

企业数字化基石

阿里巴巴云计算基础设施实践

高山渊 蔡德忠 赵晓雪 刘礼寅 刘水旺 陈义全 徐波◎编著

基础设施技术
业务发展的坚实基座

全面剖析阿里巴巴基础设施实践

- 数据中心工程实践 ·
- 数据中心网络架构和技术 ·
- 服务器技术 ·
- 计算硬件技术 ·
- 存储硬件技术 ·
- 智能化数据中心运营 ·

企业数字化基石

阿里巴巴云计算基础设施实践

高山渊 蔡德忠 赵晓雪 刘礼寅 刘水旺 陈义全 徐波◎编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

在互联网行业中，云计算基础设施技术具有非常重要的作用与影响，一方面，在电商、金融科技、云计算、大数据、AI等重要战略领域发展的过程中，云计算基础设施技术的重要支撑作用已经无可替代；另一方面，云计算基础设施技术已经成为许多类型业务的直接竞争力，甚至在云计算、AI等领域中，计算能力、存储能力、网络能力都成为业务的一部分。本书将详细地阐述阿里巴巴在基础设施上的实践细节及先进技术，包括数据中心、网络、服务器、计算硬件、存储硬件，以及智能化运营与自动化运维等内容。

本书是由AIS（Alibaba Infrastructure Service）组织编写的，历时三年，AIS是负责阿里巴巴集团所有基础设施的研发、运营、维护等工作的部门。本书梳理了基础设施演进过程中的思路、得失、最佳实践、未来路径规划等，可以让IT企业的管理者、软件工程师、测试工程师、运维工程师、软件架构师、技术经理能够更清晰地知道基础设施的根本在哪里，而且还可以让IT企业客户对基础设施的发展和未来有一定的了解，以便更好地使用基础设施提供的技术和服务，让阿里巴巴在基础设施上的经验成为国家和行业基础设施发展的参考。

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

企业数字化基石：阿里巴巴云计算基础设施实践 / 高山渊等编著. —北京：电子工业出版社，2020.1

（阿里巴巴集团技术丛书）

ISBN 978-7-121-37388-6

I. ①企… II. ①高… III. ①云计算-应用-电子商务-商业企业管理-中国 IV. ①F724.6-39

中国版本图书馆 CIP 数据核字（2019）第 201707 号

责任编辑：孙学瑛

印 刷：中国电影出版社印刷厂

装 订：中国电影出版社印刷厂

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：720×1000 1/16 印张：22 字数：438 千字

版 次：2020 年 1 月第 1 版

印 次：2020 年 3 月第 2 次印刷

定 价：109.00 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，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

致 谢

第1章：张世长

第2章：蒋林泉 陈 绪

第4章：沈烨烨 杜 鹏 陈 贇 王 锋 周天宇 闫 昆 薛国庆

第5章：谢崇进 姚志平 王海勇 程 刚 曹 捷 席永青 陈俊耿
包贵新 曾 军 周道龙 王 超 朱芳波

第6章：文芳志 郭 锐 钟杨帆 曹洪浩 刘向东 张 胜 吴剑锋
严 勇 王式文 方海宾 王 波 张学刚 黄 征 李人可
宋文俊 赵 兵

第7章：徐凌杰 蒋晓维 虞旭林 韩 天 王文强

第8章：李 舒 王一静 黎安宇 朱 峰 张 袁

第9章：朱婉怡 金伟安 陆增义 黄江伟 张 铭 何 诚 李 湛

推荐序1

本书作为阿里巴巴集团技术丛书中的一部专著，系统地讲述了阿里巴巴云计算基础设施技术的发展史，同时也是阿里经济体成长过程中对基础设施需求的演进史。

基础设施技术伴随了阿里巴巴的整个发展历程。在业务发展的早期，我们以电商、支付为主，采用的是当时业界较为流行的“IOE架构”（指IBM、Oracle和EMC的产品），然而随着阿里巴巴业务的蓬勃发展，我们发现这样的架构无论从技术可扩展性还是从成本效率维度考量，都无法支撑我们走得更远。于是在多年前我们启动了“去IOE”项目，本质上意味着小型机时代的终结，也同时成为阿里巴巴集团云计算基础设施技术发展的开端。

