大型网站的访问负载前所未有，高性能、高可靠成为现实挑战。本书凝聚了作者在淘宝高速发展和架构变迁中积累的宝贵经验，网站开发、架构人员不可不读。

——章文嵩 阿里巴巴集团高级研究员/核心系统研发部负责人

Broadview®
www.broadview.com.cn

大型网站系统与Java中间件实践

曾宪杰 著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

大型网站系统与Java中间件实践

曾宪杰 著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书围绕大型网站和支撑大型网站架构的 Java 中间件的实践展开介绍。从分布式系统的知识切入，让读者对分布式系统有基本的了解；然后介绍大型网站随着数据量、访问量增长而发生的架构变迁；接着讲述构建 Java 中间件的相关知识；之后的几章都是根据笔者的经验来介绍支撑大型网站架构的 Java 中间件系统的设计和实践。希望读者通过本书可以了解大型网站架构变迁过程中的较为通用的问题和解法，并了解构建支撑大型网站的 Java 中间件的实践经验。

对于有一定网站开发、设计经验，并想了解大型网站架构和支撑这种架构的系统的开发、测试等的相关工程人员，本书有很大的参考意义；对于没有网站开发设计经验的人员，通过本书也能宏观了解大型网站的架构及相关问题的解决思路和方案。

图书在版编目（CIP）数据

大型网站系统与 Java 中间件实践 / 曾宪杰著. 北京：电子工业出版社，2014.4
ISBN 978-7-121-22761-5

Ⅰ. ①大… Ⅱ. ①曾… Ⅲ. ①网站—建设②JAVA 语言—程序设计 Ⅳ. ①TP393.092②TP312

中国版本图书馆 CIP 数据核字(2014)第 059272 号

策划编辑：张春雨
责任编辑：徐津平
印　　刷：北京丰源印刷厂
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：787×980　1/16　印张：21　字数：338.7 千字
印　　次：2014 年 5 月第 2 次印刷
印　　数：4001～8000 册　　定价：65.00 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 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

推荐序一

从事互联网系统开发的人员大多希望成为资深的架构师或领域专家。但大部分人员由于自身工作环境及条件的限制，缺少大型系统实践经验，或者对核心的案例缺乏真实的了解，因此很难有机会理解分布式设计中的关键问题及应对方案。如何才能找到有效的方法并早日成为资深系统架构师呢?

《大型网站系统与 Java 中间件实践》一书介绍了大型网站分布式领域的各种问题，并且以互联网语言 Java 语言为主。这对于希望提升架构能力的技术人员来说，一方面有助于他们了解理论层面体系，掌握大型系统的全貌；另一方面，由于作者具有淘宝平台的丰富的架构及中间件开发经验，因而书中的要点都是大型网站在实际运行中的精华经验，不管你是使用一个已有的分布式开源解决方案，还是自行开发分布式组件，了解这些关键点都会帮助你快速深入地驾驭分布式领域的核心架构。

书中内容尽是实战经验，虽不布道，但所述内容却不乏硝烟——因为是作者在分布式系统的构建、拆分、服务化、部署、实战过程中所经历的教训、积累的经验。书中还有很多性能优化分析、多种方案选择时的 tradeoff 及实战中的方案。方案选择无所谓最佳，只有最适合，这本书不仅给出了方案选择的方法，更给出了方案选择的原因。本书除了适合希望提升架构能力的技术人员阅读，对于正在从事大数据、高并发、中间件使用或研发的一线开发人员也很有价值。

——杨卫华（@TimYang）

新浪网技术总监

推荐序二

看了华黎寄给我的样章有很深的感触，时间仿佛又回到两年多前，当时“去哪儿”网的业务飞速发展，系统遇到了各种各样的问题。

首先是系统无节制地变得臃肿庞大，大量的 web service 的调用将我们的系统变成了一个蜘蛛网，新进入的工程师需要很长时间的熟悉才能对原有系统做出修改。

其次系统随着业务量的不断增大变得不堪重负，开始还能通过增加硬件来扩容，后来增加硬件能够带来的效果已无济于事。

还有，质量越来越难以保证，测试的时间变得越来越长，无法跟上和满足业务发展和变化的需要，团队的压力也越来越大，各个团队都需要增加人员，但是生产力的提升并不明显。

回顾那段时间，故障频发，效率低下，团队人困马乏，成就感变得越来越低。于是我们参考了国内外经历过这个阶段的公司的做法，引入了服务化框架，将系统拆小，重视了系统层次，控制了系统之间的调用关系，也采用了可靠消息系统来应对业务系统之间的强耦合问题。经过两年的努力，现在终于看到了胜利的曙光。

