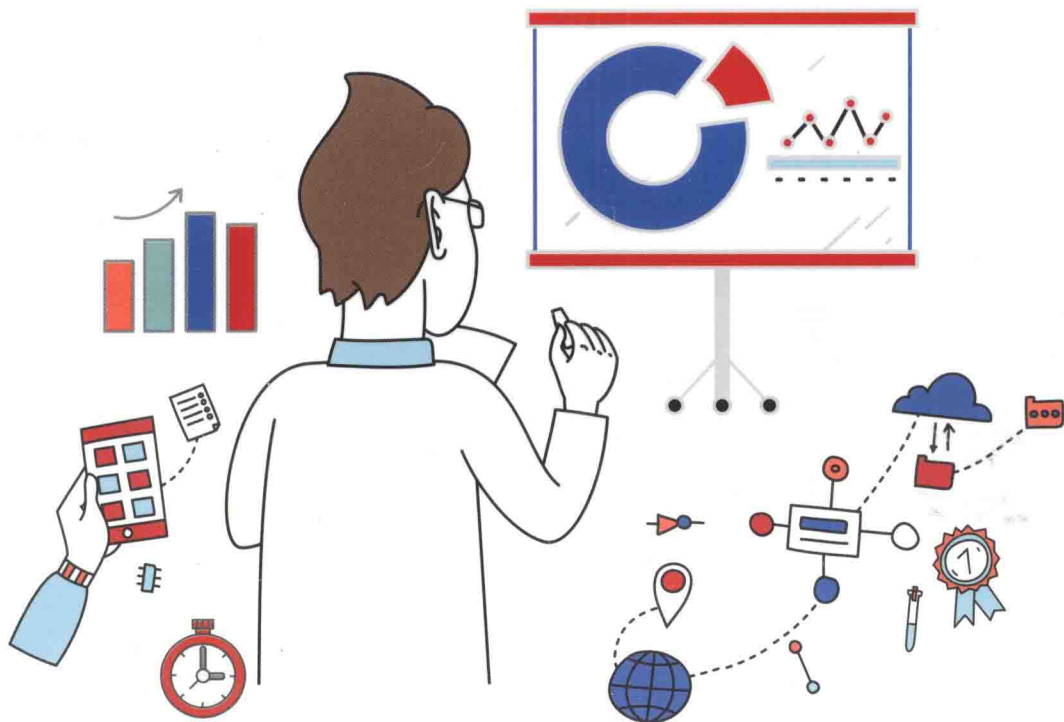


大数据的本质

大数据现状及挑战 驱动因素 未来趋势 Spark原理及应用

探针 爬虫 日志采集 Flink 深度学习 数据分发中间件

Broadview®  
www.broadview.com.cn



# 大数据架构详解

## 从数据获取到深度学习

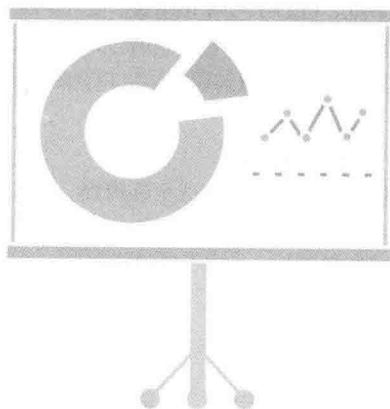
朱洁 罗华霖 / 编著



中国工信出版集团



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
http://www.phei.com.cn



# 大数据架构详解

## 从数据获取到深度学习

朱洁 罗华霖 / 编著

电子工业出版社  
Publishing House of Electronics Industry  
北京·BEIJING

## 内 容 简 介

本书从架构、业务、技术三个维度深入浅出地介绍了大数据处理领域端到端的知识。主要内容包括三部分：第一部分从数据的产生、采集、计算、存储、消费端到端的角度介绍大数据技术的起源、发展、关键技术点和未来趋势，结合生动的业界最新产品，以及学术界最新的研究方向和成果，让深奥的技术浅显易懂；第二部分从业务和技术角度介绍实际案例，让读者理解大数据的用途及技术的本质；第三部分介绍大数据技术不是孤立的，讲解如何与前沿的云技术、深度学习、机器学习等相结合。

