

Hadoop 学习总结

一、 背景

随着信息时代脚步的加快，各类数据信息越来越多，海量数据的来源列举如下：

纽约证券交易所每天产生 1TB的交易数据。

Facebook 存储着约 100 亿张照片，约 1PB 数据。

Ancestry.com ，一个家谱网站，存储着 2.5PB 数据。

The Internet Archive （互联网档案馆）存储着约 2PB的数据，并以每月至少 20TB 的速度增长。

瑞士日内瓦附近的大型强子对撞机每年产生约 15PB数据。

面对海量数据，如何存储和分析，从中获取有价值信息，变得十分重要。 Hadoop 正是在这样的背景下产生的， 它提供了一个可靠的共享存储和分析系统。 由于具备低成本和前所未有的高扩展性， Hadoop已被公认为是新一代的大数据处理平台， 就像 30 年前的 SQL出现一样， Hadoop 正带来了新一轮的数据革命。

二、 Hadoop相关概念

1、 Hadoop 简述

Hadoop是 Apache 的一个分布式计算开源框架， 它可以运行于大中型集群的廉价硬件设备上， 为应用程序提供了一组稳定可靠的接口。同时它是 Google 集群系统的一个开源项目总称。

底层是 Google 文件系统（ GFS）。

基于 java 语言构建的 Hadoop框架实际上一种分布式处理大数据平台， 其包括软件和众多子项目。在近十年中 Hadoop 已成为大数据革命的中心。

2、 其子项目简述：

MapReduce 分布式数据处理模型和执行环境，运行于大型商用机集群。

HDFS 分布式文件系统，运行于大型商用机集群。

Pig 一种数据流语言和运行环境， 用以检索非常大的数据集。 Pig 运行在 MapReduce和 HDFS 的集群上。

Hive 一个分布式、 按列存储的数据仓库。 Hive 管理 HDFS中存储的数据， 并提供基于 SQL 的查询语言（由运行时引擎翻译成 MapReduce作业）用以查询数据。

ZooKeeper 一个分布式、 可用性高的协调服务。 ZooKeeper 提供分布式锁之类的基本服务用于构建分布式应用。

Sqoop 在数据库和 HDFS之间高效传输数据的工具。

Common 一组分布式文件系统和通用 I/O 的组件与接口 (序列化、 javaRPC 和持久化数据结构)。

Avro 一种支持高效, 跨语言的 RPC以及永久存储数据的序列化系统。

3、Hadoop 的优点

可扩展性 : 不论是存储的可扩展还是计算的可扩展都是 Hadoop设计的根本。

经济 : 框架可以运行在任何普通的 PC上。

可靠 : 分布式文件系统的备份恢复机制以及 MapReduce的任务监控保证了分布式处理的可靠性。

高效 : 分布式分拣系统的高效数据交互实现以及 MapReduce结合 Local Data 处理的模式, 为高效处理海量的信息作了基础准备。

三、Hadoop 的核心

Hadoop 的核心是 HDFS和 MapReduce。HDFS实现存储, 而 MapReduce实现分析处理。HDFS和 MapReduce实现是完全分离的, 并不是没有 HDFS就不能 MapReduce运算。

1、HDFS

1.1HDFS 的基本概念

Hadoop 跟其他云计算项目有共同点和目标: 实现海量数据的计算。而进行海量计算需要一个稳定的, 安全的数据容器, 才有了 Hadoop 分布式文件系统 (HDFS,Hadoop Distributed File System)。HDFS是 GFS的开源实现, 是 Hadoop的底层实现部分, 存储 Hadoop 集群中所有存储节点上的文件。

数据块 (block)

HDFS默认的最基本的存储单位是 64M的数据块。和普通文件系统相同的是, HDFS中的文件是被分成 64M 一块的数据块存储的。不同于普通文件系统的是, HDFS中, 如果一个文件小于一个数据块的大小, 并不占用整个数据块存储空间。

