

《大型网站技术架构》

图书基本信息

书名：《大型网站技术架构》

13位ISBN编号：9787121212000

10位ISBN编号：7121212005

出版时间：2013-9-1

出版社：电子工业出版社

作者：李智慧

页数：218

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu111.com

《大型网站技术架构》

内容概要

《大型网站技术架构：核心原理与案例分析》通过梳理大型网站技术发展历程，剖析大型网站技术架构模式，深入讲述大型互联网架构设计的核心原理，并通过一组典型网站技术架构设计案例，为读者呈现一幅包括技术选型、架构设计、性能优化、Web 安全、系统发布、运维监控等在内的大型网站开发全景视图。

《大型网站技术架构：核心原理与案例分析》不仅适用于指导网站工程师、架构师进行网站技术架构设计，也可用于指导产品经理、项目经理、测试运维人员等了解网站技术架构的基础概念；还可供包括企业系统开发人员在内的各类软件开发从业人员借鉴，了解大型网站的解决方案和开发理念。

《大型网站技术架构》

作者简介

李智慧

曾在阿里巴巴担任技术专家，参与阿里巴巴基础技术平台开发和www.alibaba.com架构设计。目前就职英特尔亚太研发中心从事云计算与大数据方面的研发工作。

书籍目录

第1篇 概述

1 大型网站架构演化

2

1.1 大型网站软件系统的特点

3

1.2 大型网站架构演化发展历程

4

1.2.1 初始阶段的网站架构

4

1.2.2 应用服务和数据服务分离

4

1.2.3 使用缓存改善网站性能

5

1.2.4 使用应用服务器集群改善网站的并发处理能力

6

1.2.5 数据库读写分离

7

1.2.6 使用反向代理和CDN加速网站响应

8

1.2.7 使用分布式文件系统和分布式数据库系统

9

1.2.8 使用NoSQL和搜索引擎

10

1.2.9 业务拆分

11

1.2.10 分布式服务

11

1.3 大型网站架构演化的价值观

13

1.3.1 大型网站架构技术的核心价值是随网站所需灵活应对

13

1.3.2 驱动大型网站技术发展的主要力量是网站的业务发展

13

1.4 网站架构设计误区

14

1.4.1 一味追随大公司的解决方案

14

1.4.2 为了技术而技术

14

1.4.3 企图用技术解决所有问题

14

1.5 小结

15

2 大型网站架构模式

16

2.1 网站架构模式

16

2.1.1 分层	17
2.1.2 分割	18
2.1.3 分布式	18
2.1.4 集群	19
2.1.5 缓存	20
2.1.6 异步	20
2.1.7 冗余	21
2.1.8 自动化	22
2.1.9 安全	23
2.2 架构模式在新浪微博的应用	23
2.3 小结	25
3 大型网站核心架构要素	26
3.1 性能	27
3.2 可用性	28
3.3 伸缩性	29
3.4 扩展性	30
3.5 安全性	30
3.6 小结	31
第2篇 架构	
4 瞬时响应：网站的高性能架构	34
4.1 网站性能测试	35
4.1.1 不同视角下的网站性能	35
4.1.2 性能测试指标	36
4.1.3 性能测试方法	39
4.1.4 性能测试报告	41

4.1.5 性能优化策略	41
4.2 Web前端性能优化	42
4.2.1 浏览器访问优化	42
4.2.2 CDN加速	43
4.2.3 反向代理	44
4.3 应用服务器性能优化	45
4.3.1 分布式缓存	45
4.3.2 异步操作	52
4.3.3 使用集群	53
4.3.4 代码优化	54
4.4 存储性能优化	58
4.4.1 机械硬盘vs. 固态硬盘	58
4.4.2 B+树vs. LSM树	59
4.4.3 RAID vs. HDFS	61
4.5 小结	64
5 万无一失：网站的高可用架构	66
5.1 网站可用性的度量与考核	67
5.1.1 网站可用性度量	67
5.1.2 网站可用性考核	67
5.2 高可用的网站架构	69
5.3 高可用的应用	71
5.3.1 通过负载均衡进行无状态服务的失效转移	72
5.3.2 应用服务器集群的Session管理	73
5.4 高可用的服务	76
5.5 高可用的数据	