本书内容深入浅出，技术结合实践，从实践中理解架构和技术的本质，适合大数据技术领域的从业人员如架构师、工程师、产品经理等，以及准备学习相关领域知识的学生和老师阅读。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。  
版权所有，侵权必究。

### 图书在版编目（CIP）数据

大数据架构详解：从数据获取到深度学习 / 朱洁，罗华霖编著. —北京：电子工业出版社，2016.10  
ISBN 978-7-121-30000-4

I. ①大… II. ①朱… ②罗… III. ①数据处理—研究 IV. ①TP274

中国版本图书馆 CIP 数据核字（2016）第 233622 号

策划编辑：安 娜

责任编辑：徐津平

特约编辑：赵树刚

印 刷：北京中新伟业印刷有限公司

装 订：北京中新伟业印刷有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱

邮编：100036

开 本：787×980 1/16 印张：23.25 字数：566 千字

版 次：2016 年 10 月第 1 版

印 次：2016 年 10 月第 1 次印刷

印 数：3000 册 定价：69.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 [zllts@phei.com.cn](mailto:zllts@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

本书咨询联系方式：010-51260888-819 [faq@phei.com.cn](mailto:faq@phei.com.cn)。

# 前 言

大数据这几年真的很火，于是有越来越多的人开始学习大数据技术。很多人会误以为大数据是一门技术，其实不然，大数据更多的是一门市场宣传语言，也可以理解为一种思考方式。从技术角度来看，大数据是一系列技术的组合，所以真正全面掌握大数据技术也是一件很困难的事情。编写这本书的初衷就是总结这些年的工作和学习经验，希望可以分享给更多人，同时对自己而言也是一个提高、总结和升华的过程。

总的来说，本书围绕一个通用技术栈来组织章节，主要聚焦大数据平台的一些知识。主要分为三部分。

第一部分：第 1~3 章，主要讲述大数据的本质、运营商大数据的架构和一些基本的业务知识。

- 第 1 章：阐述大数据的本质和面临的挑战。
- 第 2 章：概述大数据架构及背后的驱动因素，以及未来发展的趋势。
- 第 3 章：介绍运营商领域的业务，让读者对大数据能做什么有一个直观的感受。

第二部分：第 4~11 章，围绕大数据平台技术栈来阐述数据获取、处理、分析和应用平台涉及的技术。

- 第 4 章：介绍数据获取涉及的探针、爬虫、日志采集、数据分发中间件等技术。
- 第 5 章：介绍流式数据处理引擎、CEP、流式应用。
- 第 6 章：介绍交互式分析技术、MPP DB、热门的 SQL on Hadoop 技术。
- 第 7 章：介绍批处理技术、Spark，以及大规模机器学习的 BSP 技术等。
- 第 8 章：探讨机器学习、深度学习相关技术。
- 第 9 章：统一资源管理是趋势，本章介绍资源管理的核心技术和算法。
- 第 10 章：存储是基础，本章介绍存储的关键技术。
- 第 11 章：探讨大数据技术怎么云化，以及关键技术是什么。

第三部分：第 12 章，技术和文化息息相关，技术影响文化，文化影响技术。

第 12 章：介绍大数据开发文化、开源、DevOps，探讨理念和文化对技术的冲击。

由于编者水平有限，书中疏漏之处在所难免，敬请谅解。

最后以乔布斯的经典名句结尾：Stay hungry, Stay foolish。

朱 洁  
2016 年 5 月于深圳

# 目 录

## 第一部分 大数据的本质

第 1 章 大数据是什么.....	2
1.1 大数据导论.....	2
1.1.1 大数据简史.....	2
1.1.2 大数据现状.....	3
1.1.3 大数据与 BI.....	3
1.2 企业数据资产.....	4
1.3 大数据挑战.....	5
1.3.1 成本挑战.....	6
1.3.2 实时性挑战.....	6
1.3.3 安全挑战.....	6
1.4 小结.....	6
第 2 章 运营商大数据架构.....	7
2.1 架构驱动的因素.....	7
2.2 大数据平台架构.....	7
2.3 平台发展趋势.....	8
2.4 小结.....	8
第 3 章 运营商大数据业务.....	9
3.1 运营商常见的大数据业务.....	9
3.1.1 SQM (运维质量管理).....	9
3.1.2 CSE (客户体验提升).....	9
3.1.3 MSS (市场运维支撑).....	10
3.1.4 DMP (数据管理平台).....	10
3.2 小结.....	11

