

基于 Hadoop 的云计算与存储平台研究与实现

唐世庆¹, 李云龙¹, 田凤明², 胡海荣¹

(1. 装甲兵工程学院 信息工程系, 北京 100072; 2. 65571 部队保障部, 吉林 四平 136000)

摘要:传统数据计算与存储成本高, 并行程序编写困难, 针对该问题, 在综述 Hadoop 核心技术的基础上, 利用 Hadoop 分布式处理技术和虚拟化技术设计、搭建了云数据计算与存储集群平台, 实验证明, 此平台与传统单机计算方式以及物理机集群计算方式相比具有较高性能, 同时资源利用率也得到了提高。

关键词: Hadoop; 云计算; 云存储; 虚拟化技术

本文引用格式:唐世庆, 李云龙, 田凤明, 等. 基于 Hadoop 的云计算与存储平台研究与实现[J]. 四川兵工学报, 2014(8): 97-100.

中图分类号: TP391

文献标识码: A

文章编号: 1006-0707(2014)08-0097-04

Research on Cloud Computing and Storage Platform Based on Hadoop and its Implementation

TANG Shi-qing¹, LI Yun-long¹, TIAN Feng-ming², HU Hai-rong¹

(1. Department of Information Engineering, Academy of Armored Force Engineering, Beijing 100072, China;

2. The Security Department of the 65571 Army, Siping 136000, China)

Abstract: The cost of the traditional calculation and storage is very high, and the parallel programming is very difficult to be programmed. To solve the question, after telling about the core technologies of Hadoop, this paper used the distributed technology and virtual technology to build a computing and storage platform of cloud data. Finally, the experimental result shows that, the platform has the high performance, meanwhile, the coefficient of physical resource utilization has been improved comparing to the traditional computing way of single computer and multi-node physical computers.

Key words: Hadoop; cloud computing; cloud storage; virtual technology

Citation format: TANG Shi-qing, LI Yun-long, TIAN Feng-ming, et al. Research on Cloud Computing and Storage Platform Based on Hadoop and its Implementation [J]. Journal of Sichuan Ordnance, 2014(8): 97-100.

随着互联网技术的发展, 数字信息正在成指数增加, 根据 Internet Data Center 发布的 Digital Universe 报告显示, 在未来 8 年中所产生的数据量将达到 40 ZB, 相当于每人产生 5 200 G 的数据^[1], 如何高效地计算和存储这些海量数据成为互联网企业所要面临的挑战。传统的大规模数据处理大多采用并行计算、网格计算、分布式高性能计算等, 耗费昂贵的存储与计算资源, 而且对于大规模数据计算任务的有效分配和数据合理分割都需要复杂的编程才可以实现^[2]。基于 Hadoop 分布式云平台的出现成为解决此类问题的良好途

径, 本文将在综述 Hadoop 核心技术: HDFS 和 MapReduce 基础上, 利用 VMware 虚拟机搭建一个基于 Hadoop 分布式技术的高效、易扩展的云数据计算与存储平台, 并通过实验验证分布式计算与存储的优势。

1 Hadoop 及其相关技术

Hadoop^[3]是并行技术、分布式技术和网格计算技术发展的产物, 是一种为适应大规模数据计算和存储而发展起来的

收稿日期: 2014-04-06

基金项目: 军队科研计划项目。

作者简介: 唐世庆(1965—), 男, 硕士, 副教授, 主要从事指挥信息系统研究。

模型架构。Hadoop 是 Apache 公司旗下的一个分布式计算和存储的框架平台,能够高效存储大量数据,而且可以编写分布式应用程序来分析计算海量数据。Hadoop 可在大量廉价硬件设备集群中运行程序,为各应用程序提供可靠稳定的接口来构建高扩展性和高可靠行的分布式系统。Hadoop 具有成本低廉、可靠性高、容错性高、扩展性强、效率高、可移植性强、免费开源的优点^[4]。

Hadoop 集群为典型 Master/Slaves 结构,基于 Hadoop 的云计算与存储架构模型如图 1 所示。

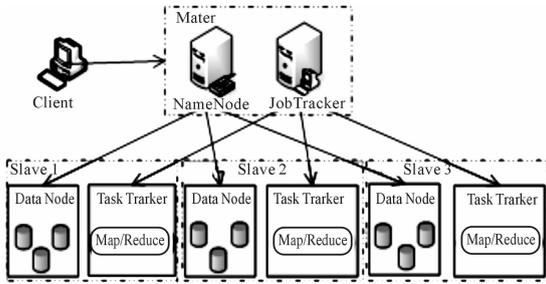


图 1 基于 Hadoop 的云计算与存储架构模型

1.1 Hadoop 分布式文件系统 HDFS

HDFS^[5]是一个运行在大量廉价硬件之上的分布式文件系统,它是 Hadoop 平台的底层文件存储系统,主要负责数据的管理和存储,对于大文件的数据访问具有良好性能。HDFS 与传统的分布式文件系统相似,但是也存在着一一定的不同,具有硬件故障、大数据集、简单一致性、数据流式访问、移动计算的便捷性等特点^[6]。HDFS 的工作流程及架构如图 2 所示。