78	
5.5.1	CAP原理
79	
5.5.2	数据备份
82	
5.5.3	失效转移
84	
5.6	高可用网站的软件质量保证
85	
5.6.1	网站发布
85	
5.6.2	自动化测试
86	
5.6.3	预发布验证
87	
5.6.4	代码控制
88	
5.6.5	自动化发布
90	
5.6.6	灰度发布
91	
5.7	网站运行监控
91	
5.7.1	监控数据采集
92	
5.7.2	监控管理
93	
5.8	小结
94	
6	永无止境：网站的伸缩性架构
95	
6.1	网站架构的伸缩性设计
97	
6.1.1	不同功能进行物理分离实现伸缩
97	
6.1.2	单一功能通过集群规模实现伸缩
98	
6.2	应用服务器集群的伸缩性设计
99	
6.2.1	HTTP重定向负载均衡
100	
6.2.2	DNS域名解析负载均衡
101	
6.2.3	反向代理负载均衡
102	
6.2.4	IP负载均衡
103	
6.2.5	数据链路层负载均衡
104	

6.2.6 负载均衡算法	105
6.3 分布式缓存集群的伸缩性设计	106
6.3.1 Memcached分布式缓存集群的访问模型	107
6.3.2 Memcached分布式缓存集群的伸缩性挑战	107
6.3.3 分布式缓存的一致性Hash算法	109
6.4 数据存储服务器集群的伸缩性设计	112
6.4.1 关系数据库集群的伸缩性设计	113
6.4.2 NoSQL数据库的伸缩性设计	117
6.5 小结	119
7 按需应变：网站的可扩展架构	121
7.1 构建可扩展的网站架构	122
7.2 利用分布式消息队列降低系统耦合性	123
7.2.1 事件驱动架构	123
7.2.2 分布式消息队列	124
7.3 利用分布式服务打造可复用的业务平台	126
7.3.1 Web Service与企业级分布式服务	128
7.3.2 大型网站分布式服务的需求与特点	129
7.3.3 分布式服务框架设计	130
7.4 可扩展的数据结构	131
7.5 利用开放平台建设网站生态圈	132
7.6 小结	134
8 固若金汤：网站的安全架构	135
8.1 道高一尺魔高一丈的网站应用攻击与防御	136
8.1.1 XSS攻击	136
8.1.2 注入攻击	

138	
8.1.3	CSRF攻击
139	
8.1.4	其他攻击和漏洞
140	
8.1.5	Web应用防火墙
141	
8.1.6	网站安全漏洞扫描
142	
8.2	信息加密技术及密钥安全管理
142	
8.2.1	单向散列加密
143	
8.2.2	对称加密
144	
8.2.3	非对称加密
144	
8.2.4	密钥安全管理
145	
8.3	信息过滤与反垃圾
146	
8.3.1	文本匹配
147	
8.3.2	分类算法
148	
8.3.3	黑名单
149	
8.4	电子商务风险控制
150	
8.4.1	风险
151	
8.4.2	风控
151	
8.5	小结
153	
第3篇 案例	
9	淘宝网的架构演化案例分析
156	
9.1	淘宝网的业务发展历程
157	
9.2	淘宝网技术架构演化
158	
9.3	小结
162	
10	维基百科的高性能架构设计分析
163	
10.1	Wikipedia网站整体架构
163	
10.2	Wikipedia性能优化策略