## 第二部分 大数据技术

第 4 章 数据获取.....	14
4.1 数据分类.....	14
4.2 数据获取组件.....	14
4.3 探针.....	15
4.3.1 探针原理.....	15
4.3.2 探针的关键能力.....	16
4.4 网页采集.....	26
4.4.1 网络爬虫.....	26
4.4.2 简单爬虫 Python 代码示例.....	32
4.5 日志收集.....	33
4.5.1 Flume.....	33
4.5.2 其他日志收集组件.....	47
4.6 数据分发中间件.....	47
4.6.1 数据分发中间件的作用.....	47
4.6.2 Kafka 架构和原理.....	47
4.7 小结.....	82
第 5 章 流处理.....	83
5.1 算子.....	83
5.2 流的概念.....	83
5.3 流的应用场景.....	84
5.3.1 金融领域.....	84
5.3.2 电信领域.....	85
5.4 业界两种典型的流引擎.....	85
5.4.1 Storm.....	85
5.4.2 Spark Streaming.....	89
5.4.3 融合框架.....	102
5.5 CEP.....	108
5.5.1 CEP 是什么.....	108
5.5.2 CEP 的架构.....	109
5.5.3 Esper.....	110

5.6	实时结合机器学习 .....	110
5.6.1	Eagle 的特点 .....	111
5.6.2	Eagle 概览 .....	111
5.7	小结 .....	116
第 6 章	交互式分析 .....	117
6.1	交互式分析的概念 .....	117
6.2	MPP DB 技术 .....	118
6.2.1	MPP 的概念 .....	118
6.2.2	典型的 MPP 数据库 .....	121
6.2.3	MPP DB 调优实战 .....	131
6.2.4	MPP DB 适用场景 .....	162
6.3	SQL on Hadoop .....	163
6.3.1	Hive .....	163
6.3.2	Phoenix .....	165
6.3.3	Impala .....	166
6.4	大数据仓库 .....	167
6.4.1	数据仓库的概念 .....	167
6.4.2	OLTP/OLAP 对比 .....	168
6.4.3	大数据场景下的同与不同 .....	168
6.4.4	查询引擎 .....	169
6.4.5	存储引擎 .....	170
6.5	小结 .....	171
第 7 章	批处理技术 .....	172
7.1	批处理技术的概念 .....	172
7.2	MPP DB 技术 .....	172
7.3	MapReduce 编程框架 .....	173
7.3.1	MapReduce 起源 .....	173
7.3.2	MapReduce 原理 .....	173
7.3.3	Shuffle .....	174
7.3.4	性能差的主要原因 .....	177
7.4	Spark 架构和原理 .....	177
7.4.1	Spark 的起源和特点 .....	177

7.4.2 Spark 的核心概念.....	178
7.5 BSP 框架.....	217
7.5.1 什么是 BSP 模型.....	217
7.5.2 并行模型介绍.....	218
7.5.3 BSP 模型基本原理.....	220
7.5.4 BSP 模型的特点.....	222
7.5.5 BSP 模型的评价.....	222
7.5.6 BSP 与 MapReduce 对比.....	222
7.5.7 BSP 模型的实现.....	223
7.5.8 Apache Hama 简介.....	223
7.6 批处理关键技术.....	227
7.6.1 CodeGen.....	227
7.6.2 CPU 亲和技术.....	228
7.7 小结.....	229
<b>第 8 章 机器学习和数据挖掘.....</b>	<b>230</b>
8.1 机器学习和数据挖掘的联系与区别.....	230
8.2 典型的数据挖掘和机器学习过程.....	231
8.3 机器学习概览.....	232
8.3.1 学习方式.....	232
8.3.2 算法类似性.....	233
8.4 机器学习&数据挖掘应用案例.....	235
8.4.1 尿布和啤酒的故事.....	235
8.4.2 决策树用于电信领域故障快速定位.....	236
8.4.3 图像识别领域.....	236
8.4.4 自然语言识别.....	238
8.5 交互式分析.....	239
8.6 深度学习.....	240
8.6.1 深度学习概述.....	240
8.6.2 机器学习的背景.....	241
8.6.3 人脑视觉机理.....	242
8.6.4 关于特征.....	244
8.6.5 需要有多少个特征.....	245
8.6.6 深度学习的基本思想.....	246