如今，阿里巴巴经济体的业务涵盖了电子商务、在线支付、云计算、物流、视频娱乐、移动导航等诸多领域，也带来了不同的挑战，尤其是云计算的蓬勃发展，不但对基础设施提出了稳定性、成本的要求，还提出了具备持续的高性能计算能力、存储能力、网络互联能力，以及用户几乎无感知的调度能力的要求。我们在定制计算、异构计算、自研存储、大规模网络架构、自研交换机、自建数据中心等方面进行了一系列的实践，积累了一定的经验。

在架构上，2016年，我们提出了存储计算分离、在离线计算混部；2018年，我们提出了阿里巴巴核心系统全面上云，2019年的“双11”以54.4万笔/秒的峰值交易能力、顺畅的用户体验，完美地宣告了目标的达成、架构的成功。

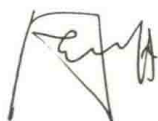
阿里巴巴不仅是一家商业公司，同时也是一家技术公司。我认为本书对阿里巴巴在基础设施领域的工程技术实践进行了一个有益的回顾与总结，不仅有助于我们自身精益求精，也为业界同仁提供了一个可供借鉴和评判的参考，让我们能够共同为云计算基础设施技术演进添砖加瓦，推动技术革新并促进互联网业务发展。

张建锋 (行癫)

阿里巴巴集团 CTO 兼阿里云智能事业群总裁
2019年12月于杭州

到运维管控的“3分钟故障恢复”(1分钟故障发现、1分钟故障定位、1分钟故障恢复),再到数据中心的运营大脑(包括供应链等),我们收获了不少的成功,也有很多血和泪的教训。

技术基础设施是如此的重要,遗憾的是,因为专业领域众多,一直没有一本跨领域、跨专业的书,因此,我们决定在公司内部材料基础设施白皮书的基础上,编撰一本适合公开出版发行的书。本书也融入了这些年来我们在关键技术决策中的思考与实践,也有许多沉淀的经验。期望本书的出版能够给业界带来一块璞玉,以供雕琢,同时也鞭策我们自己在未来更进一步。



阿里巴巴集团副总裁基础设施事业部负责人

2019年12月于杭州

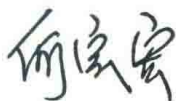
推荐序3

随着越来越多的用户选择云而不是自建基础设施，云计算基础设施技术已经成为各类业务的直接竞争力，得到了社会各界的广泛关注，而阿里巴巴作为国内最大的云服务供应商，无疑在云计算基础设施方面有着丰富的经验和思考。随着众多云服务的出现和各种新技术的层出不穷，云计算基础设施技术如何定义和发展，毫无疑问会成为众多企业，包括阿里巴巴未来业务能否成功的关键因素。

本书从数据中心 TCO 建模和能效的分析入手，针对供配电、制冷、网络架构、服务器、计算硬件、存储硬件、智能化运营、一体化等，从多个维度详细地讲述了基础设施技术的组成和架构，每一项技术都体现了阿里巴巴云计算基础设施技术专家们的实践、思考、创新和沉淀。

在本书中，不仅有对 20 年逐渐成熟沉淀下来的云计算基础设施技术的解读，也包含了浸没式液冷、AI 技术、异构技术、量子通信等新技术的研究和布局，更展望了未来 5 年可能出现的新业务场景、新挑战，描绘了阿里巴巴未来 5 年的奋斗目标，更表现出阿里巴巴对云计算基础设施技术的深刻思考和长远布局。希望阿里巴巴这 20 年云计算基础设施技术上的理论和实践经验，能够为行业和国家，乃至世界的云计算基础设施技术发展提供一个参考。