165	
10.2.1	Wikipedia前端性能优化
165	
10.2.2	Wikipedia服务端性能优化
166	
10.2.3	Wikipedia后端性能优化
167	
11	海量分布式存储系统Doris的高可用架构设计分析
169	
11.1	分布式存储系统的高可用架构
170	
11.2	不同故障情况下的高可用解决方案
171	
11.2.1	分布式存储系统的故障分类
172	
11.2.2	正常情况下系统访问结构
172	
11.2.3	瞬时故障的高可用解决方案
173	
11.2.4	临时故障的高可用解决方案
174	
11.2.5	永久故障的高可用解决方案
175	
12	网购秒杀系统架构设计案例分析
176	
12.1	秒杀活动的技术挑战
177	
12.2	秒杀系统的应对策略
177	
12.3	秒杀系统架构设计
178	
12.4	小结
182	
13	大型网站典型故障案例分析
183	
13.1	写日志也会引发故障
184	
13.2	高并发访问数据库引发的故障
184	
13.3	高并发情况下锁引发的故障
185	
13.4	缓存引发的故障
185	
13.5	应用启动不同步引发的故障
186	
13.6	大文件读写独占磁盘引发的故障
186	
13.7	滥用生产环境引发的故障
187	

13.8 不规范的流程引发的故障	187
13.9 不好的编程习惯引发的故障	188
13.10 小结	188
第4篇 架构师	
14 架构师领导艺术	190
14.1 关注人而不是产品	191
14.2 发掘人的优秀	191
14.3 共享美好蓝图	192
14.4 共同参与架构	193
14.5 学会妥协	194
14.6 成就他人	194
15 网站架构师职场攻略	196
15.1 发现问题，寻找突破	197
15.2 提出问题，寻求支持	199
15.3 解决问题，达成绩效	201
16 漫话网站架构师	203
16.1 按作用划分架构师	203
16.2 按效果划分架构师	204
16.3 按职责角色划分架构师	205
16.4 按关注层次划分架构师	205
16.5 按口碑划分架构师	206
16.6 非主流方式划分架构师	207
附录A 大型网站架构技术一览	208
附录B Web开发技术发展历程	215
后记	218

《大型网站技术架构》

精彩短评

- 1、介绍了主流网站一般架构的设计，适合想了解大型网站架构设计的读者。当然如果要深入具体的实施细节，基本上每一章都可以独立成书了，但由于技术更新得快，这方面的知识可以从互联网上获取。
- 2、所有互联网产品经理都应该读一读再去和技术讨论
- 3、入门挺好
- 4、对后台架构感兴趣的话，入门还行
- 5、讲的比较泛，能大概了解大型网站的演进过程和相关技术
- 6、值得借鉴
- 7、科普书。
- 8、作为互联网网站的架构演化和设计的入门读物。作者写的比较浅显易懂。
- 9、入门的架构书籍，大型网站设计的科普读物。这本书讲了很多关于网站设计和架构的知识，极大开拓了我的视野
- 10、如果你是个新手 那可以用这本书作为入门：如果你有一定的编程经验 或者经历过一个网站从后到前的开发历程了 那这本书只能算是泛泛而谈 罗列些条目而已。
- 11、网站架构入门科普，简单易懂完整形成了体系。总结归纳性的书，页数很少，两三天就看完了，内容当然泛泛而谈了
- 12、虽然有点年头了，但是对完全不同网站架构的新手来说，还是把概念讲的很明白，书的整体逻辑也很清晰，赞。
- 13、虽然属于程序员高阶知识，但是作者用简单通俗的语言表达出来。内容不多，花个两三个小时就能读完，很适合架构入门。
- 14、比较范着讲的大型网站技术架构方式，分多个方面讲了大型网站可以才用的架构，也举了一些作者自己了解的大型公司案例，简明易懂，适合想了解架构的人和小型互联网公司的技术人员看看
- 15、主要提供思路和方向，具体实施细节还要自己钻研啊
- 16、好的科普书
- 17、菜鸟当科普了
- 18、网站架构_科普！
- 19、这本书不会探讨很深的技术点，但很全面包括了大型交互式网站的方方面面，提供很多值得借鉴的解决方案。
- 20、扫盲书，地铁上和厕所读物
- 21、能对常见的网站架构及其技术有个简单了解。
- 22、对于一个想了解架构所设计的方方面面的人而言，本书做了很好的介绍，不会觉得太过专业晦涩而放弃。相反，花一点时间，就可以对网站架构有一个全面的了解，虽然不一定有机会从事，但是对于我们在对于自己技术的定位以及未来的发展有一定的指导意见。
- 23、内容3星，三观加1星
- 24、新东西不多
- 25、作者对于网站的架构分析很是高屋建瓴，各方面的内容也说得很恰当，适合有一定工作经验的人看
- 26、#读书#最近把大型架构的书看完，虽然看得蛮快的，但是正在受用的不多。整本书讲得很多都是概念性的东西，感觉就是一本大型网站相关名词介绍，然后加一些实例。而后面几章关于架构师角色和工作成长也是如此。整体来说，此书可以作为一本工具书，具体的实施方案，各种细节作者并为写出来。有一种感觉就是作者不够真诚，感觉他接触过很多，但是都是一笔带过，没有具体的细节。到最后没有一个真正的实例，从无到有开始利用他构架出来的理论来做一个大型网站的设计。虽然他说，大型网站不是说一开始就有的，是逐步演进的。然而却木有实例一步一步构建出来。
- 27、虽然都是泛泛而谈但是用来作为架构过程中需要考虑的清单索引还是够用的
- 28、入门看看还行
- 29、也就是读读来面试了，没啥大作用
- 30、一部介绍大型网站架构技术的手册，了解主流技术和手段，可以作为入门参考