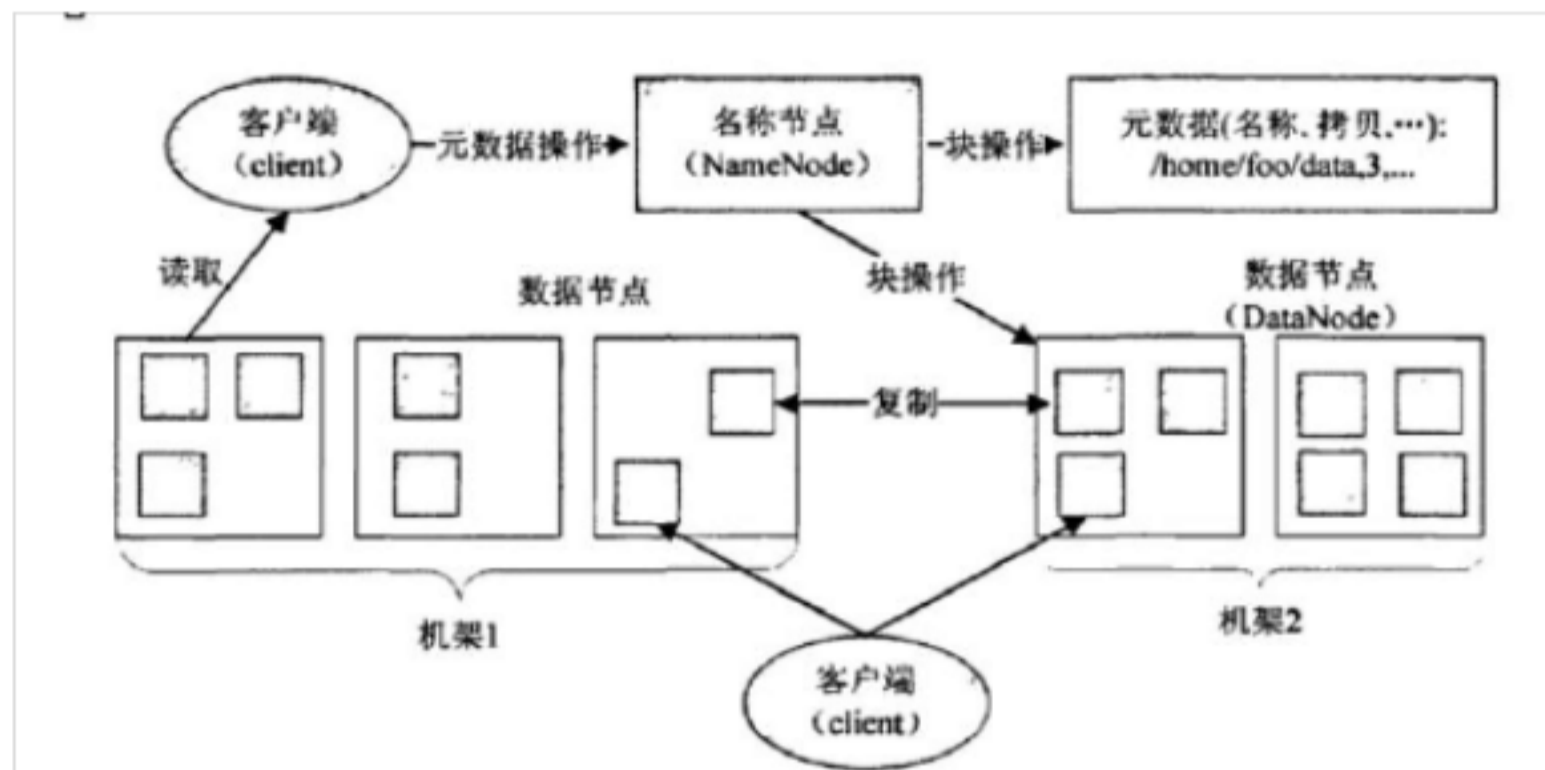
NameNode和 DataNode

NameNode用来管理文件系统的命名空间, 其将所有的文件和文件夹的元数据保存在一个文件系统树中。它还保存了一个文件包括哪些数据块, 分布在哪些 DataNode 上。这些是在系统启动的时候从 DataNode 收集而成的。