“不谋万世者，不足谋一时；不谋全局者，不足谋一域。”阿里巴巴作为一家“要活 102 年”（马云原语）的好公司，不仅要做好自己的云计算基础设施技术，更要推动世界云计算基础设施技术的蓬勃发展。为了实现这一目标，不但要结合阿里巴巴特有的业务场景去提高云计算基础设施技术的综合竞争力，更要不断地探索实践、努力创新，提前对“黑天鹅”类技术进行布局和研究，才能保持活力和实现目标。这本书无疑表现了阿里巴巴在长远技术发展的深刻思考和谋略，对当前国内外云计算基础设施技术发展具有里程碑式的指导意义。



中国信息通信研究院云计算与大数据研究所所长
开放数据中心委员会名誉主席

2019 年 12 月于北京

前 言

随着云计算的广泛应用，特别是为创新企业带来的低成本、高效率赋能创新的优势，云计算基础设施技术成为当下创新非常活跃的领域。

为什么写本书

阿里巴巴的云计算基础设施技术的演进和发展，是阿里巴巴集团整体发展的一部分，也是阿里巴巴未来业务能否成功的核心因素。在这个过程中，我们有很多思考和实践、尝试和创新，以及经验和教训，并取得了一定的成果，面向未来也有了越来越明确的思路和规划。我们希望通过这本书将云计算基础设施技术的过去、现在和将来呈现给读者，让阿里巴巴在云计算基础设施技术上的实践经验成为国家和行业云计算基础设施发展的重要参考。

本书主要内容

本书针对云计算基础设施技术的各方面做了概要性的介绍和说明，包括互联网数据中心（IDC）、网络、服务器计算、存储、智能化等。这些方面的每一项都是一个非常专业和复杂的技术领域，本书力图能简明扼要地加以说明。

第1章 基础设施架构综述，简要介绍云计算基础设施技术的范围、价值、挑战等。

第2章 TCO建模，介绍TCO的概念，数据中心、服务器、网络设备的TCO组成，以及如何在系统层面进行TCO分析和建模，并做出最优的选择。

第3章 数据中心能效，介绍如何系统地评估数据中心能效水平，如PUE、SPUE、IUE等。

第4章 数据中心，介绍狭义数据中心的核心理技术，包括供电、制冷等的挑战、发展趋势。

第5章 数据中心网络架构和关键技术，介绍阿里巴巴数据中心的网络架构和关键技术，包括历史回顾、架构理念、高性能网络、网络可视化等。

第6章 服务器技术，介绍服务器的硬件组成、演进、挑战等。

第7章 计算硬件技术，介绍通用计算、异构计算、总线互联等。

第8章 存储硬件技术，介绍内存、HDD、SSD 及阿里巴巴自研存储产品 AliFlash 等。

第9章 基础设施智能化运营与自动化运维，介绍大规模云计算基础设施的运营和运维的自动化、数据化、智能化等内容。

第10章 IDC、网络、服务器的一体化，介绍如何在技术演进中让 IDC 建筑、网络大集群、服务器协同发展，并提供最优的解决方案。

第11章 总结及展望，面向未来，云计算基础设施技术如何进一步提升能力、支持创新。

致谢

本书从策划到成稿，再到出版经历了三年多的时间，前后有百余位同事参加了规划、构思、写作、校对等工作，署名的作者及致谢名单未能将每一位参与者一一列出，在这里一并致谢。