《大型网站技术架构》

- 31、特别棒 长见识 想去做类似的工作
- 32、很耐读，知识点都涵盖到了，可以作为一本大型互联网企业技术索引手册
- 33、面试的嘴炮神器！嘴炮架构师速成秘籍！
- 34、用来了解大型网站的技术发展路线还是挺合适的，不含有可供实操的内容，最后一章讲架构师这个角色的定位等比较有趣的内容，还挺不错的
- 35、理清了我对网站技术演变的看法，如果从技术角度考虑，知道大概在哪个阶段需要做点什么事情。
- 36、对主流互联网技术的一个科普吧，没有深入的技术知识点，从事相关工作的都能看明白~
- 37、技术性太低，都是一些理论知识，如果架构大系统的大家都知道要服务器越多越好的，可如果处理好多台服务器间性能合理，服务器分类并没有实际的说明到。
- 38、一本正经的科普了网站架构，捡起了不少零碎的概念和原理，还是那句，知易行难。推荐给对网站开发，分布式系统，大数据感兴趣的同学。
- 39、简单易懂
- 40、给服务端小白涨姿势还是足够了。
- 41、还行吧，虽然写的不深入，但也算有条理，适合入门
- 42、架构入门
- 43、技术不是凭空想象出来的，技术是不断遇见问题，不断解决问题的过程中总结出来的。最后附上本书的结尾词:互联网正在并将继续改变这个世界，一切才刚刚开始，你我正生逢其时！
- 44、技术原理，科普读物，拓宽视野
- 45、比较通俗易懂，对于网站架构的演化，涉及方面均有介绍。可靠性，伸缩性，扩展性方面撰述的较多。后半部分太过于简单，书中有些案例可简单翻翻。
- 46、大而全，适合入门
- 47、Weber 必读，作为一个前端可以把后端的一些架构，解耦思想用到前端里面来，代码不一样，思想是一样了
- 48、技术科普类书籍，没有深度，每个章节都是一个技术点，要想深入了解还是要另外找资料。适合新兵入门全面了解，然后再确定自己的发展方向。老兵也可以迅速看看，然后把目录和附录贴在办公桌前，工作的时候提醒自己别忘记什么。
- 49、罗列了一个架构师应该知道的一切，对于新手入门来说非常有用，对于架构师查漏补缺应该也不错。总之，书里内容很简单，随便读读也没有多大负担，开卷有益说的就是这种书吧
- 50、对形成大体认识很有帮助

