

# 淘宝技术 Taobao Technology 这十年

子柳 著



013948556

F724.6

37

# 淘宝技术这十年

子柳 著



北航

C1656612

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

F724.6

37

## 内 容 简 介

任何网站的发展都不是一蹴而就的。它在发展过程中会遇到各种各样的问题和业务带来的压力。正是这些问题和压力推动着技术的进步和发展，而技术的发展反过来又会促进业务的更大提升。如今淘宝网的流量排名已是全球前15名、国内前3名，其系统服务器也从一台发展到万台以上。

本书从工程师的角度讲述淘宝这个超大规模互联网系统的成长历程，及其所有主动和被动的技术变革的前因后果。书中有幕后故事、产品经验、架构演进、技术启蒙，也有大牛成长、业内八卦、失败案例、励志故事。全文文风流畅，有技术人员特有的幽默感；内容积极正面，有现场感，全部是作者亲身经历。

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

淘宝技术这十年 / 子柳著. —北京：电子工业出版社，2013.5  
ISBN 978-7-121-20191-2

I. ①淘… II. ①子… III. ①电子商务—商业经营—中国 IV. ①F724.6

中国版本图书馆 CIP 数据核字（2013）第 077294 号

策划编辑：张春雨

责任编辑：李利健

印 刷：中国电影出版社印刷厂

装 订：三河市鹏成印业有限公司

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：900×1280 1/32 印张：7.875 字数：252 千字

印 次：2013 年 7 月第 3 次印刷

印 数：9001~19000 册 定价：45.00 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 [zltz@phei.com.cn](mailto:zltz@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

服务热线：（010）88258888。

# 序

这是一本好玩的书，从前面两章看，主要是讲故事；从后面两章看，主要是讲人物；从中间几章看，主要是讲技术。我在写作的过程中，把部分章节的内容放到了我的博客上，有人看了说很励志，有人看了说对技术有帮助，有人看了说对产品有帮助，有人看了说对创业有帮助，有人看了说其中的内容特别逗，现在我也不知道这是人文读物还是技术书了。

其实技术书很容易写成催眠的读物，对于这本书，我一开始就没有当作技术书来写，最早是我写给自己看的。在2011年年底的时候，我突然意识到在淘宝度过了七年的时光，七年说长不长，说短不短，回忆起来有些往事历历在目，有些却已开始模糊。为了给自己留点回忆，我就开始记录自己这几年的故事，七年之间发生了很多有趣的事情，越写越多，写着写着，我就希望能够挖掘更多的内容了。当时我还担任着“淘宝技术大学”的校长，这给我的写作提供了得天独厚的优势，我可以听不同的人讲淘宝技术所有的事情。于是我就开始凭记忆去写淘宝在过去几年的技术进展，再去找当事人求证内容的细节。然后又由于有些当事

人实在太有料了，我干脆把跟他交谈的内容也整理出来，当作《牛P列传》来呈现给读者（淘宝的技术人员以P级来定义，牛P就是很高级别的专家）。至此，本书包含了三种行文格式的内容：一个是笔者自己经历的美好时光，一个是淘宝技术的发展历程，一个是那些牛人的牛事。希望能够通过这三个维度，让读者对淘宝的技术有一个概括的了解，进而通过淘宝对互联网的技术有一个概览。

书名取为《淘宝技术这十年》，口气有点大了，但我很难用一个更确切的名字来概括本书的内容，淘宝在这十年里的技术是很难用一本书来写完的，对于笔者不太熟悉的领域（例如运维、云计算、大数据），没敢着墨太多，书中主要围绕网站的业务和系统架构之间的关系展开论述。我希望能够抛砖引玉，未来有更多的人来写《淘宝运维这十年》《淘宝数据这十年》……

写完本书之后，再回过头来读，我意识到自己经历了一场伟大的变革，这个时代把电子商务推到了一个浪潮之巅，我是这个浪潮中的一滴水珠，我无意描述这场伟大的浪潮，只希望通过一滴水珠映射出浪潮底下的故事。这不是淘宝官方的史书，我不具备这样的高度和视野，这只是一个小人物的所见所想。