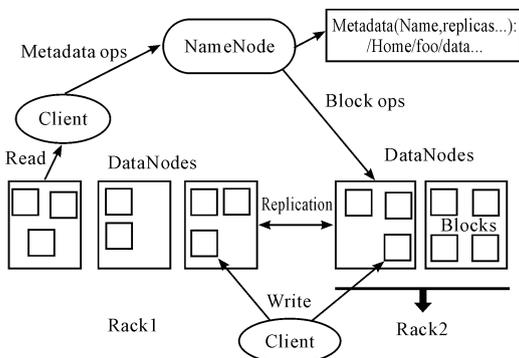


图 2 HDFS 的工作流程及架构结构

一个 HDFS 集群中有一个 NameNode 和多个 DataNode。如图 2 所示,NameNode 是中心服务器,它用来管理文件系统的元数据信息以及客户端对文件的读写访问,维护文件系统的元数据信息以及客户端对文件的读写访问,维护文件系统的元数据信息以及客户端对文件的读写访问,维护文件系统的元数据信息以及客户端对文件的读写访问。这些信息以编辑日志文件(Editlog)和命名空间镜像文件(FsImage)的形式保存在磁盘中。NameNode 还暂时记录着各个块(Block)所在的 DataNode 信息。其功能主要有:管理元数据和文件块;简化元数据更新操作;监听和处理请求^[7]。

DataNode 通常在集群中一个节点一个,用来存储、检索

数据块,响应 NameNode 下达的创建、复制、删除数据块的命令,并定时向 NameNode 发送“心跳”,通过心跳信息向 NameNode 汇报自己的负载情况,同时通过心跳信息来接受 NameNode 下达的指令信息;NameNode 通过“心跳”信息来确定 DataNode 是否失效,它定时 ping 每个 DataNode,如果在规定的时间内没有收到 DataNode 的反馈就认为此节点失效,然后对整个系统进行负载调整。在 HDFS 中,每个文件划分成一个或多个 blocks(数据块)分散存储在不同的 DataNode 中,DataNode 之间进行数据块的相互复制而形成多个备份。

1.2 Map/Reduce 编程框架

Map/Reduce^[8]是 Hadoop 用来处理云计算中海量数据的编程框架,简单易用,程序员在不必了解底层实现细节的基础上便可写出程序来处理海量数据。利用 Map/Reduce 技术可以在数千部服务器上同时开展广告业务和网络搜索等任务,并可以方便地处理 TB、PB,甚至是 EB 级的数据。

Map/Reduce 框架由 JobTracker 和 TaskTracker 组成。JobTracker 只有一个,它是主节点,负责任务的分配和调度,管理着几个 TaskTracker;TaskTracker 一个节点一个,用来接受并处理 JobTracker 发来的任务。

MapReduce 针对集群中的大型数据集进行分布式运算,它的整个框架由 Map 和 Reduce 函数组成,处理数据时先执行 map 再执行 reduce。具体执行过程如图 3 所示。执行 map 函数前先对输入数据进行分片;然后将不同的片段分配给不同的 map 执行,map 函数处理之后以<key, value>的形式输出;在进入 reduce 阶段前,map 函数先将原来的<key, value>分成多组中间的键值对再发给一个 reducer 进行处理;最后 reduce 函数合并 key 相同的 value,并输出结果到磁盘上。

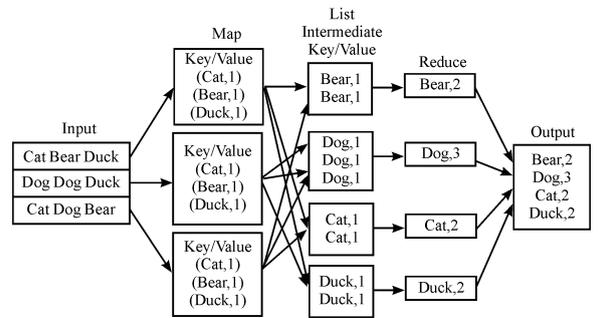


图 3 MapReduce 计算过程

2 基于 Hadoop 的云计算与存储平台设计

目前,多核计算机的广泛使用使其在搭建 Hadoop 集群系统时,分给各 DataNode 节点的多个任务会产生对资源的竞争,例如:内存、CPU、输入输出带宽等,这会导致暂时用不到的资源处于闲置状态,致使一些资源的浪费以及响应时间的延长,资源开销的增加最终会导致系统性能的降低。为解决此问题,本研究提出一种基于 VMware 虚拟机和 Hadoop 相结合的集群环境模型,如图 4 所示,即在一台计算机中搭建多台虚拟操作系统^[9],此种做法的优点是可以增加 DataNode