- 1、对整个技术体系缺少整体认知的我，这本书给我一个整体的骨架感。以前只做后端技术，喜欢研究算法，知识面比较窄。花1个小时看完第一章就有种神清气爽的感觉，忽然明白了那些大型网站，海量的淘宝是怎么架构出来的。性能、可用性、伸缩性、扩展性和安全五大方面来阐述网站架构，从纵向的前端、应用、服务、数据，到横向的各模块交互，全面丰富的解构。作者把复杂的问题简单的概括总结出来，对于互联网经验不够充足的我来说，这些知识整体、全面，学到作者的分析和理解能力，弥足珍贵。最后不忘说说架构师应该如何和大家合作，如何更好工作的话题，说的相当在理，作者是想明白了的聪明过来人。
- 2、这年头网络很发达，几乎什么东西都可以down。有些聪明人，直接根据关键字google一下，信息也能获取个十之八九。这年头知识不值钱，尤其是冷门小众的东西，国外分享的那么多，直接拿来用就是了。而且，程序员大多内(闷)秀(骚)，写点漂亮的代码可以，追求妹纸这么的事情都表达不好，更别说出书了。所以就冲着这本书深入浅出我也得掏钱买了，更何况里面的内容还真不是不错呢。大型网站技术架构，涉及方面很多，想全部讲个清楚，那至少得2000页，更何况，讲清楚你就能看懂了吗？任何NB的人，都是项目实战而来的。所以，那些认为内容太泛泛的人，可以you can you up做事情最怕的是方向错了，做程序员也一样。而这本书，就是一个指明大方向的参考。
- 3、前几天，犹豫了很久，同时入手了李智慧的《大型网站技术架构 核心原理与案例分析》和杨传辉的《大规模分布式存储系统 原理解析与架构实践》。一到手，看到李智慧的《大型网站技术架构 核心原理与案例分析》，没有开读，看到【“电子工业出版社”】印的书就后悔了。标价同样是59的书。李智慧的《大型网站技术架构 核心原理与案例分析》218页，封面还没开始用，就皱了。封面，一前，以后，扉页处的纸那是一个差，像是以前八九十年代，上厕所用的那种纸一样。刚送过来，封面就有点脱胶。【电子工业出版社】，你还敢更没有节操一点么？这样的书，真心担心看一次，就破烂不堪了。内容还未看，不过看过封面，讲解的内容还是不错。虽然可能都是皮毛，但是提供了一个深入下去的方向。而同样是阿里系出的，杨传辉的《大规模分布式存储系统 原理解析与架构实践》，【机械工业出版社】共293页，比起李智慧的《大型网站技术架构 核心原理与案例分析》，纸张和包装完全是属于精美的级别。书到手时，还有塑料封装。阿里的这两本书，总的来说，书中讨论的内容还是不错。但是强烈鄙视【“电子工业出版社”】这种无良出版商。后面阅读完了，在进行后续的评价.....
- 4、读这本书很容易睡觉，读着读着就睡着了，好吧，我承认书中很多内容之前都了解过。内容比较泛，比较适合刚接触网站架构的童鞋。很多知识点到为止。但基本上关于网站架构的各个知识点都能有所涉及。不过一本200多页的书也很难深入讲解每个知识点。如果想对某些知识点有更深入了解，还需要辅助得看其它方面的资料。但还是特别感谢作者在百忙之中能抽空分享这些心得，希望有更多像阿里巴巴、腾讯、百度或其它互联网公司的工程师能做更多同样的分享，如果能就某一主题深入介绍就更好了。
- 5、评价一本书需要从这本书不同目标受众的视角，对于资深架构师来说，这本书可能没有什么；对于初中级WEB运维和开发工程师，这本书可以让你节省很多摸索的时间，少走很多弯路；对于项目经理和产品经理，这本书是很好的互联网技术“科普”读物。
- 6、基本就是互联网公司标准架构设计和开发流程。架构的经验是很多故障堆起来的，需要较好的计算机基础，知道系统瓶颈在哪里，还需要开阔的视野，知道什么工具能解决什么问题。很赞同书里一句话“是业务成就了技术，是事业成就了人”。内容排版不错，属于科普书籍，两百页，字不多，没有太多深入，毕竟每个专题都可以讲太多，属于大纲型的书籍，两天抽空看完了。
- 7、本文并没有什么特别的东西，但是都很实在，而且能很好的组织起来，也可以看为一个架构。何为架构，要有大局观，大局观就是提前预防掉那些通用的问题：高可用，工程化，伸缩性，扩展性。对应需要的能力：了解分布式的一些东西，了解项目的业务和流程和运维使之工程化，了解负载均衡，能够对业务的分割和代码的分层。文中其他的一些观点我倒是很喜欢：1 先成就他人，再成就自己 2 刚开始加入的时候不要急于证明自己，要先融入。 3 最好的奖励就是目标的达成，最大的惩罚就是目标没实现 4 技术是要解决问题，但是我们要关心的是解决问题的人。 5 学会妥协 6 越激烈的争辩代表越关心这个问题 文中有些例子有点意思1 wiki的实现中就是业务退一小步，技术进一大步。这个他们能够那么省钱的原因啊。 2 秒杀从根本上来讲并不是很难，首先是页面的静态化，开始秒杀的按钮通