8.6.7	浅层学习和深度学习 .....	246
8.6.8	深度学习与神经网络 .....	247
8.6.9	深度学习的训练过程 .....	248
8.6.10	深度学习的框架 .....	248
8.6.11	深度学习与 GPU .....	255
8.6.12	深度学习小结与展望 .....	256
8.7	小结 .....	257
第 9 章	资源管理 .....	258
9.1	资源管理的基本概念 .....	258
9.1.1	资源调度的目标和价值 .....	258
9.1.2	资源调度的使用限制及难点 .....	258
9.2	Hadoop 领域的资源调度框架 .....	259
9.2.1	YARN .....	259
9.2.2	Borg .....	260
9.2.3	Omega .....	262
9.2.4	本节小结 .....	263
9.3	资源分配算法 .....	263
9.3.1	算法的作用 .....	263
9.3.2	几种调度算法分析 .....	263
9.4	数据中心统一资源调度 .....	271
9.4.1	Mesos+Marathon 架构和原理 .....	271
9.4.2	Mesos+Marathon 小结 .....	283
9.5	多租户技术 .....	284
9.5.1	多租户概念 .....	284
9.5.2	多租户方案 .....	284
9.6	基于应用描述的智能调度 .....	287
9.7	Apache Mesos 架构和原理 .....	288
9.7.1	Apache Mesos 背景 .....	288
9.7.2	Apache Mesos 总体架构 .....	288
9.7.3	Apache Mesos 工作原理 .....	290
9.7.4	Apache Mesos 关键技术 .....	295
9.7.5	Mesos 与 YARN 比较 .....	304
9.8	小结 .....	305

第 10 章 存储是基础 .....	306
10.1 分久必合，合久必分 .....	306
10.2 存储硬件的发展 .....	306
10.2.1 机械硬盘的工作原理 .....	306
10.2.2 SSD 的原理 .....	307
10.2.3 3DXPoint .....	309
10.2.4 硬件发展小结 .....	309
10.3 存储关键指标 .....	309
10.4 RAID 技术 .....	309
10.5 存储接口 .....	310
10.5.1 文件接口 .....	311
10.5.2 裸设备 .....	311
10.5.3 对象接口 .....	312
10.5.4 块接口 .....	316
10.5.5 融合是趋势 .....	328
10.6 存储加速技术 .....	328
10.6.1 数据组织技术 .....	328
10.6.2 缓存技术 .....	335
10.7 小结 .....	336
第 11 章 大数据云化 .....	337
11.1 云计算定义 .....	337
11.2 应用上云 .....	337
11.2.1 Cloud Native 概念 .....	338
11.2.2 微服务架构 .....	338
11.2.3 Docker 配合微服务架构 .....	342
11.2.4 应用上云小结 .....	348
11.3 大数据上云 .....	348
11.3.1 大数据云服务的两种模式 .....	348
11.3.2 集群模式 AWSEMR .....	349
11.3.3 服务模式 Azure Data Lake Analytics .....	352
11.4 小结 .....	354

## 第三部分 大数据文化

第 12 章 大数据技术开发文化.....	356
12.1 开源文化.....	356
12.2 DevOps 理念.....	356
12.2.1 Development 和 Operations 的组合 .....	357
12.2.2 对应用程序发布的影响 .....	357
12.2.3 遇到的问题 .....	358
12.2.4 协调人 .....	358
12.2.5 成功的关键 .....	359
12.3 速度远比你想的重要.....	359
12.4 小结.....	361

第一部分  
**大数据的本质**

# 第 1 章

## 大数据是什么

---

首先提一个问题：“大数据”是一项专门的技术吗？有的人可能会以为大数据是一项专门的技术，其实不是。“大数据”这三个字只是一门市场语言（Marketing Language），其背后是硬件、数据库、操作系统、Hadoop 等一系列技术的综合应用，所以本书我们希望从一个端到端的架构展开讲解典型的大数据技术。

