# 大数据处理

# Hadoop

# Big Data Processing

- 本书以大数据处理系统的三大关键要素——“存储”、“计算”与“容错”为起点，深入浅出地介绍了如何使用Hadoop这一高性能分布式技术完成大数据处理任务。
- 本书不仅包含了使用Hadoop进行大数据处理的实践性知识和示例，还以图文并茂的形式系统性地揭示了Hadoop技术族中关键组件的运行原理和优化手段，为读者进一步提升Hadoop使用技巧和运行效率提供了颇具价值的参考。

编著 刘军

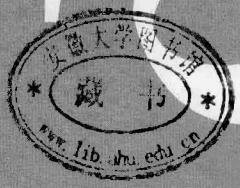
人民邮电出版社
POSTS & TELECOM PRESS

# Hadoop

# 大数据处理

编著 刘军

Big Data Processing




人民邮电出版社
北京

**图书在版编目（CIP）数据**

Hadoop大数据处理 / 刘军编著. -- 北京 : 人民邮电出版社, 2013.9
ISBN 978-7-115-32324-8

Ⅰ. ①H… Ⅱ. ①刘… Ⅲ. ①数据处理软件 Ⅳ. ①TP274

中国版本图书馆CIP数据核字(2013)第133366号

## 内 容 提 要

本书以大数据处理系统的三大关键要素——“存储”、“计算”与“容错”为起点，深入浅出地介绍了如何使用 Hadoop 这一高性能分布式技术完成大数据处理任务。本书不仅包含了使用 Hadoop 进行大数据处理的实践性知识和示例，还以图文并茂的形式系统性地揭示了 Hadoop 技术族中关键组件的运行原理和优化手段，为读者进一步提升 Hadoop 使用技巧和运行效率提供了颇具价值的参考。

本书共 10 章，涉及的主题包括大数据处理概论、基于 Hadoop 的大数据处理框架、MapReduce 计算模式、使用 HDFS 存储大数据、HBase 大数据库、大数据的分析处理、Hadoop 环境下的数据整合、Hadoop 集群的管理与维护、基于 MapReduce 的数据挖掘实践及面向未来的大数据处理技术。最后附有一个在 Windows 环境下搭建 Hadoop 开发及调试环境的参考手册。

本书适合需要使用 Hadoop 处理大数据的程序员、架构师和产品经理作为技术参考和培训资料，也可作为高校研究生和本科生教材。

◆ 编　　著　刘　军
责任编辑　刘　洋
责任印制　彭志环　焦志炜

◆ 人民邮电出版社出版发行　　北京市崇文区夕照寺街 14 号
邮编　100061　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京鑫正大印刷有限公司印刷

◆ 开本：787×1092　1/16
印张：18.75
字数：386 千字　　2013 年 9 月第 1 版
印数：1 – 3 500 册　　2013 年 9 月北京第 1 次印刷

定价：59.00 元

**读者服务热线：(010)67132692　印装质量热线：(010)67129223**
**反盗版热线：(010)67171154**
**广告经营许可证：京崇工商广字第 0021 号**

# 前　言

毫无疑问，Hadoop 已经成为当下大数据处理领域的王者技术。经过开源社区无数贡献者的强大推动，Hadoop 用其显著的低成本、高性能特性，成功地征服了众多具有大数据处理需求的商业机构和科研团体。它的拥趸者中既有 Google、Facebook、Yahoo 这样的知名企业，还有数以万计的中小企业和高校研究团体。这其中也包括了笔者所在的这样长期从事宽带网络领域大数据分析的研究团队。