和 TaskTracker 节点,而且可以充分利用物理资源,提高运算和存储的效率。

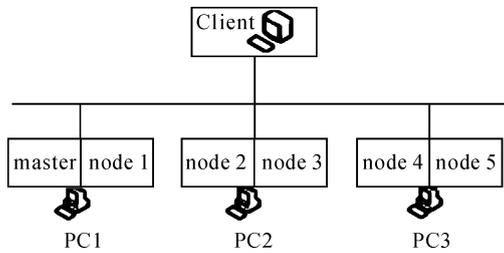


图4 基于 VMware 虚拟机和 Hadoop 结合的模式

3 实验平台搭建

3.1 硬件环境配置

准备 3 台双核计算机,并分别安装 2 台 VMware 虚拟机软件,在虚拟机中装入 Linux OS,从而将 3 台计算机扩展成为 6 台计算机,3 台计算机具有相同的配置,配置具体如表 1 所示。

表1 软硬件配置信息表

	处理器: Intel Core i5
硬件	CPU: 2.4 GHz
	内存: 4 G
	OS: Red Hat Linux OS5.0
软件	JDK 版本: jdk_1.6
	Hadoop 版本: hadoop - 1.0.3

Hadoop 集群包括 1 个 NameNode 服务器和 5 个 DataNode 服务器,配置信息如表 2 所示。

表2 Hadoop 集群设备配置信息

主机名	IP 地址	作用	内存/ GB	硬盘/ GB
master	192.168.1.100	NameNode, JobTracker	1	20G
node1	192.168.1.101	DataNode, TaskTracker	1	20
node2	192.168.1.102	DataNode, TaskTracker	1	20
node3	192.168.1.103	DataNode, TaskTracker	1	20
node4	192.168.1.104	DataNode, TaskTracker	1	20
node5	192.168.1.105	DataNode, TaskTracker	1	20

3.2 Hadoop 环境搭建

Hadoop 环境搭建过程为:配置集群 hosts 列表、安装 JAVA JDK 系统软件、配置环境变量、生成登陆密钥、创建用户帐号和 Hadoop 部署目录及数据目录、配置 hadoop-env. sh 环境变量、配置 core-site. xml、hdfs-site. xml、mapred-site. xml。

配置完毕之后进行格式化文件,命令为:

```
/opt/modules/hadoop/hadoop-1.0.3/bin/hadoop namenode -format
```

然后启动所有节点,输入命令: start-all. sh。通过界面查看集群是否部署成功,首先检查 NameNode 和 DataNode 节点是否正常,打开浏览器输入网址: http://master:50070,若 Live Nodes 有 6 个,说明全部节点成功启动。然后检查 JobTracker 和 TaskTracker 节点,输入网址: http://master:50030,若 Nodes 节点有 6 个说明节点启动成功。

4 实验内容及结果分析

在部署好的 Hadoop 云数据计算与存储平台上进行实验来验证基于分布式数据计算和存储的方法在数据计算和存储上存在优势。

1) 实验一:运行 Hadoop 自带的蒙特卡洛求 PI 程序验证基于 Hadoop 分布式云计算的高效性。计算任务设为 10 个,计算量为 10 的 3、4、5、6 次方。

环境一:单机情况下运行;

环境二:3 台物理机搭建的集群系统中运行;

环境三:6 台虚拟机搭建的集群系统中运行。集群环境运行日志如图 5 所示。

```
[hadoop@master hadoop-1.0.3]$ hadoop jar hadoop-examples-1.0.3.jar pi 10 10000
Warning: $HADOOP_HOME is deprecated.

Number of Maps = 10
Samples per Map = 10000
Wrote input for Map #0
Wrote input for Map #1
Wrote input for Map #2
Wrote input for Map #3
Wrote input for Map #4
Wrote input for Map #5
Wrote input for Map #6
Wrote input for Map #7
Wrote input for Map #8
Wrote input for Map #9
Starting Job
13/12/08 11:34:15 INFO mapred.FileInputFormat: Total input paths to process :
13/12/08 11:34:15 INFO mapred.JobClient: Running job: job_201312081126_0002
13/12/08 11:34:16 INFO mapred.JobClient: map 0% reduce 0%
13/12/08 11:34:30 INFO mapred.JobClient: map 28% reduce 0%
13/12/08 11:34:33 INFO mapred.JobClient: map 60% reduce 0%
```

图5 蒙特卡洛求 PI 程序运行日志

每组实验运行 5 次求所需时间的平均值,计算执行时间结果如图 6 所示,纵轴为时间/s,横轴是计算量/次方。从图 6 中可以看出单机环境下的运算时间远远大于分布式系统下的运算时间,而且集群系统中的节点越多计算速度越快。