DataNode 是文件系统中真正存储数据的地方。

1.2 HDFS 的设计原理

构建一个非常庞大的分布式文件系统，在集群中节点失效是正常的。节点的数量在 Hadoop 中不是固定的。单一的文件命名空间，保证数据的一致性，写入一次多次读取。典型的数据块大小，每一个数据块在多个 DataNode 有复制。客户端通过 NameNode 得到数据块的位置，直接访问 DataNode 获取数据。



HDFS体系结构图

1.3 文件操作

HDFS并不是一个万能的文件系统。它的主要目的是支持以流的形式访问写入的大型文件。

如果客户机想将文件写到 HDFS 上，首先需要将该文件缓存到本地的临时存储。如果缓存的数据大于所需的 HDFS 块大小，创建文件的请求将发送给 NameNode。NameNode 将以 DataNode 标识和目标块响应客户机。同时也通知将要保存文件块副本的 DataNode。当客户机开始将临时文件发送给第一个 DataNode 时，将立即通过管道方式将块内容转发给副本 DataNode。客户机也负责创建保存在相同 HDFS 名称空间中的校验和 (checksum) 文件。在最后的文件块发送之后，NameNode 将文件创建提交到它的持久化元数据存储 (在 EditLog 和 FsImage 文件)。

1.4 HDFS 的适用情况

适用情况：

为以流式数据访问模式存储超大文件而设计的文件系统。

流式数据访问：HDFS 建立的思想是一次写入、多次读取模式是最高效的。

商用硬件：不需要运行在昂贵并且高可靠的硬件上。

不适用情况：

大量的小文件。

低延迟数据访问。

多用户写入，任意修改。

2、MapReduce

2.1 MapReduce 概念

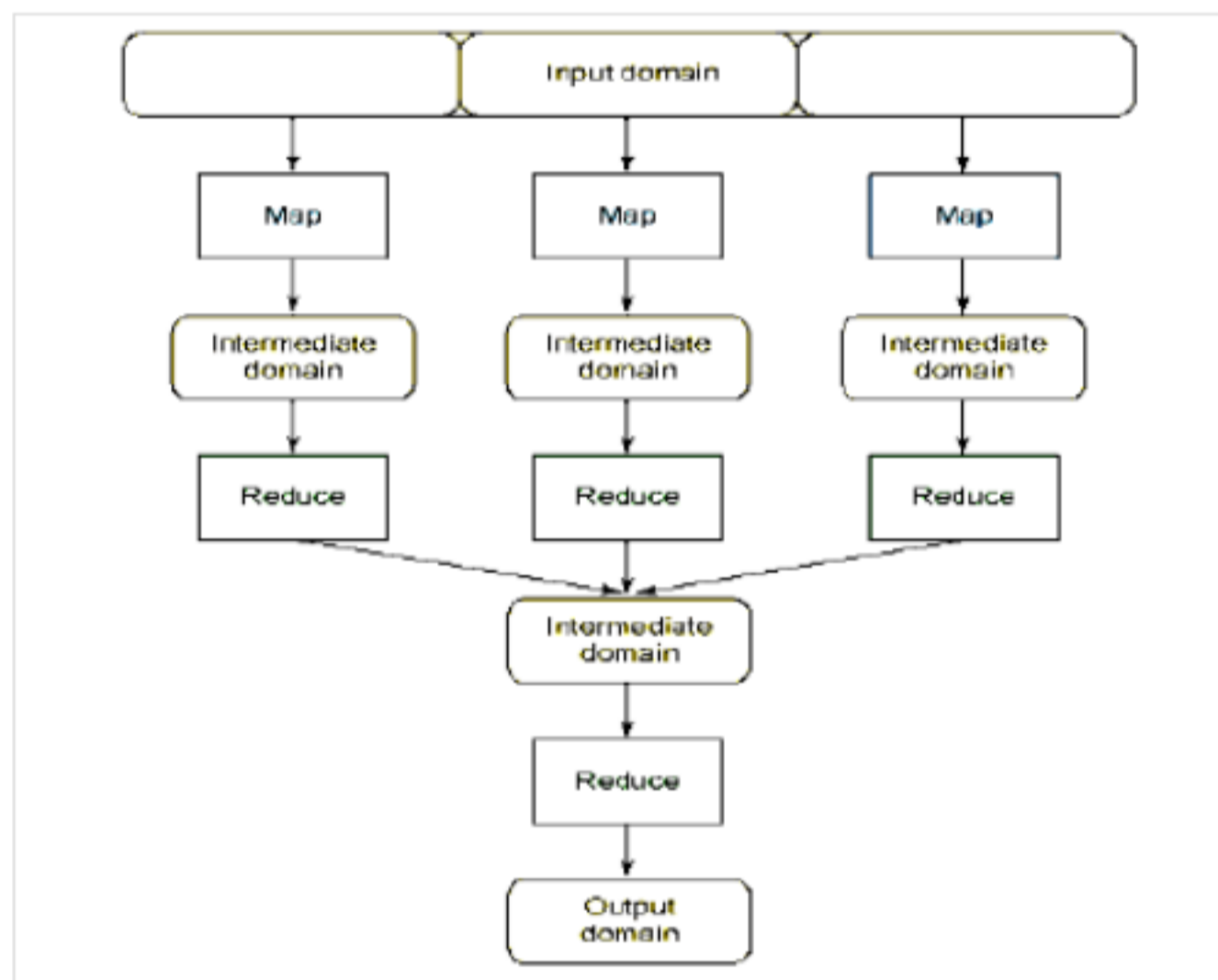
MapReduce是一种编程模型。 Map/Reduce 主要包括两个步骤： Map和 Reduce。 一个 MapReduce 作业通常会把输入的数据集切分为若干独立的数据块，由 map任务以完全并行的方式处理。