随着宽带网络技术的飞速发展、用户规模的不断扩大和网络应用的日益丰富，网络中时时刻刻都在产生着蕴含丰富价值的流量数据。如果能对这些海量网络流量数据进行准确而高效的分析，将可以极大地挖掘网络资源潜力、优化网络结构和提高用户体验，为网络运营者和应用开发者带来丰厚的回报。而与此诱人机遇同时到来的，还有由于大数据带来的存储与计算挑战。为了迎接这一挑战，笔者也将多年积累的网络流量分析经验与 Hadoop 技术进行结合，完成了一些典型的海量流量数据分析工作。在此过程中，笔者感觉到目前已有的 Hadoop 相关书籍大多集中在讲授 Hadoop 技术的基础使用方法上。在需要使用 Hadoop 进行一些相对复杂的工作，例如架构设计、技术路线选择、程序调优、算法设计和集群管理等时，只能以大浪淘沙的方式从散落在各处的资料中寻求答案。在此过程中笔者时常期望能有一本系统且深入地展示 Hadoop 关键技术原理，并能与实践相结合的书籍，这也就是笔者创作此书的原动力。

在本书中，笔者以图文并茂的形式，深入浅出地阐述了 Hadoop 中各项关键组件的技术原理和内部结构，并结合实践经验介绍了一些重要的使用技巧和优化方法，分为 10 章分别阐述。第 1 章为大数据处理概论，对大数据处理给出了一个多维度定义，梳理了大数据处理平台的基础架构，介绍了完成大数据处理任务要解决的 3 个关键问题——存储、计算和容错，并归纳性地总结了 Hadoop 技术的关键性思路。第 2 章以 Hadoop 技术的来源，Google 的三大关键技术为引子，介绍了 Hadoop 整体架构、基本原理和发展历程，在此基础上展示了一个使用 Hadoop 技术完成大数据处理工作的简明框架，同时简要介绍了目前 Hadoop 技术在国内知名企业中的应用情况。第 3 章深入剖析了 MapReduce 计算模式，包括原理和工作机制，介绍了实用性的 MapReduce 应用开发方法，并结合简单的实例讲解了几类常用的 MapReduce 设计模式，同时以 3 个经典算法为例讲解了 MapReduce 算法的设计精髓，最后给出了一些重要的 MapReduce 程序优化的方法。第 4 章全面讲解了 HDFS 分布式文件存储系统的工作原理和机制，说明了使用命令行和代码对 HDFS 文件进行操作的方法，介绍了提高文件访问效率的若干重要优化方法，并梳理了目前解决 HDFS 中小文件存储和高可用性问题的解决方法。第 5 章介绍了可用于快速数据检索的 HBase 技术，包括

核心原理与架构，以及管理 HBase 中数据的方法，同时从传统关系型数据库使用者的角度讲解了如何在新型列存储数据库上进行设计与开发的方法，并梳理了提高 HBase 性能的重要方法。第 6 章介绍了 Hadoop 中的高层数据分析工具 Hive 和 Pig，结合具体实例讲解了这两个工具的使用方法，并从实际应用的角度对具有相似性的两者进行了差异对比，给出了选择建议。第 7 章阐述了复杂数据应用环境下的数据整合问题，详细介绍了适用于 Hadoop 与外部 SQL 数据整合的工具 Sqoop，以及 Hadoop 平台内部的数据整合工具 HCatalog。第 8 章介绍了集群管理者所关注的管理与维护体系和相关工具，从配置管理、集群监控、故障处理与安全性等角度梳理了目前常用的 Hadoop 集群管理工具，并进行了总结与对比。第 9 章重点介绍基于 MapReduce 的数据挖掘技术，包括基本原理和方法、若干经典算法的实例，以及目前已有的基于 MapReduce 的数据挖掘工具。最后，在第 10 章介绍了将对未来大数据处理技术产生深远影响的一些新型大数据处理技术，包括 Hadoop 的下一代计算框架 YARN、大数据的实时交互式分析工具 Dremel 和 Impala、大数据的图运算模型和工具 Pregel 和 Hama。

在本书的编写过程中，笔者的导师雷振明教授给予了细心的指导和修改建议，也感谢笔者所在团队成员的大力支持，同时学院的领导与同事也为笔者撰写此书提供了写作环境的支持，人民邮电出版社的同仁也为本书的出版做了大量的工作，在此一并表示感谢。最后，本书的写作也是一项艰巨的任务，没有家人的长期支持，我是不可能在漫漫长夜中完成此书的，感谢她们对我的理解与照顾。