总结下来系统发展的困难也是演进推动力，主要来自于三个方面：一是系统的

负载规模，二是系统的复杂度，三是由前两个方面带来的开发团队的规模扩张。而中间件技术是解决上述三个问题的重要方法。

如果在两年甚至三年前华黎的这本书就已经出版，那么去哪儿网的系统发展就能少走很多弯路。过去两年中，我们为了概念和做法进行了无数次的讨论、争执、尝试、修正。因为我们当时获得经验的途径主要是通过阅读国内外各大网站的同行在各种技术会议上的演讲、PPT，或者与他们交流过程中得到各种启示，这对于一个快速成长中的系统来讲太不成体系了，无法对日常的工作进行指导。而华黎写的这本书融合了他过去在淘宝的经验，书中的做法、理念经过了淘宝系统的爆炸性增长的检验，详实地阐述了 Java 中间件技术在大型网站，尤其是大型交易类网站的建设和应用经验。

书若其人，这本书很实在，用现在流行的话语来讲，就是干货多。我认识华黎有三年了，三年内见过几面，每次见面我都有很多收获。这次他把他的经验和领悟集结成书，相信对很多正在投身于互联网系统开发，特别是高负载、高复杂度的系统开发的工程师们会有很大帮助。也衷心祝福华黎在未来的日子里，儿子健康成长，家庭幸福，工作顺利。

——吴永强（@吴永强去哪）
去哪网 CTO

前言

由于2007年一个很偶然的机会，我加入了淘宝平台架构组，职位是C++工程师。然后我就在只完成了C语言的一个小功能后，开始了Java中间件的研究生涯。从2007年下半年到2013年年初，近6年时间我都在和支撑整个网站应用的Java中间件打交道——从设计实现消息中间件到参与数据访问层设计，再到负责整个Java中间件团队，我也从一个不太懂Java的C++工程师成长为对Java中间件有一定了解和积累的工程负责人。在这个过程中，我也有幸参与了淘宝从集中式的Java应用到分布式Java应用的架构变迁。

本书从分布式系统说起，然后介绍大型网站的变迁中遇到的挑战和应对策略，接着讲解Java中间件的内容，重点介绍了笔者在实践中自主开发的支撑大型网站应用的几个Java中间件产品，包括对它们的思考及其设计和实现原理。最后介绍了支撑大型网站的其他基础要素，包括CDN、搜索、存储、计算平台，以及运维相关的系统等内容。

通过阅读本书，笔者希望读者能够尽量完整地了解大型网站的挑战和应对办法，并且能够了解淘宝在大型网站变迁过程中产生的这几个中间件的具体产品及其背后的思考和设计，并能够对除中间件之外的支撑大型网站的其他系统有一定的了解。希望初学者能够更多地关注全貌，也希望有相关经验的人士可以从本书中得到一些启发，汲取一些经验。

2013 年 5 月，我的岗位有了调整，在接下来的时间中我将带领淘宝技术部承担淘宝业务应用的开发工作。这本书也是对自己淘宝中间件 6 年工作生涯的一份纪念。

最后要说的是，能够完成本书有很多的人要感谢，首先要感谢淘宝给我这么好的平台和机会，没有这个机会就不会有本书。然后也非常感谢太太王海凤对我的支持，4 年前和林昊合著《OSGi 原理与最佳实践》一书的时候，我们刚谈恋爱，我把很多本应陪你的时间用在了写作上；4 年后，我又把本应陪你和儿子的时间用在了写作上，没有你的支持和理解，我不可能完成这次写作。最后也要感谢我的父母、岳父母、姑姑和小表妹，有你们照顾宸宸，我才能专心地写作本书。

曾宪杰

2013 年 11 月于杭州

目录

1 第1章 分布式系统介绍

1.1 初识分布式系统

我第一次听说分布式系统，大约是在2000年的时候。当时很偶然地了解到，1997年版本的电影《泰坦尼克号》中的特效就是通用多台运行Linux的机器组成的系统来共同完成的。整个系统的规模有多大，我没有确切数字，印象中是一百多台机器。这个集群的规模现在看不算大，但在当时深深地震撼了我。更加让我感慨的是，那个时候身边正好有一位同学在用3D软件做特效，因为过于复杂，在寝室要熄灯时总是不能全部完成。如果能够把其他人的电脑拿过来一起分担工作，同时渲染，应该就能在熄灯前完成，那样就不需要把电脑放到负责管理宿舍楼的大爷那边来保证它一直有电了。