### 1.1 大数据导论

#### 1.1.1 大数据简史<sup>①</sup>

大数据（Big Data）现在可以说是人尽皆知，其实真正回溯起来，其是由 SGI 的首席科学家 John R. Masey 于 1998 年在 USENIX 大会上首次提出的。他在其发表的一篇名为 *Big Data and the Next Wave of Infrastrress*<sup>②</sup> 的论文中首次提出这个词，用来描述数据爆炸的现象。估计他当时未必能想到十几年后 Big Data 能这么火。

如果追溯大数据的概念，则是阿尔文·托夫勒（Alvin Toffler）于 1980 年在《第三次浪潮》一书中预言了信息时代的到来会带来数据爆发，所以科学家很早就预见到了大数据。大数据的历史由来已久，但是技术需要持续积累，才能由量变到质变。

对于工业界来说，不得不提 Google 在 2003—2006 年公布的关于 GFS、MapReduce 和 BigTable 的三篇技术论文，正是这三篇论文奠定了大数据发展的基石。Hadoop 之父——Doug Cutting<sup>③</sup>正是参考论文，后来才实现了当前鼎鼎大名的 Hadoop，而 Hadoop 的诞生极大地促进了大数据技术的蓬勃发展。

当然，这里特别要指出，Hadoop 并不等同于大数据，大数据也并不特指 Hadoop，大数据只是一门市场语言，代表的是一种理念、一种问题解决思路、一系列技术的集合，Hadoop 只是其中一种具体的处理数据的框架技术。

---

① 参考 forbes <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/>。

② *Big Data and the Next Wave of InfraStress* (<https://www.usenix.org/conference/1999-usenix-annual-technical-conference/big-data-and-next-wave-infrastrress-problems>)。

③ Doug Cutting 是 Lucene、Nutch、Hadoop 等项目的发起人。

## 1.1.2 大数据现状

Gartner 发布的 2016 技术成熟度曲线 (见图 1.1) 首次将云计算、大数据及相关技术移除。Gartner 指出这些技术不是不重要, 而是不再“新兴”, 大家虽然对大数据的兴趣依然不减, 但是这个市场已经安定下来, 有了一整套合理的方法, 新的技术和实践被添加进现有方案。所以大数据度过了技术的期望膨胀高峰期, 到了真正使用大数据解决问题的时候。未来大数据相关技术的演进在很长一段时间仍将展现出强大的生命力, 相关市场的营收也将不断放大。