由于 Hadoop 技术发展迅速，加之作者水平有限，书中难免存在写作不到位或错漏之处，敬请读者批评指正。意见建议或交流请发邮件至 liujun@bupt.edu.cn。

**北京邮电大学　刘军**

**2013 年 7 月**

# 目 录

# 第1章

# 大数据处理概论

It was the best of times, it was the worst of times. （这是最好的时代，也是最坏的时代。）

——《A Tale of Two Cities》, Charles Dickens (1812～1870)

大数据（Big Data），也称为海量数据（Massive Data），是随着计算机技术及互联网技术的高速发展而产生的独特数据现象。现代社会正以不可想象的速度产生大数据，手机通信、网站访问、微博留言、视频上传、商品生成、物流运送、科学实验……无处不在的社会和商业活动源源不断地产生各种数据，人们已经进入数据爆炸性增长的全新时代——大数据时代。

“这是最好的时代，也是最坏的时代”，著名作家狄更斯 150 年前在《双城记》中留下的名句，预言般准确地描述了人们今天面临的大数据时代的两面性。一方面，网络和数据库中所记载的各种数据，真实地记录和反映了现实世界中的各类活动信息，这些信息就如巨大的宝藏等待人们去挖掘。如能善加利用，这些数据将如航道中的灯塔，指引现代社会的科研和商业活动，进入下一个黄金时代。而另一方面，不可控的持续爆炸的大数据，如海啸般涌向传统的 IT 世界，这股迅猛的浪潮，挑战着包括数据中心基础设施和数据分析基础架构在内的数据处理的各个环节，稍有不慎，数据的拥有者和使用者将迷失在这大数据中。

幸运的是，计算机技术与互联网技术的发展，在产生大数据的同时，也为人们带来了全新的云计算技术。云计算技术带来的大数据处理能力，使得分析和掌握大数据中蕴藏的无尽信息、知识和智慧成为可能。在本书中，将结合大数据带来的技术挑战和云计算技术带来的全新能力，从数据处理发展进程、云计算算法和架构、云存储与数据仓库等多个方面，全面地介绍以 Hadoop 为代表的云计算技术为大数据处理带来的变革。

本章接下来的内容组织如下：1.1 节给出了大数据的多维度定义；1.2 节阐述了大数据带来的技术挑战；1.3 节介绍了已有的大数据存储技术；1.4 节介绍了针对大数据环境的计算技术；1.5 节对大数据处理平台的容错问题进行了讨论；1.6 节介绍了云计算技术引入后对大数据处理领域的影响。

## 1.1 什么是大数据

在本书中，大数据（Big Data）与海量数据（Massive Data）是同一个术语。目前在学术研究领域和产业界，对大数据并没有一个严格的定义。通常来说，凡是数据量超过一定大小，导致常规软件无法在一个可接受的时间范围内完成对其进行抓取、管理和处理工作的数据即可称为大数据。现实世界中的大数据的实例包括互联网上的网页数据、社交网站上的用户交互数据、物联网中产生的活动数据、电信网络中的话单数据等。

从以上对大数据相对宽泛的描述中可以看出，大数据并不是一个简单的定义即可准确概括的概念。为了能更清楚地了解大数据的各个特征，下面从3个维度对大数据进行分析：数据量大小、数据类型、数据时效性。