1.1.1 分布式系统的定义

对于分布式系统的定义，一直以来我都没有找到或者想到特别简练而又合适的

定义。这里引用一下 *Distributed Systems Concepts and Design*（*Third Edition*）中的一句话：“A distributed system is one in which components located at networked computers communicate and coordinate their actions only by passing messages”。从这句话我们可以看到几个重点，一是组件分布在网络计算机上，二是组件之间仅仅通过消息传递来通信并协调行动。

图 1-1 是一个分布式系统的示意图，从用户的视角看，用户面对的就是一个服务器，提供用户需要的服务，而实际上是靠背后的众多服务器组成的一个分布式系统来提供服务。分布式系统看起来就像一个超级计算机一样。

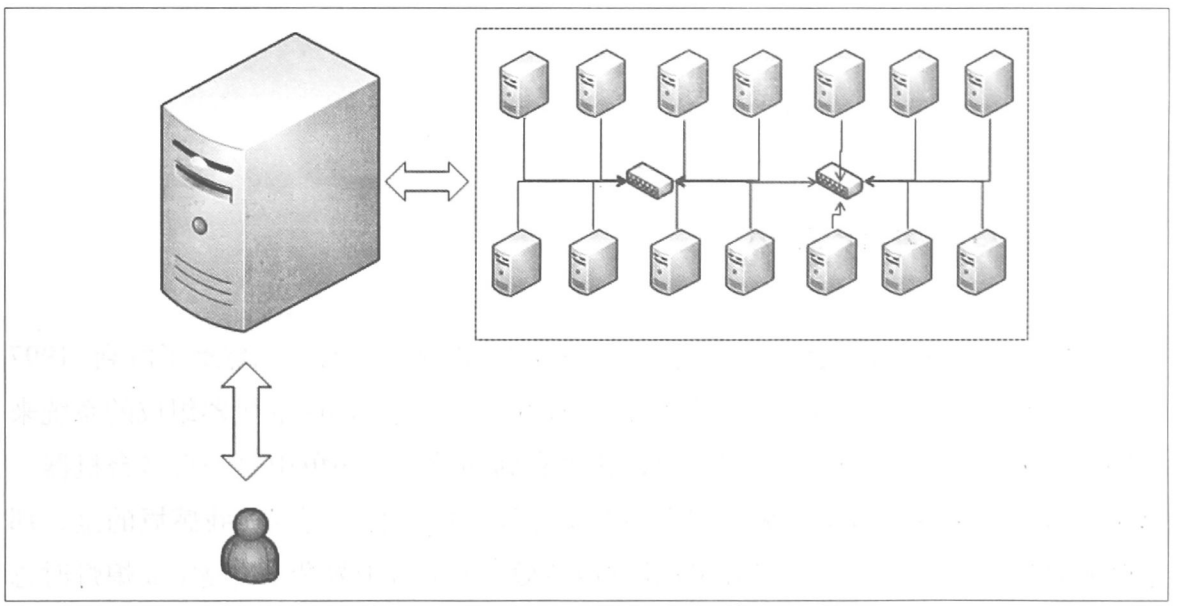

图 1-1 分布式系统示意图

我们来理解一下分布式系统的定义。首先分布式系统一定是由多个节点组成的系统，一般来说一个节点就是我们的一台计算机；然后这些节点不是孤立的，而是互相连通的；最后，这些连通的节点上部署了我们的组件，并且相互之间的操作会有协同。有了这样的原则，我们就可以看看身边都有哪些分布式系统了。像大家平时都会使用的互联网就是一个分布式系统，我们通过浏览器去访问某一个网站（例

如淘宝），在对浏览器发出请求的背后是一个大型的分布式系统在为我们提供服务，整个系统中有的负责请求处理，有的负责存储，有的负责计算，最终通过相互的协同把我们的请求变成了最后的结果返回给浏览器，并呈献给我们。

1.1.2 分布式系统的意义

从单机单用户到单机多用户，再到现在的网络时代，应用系统发生了很多的变化。而分布式系统依然是目前很热门的讨论话题。那么，分布式系统给我们带来了什么，或者说为什么要有分布式系统呢？下面从三个方面来介绍一下其中的原因：

