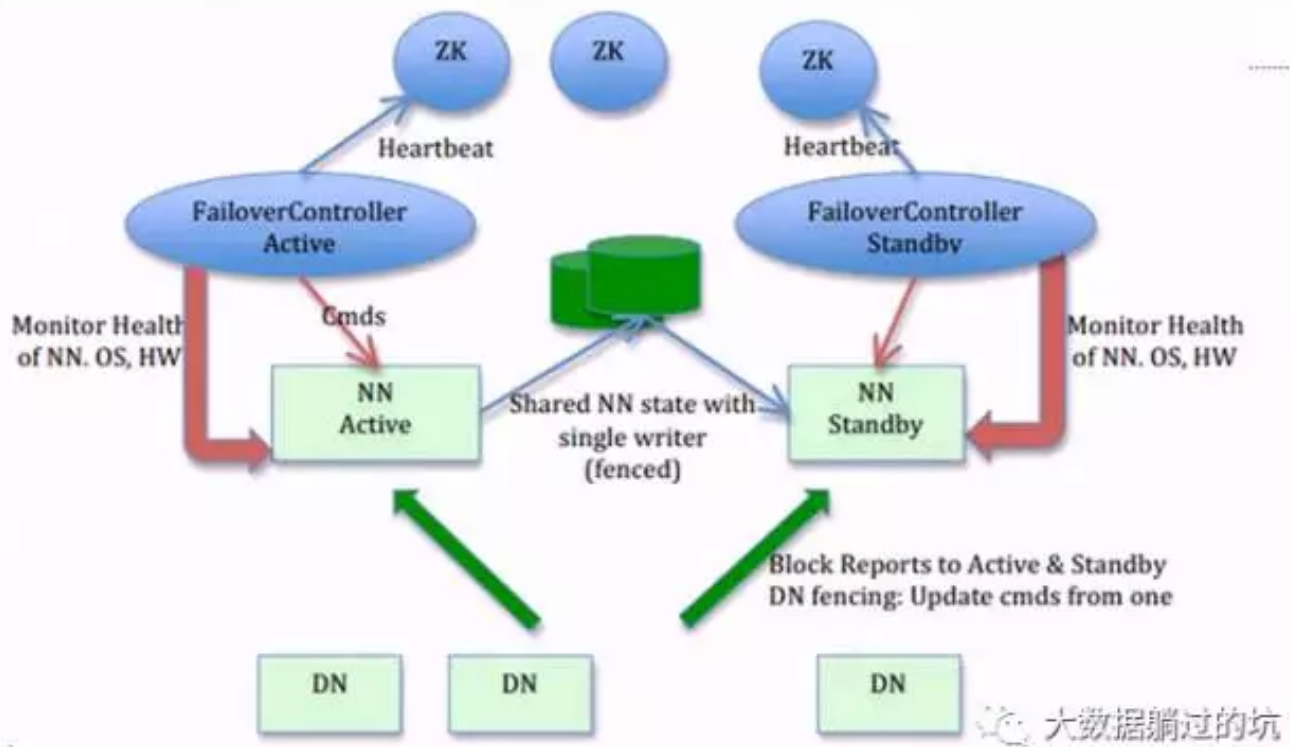
下图中，我们可以看到一个NameNode(Active)还有一个NameNode(Standby)。这两者中间有JournalNode，放镜像文件，编辑日志文件。使用Zookeeper来保证故障转移，从而实现HA。还具有联邦（Federation）的功能，具体见如下。



## ( 8 ) HDFS HA的自动failover

FailoverController用来监控NameNode的状态，如果NameNode的状态有变化，那么会通知Zookeeper的，FailoverController和Zookeeper的交互很多，FailoverController属于中间的一个中转管理者，FailoverController是HDFS2中的进程。