**1．数据量大小——大容量**

为了从数据量大小维度进行观察，首先将时间刻度放在一个很小的尺度，以1分钟为单位，看看在爆炸的数据世界中发生了什么。

（1）E-mail：全球所有电子邮件用户发出了2.04亿封电子邮件。

（2）搜索：全球最大的搜索引擎Google处理了200万次搜索请求。

（3）图片：图片分享网站Flickr的用户上传了3 125张新照片，2 000万张照片被浏览。

（4）音频：在Pandora音乐网站上，播放的音乐时长超过61 000小时。

（5）视频：YouTube的用户上传了总计时长48小时的视频，130万个视频被观看。

（6）社交网站：Facebook网站的用户分享了684 478篇文章，超过600万页面被点击。

（7）微博：Twitter的用户发出了10万条微博。

（8）应用：Apple的应用商店完成了4.7万次应用下载。

（9）电子商务：eBay上产生了7万次页面访问，新增了35GB的数据。

（10）通信：在中国产生了时长531万分钟的移动通话，发出了165万条短信[1]。

为了更准确地理解人们现在面临的数据量大小，再来看一组公式：1 024GB=1TB；1 024TB=1PB；1 024PB=1EB；1 024 EB=1ZB；1 024ZB=1YB。在电子商务平台eBay上，每天新增的数据量达到50TB，1年累计的数据量即达到18PB。与之相对地，根据IDC的研究报告[2]，自人类开始记录历史以来，到2006年为止全人类全部的印刷书本文字加起来大约50PB。也就是说，仅eBay平台3年的新增数据，就超过了全人类全部书本的数据量。同时，在社交网站Facebook的计算机集群的磁盘空间中，目前已存储了超过100PB的数据，也就是说，仅Facebook一个网站存储的数据，就已经是人类书本数据量的2倍之多。

与海量的数据同时存在的还有越来越快的数据增长速度。根据IDC的统计，2012年全球产生的数据内容将达到2.7ZB之巨，相比2011年增长48%，这相当于全球70

亿人口每人手持一个 420GB 的硬盘所能容纳的数据总和。而且这样的增长速度，还将随着存储成本的降低和互联网活动的增加而加快。预计到 2015 年，每年产生的数据将达到 8ZB。

也许你会觉得这些海量的数据是那些巨型网站或机构所拥有的数据，离我们太过遥远，那么来看看也许就在你身边发生的例子。以大家都会接触的数码影像为例，我并不是一个狂热的摄影爱好者，但在撰写本书的同时，我将过去几年的数码影像数据简单地进行了整理，如图 1.1 所示。从图中可以看出，随着摄影设备的能力增长，我所产生的数码影像数据大小，以年均 69.4%的增长率逐年增长，并且在每次设备更新换代时，都会产生 100%以上的数据爆发。这些快速增长的个体数据汇集在一起，就形成了网络中的滚滚数据洪流。

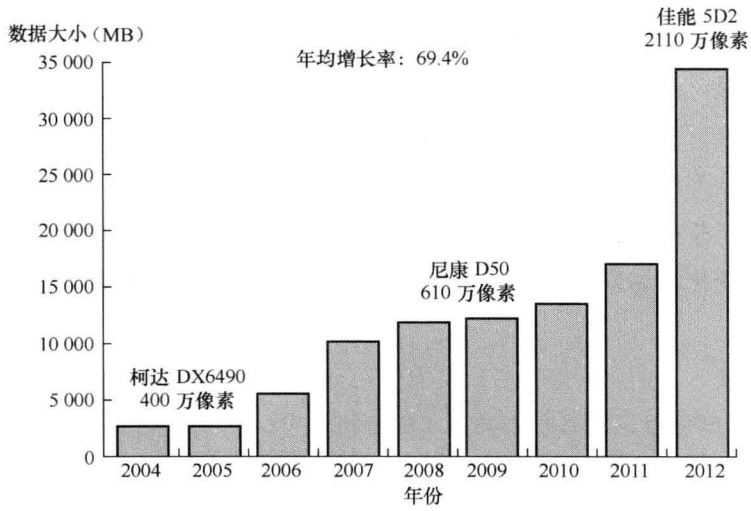



图 1.1 个人影像文件统计图

**2．数据类型——多类型**

仅从字面上理解，大数据似乎仅仅强调的是数据量的大小。然而实际上大数据不仅仅是“大”，还关系到数据类型的改变。

从数据组织形式的角度，可以把数据类型简单地分为两种，即结构化数据（Structured Data）和非结构化数据（Unstructured Data）。结构化数据是可以用二维表结构来逻辑表达实现，并可存储在数据库中的数据，例如银行交易数据、民航航班信息数据等格式严谨的数据库数据。而非结构化数据则是指那些无法通过预先定义的数据模型表述或无法存入关系型数据库表中的数据，例如办公文档、图片、音频和视频等。