感谢在本书写作过程中基础设施事业部的各位管理者给予的支持，给予本书大方向的把控及人力上的支持。

感谢电子工业出版社的认可，让本书得以出版，感谢孙学瑛老师及其团队细致、艰苦的工作，让本书最终能够与读者见面。

由于作者水平有限，书中不足及错误之处在所难免，敬请专家和读者给予批评指正。



2019年12月



读者服务

微信扫码回复：37388

- 获取博文视点学院 20 元优惠券。
- 获取免费增值资源。
- 加入读者交流群，与更多读者互动。
- 获取精选书单推荐。

目 录

第1章 基础设施架构综述 / 1

- 1.1 基础设施的概念 / 1
- 1.2 基础设施作为技术驱动力 / 3
- 1.3 基础设施的技术价值 / 7
 - 1.3.1 技术安全 / 8
 - 1.3.2 稳定性 / 9
 - 1.3.3 TCO / 10
 - 1.3.4 稳定性、性能、成本的平衡 / 11
 - 1.3.5 追求效率、标准化、规范化 / 11
- 1.4 基础设施的核心竞争力特征 / 12

第2章 TCO建模 / 13

- 2.1 静态TCO模型 / 14
 - 2.1.1 数据中心CapEx / 15
 - 2.1.2 服务器及网络设备CapEx / 16
 - 2.1.3 OpEx / 16
 - 2.1.4 产能/TCO / 16
- 2.2 动态TCO模型 / 17
- 2.3 TCO最优 / 18

第3章 数据中心能效 / 22

- 3.1 数据中心电能利用率 / 22
- 3.2 数据中心服务器电能利用率 / 30
 - 3.2.1 服务器电源的能效 / 31
 - 3.2.2 服务器风扇的能效 / 33
- 3.3 数据中心基础设施使用率 / 35
- 3.4 小结 / 37

第4章 数据中心 / 39

- 4.1 数据中心面临的挑战 / 39
 - 4.1.1 中国能源供给结构对数据中心的影
响 / 39
 - 4.1.2 新政策环境下数据中心面临的
技术升级挑战 / 41
 - 4.1.3 云计算业务给数据中心带来的
机遇与挑战 / 41
- 4.2 数据中心技术的发展趋势 / 43
 - 4.2.1 数据中心技术演进的路线与
趋势 / 43

- 4.2.2 阿里巴巴数据中心的演进路线与趋势 / 46
 - 4.3 数据中心能源与供电技术 / 49
 - 4.3.1 数据中心能源的选择与引入 / 49
 - 4.3.2 数据中心电气系统架构演进的趋势与路线 / 54
 - 4.3.3 数据中心电气关键技术与设备 / 59
 - 4.4 阿里巴巴数据中心供电工程实践案例 / 68
 - 4.4.1 分布式不间断供电和集中式不间断供电的比较 / 68
 - 4.4.2 240V+PSU集中供电系统 / 70
 - 4.4.3 天蝎BBS / 71
 - 4.4.4 服务器BBU / 73
 - 4.4.5 机柜DPS / 74
 - 4.5 数据中心制冷技术 / 76
 - 4.5.1 数据中心制冷系统架构的选择与设计 / 76
 - 4.5.2 连续制冷 / 89
 - 4.5.3 制冷系统的节能措施 / 91
 - 4.6 面向下一代IT系统的冷却技术 / 96
 - 4.6.1 液冷技术简介 / 96
 - 4.6.2 冷板式液冷技术 / 98
 - 4.6.3 浸没式液冷技术 / 99
 - 4.6.4 不同冷却技术对比 / 100
 - 4.7 数据中心制冷技术的发展趋势 / 102
 - 4.8 阿里巴巴数据中心冷却工程实践案例 / 103
 - 4.8.1 千岛湖数据中心 / 103
 - 4.8.2 张北新风自然冷数据中心 / 105
- ## 第5章 数据中心网络架构和关键技术 / 112
- 5.1 超大规模数据中心网络概述 / 112
 - 5.1.1 超大规模数据中心的网络架构及其技术演变 / 112
 - 5.1.2 超大规模数据中心网络的未来发展趋势 / 115
 - 5.2 阿里巴巴数据中心网络 / 118
 - 5.2.1 云计算对数据中心网络的核心要求 / 118
 - 5.2.2 阿里巴巴数据中心网络架构演进 / 120
 - 5.2.3 阿里巴巴数据中心网络HAIL架构的设计理念 / 122
 - 5.2.4 阿里巴巴数据中心网络HAIL架构简介 / 130
 - 5.2.5 阿里巴巴数据中心交换机 / 131
 - 5.2.6 阿里巴巴 vSwitch / 138
 - 5.2.7 阿里巴巴数据中心网络光互联技术 / 141
 - 5.3 数据中心网络协议和高性能网络 / 142
 - 5.3.1 数据中心网络互联协议 / 142
 - 5.3.2 高性能网络 RDMA技术 / 145