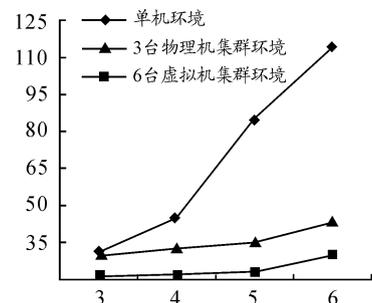


图6 蒙特卡洛程序在 3 种环境中运行的性能对比

2) 实验二:通过运行字符统计程序 (wordcounter. jar) 测试基于 Hadoop 分布式云数据读写的高效性来验证其存储性

能。有4组数据,大小分别为400MB、600MB、1GB和1.5GB。

本组实验设置 Hadoop 块大小为 16M (默认情况下是 64 M),冗余备份参数设置为 3 (默认值),实验环境同实验一,程序运行 5 次,记录时间并计算平均值,运行日志如图 7 所示。

```
(hadoop@master hadoop-1.0.3)$ hadoop jar wordcounter.jar com.hadoop.mapred
anyword.txt /user/hadoop/output20131208
Warning: $HADOOP_HOME is deprecated.

13/12/08 19:11:55 INFO input.FileInputFormat: Total input paths to process:
13/12/08 19:11:55 INFO mapred.JobClient: Running job: job_201312081908_0005
13/12/08 19:11:56 INFO mapred.JobClient: map 0% reduce 0%
13/12/08 19:19:09 INFO mapred.JobClient: map 100% reduce 0%
13/12/08 19:28:21 INFO mapred.JobClient: map 100% reduce 100%
13/12/08 19:28:26 INFO mapred.JobClient: Job complete: job_201312081908_0005
```

图 7 字符统计程序运行日志

运行结果如图 8 所示,纵轴为执行时间/s,横轴为数据量/MB。从图 8 中可以得出单机环境下的数据读写速度明显低于分布式环境下的速度,而且节点越多读写速度越快。

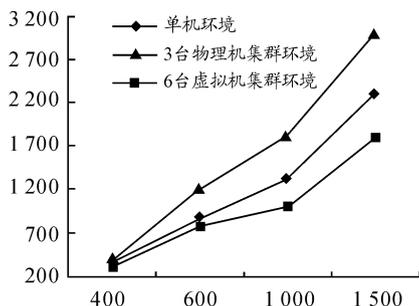


图 8 字符统计程序在 3 种环境中的性能对比

可以看出,与传统数据计算与读写方式相比,本文提出的在虚拟化环境下搭建的基于 Hadoop 分布式技术的云计算与存储平台,有效地提高了海量数据分析与读写的速度和效率;而且利用虚拟化技术搭建的集群比物理机集群效率更高,速度更快,从而大大提高了资源的利用率。

5 结束语

本文通过对 Hadoop 分布式文件系统 HDFS、MapReduce

编程框架进行研究,利用 VMware 虚拟机搭建基于 Hadoop 的云数据计算与存储平台,并通过实验验证其相对于传统数据处理方式具有高效、快速的特点,满足云计算领域的相关需求;而且通过应用虚拟化技术来扩展节点数量,既提高了运行效率又提高了硬件资源的利用率,为今后云计算的研究方向打下了基础。

参考文献:

- [1] 张波. HDFS 下文件存储研究与优化[D]. 广州: 广东工业大学, 2013.
- [2] 崔杰, 李陶深, 兰红星. 基于 Hadoop 的海量数据存储平台设计与实现[J]. 计算机研究与发展, 2012(49): 12-18.
- [3] Konstantin Shvachko, Hairong Kuang, Sanjay Radia, et al. The Hadoop Distributed File System[Z]. Sunnyvale, California USA, IEEE 2010: 1-10.
- [4] Taylor R C. An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics [C]//BMC Bioinformatics. Washington: BMC Bioinformatics, 2010.
- [5] 王宏宇. Hadoop 平台在云计算中的应用[J]. 软件, 2011, 32(4): 36-39.
- [6] 黄晓云. 基于 HDFS 的云存储服务系统研究[D]. 大连: 大连海事大学, 2010.
- [7] 王永洲. 基于 HDFS 的存储技术的研究[D]. 南京: 南京邮电大学, 2013.
- [8] 高蓟超. Hadoop 平台存储策略的研究与优化[D]. 北京: 北京交通大学, 2012.
- [9] Clark C, Fraser K, Hansen JG, et al. Live migration of virtual machines[C]//Proc. Of the 2nd Symp. on Networked Systems Design and Implementation. Berkeley: USENIX Association. [S. l.]: [s. n.], 2005: 273-286.

(责任编辑 杨继森)