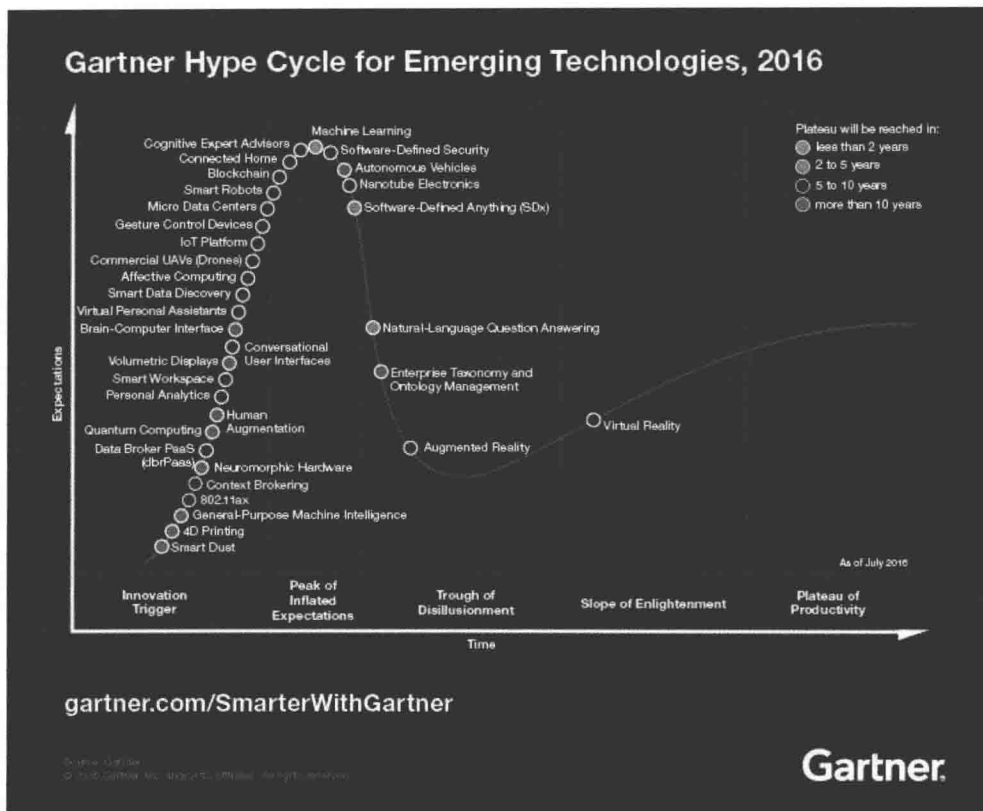


图 1.1

## 1.1.3 大数据与 BI<sup>①</sup>

前面说了大数据是一种理念、一种问题解决思路和一系列技术的集合, 它与传统的 BI 既有相同之处, 也有不同之处。

相同之处, 都是从数据中挖掘价值, 促进商业成功。不同之处, 核心是分布式技术的发展、处理能力的极大提高, 以前想都不敢想的处理变成了可能。所以在对数据的处理理念上也得到了扩展:

<sup>①</sup> Business Intelligence, 商业智能。

(1) 不局限于传统的 BI 从数据中抽样建模，再回 DW<sup>①</sup>实施，大数据可以直接从全量数据中找出规律，通过数据的样本多样化弥补模型的准确性。

(2) 不局限于传统的 BI 简单地通过汇总、统计分析找出群体共性从而输出报表，大数据可以直接通过足够多的数据对个体进行刻画。

虽然有种种不同，但未来大数据和 BI 的界限会模糊，企业的核心驱动目标是从数据资产中找出商业价值，而不关心构建和分析的方法论。

## 1.2 企业数据资产

有了大数据的光环，有了从数据中挖掘商业价值的方法和工具之后，那些原本存放在服务器上平淡无奇的陈年旧数一夜之间身价倍增。按照世界经济论坛报告的看法，“大数据为新财富，价值堪比石油”。《大数据时代》一书的作者维克托则乐观地预测，数据列入企业资产负债表只是时间问题。

本质上，任何企业在生产活动中都会产生数据，数据都有分析的价值。我们来看看典型的运营商会产生哪些数据。

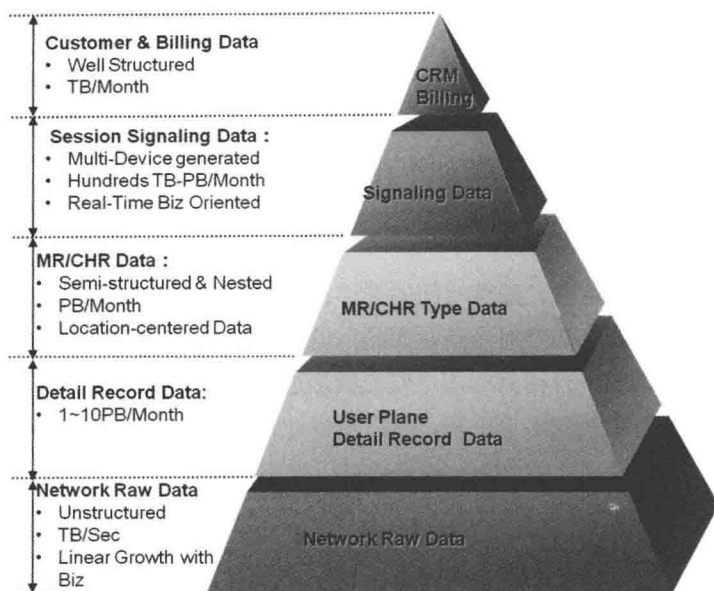
图 1.2 是典型的运营商产生的数据，从下到上分为如下几类。

- **Network Raw Data:** 电信网络里任何一个呼叫或者上网行为都会引起电信设备之间的数据进行交换，这个数据就是网络底数据。
- **User Plane Detail Record Data:** 从网络原始数据里面提取出来的用户行为数据，如打电话数据、上网行为等。
- **MR/CHR Type Data:** 无线测量数据、呼叫历史记录单据数据。用户的位置信息就是从 MR/CHR 数据里面通过算法得出的。
- **Signaling Data:** 信令数据。电信网络分控制面、数据面、用户面。在控制面上设备之间相互按照协议协商通信的数据叫信令数据。
- **CRM Billing:** 电信设备厂商登记的开户信息、账单信息。