框架会对 map的输出进行排序，然后把结果输入给 reduce 任务。 每一步都有 Key-Value 对作为输入和输出：

map阶段的 Key-Value 对的格式是由输入的格式决定的。

map阶段的输出 Key-Value 对的格式必须同 reduce 阶段的输入 key-value 对的合适相对应。

reduce 过程中，将 map过程中的输出作为其输入。

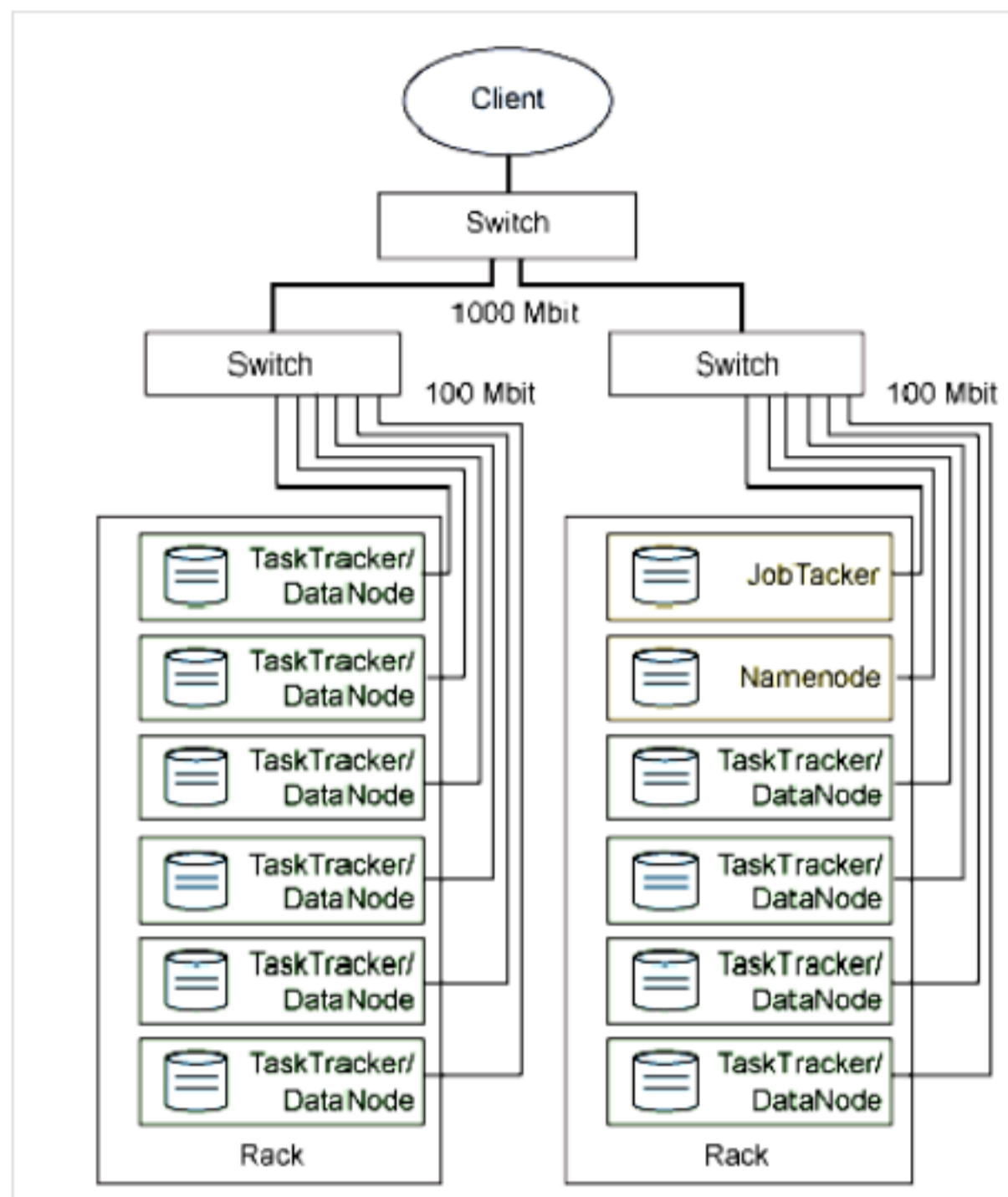


MapReduce数据流

2.2 实现过程

一个代表客户机在单个主系统上启动的 MapReduce 应用程序称为 JobTracker 。类似于 NameNode, 它是 Hadoop 集群中惟一负责控制 MapReduce 应用程序的系统。在应用程序提交之后，将提供包含在 HDFS 中的输入和输出目录。 JobTracker 使用文件块信息（物理量和位路）确定如何创建其他 TaskTracker 从属任务。 MapReduce 应用程序被复制到每个出现输入文件块的节点。将为特定节点上的每个文件块创建一个惟一的从属任务。每个

TaskTracker 将状态和完成信息报告给 JobTracker 。下图显示一个示例集群中的工作分布。



集群中的工作分布

client : 提交 MapReduce作业。

JobTracker : 协调作业的运行。

TaskTracker : 运行作业划分后的任务。

Hadoop 并没有将存储移动到某个位路以供处理，而是将处理移动到存储。这通过根据集群中的节点数调节处理，因此支持高效的数据处理。

2.3MapReduce 与关系数据库管理系统的对比

MapReduce视为关系数据库管理系统的补充。两个系统的 差异 主要表现如下：

MapReduce 比较适合以批处理的方式处理需要分析整个数据集的问题，尤其是即席分析。

RDBMS适用于“点查询”和更新，数据集被索引后，数据库系统能够提供低延迟的数据检索和快速的少量数据的更新。

MapReduce适合一次写入、多次读取数据的应用，而关系型数据库更适合持续更新的数据集。

它们所操作的数据集的结构化程度。 MapReduce 对于非结构化或半结构化数据非常有效，因为在处理数据时才对数据进行解析（ MapReduce 属于的键和值不是数据固有的属性，而是由分析数据的人员选择的）。关系型数据往往是规范的。