- 升级单机处理能力的性价比越来越低。
- 单机处理能力存在瓶颈。
- 出于稳定性和可用性的考虑。

那么我们先来看单机处理能力包括什么。一般来说，我们关注的是单机的处理器（CPU）、内存、磁盘和网络。下面我们就用处理器来举例说明与单机处理能力相关的问题。

我们都知道摩尔定律：当价格不变时，每隔 18 个月，集成电路上可容纳的晶体管数目会增加一倍，性能也将提升一倍，如图 1-2 所示。

这个定律告诉我们，随着时间的推移，单位成本的支出所能购买的计算能力在提升。不过，如果我们把时间固定下来，也就是固定在某个具体时间点来购买单颗不同型号处理器，那么所购买的处理器性能越高，所要付出的成本就越高，性价比就越低。那么，就是说在一个确定的时间点，通过更换硬件做垂直扩展的方式来提升性能会越来越不划算。除此之外，同样是在某个固定的时间点，单颗处理器有自己的性能瓶颈，也就是说即使你愿意花更多的钱去买计算能力也买不到了，这就是前面提到的第二点。而第三点，强调的是分布式系统带来的稳定性、可用性的提升。如果我们采用单机系统，那么在这台机器正常的时候一切 OK，一旦出问题，那么系

统就完全不能用了。当然，可以考虑做容灾备份等方案，而这些方案就会让你的单机系统演变成分布式系统了。

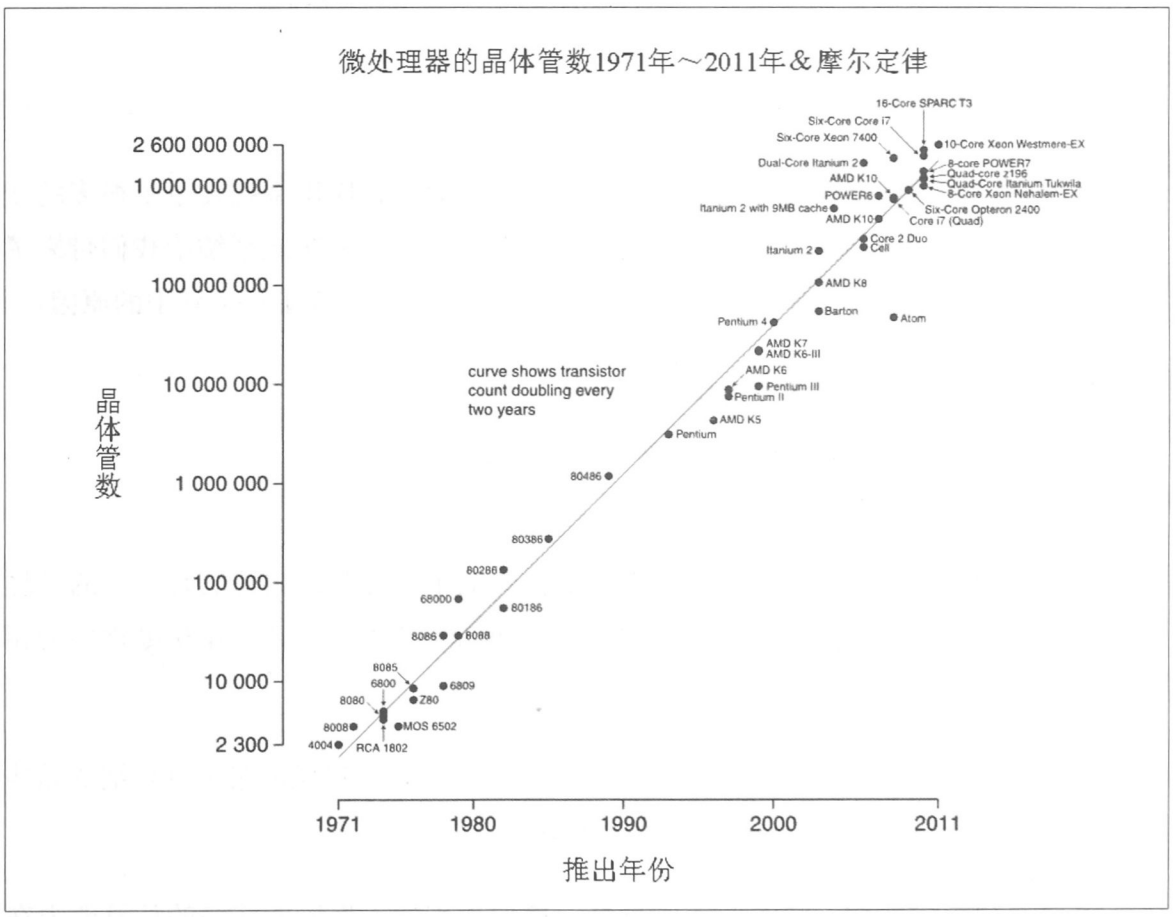

图 1-2 摩尔定律

多年以来分布式系统相关技术一直是技术方面的热点，在接下来的一节中我们看一些分布式系统的基础知识。