

doi:10.3969/j.issn.1673-9833.2013.01.017

面向 Hadoop 的云计算核心技术分析

吴岳忠, 周训志

(湖南工业大学 研究生处, 湖南 株洲 412007)

摘要: 在企业以海量数据处理能力为核心竞争力之一的时代, 云计算技术的出现为此提供了良好的解决方案。首先, 研究分析了云计算技术的开源实现 Hadoop 的核心内容: 分布式系统文件 HDFS 和 MapReduce 算法模型; 然后, 通过实例阐述了面向 Hadoop 的分布式云计算平台的构建, 证实了 Hadoop 是一个更容易开发和并行处理大规模数据的分布式云计算平台; 最后, 对如何实现 Hadoop 的高可用性和高效性进行了展望。

关键词: Hadoop; 云计算; 集群

中图分类号: TP338.8

文献标志码: A

文章编号: 1673-9833(2013)01-0077-04

The Core Technology of Hadoop-Oriented Cloud Computing

Wu Yuezhong, Zhou Xunzhi

(Department of Postgraduates Administration, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: In the era of using mass data processing capacity as one of the core competitiveness, the emergence of cloud computing technology provides a good solution for the enterprise. Firstly, analyzes the core content of the Hadoop as the open source realization of cloud computing technology: Hadoop distributed file system (HDFS) and algorithm model MapReduce. Then, elaborates the construction of Hadoop-oriented distributed cloud computing platform through examples and confirms that the Hadoop as a distributed cloud computing platform is more likely to develop and process large data in parallel. Finally, prospects the implementation of high availability and efficiency of Hadoop.

Keywords: Hadoop; cloud computing; the cluster

在当今这个信息爆炸的时代, 无论是传统企业, 还是互联网企业都越来越重视对所有与企业相关的各种数据的分析和利用, 对海量数据的处理能力已成为现代企业的核心竞争力之一。云计算 (cloud computing) 技术的出现为海量数据的处理提供了良好的解决方案^[1]。云计算是分布式计算 (distributed computing)、并行计算 (parallel computing)、效用计算 (utility computing)、网络存储 (network storage

technologies)、虚拟化 (virtualization)、负载均衡 (load balance) 等传统计算机和网络技术发展融合的产物。它是一种通过网络统一组织和灵活调用各种 ICT (information communication technology, ICT) 信息资源, 实现大规模计算的信息处理方式。它利用分布式计算和虚拟资源管理等技术, 通过网络将分散的 ICT 资源 (包括计算与存储、应用运行平台、软件等) 集中起来形成共享的资源池, 并以动态按需和

收稿日期: 2012-11-05

基金项目: 湖南省科技厅科研基金资助项目 (2011TP4004-1), 湖南省教育厅科研基金资助项目 (11C0398, 12C0071), 湖南工业大学科研基金资助项目 (2011HZX30)

作者简介: 吴岳忠 (1981-), 男, 江苏江阴人, 湖南工业大学讲师, 中南大学博士生, 主要研究方向为云计算, 可信软件和信息安全, E-mail: 5174979@qq.com

可度量的方式向用户提供服务。用户可以使用各种形式的终端（如PC、平板电脑、智能手机甚至智能电视等）通过网络获取ICT资源服务^[2]。而Hadoop是云计算技术的一种开源实现，具有高容错、跨平台等优势，用户可以利用Hadoop轻松地组织计算机资源来搭建自己的分布式云计算平台，从而充分地利用集群的计算和存储能力，完成对海量数据的处理工作。

1 Hadoop的平台概述

Hadoop是Apache软件基金会旗下的一个开源分布式计算平台^[3]。它的高容错性、高扩展性等优点可以让用户在普通廉价硬件设备上搭建成分布式系统，实现对集群的控制和管理，同时其提供了系统底层细节透明的分布式基础架构，允许用户轻松快捷开发并行应用程序，实现海量数据的管理和分布式数据处理。

Hadoop核心内容为：分布式系统文件（Hadoop distributed file system, HDFS）和MapReduce（Google MapReduce的开源实现）算法模型。HDFS是分布式计算中数据存储管理的基础。它所具有的高容错高可靠性、高可扩展性、高获得性、高吞吐率等特征为海量数据提供了不怕故障的存储，为超大数据集的应用带来了许多便利。MapReduce是一种编程模型，主要用于对大规模数据集（通常大于1TB）的并行运算。它借用了函数式编程语言Map（映射）和Reduce（化简）概念，极大地方便软件开发者编写自己的程序并在分布式系统上运行。

Hadoop还包括为子项目提供支持的常用工具Common、用于数据序列化的系统Avro、开源的数据收集系统Chukwa、数据仓库的基础设施Hive、分布式并面向列的开源数据库HBase、对大型数据集进行分析和评估的平台Pig等子项目。图1展现了Hadoop的项目结构图。