这本书写得很轻松，我大概花了3个月时间就写完了，这主要得益于其中的事情都是笔者亲身经历的。同时，这种写作状态也直接导致了本书的行文风格天马行空，一边讲业务，一边讲技术，时不时地还做一些不太高明的点评。从2011年年底开始写，到2012年年初，书中的主要内容都有了，接下来就是求证各种细节的过程，这比初稿要艰难很多，我访谈了书中提到的几乎所有人，有时候还要请他们帮忙补充一些内容，

由于人员众多，这里不一一列出，在此一并表示感谢——亲，这是我们共同的作品。

这里特别要提出感谢的是我的师父岳旭强，他比我更早加入淘宝，行文的时间线是根据他给我讲的故事来写的。 he 现在是蘑菇街的CTO，属于自己创业。他离开淘宝也是触动我写本书的原因之一，铁打的营盘，流水的兵，虽然淘宝的人员流失率总体较低，但是每离开一个人就失去了一段故事，如果我现在不写，未来也许很难呈现出过往的这一切了。老天选择了我做技术，选择了我能写点文字，选择了我经历淘宝的发展，选择了我做淘宝技术大学的校长，似乎也选择了我要把这些故事交代给读者。

其次要感谢对本书贡献最大的一个人，书里的故事是我在“淘宝技术大学”给学员上课时讲的，但每个故事的第一个听众都不是技术人员，也不是淘宝员工，是我家里的CEO，她经常被我当作学生，听我讲一个多小时的课，然后给我指出各种不足。同时，为了让我安心写作，她也赦免了我很多家务活，谢谢亲。

最后要感谢本书的编辑张春雨先生，他花费了很大精力帮我修改书稿，在他的帮助下，本书从网文变得像一本书了。另外，虽然本书经历了笔者和编辑的多次修改，但这如同开发软件一样，上线之后发现bug的概率还是很大的，如有错漏之处，敬请读者不吝赐教。

子柳@杭州

2013年4月1日

## 第0章 引言：光棍节的狂欢 / 1

淘宝网不就是一个网站吗？是的，但淘宝网不是一个简单的网站，它的规模排在全球前十几名，顶尖的网站需要顶尖的技术作为支撑。在用户享受淘宝的乐趣的背后，有哪些神秘的技术在支撑它呢？

## 第1章 个人网站 / 13

创建一个小小的网站需要什么技术？在群敌环伺的状况下怎么走出网站的第一步？淘宝网独特的文化和价值观又是怎么形成的？本章试图给你一个解读。

LAMP架构的网站 / 14

武侠和倒立文化的起源 / 20

## 第2章 个人网站的升级 / 23

小网站都想变大，变大总会遇到成长的烦恼，淘宝网的青春期有哪些烦心事？淘宝网的工程师用什么思路来解决遇到的问题？在解决问题之余，他们又用了哪些创新的手段来超越竞争对手？

数据库从mySQL到Oracle / 25

支付手段的创新——支付宝 / 28

交流方式的创新——淘宝旺旺 / 29

## 第3章 企业级Java网站 / 33

给一个网站更换开发语言，这种事情想想都恐怖，淘宝网在2004年就从PHP语言转换成了Java语言，这是怎么做到的？一个企业级的Java系统应该采用什么样的架构？归纳网站遇到的问题，主要是对性能、容量和成本的控制，这有哪些通用的解决方案？在技术实力得到提升之后，淘宝网的哪些业务有了突飞猛进的发展？哪些业务又会遇到技术的障碍？

脱胎换骨的升级——更换开发语言 / 34

坚若磐石——围绕性能、容量和成本的进化 / 41



## 第4章 创造技术 / 33

用钱能解决的问题都不是问题，我们花钱可以购买更好的机器和更好的服务。但当你变成业内最强之后，你的问题就会独特到没有人碰到过，这就意味着你必须自己动手解决问题。幸运的是，淘宝网的发展很快，不幸的是，淘宝网遇到了前所未有的问题。于是，淘宝网被迫开始走向了技术创新的道路。而技术的创新和业务的创新到底又是谁推动了谁？

淘宝文件系统——TFS / 53

淘宝KV缓存系统——Tair / 64

## 第5章 分布式电子商务操作系统 / 77

在飞速发展的背后，隐患已经埋下。在技术架构的制约下，团队协作越来越艰难，代码越来越臃肿，开发的效率越来越低，新业务的需求越来越多，老业务的压力眼看就要超过系统的容量了。这时候，架构该做怎样的调整？又一次的脱胎换骨，浴火重生。重生后，在分布式的电子商务操作系统下，我们才敢举办一次又一次的“双十一”活动。