5.4 可编程芯片和高性能网关平台 / 146

5.4.1 基于x86高性能软件网关平台 Netframe / 146

5.4.2 基于可编程交换机与SDN技术的硬件交换平台 / 147

5.5 网络可视化技术 / 148

5.5.1 网络可视化前端技术 / 149

5.5.2 网络可视化后端技术 / 150

5.6 城域网光传输技术 / 151

5.6.1 城域网架构简介 / 151

5.6.2 光传输技术 / 154

第6章 服务器技术 / 159

6.1 服务器简介 / 159

6.1.1 服务器系统分级 / 160

6.1.2 服务器行业生态 / 161

6.1.3 服务器应用场景 / 162

6.1.4 服务器面临的挑战 / 162

6.1.5 服务器设计考虑 / 163

6.2 服务器研发和技术体系 / 164

6.2.1 数据中心服务器研发模式 / 165

6.2.2 数据中心服务器工程技术之主板 / 167

6.2.3 数据中心服务器工程技术之散热技术 / 170

6.2.4 服务器结构工程技术 / 174

6.2.5 服务器监控及运维体系 / 189

6.2.6 服务器测试及性能评估体系 / 191

6.3 软硬件优化 / 197

6.4 服务器演进展望 / 198

6.4.1 阿里巴巴的服务器演进 / 198

6.4.2 计算/存储分离和部件化 / 199

6.4.3 进一步解耦CPU-Storage, CPU-GPU, CPU-MEM和架构一体化 / 201

6.5 硬件安全体系 / 203

6.5.1 可信计算 / 204

6.5.2 可信执行环境 / 205

6.5.3 内存安全 / 208

6.5.4 构建完善的硬件安全体系 / 209

第7章 计算硬件技术 / 211

7.1 通用计算 / 212

7.1.1 x86 / 212

7.1.2 Non-IA / 215

7.1.3 Non-IA与x86的比较 / 218

7.1.4 典型业务与通用计算 / 219

7.2 异构计算 / 221

7.2.1 GPU / 221

7.2.2 FPGA / 225

7.2.3 ASIC / 227

7.2.4 CPU、GPU、FPGA、ASIC简单比较 / 230

7.3 边缘计算 / 230

- 7.4 总线互联 / 233
 - 7.4.1 QPI/UPI / 233
 - 7.4.2 NV-Link / 234
 - 7.4.3 CAPI / 234
 - 7.4.4 CCIX / 235
- 7.5 计算与业务软硬件结合 / 236
- 7.6 小结 / 238

第8章 存储硬件技术 / 239

- 8.1 内存 / 240
 - 8.1.1 DDR / 240
 - 8.1.2 NVDIMM与UltraDIMM / 241
 - 8.1.3 AEP / 243
 - 8.1.4 HBM与HMC / 243
- 8.2 HDD / 245
 - 8.2.1 基本原理 / 245
 - 8.2.2 大容量硬盘应用 / 247
 - 8.2.3 HDD的发展趋势 / 251
- 8.3 SSD / 251
 - 8.3.1 NAND Flash / 254
 - 8.3.2 2D/3D NAND / 254
 - 8.3.3 SSD的使用 / 259
- 8.4 新型存储介质 / 261
 - 8.4.1 3D XPoint / 261
 - 8.4.2 ReRAM / 264
 - 8.4.3 NVMe / 265
- 8.5 新型存储介质应用 / 265