在早期出现的大数据场合中，大多数都是银行、民航等数据格式严谨的应用场景，数据基本都是以结构化的表形式存放在数据库中。这些数据通常比较利于处理，在数据量增加的时候，通过采用提升计算或存储节点的处理能力，即可很好地适应数据量的增长。但是随着技术的飞速发展，非结构化数据在整个数据量中所占的比例快速上升。根据 IDC 的统计，在企业数据中，

目前已有超过80%的数据是以非结构化的形式存在的，结构化数据仅占20%不到。而在整个互联网领域，非结构化数据已占到整个数据量比例的75%以上，并且非结构化数据超越结构化数据的速度仍在加速中，现在整个数据领域，非结构化数据的年增长速度大约为63%，远超过结构化数据32%的增长速度。

这个“数字宇宙”中已占统治地位的非结构化数据，正如现实宇宙中的暗物质（暗能量）一样，它们是构成这个宇宙的主要成分，蕴藏了无尽的知识和能量，但却如藏于黑幕中的钻石，不为人们所知。传统数据库处理技术的数据处理方式，已不能适应不遵从二维表格规律展示的非结构化数据，这就要求现代数据处理技术从算法到架构的各个层面进行革新，以应对非结构化数据带来的挑战。

**3．数据时效性——高时效**

在大数据时代，随着数据量的剧增和数据类型的多样化，数据中所蕴藏价值的时效性特征也随之愈加凸显。在传统的数据分析或商业智能（BI）中，数据处理的工作重点更多地放在对历史数据的分析和挖掘。例如以常见的客户关系管理（CRM）或企业资源规划（ERP）应用系统中的数据为例，几乎所有分析报表的产生都是以过去若干周或若干月的数据为基准产生，然后企业管理者根据这样的分析报表对下一步的生产计划进行决策。随着技术更新的周期加快和市场变化的加剧，这种以过长历史数据为依据的决策方式，已经越来越跟不上用户的脚步。

在这样的背景下，企业或组织必须具有实时分析所拥有的最新数据，并提取其中有价值的信息的能力，以产生对未来具有指导意义的分析结果。例如搜索引擎需要将几分钟前上线的新闻归并到检索索引中，如果一个搜索引擎不能及时建立搜索结果，用户必将流失到时效性更高的其他搜索引擎去；电子商务网站必须在当天分析用户的购买行为并预测第二天的货物短缺状况，如果不能达到这样的处理速度，第二天的缺货状况必将引来不可估量的用户流失和收入损失；地质管理机构必须在地震发生后的几分钟内发布海啸或其他次生灾害的预警，如果不能及时发布，无数宝贵的生命将很快面临巨大的威胁。

然而，在面对大数据的大容量、多类型天然特性时，尤其是处理PB级数据及以非结构化为主的数据时，要满足这样的高时效性变得尤为困难。在稍纵即逝的市场机会和变幻莫测的大自然面前，大数据的高时效性犹如皇冠上那颗最炫耀夺目的宝石，吸引了无数从业者的目光。

这里将大数据的3大特征：大容量、多类型和高时效归纳为图1.2，以便于理解和掌握。在本书的后面章节将会看到，大数据的这3大特征给数据处理技术带来了极大挑战，带来了一场数据处理技术的革命，并由此开创了一个大数据处理的新时代。

