

BIG DATA

大数据日知录

架构与算法

——张俊林 著——

全面梳理大数据相关技术，从数据、算法、策略、应用和系统架构等多个维度进行剖析，既包罗万象，又深入浅出。

《大数据》《信息检索导论》译者**王斌** 机器学习专家**张栋** 新浪微博平台及大数据总经理**刘子正**
盛大文学首席数据官**陈运文** CSDN/《程序员》创始人**蒋涛**

——联袂力荐——

大数据丛书

十二五国家重点图书出版规划项目

大数据日知录

架构与算法

张俊林 著

电子工业出版社

Publishing House of Electronics Industry

内 容 简 介

大数据是当前最为流行的热点概念之一,其已由技术名词衍生到对很多行业产生颠覆性影响的社会现象,作为最明确的技术发展趋势之一,基于大数据的各种新型产品必将会对每个人的日常生活产生日益重要的影响。

本书从架构与算法角度全面梳理了大数据存储与处理的相关技术。大数据技术具有涉及的知识点异常众多且正处于快速演进发展过程中等特点,其技术点包括底层的硬件体系结构、相关的基础理论、大规模数据存储系统、分布式架构设计、各种不同应用场景下的差异化系统设计思路、机器学习与数据挖掘并行算法以及层出不穷的新架构、新系统等。本书对众多纷繁芜杂的相关技术文献和系统进行了择优汰劣并系统性地对相关知识分门别类地进行整理和介绍,将大数据相关技术分为大数据基础理论、大数据系统体系结构、大数据存储,以及包含批处理、流式计算、交互式数据分析、图数据库、并行机器学习的架构与算法以及增量计算等技术分支在内的大数据处理等几个大的方向。通过这种体系化的知识梳理与讲解,相信对于读者整体和系统地了解、吸收和掌握相关的优秀技术有极大的帮助与促进作用。

本书的读者对象包括对 NoSQL 系统及大数据处理感兴趣的所有技术人员,以及有志于投身到大数据处理方向从事架构师、算法工程师、数据科学家等相关职业的在校本科生及研究生。

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

大数据日知录:架构与算法/张俊林著. —北京:电子工业出版社,2014.9
(大数据丛书)

ISBN 978-7-121-24153-6

I. ①大… II. ①张… III. ①数据处理—IV. TP294

中国版本图书馆CIP数据核字(2014)第194446号



策划编辑:付睿

责任编辑:李利健

特约编辑:顾慧芳

印刷:北京京师印务有限公司

装订:北京京师印务有限公司

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:787×1092 1/16 印张:25.25 字数:587千字

版次:2014年9月第1版

印次:2014年10月第2次印刷

印数:3001~6000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。

质量投诉请发邮件至 zllts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

前 言

像移动互联网、O2O、可穿戴设备等概念一样，“大数据”从甫一提出到飓风般席卷并风靡全球，从最初的技术名词到形成渗透各行各业的社会现象，所耗时间仅几年而已，其兴也勃焉。

那么，大数据是否会像很多曾经火热现在已难觅踪迹的流行概念一样，将来某日，人们静心抬眼，发现风已去而水波不兴，徒留夕阳下波光粼粼的涟漪，让人不禁哀叹其亡也忽焉？

本书的背景

目前看仿佛有此迹象，当一个概念火爆到从街头随便抓一个路人，他都能跟你滔滔不绝地侃侃而谈；当一个新名词铺天盖地而来，让你熟悉到再见一次就要吐的程度，这些确实是典型的泡沫将破的征兆。目前业已出现越来越多的质疑声音，在这种狂热的氛围下，理性的质疑是最难能可贵的，毕竟大数据归根结底是少数派的游戏，而现状好像是人人都处于大数据进行时的状态，这让其看上去显得不甚理性。

但是，从社会发展趋势的角度，很明显大数据会是目前肉眼可及的视野范围里能看到的最大趋势之一。从传统 IT 业到互联网、互联网到移动互联网，从以智能手机和 Pad 为主要终端载体的移动互联网到可穿戴设备的移动互联网，然后再到万物互联的物联网，这一定是不可违抗的发展规律和前进方向。伴随着这个趋势必然有越来越多、形态越来越丰富的超量数据不断产生，而大数据明显是由此衍生出来的明确且必然的发展趋势。

所以，归根结底，大数据概念是个短期内炒得过热，但是从长期来看炒得不足的领域。再过十年如果回望现在，也许会发现：我们今天正在巍峨群山的山脚徘徊，试图找出一条通往山顶的羊肠小道，如此而已。我们当然不能盲目跟风不断追逐切换热点，但是忽视趋势的力量同样也不是一个理性的选择。