Pig	Chukwa	Hive	HBase
MapReduce		HDFS	ZooKeeper
Core		Avro	

图1 Hadoop的项目结构图

Fig. 1 Project structure of Hadoop

2 面向Hadoop的云计算核心技术分析

2.1 分布式文件系统HDFS

HDFS是一个分布式文件系统，开始是开源的Apache项目Nutch的基础结构，最后成为了Hadoop

的基础架构之一。其设计前提和目标^[4]：1) 硬件错误是常态而不是异常；2) 流式数据访问；3) 大规模数据集；4) 简单一致性模型；5) 移动计算比移动数据更划算。

HDFS是一个主从（Master/Slave）结构模型，见图2。一个HDFS集群是由一个NameNode和若干个DataNode组成。其中NameNode作为主服务器，是整个文件系统的管理节点，负责管理文件系统的命名空间，记录文件数据块在每个DataNode上的位置和副本信息，协调客户端对文件的访问，以及记录命名空间内的改动或命名空间本身属性的改动；集群中的DataNode提供真实文件数据的存储服务，管理其所在物理节点上的数据存储，并负责处理客户的读写请求，依照NameNode的命令，执行数据块的创建、复制、删除等工作。从最终用户的角度来看，HDFS就像传统文件系统一样，可以通过目录路径对文件执行CRUD（create, read, update和delete）操作。而根据分布式存储特性，实际上客户端访问文件系统是通过与NameNode和DataNode的交互实现。例如客户端要访问一个文件，首先，客户端会从NameNode中获得组成该文件的数据块位置详细列表，即可知道数据块被存储在哪些DataNode上；然后，客户端就直接从DataNode上读取文件数据。NameNode在整个过程中不参与文件的传输。

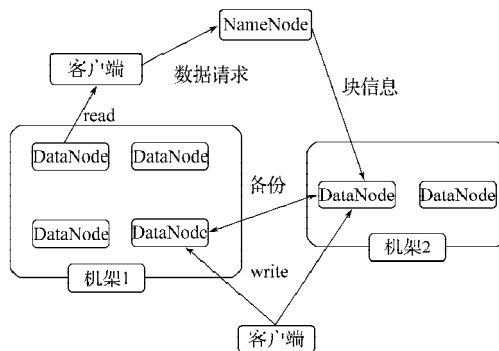


图2 HDFS的体系结构图

Fig. 2 Architecture of HDFS

2.2 MapReduce计算模型

MapReduce是一种处理海量数据的并行编程模型和计算框架。MapReduce框架由一个单独运行在主节点上的Master JobTracker和运行在每个集群从节点上的Slave TaskTracker共同组成。主节点负责调度构成一个作业的分布在从节点上的所有任务。主节点监控它们的执行情况，并且重新执行之前失败的任务；从节点仅负责由主节点指派的任务。当1个Job被提交时，JobTracker接收到提交作业和配置信息后，就会将配置信息等分发给从节点，同时

调度任务并监控 TaskTracker 的执行。MapReduce 编程模型的原理是：利用 1 个输入的 key/value 对集合产生 1 个输出的 key/value 对集合。MapReduce 高度抽象出 2 个函数来表达这个计算：Map 和 Reduce，Map 负责把 1 个任务分解成多个任务，Reduce 负责把分解后的多任务处理结果汇总起来。这 2 个函数的具体功能可以由用户根据需要自己设计实现，只要能够按照用户自定义的规则，将输入的<key, value>对转换成另一个或一批<key, value>对输出即可。图 3 为 MapReduce 大数据集处理过程图。

在 Map 阶段，MapReduce 框架将任务的输入大数据集分割成固定大小的小数据集片段 (splits)，随后将每一个 split 进一步分解成一批键值对<K1, V1>。Hadoop 为每一个 split 创建一个 Map 任务 (以下简称 Mapper) 用于执行用户自定义的 Map 函数，并将对应 split 中的<K1, V1>对作为输入，得到计算的中间结果<K2, V2>。接着将中间结果按照 K2 进行排序，并将 key 值相同的 value 放在一起形成一个新列表，形成<K2, list (V2)>元组。最后再根据 key 值的范围将这些元组进行分组，对应不同的 Reduce 任务 (以下简称 Reducer)。

在 Reduce 阶段，Reducer 把从不同 Mapper 接收来的数据整合在一起并进行排序，然后调用用户自定义的 Reduce 函数，对输入的<K2, list (V2)>对进行相应的处理，得到键值对<K3, V3>并输出到 HDFS 上。在 MapReduce 框架下并程序中会出现 3 个主要函数：Map, Reduce 和 Main，在这个结构中，需要用用户完成的工作仅仅是根据任务编写自定义的 Map 和 Reduce 2 个函数。