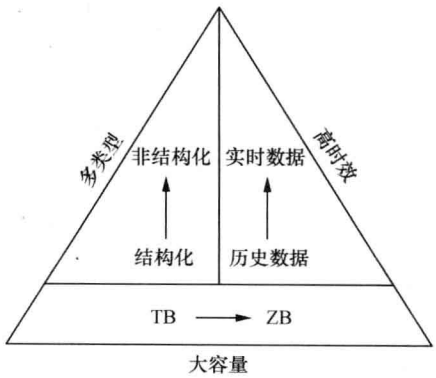



图1.2 大数据特征图

## 1.2 数据处理平台的基础架构

自人类诞生以来，数据处理就一直伴随在人们左右。在经历了笔纸人工处理（公元前 4000 年～公元 1900 年）、机器打孔卡处理（1900～1955 年）、电子存储计算器（1955～1970 年）、在线数据库（1965～1980 年）、关系型数据库（1980～1995 年）、多类型数据处理（1995～2010 年）6 个阶段后[4]，人们正式迈入了大数据处理阶段。自关系型数据库阶段起，可以称之为现代数据处理，其基础技术组件可归纳为如图 1.3 所示的结构，其中包含 6 个主要的基本能力组件（Capability Component）：数据集成、文件存储、数据存储、数据计算、数据分析、平台管理。下面结合现代数据处理基础架构逐一解析大数据带来的挑战。

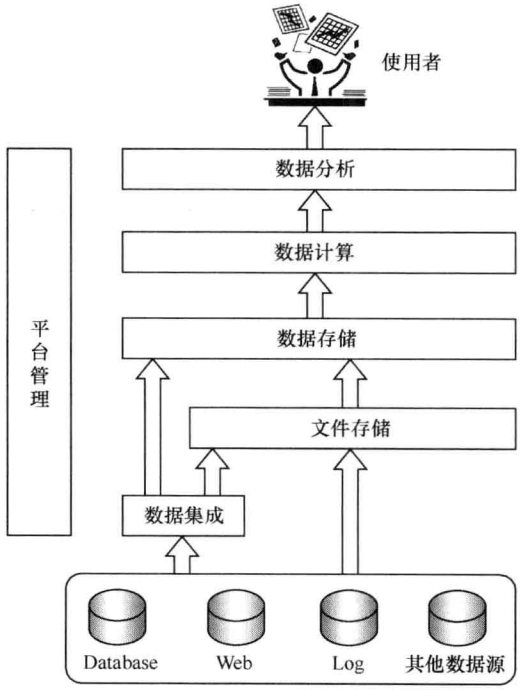



图 1.3 现代数据处理基本能力组件结构图

### 1. 文件存储

文件存储能力组件是数据处理架构中进行数据管理的基本单元，其功能是提供快速、可靠的文件访问能力，以满足大数据量计算的需求，例如 Linux 的 EXT2[5]就是一个文件存储能力组件的实例。大数据的数据量由 TB 级别迈向 PB 甚至 ZB 的级别，同时还存在类型多样的数据文件，以著名的 Facebook 网站为例，其文件系统需要存储文字、图片、视频等各种类型的数据，其中仅图片一类就超过 600 亿张，且每天增加 3 亿张以上。在这样的应用环境下，文件存储能力组件面临着以下方面的挑战：如何以较低成本适应待存储文件的快速增

长及可靠备份的需求；如何实现高效率的大规模小文件及大文件并存的管理机制；如何同时支持结构化数据、半结构化数据及非结构化数据的存储；在海量文件数量的情况下如何实现高效率的名字空间、属性数据、元数据的管理。

**2．数据存储**

数据存储能力组件是数据处理架构中进行数据管理的高级单元，其功能是存储按照特定的数据模型组织起来的数据集合，并提供独立于应用的数据增加、删除、修改能力，例如IBM的DB2[6]就是一个数据存储能力组件。在面对大数据的大数据量、多类型、高时效性的需求时，数据存储能力组件主要需解决以下问题：如何以较低成本适应海量存储数据的快速增长；如何同时支持结构化数据、半结构化数据及非结构化数据的存储；数据组织形式如何适应业务模型和分析维度的多样化；如何实现大并发量大数据量环境下的高效数据读写操作；如何在大规模计算机机器的环境下保证数据的可靠性；如何在大数据的环境下很好地支持复杂查询。