本书的内容

以上所述是本书诞生的大背景，目前市面上陆陆续续已经有不少讲大数据的书，有面向大众的概念普及类图书，也有讲解大数据技术的书，本书属于第二类，专注于与大数据处理有关的架构与算法，我相信这是一本比较全面地分门别类梳理大数据技术的书籍。从大约 2010 年年底起我开始关注并收集整理这方面的技术资料，当然那时还没有听说过大数据的概念，现在的所谓大数据已经是后来的事情了。最初引起我关注的是与 NoSQL 相关的技术，尤其是 Google 和亚马逊的一系列相关工作，当时隐约地觉得这是一种新的技术发展趋势，甚至是一次技术范型的大转换，所以逐步开始投入越来越多的精力到其中，这包括将相关技术资料分门别类地进行收集、阅读和整理，利用业余时间开始一章一节地慢慢写本书，以及在实际工作中尽可能地应用这些技术和系统等若干方面。

本书的写作

断断续续完成本书花了大约 3 年左右的时间，和最初预估的时间还是比较吻合的，这一方面是由于我可投入的时间本身就不多，但是还是希望能够写出一本高质量的技术书籍，所以不得不慢工出细活；另外一方面是由于大数据处理作为一个新领域，其涉及的方方面面的技术点实在太多，而且正处于快速的发展过程中。这个领域是我接触过的领域中知识涵盖面最广的，从底层的硬件开始，到涉及基础理论、大规模数据存储系统、分布式架构设计、各种不同适用场景下的差异化系统设计思路、机器学习与数据挖掘并行算法、层出不穷的新架构和新系统等，说无所不包有些夸张但是所需掌握知识点之多确实是很少见的。而且因为其处于快速发展的过程当中，所以各种技术纷繁芜杂，并无一个成熟的知识分类体系可供参考，需要不断梳理相关知识点之间的相互区别和联系并进行分门别类，如何将纷杂多样的技术梳理成清晰合理的章节内容曾让我头疼不已，当然发展到目前，整个大数据技术体系脉络已经日渐明晰，我相信参考本书目录可以清晰地发现这一点。另外一个阻碍是可参考的资料和系统多而杂，且质量良莠不齐，需要从中汰劣余优，尽可能选出有代表性及有发展潜力的理论、方案和系统，这也是很耗费精力的过程，尽管本书每章后只列出了很少一部分参考资料，但是实际参考的文献与系统要数倍于此，只列出精华部分是为了缩短读者选择高质量文献的过程。

十几年前当我还在中国科学院读书的时候，就对互联网的蓬勃发展及其对生活和工作各方面的影响感慨不已。当时最直观的感受是最新的国际会议论文很容易获得，往往是会议一开完就能从网上下载到感兴趣的文献，有时候会议没开有些作者就把论文先放在网上供人参考。觉得有了互联网这么强大的全球范围信息便捷分享工具，尽管当时国内科研水平不算高，在最好的国际顶级会议发表论文还是较难的事情，但是随着互联网的普及，科研水准应该能够获得极为快速长足的发展，因

为从追踪最新技术进展的角度看大家的起跑线是一样的，而中国人多的优势慢慢应该能够发挥出来。事实上也是如此，最近几年各种国际顶级会议中，中国人发表的论文比例越来越高可以证明这一点，而很明显这个趋势还会进一步加快。之所以提这个，是因为大数据相关技术研发道理其实也是一样的，尽管目前国内在这方面的实力和国外相比还有很大的差距，优秀的系统和技术方案往往都是 Google、亚马逊、Facebook、Linkedin 等这些国际知名互联网公司提出的，国内工业界的技术水准大部分还仅仅停留在能把开源的大数据系统应用起来解决手头碰到的问题这个阶段，但是我相信在不远的将来，国内会逐步涌现出具有国际水准的大数据系统与解决方案，其中的道理与上面所举的学术进步的例子是一样的。目前大部分优秀系统是开源的，相关技术文献也很容易找到，作为有进取心的技术人员，现在所缺乏的不是没有可参考的学习资料，相反是资料太多但良莠不齐反而让很多人无所适从，不知该如何下手。国内的技术人员只要肯下功夫、会下功夫，有好的职业发展环境和高的自我期许及技术理想，假以时日，越来越多的世界级水准的大数据处理系统出自中国人之手是完全可以指日可待的。我期望本书对于这些技术人员在全面了解吸收并掌握大数据处理的优秀技术过程中，能贡献绵薄之力。沈利也参与了本书的部分编写工作。