关于**HDFS-HA的配置-----自动Failover**，可以见

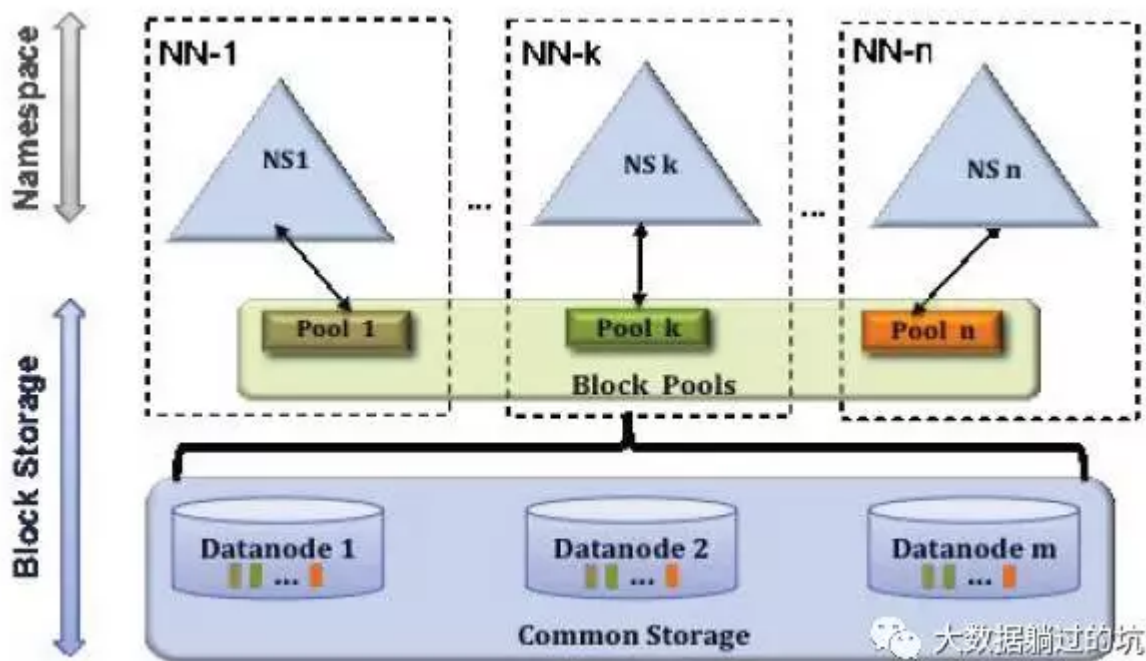
<http://blog.csdn.net/chenpingbupt/article/details/7922089>

当然，这里为了扩展，也有**HDFS-HA的配置-----手动Failover方式**

<http://blog.csdn.net/chenpingbupt/article/details/7922080>

## ( 9 ) HDFS Federation ( Hadoop2.X )

Hadoop 2.0中对HDFS进行了改进，使**NameNode可以横向扩展成多个**，每个**NameNode**分管一部分目录，进而产生了**HDFS Federation**，该机制的引入不仅增强了HDFS的扩展性，也使**HDFS具备了隔离性**。



重点详解！！！！

## 1. Hadoop 3.0简介

Hadoop 2.0是基于JDK 1.7开发的，而JDK 1.7在2015年4月已停止更新，这直接迫使Hadoop社区基于JDK 1.8重新发布一个新的Hadoop版本，而这正是hadoop 3.0。

Hadoop 3.0的alpha版已经在今年夏天发布，预计GA版本11月或12月发布。

Hadoop 3.0中引入了一些重要的功能和优化，包括HDFS 可擦除编码、多Namenode支持、MR Native Task优化、YARN基于cgroup的内存和磁盘IO隔离、YARN container resizing等。

## 2. Hadoop 3.0新特性

Apache hadoop 项目组最新消息，hadoop3.x以后将会调整方案架构，将Mapreduce 基于内存+io+磁盘，共同处理数据。

其实最大改变的是hdfs，hdfs 通过最近black块计算，根据最近计算原则，本地black块，加入到内存，先计算，通过IO，共享内存计算区域，最后快速形成计算结果。

**Hadoop 3.0在功能和性能方面，对hadoop内核进行了多项重大改进，主要包括：**

### 2.1 Hadoop Common

(1) **精简Hadoop内核，包括剔除过期的API和实现，将默认组件实现替换成最高效的实现**（比如将FileOutputCommitter缺省实现换为v2版本，废除hftp转由webhdfs替代，移除Hadoop子实现序列化库org.apache.hadoop.Records。



( 2 ) **Classpath isolation 以防止不同版本jar包冲突**，比如google Guava在混合使用Hadoop、HBase和Spark时，很容易产生冲突。见下面的链接 <https://issues.apache.org/jira/browse/HADOOP-11656>