上面的划分是从非常专业化的角度进行的，其实通俗一点可以简单归类为设备数据和用户数据两类。设备数据用来分析设备的正常与否，用于设备的维护、规划等。用户数据包括如下数据。

- **位置数据:** 无线是用一个个蜂窝来划分区域的，一个蜂窝叫一个小区，所以只要电话在线，就需要注册到一个个小区中去，知道了小区就知道了用户所在的位置。通过小区切换就能计算出用户移动的轨迹，这就是用户的位置数据。
- **上网数据:** 用户通过运营商的设备上网，所有的行为数据都可以被记录下来，如上了什么网站、网速是多少、上了多长时间。这些通过通信协议的包头就可以获取。如果继续分析内容，就可以获取更多的数据，就可以完全知道用户在干什么。

<sup>①</sup> Data Warehouse，数据仓库。



Note : In a typical network of ~30M subscribers and the data flow around 1TB/Sec.  
MR/CHR: Measurement Report/Call History Data generated during calls

图 1.2<sup>①</sup>

- 用户兴趣数据：通过用户的上网记录，就可以衍生出用户的兴趣爱好、常上什么网站、最近关注什么东西等。
- 通信数据和社交数据：例如，用户给谁打了电话、打了多长时间、给谁发了短信，这些信息都可以被记录下来。通过电话联系又可以衍生出用户的社交关系数据，如和谁是熟人、常联系谁等。
- 身份信息数据：用户到运营商开户，用户的所有个人信息就被运营商记录下来，包括姓名、年龄、身份证号码等。而且这些数据是由人工采集的，准确度远高于互联网用户自己注册的信息。
- 用户金融数据：如用户的电话、网络缴费记录，是否经常欠费等，可用于进行个人信用分析。

由此可以看出，运营商拥有从底层的设备数据到上层用户的行为数据，而且通常是全网的数据，因此可以说抱着数据的“金矿”，是其他厂商所无法媲美的。

## 1.3 大数据挑战

大数据发展到现在，有了一定的技术和商业积累，但是还有很多难题等待解答，最典型的的就是成本、实时性、安全等方面的挑战。

<sup>①</sup> 图片来源：<https://databricks.com/blog/2015/06/09/huawei-embraces-open-source-apache-spark.html>。



### 1.3.1 成本挑战

运营商普遍受到腾讯、阿里巴巴等互联网厂商的 OTT 应用的挤压，面临着管道转型、利润下降的风险。而运营商的数据量巨大，以 PB 为基本单位的数据，处理起来需要巨大的投入。外部商业环境和内部规模的双重挤压，对大数据平台提出了很高的性能和成本要求。

### 1.3.2 实时性挑战

如果从广义的数据质量角度看，随着时间的推移，数据的价值将逐渐降低，时间越久的数据，价值越低。举个例子，一家商场需要对当前在商场内的客户做一个推荐活动。但是端到端采集和处理数据的时间过长，最后推荐平台得到的用户列表都是过期的列表，列表上的名单可能已经不在商场内，而新到的用户还没有更新到名单中来，所以很多业务对大数据平台端到端的实时性提出了很高的要求。

### 1.3.3 安全挑战

安全挑战体现在两个方面：

一方面是在技术上，随着 HTTPS 的推广应用，数据在传输过程中采用管理加密的方式，运营商作为管道获取数据的难度变得越来越大。

另一方面是在法理上，用户的哪些数据是可以获取的、哪些是不允许读取的，始终存在侵犯用户隐私的法律风险。

## 1.4 小结

本章简述大数据的历史，探讨大数据的作用，作为全书的引子，让读者对大数据有一个初步的认识，带领读者进入丰富多彩的大数据世界。

后面的章节将重点讲述在构建运营商大数据系统时的业务和架构实战经验，以实际经验以飨读者。