致谢

感谢我的妻子、岳父、岳母，以及爸爸、妈妈，尽管历时 3 年才得以完成本书，但是若是没有你们的全身心支持，这本书 2020 年也无法面世。

对了，还有我的女儿，亲爱的雪晴，从蹒跚迈出人生第一步，到咿咿呀呀说出第一句话，我总是想起，小小的你，迷茫地站在这遍布新奇的世界里，东张西望时无助的样子，这总是令人心疼。你无数人生的第一次带给我太多欢乐，而我陪你的时间又太少太少。每次听到你用小手咚咚敲着书房紧闭的门，这经常让我感到惶惑，不知道这么狠心对你是对还是错，但是又不得不狠下心来。我相信将来一定会后悔没有给你更多的时间，把你介绍给这色彩斑斓而残酷的世界，嘱咐它在今后的日子里能善待你一些。我多想陪你直到你白发苍苍的岁月：在初秋的午后阳光下，看到你坐在院中的婆娑树影里，给你的儿孙讲故事，浅笑依然，仿似今天。我可以清楚地想象那时的场景，清晰得就像一切就发生在眼前。虽然我深切地知道这是永不可实现的奢望，但这个梦想会永远放在我心里，就像你清澈的目光和纯真的笑容一样。

时间是我们的朋友，也是我们的敌人，希望你们能够原谅我。

张俊林

2014 年 3 月于北京

目 录

第 0 章 当谈论大数据时我们在谈什么	1
0.1 大数据是什么	2
0.2 大数据之翼：技术范型转换	4
0.3 大数据商业炼金术	6
0.4 “大数据”在路上	7

第一篇 大数据基础理论

第 1 章 数据分片与路由	9
1.1 抽象模型	10
1.2 哈希分片 (Hash Partition)	11
1.2.1 Round Robin	11
1.2.2 虚拟桶 (Virtual Buckets)	12
1.2.3 一致性哈希 (Consistent Hashing)	13
1.3 范围分片 (Range Partition)	18
参考文献	19
第 2 章 数据复制与一致性	20
2.1 基本原则与设计理念	21
2.1.1 原教旨 CAP 主义	21
2.1.2 CAP 重装上阵 (CAP Reloaded)	23
2.1.3 ACID 原则	24
2.1.4 BASE 原则	24
2.1.5 CAP/ACID/BASE 三者的关系	25
2.1.6 幂等性 (Idempotent)	26
2.2 一致性模型分类	26

2.2.1	强一致性	27
2.2.2	最终一致性	28
2.2.3	因果一致性	28
2.2.4	“读你所写”一致性	29
2.2.5	会话一致性	29
2.2.6	单调读一致性	30
2.2.7	单调写一致性	30
2.3	副本更新策略	30
2.3.1	同时更新	30
2.3.2	主从式更新	31
2.3.3	任意节点更新	32
2.4	一致性协议	32
2.4.1	两阶段提交协议 (Two-Phrase Commit, 2PC)	33
2.4.2	向量时钟 (Vector Clock)	38
2.4.3	RWN 协议	40
2.4.4	Paxos 协议	42
2.4.5	Raft 协议	45
	参考文献	49

第 3 章 大数据常用的算法与数据结构 51

3.1	布隆过滤器 (Bloom Filter)	51
3.1.1	基本原理	52
3.1.2	误判率及相关计算	52
3.1.3	改进：计数 Bloom Filter	53
3.1.4	应用	54
3.2	SkipList	55
3.3	LSM 树	58
3.4	Merkle 哈希树 (Merkle Hash Tree)	62
3.4.1	Merkle 树基本原理	62
3.4.2	Dynamo 中的应用	63
3.4.3	比特币中的应用	63
3.5	Snappy 与 LZSS 算法	65
3.5.1	LZSS 算法	65
3.5.2	Snappy	67
3.6	Cuckoo 哈希 (Cuckoo Hashing)	67
3.6.1	基本原理	68