服务化 / 78

中间件 / 88

高性能服务框架HSF / 91

消息中间件Notify / 95

分布式数据访问层TDDL / 99

Session框架 / 107

开放平台 / 111

## 第6章 我在淘宝这八年 / 129

作为淘宝技术的亲历者，我不经间见证了一段伟大的历史，这段历史，从高层看波澜壮阔，从底层看妙趣横生。我不是高层，这里记录的是发生在底层的一些事。

## 第7章 牛P列传 / 167

前面讲述了技术和业务是怎么互相促进、共生发展的。与此类似，企业和个人也是这样一个关系：一个水平很高的人，找不到合适的平台，就难以发挥自己的价值；一个蓬勃发展的企业，找不到合适的人才，其前景也会堪忧。幸运的是，在淘宝网蓬勃发展的过程中，出现了一大批业内顶尖的高手，他们推动了淘宝的发展，淘宝也让他们实现了个人的价值。接下来让我们一起看看那些人、那些事。

正明——集团核心系统高级研究员 / 168

正祥——淘宝高级研究员，OceanBase项目负责人 / 180

毕玄——集团核心系统资深技术专家 / 191

放翁——淘宝开放平台项目负责人 / 198

吴翰清——阿里云集团信息安全中心高级安全专家 / 210

云铮——数据平台与产品部资深技术专家 / 218

小马——淘宝UED前端通用平台高级技术专家 / 225

淘宝传奇工程师多隆的程序世界 / 236

# 第0章

---

## 引言：光棍节的狂欢

淘宝网不就是一个网站吗？是的，但淘宝网不是一个简单的网站，它的规模排在全球前十几名，顶尖的网站需要顶尖的技术作为支撑。在用户享受淘宝的乐趣的背后，有哪些神秘的技术在支撑它呢？

“时间到，开抢！”坐在电脑前早已等待多时的小美一看时间已到2011年11月11日零时，便迫不及待地投身于淘宝商城一年一度的大型网购促销活动——“淘宝双11购物狂欢节”。小美打开早已收藏好的宝贝——某品牌的雪地靴，飞快的点击购买、付款，一回头发现3000双靴子已被抢购一空。

小美跳起来，大叫一声“欧耶！”

小美不知道，就在11日零点过后的这一分钟里，全国有342万人和她一起涌入淘宝商城。当然，她更不知道，此时此刻，在淘宝杭州的一间办公室里，灯火通明，这里是“战时指挥部”，淘宝技术部的一群工程师正紧盯着网站的流量和交易数据。白板上是他们刚刚下的赌注，赌谁能最准确地猜中流量峰值和全天的交易总额。他们的手边放着充足的食物和各类提神的饮料。

一阵急促的电话铃声响起，是前线部门询问数据的，工程师大声报着：“第1分钟，进入淘宝商城的会员有342万人”。过了一会儿，工程师主动拿起电话：“交易额超过1亿元人民币了，现在是第8分钟。”接下来，“第21分钟，刚突破2亿元”，“第32分钟，3亿元了”，“第1个小时，4.39亿元”。这些数据随后出现在微博上，引起了一片惊呼。

“完蛋了！”突然有人大喝一声，所有的眼睛都紧张地盯着他，只见他挠挠头，嘿嘿地笑道“我赌得少了，20亿元轻松就能

过了，我再加5亿元。”他跑到白板边上把自己的赌注擦去，写上25，接下来有人写上28，有人写上30，有人到微博上开下盘口，同事们纷纷转载下注。接下来的这24个小时，战时指挥部的工程师们都不能休息，他们盯着网站的各种监控指标，适时调整机器，增减功能。顶住第一波高峰之后，这些人开始忙里偷闲地给自己买东西，大家互相交流着哪家买的移动硬盘靠谱，哪家衣服适合自己的女朋友，不时有人哀嚎宝贝被人抢了、信用卡额度不够了。同时，旁边白板上的赌注越下越大。



11月11日，这个棍子最多的日子被网民自我调侃地变成了一个节日——“光棍节”，而淘宝网又用疯狂的折扣促销给它赋予

了另外一个意义——“购物狂欢节”。2011年11月11日这一天，淘宝商城与淘宝网交易额之和突破52亿元人民币，这个数字是“购物天堂”香港一天零售总额8.5亿元的6倍。