**3．数据集成**

数据集成能力组件的功能是将不同来源、不同格式、不同性质的数据在逻辑上或物理上进行有机集中，以支持多样数据源与文件存储和数据库存储能力组件间的数据输入和输出，例如Oracle的Oracle Data Integrator[7]产品就是一个数据集成能力组件。由于大数据处理技术发展较晚，因此在实施过程中不可避免地需要与多个异构的或其他历史遗留的信息系统同时运行并互通，这就要求数据集成组件很好地解决以下问题以支持大数据处理的平滑运行：如何实现大数据的快速、可靠的导入和导出；如何在异构环境下确保数据流动的安全性；如何实现一个灵活的数据模型和访问模式以适应不同数据系统间的差异性；如何确保足够的可扩展性以适应数据量的增长及数据类型的多样化。

**4．数据计算**

数据计算能力组件是整个架构中的核心组件，其功能是利用处理平台的计算资源解决特定的数据计算问题，例如并行计算领域常用的MPI（Message Passing Interface）[8]即是一个典型的数据计算组件。在大数据处理环境下，数据计算的核心问题是如何将一个需要巨大计算能力才能解决的问题分解为若干小的问题，并将它们分配到指定的计算资源进行处理，并将分解处理的计算结果进行归并综合，生成最终的计算结果。在解决此核心问题的同时，数据计算能力需要满足以下几方面的需求：满足大数据环境下的高时效性要求；低成本、可扩展地适应数据的快速增长；在复杂的运行环境下实现稳定、可靠的计算过程；提供一个统一的、灵活的编程模型适应复杂的大数据处理应用。

**5．数据分析**

数据分析能力组件是数据处理架构中与使用者最近的模块，其功能是提供易用的操作方

式，以支持用户从复杂的数据中提取与科学或商业目的相关的信息或内在规律，例如 SQL（Structured Query Language）[9]作为一个数据分析能力组件就提供了很好的关系型数据库分析手段。数据分析的主要目的是为用户屏蔽数据处理平台底层较为复杂的技术细节和调度逻辑，而将经过抽象化的数据访问和分析手段提供给用户使用。通过数据分析组件的协助，用户可以通过友好的交互界面在高层数据结构上进行数据处理工作，而不需要考虑数据的存取方式、数据流流向、文件存储位置等底层细节。大数据处理环境下的数据分析能力实现的关键就在于如何将抽象的数据分析方式与并行化的数据计算能力相结合，并为用户提供适应大数据、多种类的数据分析手段。

**6．平台管理**

平台管理是整个数据处理架构的管理组件，其功能是保障数据处理平台的安全稳定运行。大数据处理平台，通常规模庞大并由众多服务器构成，并且这些服务器往往还分布在不同的地点，平台上运行着数以百计的应用。在这种情况下，如何有效地管理这些服务器，保证整个系统提供不间断的服务是巨大的挑战。平台管理组件的目标就是使大量的服务器协同工作，方便地进行业务部署和开通，快速发现和恢复系统故障，通过自动化、智能化的手段实现大数据处理的可靠运营。在面对大数据处理的计算机集群时，平台管理主要面临如下调整：如何对高度虚拟化的底层 IT 资源进行动态管理；如何在问题发生时在数量众多的硬件、软件和服务组件中快速定位故障；如何以简洁高效的可视化形式展现整个平台的运行状况；如何从纷繁复杂的日志和运行数据中分析出潜在的性能瓶颈并给出优化方案。

上面介绍的 6 个基本组件在大数据环境下所面临的全新挑战虽然不尽相同，但从这些问题的本质出发，可以归纳为 3 类问题，即大数据存储、高性能计算和系统容错性。下面分别从这 3 个方面对云计算出现前已有的主要成果进行简要介绍。