( 3 ) **Shell脚本重构**。Hadoop 3.0对Hadoop的管理脚本进行了重构，修复了大量bug，增加了新特性，支持动态命令等。见下面的链接 <https://issues.apache.org/jira/browse/HADOOP-9902>

## 2.2 Hadoop HDFS

( 1 ) **HDFS支持数据的擦除编码**，这使得HDFS在不降低可靠性的前提下，节省一半存储空间。见下面的链接，<https://issues.apache.org/jira/browse/HDFS-7285>

( 2 ) 多**NameNode支持**，即支持一个集群中，一个active、多个standby namenode部署方式。注：多ResourceManager特性在hadoop 2.0中已经支持。见下面的链接，<https://issues.apache.org/jira/browse/HDFS-6440>

## 2.3 Hadoop MapReduce

( 1 ) **Tasknative优化**。为MapReduce增加了C/C++的map output collector实现（包括Spill，Sort和IFile等），通过作业级别参数调整就可切换到该实现上。对于shuffle密集型应用，其性能可提高约30%。见下面的链接<https://issues.apache.org/jira/browse/MAPREDUCE-2841>

( 2 ) **MapReduce内存参数自动推断**。在Hadoop 2.0中，为MapReduce作业设置内存参数非常繁琐，涉及到两个参数：mapreduce.{map,reduce}.memory.mb 和 mapreduce.{map,reduce}.java.opts，一旦设置不合理，则会使得内存资源浪费严重，比如将前者设置为4096MB，但后者却是“-Xmx2g”，则剩余2g实际上无法让java heap使用到。见下面的链接<https://issues.apache.org/jira/browse/MAPREDUCE-5785>

## 2.4 Hadoop YARN

( 1 ) 基于cgroup的内存隔离和IO Disk隔离 ( <https://issues.apache.org/jira/browse/YARN-2619> )

( 2 ) 用curator实现RM leader选举 ( <https://issues.apache.org/jira/browse/YARN-4438> )

( 3 ) containerresizing ( <https://issues.apache.org/jira/browse/YARN-1197> )

( 4 ) Timelineserver next generation ( <https://issues.apache.org/jira/browse/YARN-2928> )

以下是hadoop-3.0的最新参数：（翻译过来就是我下面写的总结部分）

hadoop-3.0

- HADOOP

- Move to JDK8+
- Classpath isolation on by default HADOOP-11656
- Shell script rewrite HADOOP-9902
- Move default ports out of ephemeral range HDFS-9427

- HDFS
  - Removal of hftp in favor of webhdfs HDFS-5570
  - Support for more than two standby NameNodes HDFS-6440
  - Support for Erasure Codes in HDFS HDFS-7285
- YARN
- MAPREDUCE
  - Derive heap size or mapreduce.\*.memory.mb automatically MAPREDUCE-5785

## 3. Hadoop3.0 总结

Hadoop 3.0的alpha版已经在今年夏天发布，预计GA版本11月或12月发布。Hadoop 3.0中引入了一些重要的功能和优化，包括HDFS 可擦除编码、Namenode支持、MR Native Task优化、YARN基于cgroup的内存和磁盘IO隔离、YARN container resizing等。

- 1、JDK版本的最低依赖从1.7变成了1.8；
- 2、HDFS支持Erasure Encoding；
- 3、Timeline Server v2版本；
- 4、hadoop-client这个依赖分为hadoop-client-api和hadoop-client-runtime两个依赖；
- 5、支持随机container和分布式调度；
- 6、MR进行了task级别的本地优化，性能提升30%；
- 7、支持多个Standby状态的NameNode；
- 8、多个端口被改动；
- 9、支持微软的Azure分布式文件系统和阿里的aliyun分布式文件系统；
- 10、datanode内部添加了负载均衡；

网上各种资料宣传说，比Spark快10倍的Hadoop3.0怎么怎么样，其实这种说法没有说服力不太可信，正所谓一山还有一山高。技术不会停滞不前，更是交集的火花。让热爱大数据的我们，一起来拭目以待吧！

同时，大家可以关注**我的个人博客**：

<http://www.cnblogs.com/zls1ch/> 和 <http://www.cnblogs.com/lchzls/>

以及对应本平台的QQ群：**161156071**（大数据躺过的坑）



大数据躺过的坑

扫一扫二维码，加入该群。



大数据躺过的坑