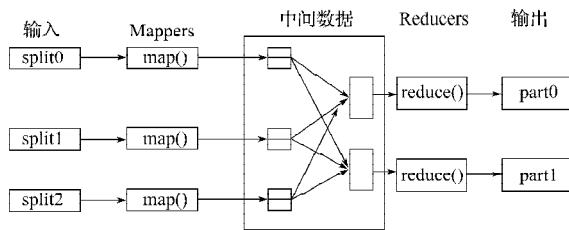


图 3 MapReduce 大数据集处理过程
Fig. 3 Large data set processing of MapReduce

3 面向 Hadoop 的云计算实例分析

3.1 试验环境

3.1.1 环境及工具

根据前面分析的 Hadoop 开源平台，Hadoop 集群分为两大角色类：Master 和 Slave，Master 主要是部署配置 NameNode 和 JobTracker 节点；Slave 是部署配

置 DataNode 和 TaskTracker 节点。在试验中，一共使用 3 台普通 PC 机来搭建 Hadoop 集群，其中 1 台部署为 Master 的 NameNode 和 JobTracker。主机上安装的操作系统为 Ubuntu 10.04 LTS，Java 运行环境为 jdk1.6.0_24，Hadoop 为 Hadoop-0.20.2 版本，程序开发平台为 eclipse-SDK-3.7.2-linux 版本。

3.1.2 环境搭建

1) 基础环境。3 台 PC 机，安装好 Ubuntu10.04 系统、JDK1.6.0_24, openssh-server，设置文件 /etc/hosts, /etc/hostname, /etc/profile，定义主机名和 IP 地址之间的关系及 JDK 环境变量。

2) 实现 Master 和 Slave 主机间的 SSH 无密码登录。3 台主机建立同一个用户，然后在 Master 主机上进行 SSH 密钥生成，并复制给 2 台 Slave，实现无密码登录。

3) Hadoop 安装和配置。在 Master 主机上安装 hadoop-0.20.2，设置文件 conf/Hadoop-env.sh, conf/core-site.xml, conf/hdfs-site.xml, conf/mapred-site.xml, conf/masters 和 conf/slaves，配置环境变量、hdfs 和 mapreduce 常用的 i/o 设置、hadoop 守护进程，然后把整个 Hadoop 安装文件复制到 2 台 Slave 上。

4) Hadoop 运行。先格式化分布式文件系统，然后启动 Hadoop 守护进程。通过 http://master:50030 及 http://master:50070 查看集群状态 (master 是指 NameNode 主机 IP 地址)。

3.2 程序实现

3.2.1 Eclipse 开发平台

在 1 台 Slave 主机上安装 eclipse-SDK-3.7.2-linux，并使用 Hadoop 在 Eclipse 中的插件，配置好 Hadoop Location (见图 4)。这样就可以看到搭建好的 Hadoop 云计算平台的分布式文件系统，并可以进行文件的上传、下载等操作。

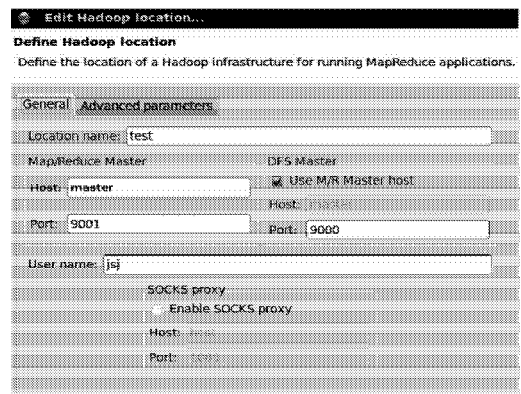


图 4 Hadoop Location 配置

Fig. 4 Configuration of Hadoop Location

本实例中，建立名字为 test 的 Hadoop Location 是

一个完整的分布式文件系统，分 home 和 user 两块。鼠标右击 user 下的 jsj，新建文件夹 in，鼠标右击 in 上传文件 text1.txt（内容为：hello hadoop）和 text2.txt（内容为：hello world）（见图 5）。

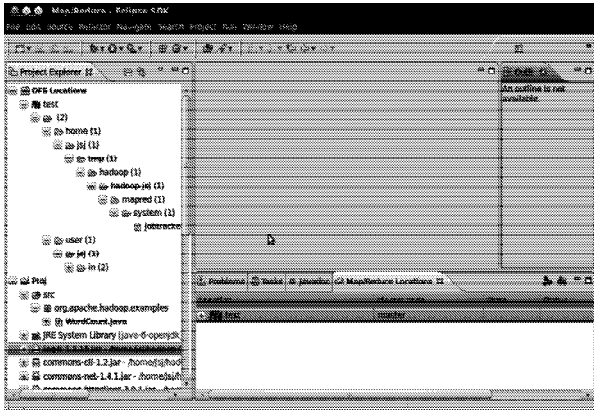


图 5 Hadoop 云计算平台的分布式文件系统
Fig. 5 Distributed file system of Hadoop-Based computing platform