3.6.2 应用: SILT 存储系统	68
参考文献	70

第二篇 大数据系统体系结构

第 4 章 集群资源管理与调度	71
4.1 资源管理抽象模型	72
4.1.1 概念模型	72
4.1.2 通用架构	73
4.2 调度系统设计的基本问题	74
4.2.1 资源异质性与工作负载异质性	74
4.2.2 数据局部性 (Data Locality)	75
4.2.3 抢占式调度与非抢占式调度	75
4.2.4 资源分配粒度 (Allocation Granularity)	76
4.2.5 饿死 (Starvation) 与死锁 (Dead Lock) 问题	76
4.2.6 资源隔离方法	77
4.3 资源管理与调度系统范型	77
4.3.1 集中式调度器 (Monolithic Scheduler)	78
4.3.2 两级调度器 (Two-Level Scheduler)	79
4.3.3 状态共享调度器 (Shared-State Scheduler)	79
4.4 资源调度策略	81
4.4.1 FIFO 调度策略	81
4.4.2 公平调度器 (Fair Scheduler)	81
4.4.3 能力调度器 (Capacity Scheduler)	82
4.4.4 延迟调度策略 (Delay Scheduling)	82
4.4.5 主资源公平调度策略 (Dominant Resource Fair Scheduling)	82
4.5 Mesos	84
4.6 YARN	87
参考文献	90
第 5 章 分布式协调系统	91
5.1 Chubby 锁服务	92
5.1.1 系统架构	93
5.1.2 数据模型	94
5.1.3 会话与 KeepAlive 机制	95
5.1.4 客户端缓存	95

5.2	ZooKeeper	96
5.2.1	体系结构	96
5.2.2	数据模型 (Data Model)	97
5.2.3	API	98
5.2.4	ZooKeeper 的典型应用场景	98
5.2.5	ZooKeeper 的实际应用	103
	参考文献	104
第 6 章	分布式通信	106
6.1	序列化与远程过程调用框架	107
6.1.1	Protocol Buffer 与 Thrift	108
6.1.2	Avro	109
6.2	消息队列	110
6.2.1	常见的消息队列系统	110
6.2.2	Kafka	111
6.3	应用层多播通信 (Application-Level Multi-Broadcast)	114
6.3.1	概述	114
6.3.2	Gossip 协议	115
	参考文献	118
第 7 章	数据通道	120
7.1	Log 数据收集	120
7.1.1	Chukwa	121
7.1.2	Scribe	122
7.2	数据总线	123
7.2.1	Databus	125
7.2.2	Wormhole	127
7.3	数据导入/导出	128
	参考文献	129

第三篇 大数据存储

第 8 章	分布式文件系统	131
8.1	Google 文件系统 (GFS)	132
8.1.1	GFS 设计原则	132
8.1.2	GFS 整体架构	133

8.1.3	GFS 主控服务器	134
8.1.4	系统交互行为	136
8.1.5	Colossus	137
8.2	HDFS	138
8.2.1	HDFS 整体架构	139
8.2.2	HA 方案	140
8.2.3	NameNode 联盟	143
8.3	HayStack 存储系统	145
8.3.1	HayStack 整体架构	146
8.3.2	目录服务	147
8.3.3	HayStack 缓存	148
8.3.4	HayStack 存储系统的实现	148
8.4	文件存储布局	150
8.4.1	行式存储	151
8.4.2	列式存储	151
8.4.3	混合式存储	156
8.5	纠删码 (Erasure Code)	158
8.5.1	Reed-Solomon 编码	159
8.5.2	LRC 编码	164
8.5.3	HDFS-RAID 架构	166
	参考文献	166
第 9 章	内存 KV 数据库	168
9.1	RAMCloud	169
9.1.1	RAMCloud 整体架构	169
9.1.2	数据副本管理与数据恢复	170
9.2	Redis	172
9.3	MemBase	173
	参考文献	175
第 10 章	列式数据库	176
10.1	BigTable	177
10.1.1	BigTable 的数据模型	177
10.1.2	BigTable 的整体结构	178
10.1.3	BigTable 的管理数据	179
10.1.4	主控服务器 (Master Server)	181

10.1.5	子表服务器 (Tablet Server)	182
10.2	PNUTS 存储系统	186
10.2.1	PNUTS 的整体架构	186
10.2.2	存储单元	187
10.2.3	子表控制器与数据路由器	187
10.2.4	雅虎消息代理	188
10.2.5	数据一致性	189
10.3	MegaStore	190
10.3.1	实体群组切分	191
10.3.2	数据模型	192
10.3.3	数据读/写与备份	193
10.4	Spanner	194
10.4.1	SpanServer 软件栈	195
10.4.2	数据模型	196
10.4.3	TrueTime	196
	参考文献	197