- 8.6 自研存储产品AliFlash / 269
 - 8.6.1 双模式SSD / 269
 - 8.6.2 开放通道模式 / 270
 - 8.6.3 开放通道模式的应用路径 / 272
 - 8.6.4 AliFlash与业务软硬件结合 / 273
- 8.7 存储路径 / 276
 - 8.7.1 传统文件系统 / 276
 - 8.7.2 KV存储 / 278
 - 8.7.3 SPDK / 279
- 8.8 NVMoF / 280
- 8.9 近存储计算 / 283
 - 8.9.1 本地近存储计算 / 283
 - 8.9.2 远端近存储计算 / 285
- 8.10 小结 / 288

第9章 基础设施智能化运营与自动化运维 / 289

- 9.1 基础设施资源运营智能化 / 290
 - 9.1.1 需求预测 / 291
 - 9.1.2 数据中心机架位容量规划 / 293
 - 9.1.3 智能制造的发展趋势 / 295
- 9.2 基础设施资源交付自动化 / 296
 - 9.2.1 通过带外网络实现的远程自动开关机 / 296
 - 9.2.2 服务器操作系统自动化安装 / 297
 - 9.2.3 自动化部署遇到的问题及解决方案 / 300
- 9.3 基础设施智能监管控 / 300

9.3.1 可用性监控分析 / 301

9.3.2 电热异常感知 / 301

9.3.3 能效优化 / 304

9.4 物理机自动化维修 / 305

9.4.1 故障发现 / 305

9.4.2 故障预测 / 307

9.4.3 故障维修 / 307

9.5 集群自动化运维 / 309

9.5.1 集群自动化运维的主要设计思想 / 309

9.5.2 物理机集群智能自愈系统 / 311

9.6 网络智能变更系统 / 314

9.6.1 变更稳定性挑战 / 314

9.6.2 解决方案 / 315

9.6.3 网络智能化变更系统 / 316

9.7 无人值守的数据中心 / 317

第10章 IDC、网络、服务器的一体化 / 319

10.1 基础设施的最优化匹配 / 319

10.1.1 大集群与业务的最优匹配 / 319

10.1.2 IDC建筑、网络大集群、服务器的宏观优化 / 320

10.1.3 机架功耗、服务器功耗、网络端口利用率的微观优化 / 321

10.2 整机架交付：OCP、天蝎与AliRack / 322

10.3 面向业务的一体化解决方案：HPC、超算、AI / 323

10.4 系统技术分层与协作 / 327

10.5 全系统模拟器 / 328

第11章 总结及展望 / 330

11.1 2022年——世界第一的基础设施 / 330

11.2 基础设施5.0——AI、异构、智能化 / 332

11.3 基础设施6.0——一体化最优基础设施架构 / 332

11.4 五年之后的基础设施——创新的机制保障 / 333

11.4.1 NASA与达摩院——阿里巴巴的科技创新 / 333

11.4.2 AIS的创新机制 / 333

11.4.3 前瞻布局与战略拖动 / 334

11.5 量子与未来 / 334

11.5.1 量子通信 / 334

11.5.2 量子计算 / 335

基础设施架构综述

从办公数字化和社会数字化到万物数字化，数据驱动智能化应用，小到和天猫精灵的一次对话、淘宝的一次商品搜索、一个包裹的运送，大到智慧城市、智慧交通、智慧医疗、智慧工业、智慧农业、新兴金融服务等，无一不是构建在大数据基础之上的。随着技术的进步和对数据的深度挖掘融合，数据量出现了爆炸式的增长。最近五年全球的数据增长量就可能超过过去所有数据的总和，而且增长的速度还在不断加快。

因此，数据的存储、计算和传输等服务就成了数字经济时代的基础设施。如果说数据是信息时代的血液，那么为所有这些数据的计算、存储和传输提供服务的“躯干”就是“基础设施”。