3.2.2 编译实例

创建一个工程 Proj，建立一个 WordCount 类（见图 6），然后进行运行参数设置（见图 7），运行程序，完成 WordCount 实例，计算出单词的数量并排序输出，结果存储在文件夹 out（结果为：1 个 hadoop 1，2 个 hello，1 个 world）（见图 8）。

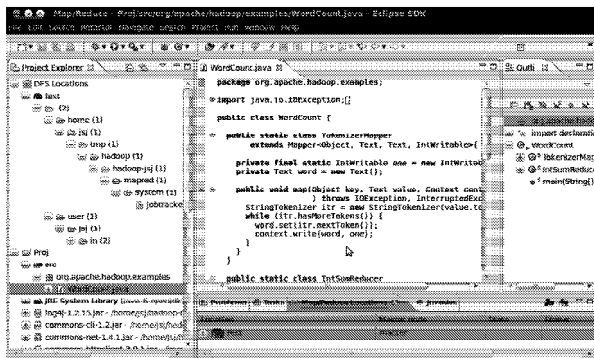


图 6 创建工程和 WordCount 类
Fig. 6 Create project and WordCount class

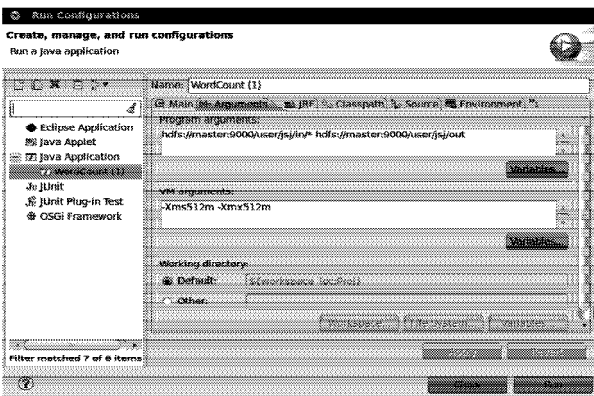


图 7 运行参数设置
Fig. 7 Operation parameters setting

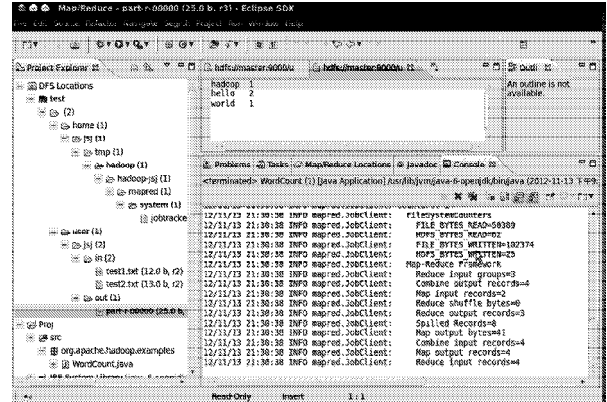


图 8 运行结果
Fig. 8 Operation result

4 结语

本文介绍了云计算的定义和 Hadoop 核心技术，详细分析了 HDFS 和 MapReduce 的工作机制，并阐述了面向 Hadoop 的云计算平台建立过程，实现了面向 Hadoop 的云计算架构应用开发，展现了云计算存储和处理数据的能力。随着云计算业务的日益发展，Hadoop 平台的研究也会越来越深入，这对 Hadoop 的高可用性和高效性挑战也会越来越高，如 NameNode 单点故障、资源管理、作业调度、筛选优化机制等多个方面，这些内容将是下一步的研究工作。

参考文献：

- [1] 陈 全, 邓倩妮. 云计算及其关键技术[J]. 计算机应用, 2009, 29(9): 2562-2567.
Chen Quan, Deng Qianni. Cloud Computing and Its Key Techniques[J]. Journal of Computer Applications, 2009, 29(9): 2562-2567.
- [2] [佚名]. 云计算白皮书: 2012年[R]. 北京: 工业和信息化部电信研究院, 2012: 1-2.
[Anon]. The White Paper of Cloud Computing: 2012[R]. Beijing: China Academy of Telecommunication Research of MIIT, 2012, 1-2.
- [3] 陆嘉恒. Hadoop实战[M]. 北京: 机械工业出版社, 2011: 2-33.
Lu Jiaheng. Hadoop in Action[M]. Beijing: China Machine Press, 2011: 2-33.
- [4] 刘 鹏. 实战 Hadoop: 开启通向云计算的捷径[M]. 北京: 电子工业出版社, 2011: 2-83.
Liu Peng. Practicing Hadoop: The Shortcut for Opening to Cloud Computing[M]. Beijing: Publishing House of Electronics Industry, 2011: 2-83.

(责任编辑: 申 剑)