第四篇 大数据处理

第 11 章	大规模批处理系统	199
11.1	MapReduce 计算模型与架构	200
11.1.1	计算模型	201
11.1.2	系统架构	203
11.1.3	MapReduce 计算的特点及不足	206
11.2	MapReduce 计算模式	206
11.2.1	求和模式 (Summarization Pattern)	207
11.2.2	过滤模式 (Filtering Pattern)	208
11.2.3	组织数据模式 (Data Organization Pattern)	210
11.2.4	Join 模式 (Join Pattern)	212
11.3	DAG 计算模型	214
11.3.1	DAG 计算系统的三层结构	214
11.3.2	Dryad	215
11.3.3	FlumeJava 和 Tez	217
	参考文献	218
第 12 章	流式计算	219
12.1	流式计算系统架构	222

12.1.1	主从架构	222
12.1.2	P2P 架构	223
12.1.3	Samza 架构	224
12.2	DAG 拓扑结构	224
12.2.1	计算节点	225
12.2.2	数据流	226
12.2.3	拓扑结构	226
12.3	送达保证 (Delivery Guarantees)	229
12.3.1	Storm 的送达保证机制	230
12.3.2	MillWheel 的“恰好送达一次”机制	233
12.4	状态持久化	234
12.4.1	容错的三种模式	234
12.4.2	Storm 的状态持久化	236
12.4.3	MillWheel 和 Samza 的状态持久化	237
	参考文献	238
第 13 章 交互式数据分析		240
13.1	Hive 系数据仓库	242
13.1.1	Hive	242
13.1.2	StingerInitiative	250
13.2	Shark 系数据仓库	251
13.2.1	Shark 架构	252
13.2.2	部分 DAG 执行引擎 (PDE)	253
13.2.3	数据共同分片	254
13.3	Dremel 系数据仓库	254
13.3.1	Dremel	255
13.3.2	PowerDrill	258
13.3.3	Impala	261
13.3.4	Presto	264
13.4	混合系数据仓库	265
	参考文献	269
第 14 章 图数据库：架构与算法		271
14.1	在线查询类图数据库	272
14.1.1	三层结构	272
14.1.2	TAO 图数据库	273

14.2	常见图挖掘问题	277
14.2.1	PageRank 计算	278
14.2.2	单源最短路径 (Single Source Shortest Path)	278
14.2.3	二部图最大匹配	279
14.3	离线挖掘数据分片	279
14.3.1	切边法 (Edge-Cut)	280
14.3.2	切点法 (Vertex-Cut)	282
14.4	离线挖掘计算模型	284
14.4.1	以节点为中心的编程模型	284
14.4.2	GAS 编程模型	285
14.4.3	同步执行模型	286
14.4.4	异步执行模型	290
14.5	离线挖掘图数据库	292
14.5.1	Pregel	292
14.5.2	Giraph	299
14.5.3	GraphChi	301
14.5.4	PowerGraph	307
	参考文献	311
第 15 章	机器学习：范型与架构	313
15.1	分布式机器学习	314
15.1.1	机器学习简介	314
15.1.2	数据并行 VS.模型并行	316
15.2	分布式机器学习范型	317
15.2.1	三种范型	318
15.2.2	MapReduce 迭代计算模型	319
15.2.3	BSP 计算模型	321
15.2.4	SSP 模型	323
15.3	分布式机器学习架构	324
15.3.1	MapReduce 系列	325
15.3.2	Spark 及 MLBase	327
15.3.3	参数服务器 (Parameter Server)	332
	参考文献	335
第 16 章	机器学习：分布式算法	337
16.1	计算广告：逻辑回归	338

16.1.1	逻辑回归 (Logistic Regression, LR)	338
16.1.2	并行随机梯度下降 (Parallel Stochastic Gradient Descent)	341
16.1.3	批学习并行逻辑回归	341
16.2	推荐系统: 矩阵分解	344
16.2.1	矩阵分解方法	344
16.2.2	ALS-WR 算法	346
16.2.3	并行版 ALS-WR 算法	347
16.3	搜索引擎: 机器学习排序	347
16.3.1	机器学习排序简介	348
16.3.2	LambdaMART	349
16.3.3	分布式 LambdaMART	351
16.4	自然语言处理: 文档相似性计算	352
16.5	社交挖掘: 谱聚类	355
16.5.1	社交挖掘实例	355
16.5.2	谱聚类	356
16.5.3	并行版谱聚类	358
16.6	深度学习: DistBelief	358
16.6.1	深度学习简介	359
16.6.2	DistBelief	360
	参考文献	364
第 17 章	增量计算	366
17.1	增量计算模式	367
17.1.1	两种计算模式	367
17.1.2	Hadoop 平台下增量计算的一般模式	368
17.2	Percolator	370
17.2.1	事务支持	371
17.2.2	“观察/通知”体系结构	373
17.3	Kineograph	374
17.3.1	整体架构	375
17.3.2	增量计算机制	375
17.4	DryadInc	376
	参考文献	377
附录 A	硬件体系结构及常用性能指标	378
附录 B	大数据必读文献	380