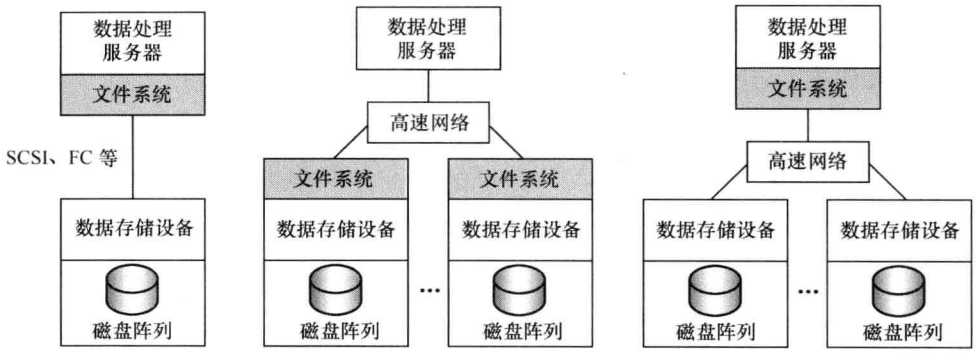
## 1.3 大数据处理的存储

数据存储是数据处理工作的基石。大数据增长带来的不仅仅是存储容量的压力，还给数据管理、存储性能带来了挑战。为了应对大数据对存储系统的挑战，数据存储领域的工作者通过不懈的努力提升了数据存储系统的能力。数据存储系统能力的提升主要体现在两方面，一方面是提升系统的存储容量；另一方面是提升系统的吞吐量。

### 1.3.1 提升容量

提升系统存储容量有两种方式，一种是提升单硬盘的容量，通过使用新的材质和新的读写技术，单个硬盘的容量已经从 MB、GB 跨入了 TB 时代。在这里主要关注在多硬盘的环境下如何提升系统的整体存储容量。经过多年的发展，系统存储技术已经由早期的直连式存储（Direct-Attached Storage，DAS），发展出网络接入存储（Network-Attached Storage，

NAS）和存储区域网络（Storage Area Network，SAN），如图 1.4 所示，并在近几年进入到云存储阶段。




（a）DAS　　（b）NAS　　（c）SAN

图 1.4　各种存储技术的结构

### 1. 直连式存储（DAS）

直连式存储（DAS）是最早出现的最直接的扩展数据存储模式，即将数据存储设备与数据使用设备（服务器或工作站）直接相连的模式。在这种模式下，数据存储设备和数据使用设备之间没有任何存储网络相连。DAS 比较典型的应用场景就是通过一个包含大量数据存储能力的设备（如磁盘阵列）与一个数据使用设备（如数据处理服务器）通过数据传输接口相连，常用的数据传输接口包括 SCSI 和 Fibre Channel（FC）。DAS 的基本架构如图 1.4（a）所示。

DAS 结构在早期数据量不是非常大，且应用场景比较简单的时候，发挥了主要作用。但是随着数据量的快速增长，数据处理的应用场景更加复杂等一系列变化，DAS 结构在多个方面表现出了严重的不足。

（1）扩展性差、成本高。由于数据使用设备与数据存储设备直接相连，在出现新的数据应用时，需要为新的数据使用设备增加单独的数据存储设备，导致投资成本加大，且随着数据量的增大，数据使用设备和数据存储设备间的传输通道很容易成为性能瓶颈。

（2）资源利用率低。用于不同数据处理服务器间的数据存储存在孤岛效应，在数据量分布出现不均衡时，数据存储能力不能实现共享，导致一些设备存储能力不足而另一些设备却有大量空间空闲。

（3）可管理性差。由于 DAS 结构下数据存储设备是相互独立的，这导致了管理功能分散、效率低下。

（4）备份、恢复和扩容过程复杂。由于数据使用设备与数据存储设备之间相连，在进行备份与恢复时，会占用正常的数据处理传输通道，这使得备份与恢复不能实时进行，必须在系统闲时执行，带来了较大风险，而在进行扩容时往往需要进行停机维护，对业务影响较大。