MapReduce 是一个线性可伸缩的编程模型。

四、 Hadoop 的应用

在大数据背景下， Apache Hadoop 已经逐渐成为一种标签形象， Hadoop 平台发挥作用的领域是不仅仅是像互联网行业，如 Google 这样的大型互联网搜索引擎，以及 Yahoo 专门的广告分析系统等。 Hadoop 能够在许多地方发挥巨大的作用，比如下列场景：

- 1、在线旅游：目前全球范围内 80% 的在线旅游网站都是在使用 Cloudera 公司提供的 Hadoop 发行版，其中 SearchBI 网站曾经报道过的 Expedia 也在其中。
- 2、移动数据：美国有 70% 的智能手机数据服务背后都是由 Hadoop 来支撑的，也就是说，包括数据的存储以及无线运营商的数据处理等，都是在利用 Hadoop 技术。
- 3、电子商务：eBay 就是最大的实践者之一。国内的电商如淘宝等。
- 4、能源开采：美国 Chevron 公司是全美第二大石油公司，他们利用 Hadoop 进行数据的收集和存储，其中这些数据是海洋的地震数据，以便于他们找到油矿的位路。
- 5、节能：另外一家能源服务商 Opower 也在使用 Hadoop, 为消费者提供节约电费的服务，其中对用户电费单进行了预测分析。
- 6、基础架构管理：这是一个非常基础的应用场景，用户可以用 Hadoop 从服务器、交换机以及其他的设备中收集并分析数据。
- 7、图像处理：创业公司 Skybox Imaging 使用 Hadoop 来存储并处理图片数据，从卫星中拍摄的高清图像中探测地理变化。
- 8、诈骗检测：这个场景用户接触的比较少，一般金融服务或者政府机构会用到。利用 Hadoop 来存储所有的客户交易数据，包括一些非结构化的数据，能够帮助机构发现客户的异常活动，预防欺诈行为。
- 9、IT 安全：除企业 IT 基础机构的管理之外，Hadoop 还可以用来处理机器生成数据以便甄别来自恶意软件或者网络中的攻击
- 10、医疗保健：医疗行业也会用到 Hadoop, 像 IBM 的 Watson 就会使用 Hadoop 集群作为其服务的基础，包括语义分析等高级分析技术等。医疗机构可以利用语义分析为患者提供医护人员，并协助医生更好地为患者进行诊断。

应用思考

- 1、基于 Hadoop 的推荐引擎中使用的算法，如 K-Means、Fuzzy K-Means，Dirichlet，EigenCuts 等，可以用在为微博的好友推荐或感兴趣的话题推荐以及政务网中对各类用户的信息推荐中。
- 2、MapReduce实现的流式计算提高了从海量数据中获取有价值信息的时效性，这在舆情分析中可以有所应用。通过使用 MapReduce网民在互联网中发布的海量数据分析，及时分析出可能存在各种舆论动向，从技术手段方面提高政府应对舆论的能力。
- 3、由于建设集群对硬件设备的要求并不是很高，而通过集群布局分布式计算的实现，可以达到很高的计算效率，所以可以考虑将一些性能不是很高端的机器集中起来，建成一个集群。通过 MapReduce任务分解，设计 Map函数，combine 函数以及 Reduce 函数，处理一些需要解决的高计算量的实际问题。
- 4、HDFS在处理批量、超大数据时有独特的优越性，并具有很高可靠性。金宏办公系统随着时间的推移，数据量不断的增加。微软正在与从雅虎拆分出来的 Hortonworks 合作开发，在已发布的 SQL2012版本中实现了对大数据的支持，这说明 SQLServer 也跨入了 NoSQL 领域。所以将来 SQL版本有可能在目前关系数据库数据存储方式的基础上实现对非结构、半结构化存储方式的兼容。这将为我们目前基于 SQL数据库系统的改造奠定基础。