0

当谈论大数据时我们在谈 什么

我是这部车
第一个乘客
我不是不快乐

天空血红色
星星灰银色
你的爱人呢

Yes I'm going home
I must hurry home
Where your life goes on

——王菲《乘客》

多年以后，当大数据应用已经无处不在地影响我们每日生活的时候，谁会想起那个面红耳赤争论大数据泡沫何时破灭的下午。

关于什么是大数据，一千个人估计有一千个说法，让我们先从一个例子谈起。

2012年伦敦奥运会赛事举行期间，全世界的观众为激烈的比赛心潮起伏，如何知道人们关于奥运是什么样的心情？很简单，只要你去看一眼世界上最大的摩天轮之一“伦敦眼”，答案就可了然于胸。

用“伦敦眼”的灯光秀来展示英国人对于奥运比赛的情绪，这是 EDF Energy 公司和麻省理工学院的合作项目。这个项目实时过滤人们在 Twitter 上发布的微博，根据一些与奥运有关的词汇比如“2012 伦敦奥运”、“奥运会”将相关微博从海量数据中过滤出来，之后采用被称为“SentiStrength”的情感计算程序自动对每条微博表达的情绪做出判断，将其分为“积极情绪”、“消极情绪”或者是“中性情绪”，每天数十万的奥运相关微博经过分析汇总后，就可以知道当天观众对于奥运的整体情绪是怎样的，将这个统计结果以伦敦眼上不同颜色灯光投射，人们一眼就可以知道其他人的心情。如果你看到四分之三的伦敦眼转轮上的灯光是橙色的，那么这代表了 75% 的奥运相关微博是积极情绪，而悲观情绪则以蓝色灯光来表示。

很神奇，不是吗？但是……别着急，你需要知道的一点是：“伦敦眼灯光秀”只是我们现在身处的大数据时代的一个小魔术而已。

0.1 大数据是什么

多大的数据量才能称得上“大数据”？这其实是颇为令人困惑的一个问题。在谈论数据大小之前，我们先来了解数据量的衡量单位，从小到大依次为 KB、MB、GB、TB、PB、EB 和 ZB，它们相互之间的转换公式为： $1024\text{KB}=1\text{MB}$ ； $1024\text{MB}=1\text{GB}$ ； $1024\text{GB}=1\text{TB}$ ； $1024\text{TB}=1\text{PB}$ ； $1024\text{PB}=1\text{EB}$ ； $1024\text{EB}=1\text{ZB}$ 。数据大小是相对而言的一个变化概念，在 20 世纪末，MB 是最常用的存储单位，那时 1GB 就可以称得上“大数据”了，但随着 IT 技术的快速发展，我们逐步迈入了 GB 时代、TB 时代，而现在正处于从 PB 到 EB 的迁移阶段。

为了对数据之大获得更感性的认识，让我们来看一组数据：Facebook 管理了超过 400 亿张图片，所需存储空间超过 100PB，每天发布的新消息超过 60 亿条，所需存储空间超过 10TB；Twitter 一天产生 1.9 亿条微博；搜索引擎一天产生的日志高达 35TB，Google 一天处理的数据量超过 25PB；YouTube 一天上传的视频总时长为 5 万小时……如果愿意的话，这个单子还可以拉得很长。

我们正处于电子数据暴涨时代，统计数据表明：2010 年全世界信息总量是 1ZB，最近 3 年人类产生的信息量已经超过了之前历史上人类产生的所有信息之和，如果再过 10 年回过头来看，这个数字其实也不算体量巨大，要知道，数字信息量正在以每 5 年增加 10 倍的速率在加速扩张。

为何数据会发生井喷式的增长？这与互联网、移动互联网、物联网大潮的高速发展以及 IT 技术的快速进步有直接关联。互联网的普及、智能手机等手持设备的广泛使用，使得越来越多的人能够将可支配时间投入到各种应用中，而物联网尚处于萌芽期，其发展的结果是：任意物品和设施都有