网民感受到的是疯抢的喜悦，而网站的技术人员感受到的却是“压力山大”。就如同你家办酒席，宴请左邻右舍，这个办起来容易，倘若宴请十里八乡所有的人，吃饭的人固然开心，但却不是一般人家能够办得起来的。能办得起来如此盛宴者，需要强大的财力和物力、组织能力、技术实力（例如做这么多菜，你的炒锅一定要是“分布式的”、“可复制的”、“可扩展的”，洗菜和切菜要有“工作流引擎”，跑堂的要计算一下最优路径，甚至连厨房的下水道都要重新设计）。

淘宝能够举办如此盛宴，网站的技术实力可见一斑。至2011年年底，淘宝网拥有全国最大的Hadoop分布式计算集群之一（2000多个节点，CPU：24000 core，Memory：48000GB，Disk：24000块），日新增数据50TB，有40PB海量数据存储，分布在全国各地80多个节点的CDN网络，支撑的流量超过800Gbps。淘宝的搜索引擎能够对数十亿的商品数据进行实时搜索，另外，还拥有自主研发的文件存储系统和缓存系统，以及Java中间件和消息中间件系统，这一切组成了一个庞大的电子商务操作系统。从商业数据上看，Amazon的财报显示2011年完成了大约480亿美元的交易额，eBay的2011年财报显示全年完成了大约600亿美元的交易额（不包括其独立的汽车交易平台）。无论从交易额、商品数量还

是从同比增速等指标上看，淘宝网均远超于此，是目前全球最大的电子商务平台。（由于淘宝是非上市公司，未公布2011年的业绩，以上内容来自淘宝网技术副总裁@\_行癫 的微博）。

以上这些技术数据可能已经让一些人产生了不适的感觉，为了让更多的人读懂这本书，我们用下面这段文字描述一下小美访问淘宝网的时候，从技术的角度来看，网站上发生了什么样的事情。

参考资料来自《你刚才在淘宝上买了一件东西》（来自阿里员工卡特）

你发现快要过年了，于是想给你的女朋友买一件毛衣，你打开了www.taobao.com，这时你的浏览器首先查询DNS服务器，将www.taobao.com转换成IP地址。不过你首先会发现，在不同的地区或者不同的网络（电信、联通、移动）下，转换后的IP地址很可能是不一样的，这首先涉及负载均衡的第一步，通过DNS解析域名时，将你的访问分配到不同的入口，同时尽可能保证你所访问的入口是所有入口中可能较快的一个（这和后文的CDN不一样）。

你通过这个入口成功地访问了www.taobao.com实际的入口IP地址，这时产生了一个PV（Page View，页面访问量。每日每个网站的总PV量是形容一个网站规模的重要指标。淘宝网全网在平日（非促销期间）的PV大概是16~25亿个之间）。同时作为一个独



立的用户，你这次访问淘宝网的所有页面均算作一个UV（Unique Visitor，用户访问）。最近臭名昭著的12306.cn的日PV量最高峰在10亿个左右，而UV量却远小于淘宝网十余倍，这其中的原因相信大家都知道。

因为同一时刻访问www.taobao.com的人数过于巨大，所以，即便是生成淘宝首页页面的服务器，也不可能仅有一台，仅用于生成www.taobao.com首页的服务器就可能有成百上千台，那么你的一次访问时生成页面给你看的任务便会被分配给其中一台服务器完成。这个过程要保证公正、公平、平均（即这成百上千台服务器每台负担的用户数要差不多），这一很复杂的过程由几个系统配合完成，其中最关键的便是LVS（Linux Virtual Server，世界上最流行的负载均衡系统之一，是由目前在淘宝网供职的章文嵩博士开发的）。

经过一系列复杂的逻辑运算和数据处理，这次用于给你看的淘宝网首页的HTML内容便成功生成了。对Web前端稍微有点常识的人都应该知道，浏览器下一步会加载页面中用到的CSS、JS（JavaScript）、图片等样式、脚本和资源文件。但是可能相对较少的人才会知道，你的浏览器在同一个域名下并发加载的资源数量是有限的，例如IE 6和IE 7是两个，IE 8是6个，chrome各版本不大一样，一般是4~6个。我刚刚看了一下，我访问淘宝网首页需要加载126个资源，那么如此小的并发连接数自然会加载很久。所以前端开发人员往往会将上述这些资源文件分布在多个域名下，变