1.1 基础设施的概念

广义上，基础设施（Infrastructure）是指为社会生产和居民生活提供公共服务的物质工程设施，是用于保障国家或地区社会经济活动正常进行的公共服务系统。它是社会赖以生存发展的一般物质条件，包括交通、邮电、水、电、煤气、园林绿化、环境保护、文化教育、卫生事业等市政公用工程设施，以及公共生活服务器设施等。

基础设施具有以下特点：

- 先行性和基础性。

- 不可贸易性。
- 整体不可分割性（规模）。
- 准公共物品性。

对于一个互联网企业来说，这些公共服务包括数据中心、网络、服务器、计算、存储等。随着云计算的发展，“基础设施即服务（Infrastructure as a Service, IaaS）”正在将这些互联网基础设施变成全社会的基础设施。

云计算作为社会的基础设施具有以下特点：

- 规模化。
- 可获得性。
- 先进性。

构建统一基础设施架构，整体数据中心从软件层面抽象为以下6个部分，如图 1-1 所示。

- 平台层软件，包括固件、驱动、操作系统、虚拟化等。
- 资源管理，主要是统一资源调度系统。
- 硬件抽象层，例如统一的分布式文件系统。
- 部署和维护，例如 Microsoft 的 AutoPilot。
- 编程框架，例如 MapReduce 等。
- 应用层软件，例如搜索、数据库、机器学习等。



图 1-1 整体数据中心核心技术分层

进一步细分，以阿里巴巴为例，将基础设施核心组件结合阿里巴巴的技术和业务现状，可以得到如图 1-2 所示的技术分层。

- 数据中心：包含风、火、水、电、土地等物理资源及其管理平台。

- **基础技术**：包含计算平台和数据平台。
- **调度平台**：统一提供各种任务需求与计算，负责存储资源的协调和管理。
- **计算**：包含计算硬件设备、操作系统和虚拟化软件及其提供的不同资源形态。
- **存储**：包含存储硬件设备、分布式存储软件及其所提供的服务接口。
- **网络**：包含交换机、网卡、线缆等设备及其运维管理和应用平台。
- **DC 大脑**：针对计算、存储、网络、数据中心、资源供应等整体的智能化运营和决策体系。

数据中心							
基础技术	MapReduce	Streaming	Mlib	Search	SQL/NoSQL	DC大脑
	计算平台 (ODPS、Spark)			数据平台			
调度平台	SIGMA	Zeus	Syreo	Suez	Fuxi		
计算	物理机		VM (KVM)	容器 (Docker)	JVM		
	OS/Virtualization						
	计算服务器/裸金属、加速器						
存储	Block Device	SDK/Lib	POSIX (FUSE、NFS)	RESTful (S3、Swift)			
	DFS						
	存储硬件 (HDD、SSD、NVMe)						
网络	Connectivity、Performance、QoS			OVSWNIC	Contiv		
	NOS/Network virtualization						
	Fabric、Switch、NIC、SmartNIC						
DC大脑	DCOS						
	物理资源						
							部署
							SCM
							监控
							预警
							PIPA
							容量
							测量
							数据

图 1-2 阿里巴巴基础设施技术分层

1.2 基础设施作为技术驱动力

基础设施技术的发展一直是互联网、大数据、AI 等技术领域的驱动力，例如计算能力的极大发展促使了已经诞生 60 年的 AI 技术的再次爆发，如图 1-3 和图 1-4 所示。

对于一家大型互联网企业而言，一方面，在电商、金融科技、云计算、大数据、AI 等重要战略领域发展过程中，云计算基础设施技术对其支持的重要性已经无可替代；另一方面，云计算基础设施技术已经成为许多类型业务的直接竞争力，甚至在云计算、AI 等领域中，计算能力、存储能力、网络能力都成为业务的一部分。