过js来实现，js不缓存，js尽量小。开始秒杀的时候使用可以秒杀的js。秒杀很少能达数据层，因为就那么几个能成功。主要的压力在应用服务器，但是用一个记数服务器，收到请求更新这个数字，大于数字的直接返回秒杀失败。所以大部分都会进入失败的逻辑，整个也很简单。只要业务服务器能抗住这些访问压力就基本ok了，如果业务服务器不够，可以直接在负载均衡那边随机失败一部分。3 负载均衡的实现：1 dns实现 2应用层实现，使用反向代理。3 网络层实现 ip负载均衡，通过网关来修改目标ip 4 数据链路层 改mac地址，如Ivs。负载均衡的策略主要有 轮询 随机 最少连接 hash。4 一致性hash的时候，用多个虚拟节点对应一台实际的服务器，如150:1 这样会大大减少负载波动。

8、这本书前两章讲了大型网站用到的一些技术和概念，属于科普性的内容。第三章的案例讲得太宽泛，与期待不合。整本书还行，作者把大型网站建站的主要技术都点到了。如果案例可以深入一点讲，这本书就比较值得读了。

9、还不错，但是感觉太浅辄止。唯一技术上深入一点是Consistent Hashing那部分，但是网上教程很多，也不是太复杂的内容。Case Study部分可以再深入一些最好，而不仅仅是丢个架构图出来。有些比较有趣的架构，例如Instagram, Facebook等，其实都可以拿出来分析一下。典型故障分析还不错，但是内容太单薄了。最后架构师部分可有可无。还有个问题是本书主要以Java生态系统为基础，对其他语言的讲解太少。即便是Java生态系统，涉及到具体的都以淘宝内部的工具为主，没有太大借鉴意义。

10、没看过，看了介绍和目录，感觉内容很丰富的说。这些都是网站建设当中实际要面对的一些问题，很有针对性，很实用。当然这些都是针对一些中大型网站来说的，对于一般企业公司网站，流量不高的网站，无所谓了。他们需要的是短平快，定位不同。不明觉厉！

11、两个视角1.技术是因业务存在而存在网站业务发展的各个阶段造就了技术支撑体系的不断发展演化。2.大型网站的5个核心要素性能、可靠性、可用性、可扩展、安全性围绕着这五点有各个篇章讲解各种应对解决方案。后续的两块1.典型系统的案例架构分析2.写给架构师的工作经验手册和建议读此书的感觉1.结构思路很清晰，可读性很高，很接地气。2.这本书涉及网站的方方面面，见识到很多可用的思路和解决方案

12、用了坐地铁的时间看完了本书，现在重新翻了前三章，梳理一下。说一本书有没有价值，不是说作者有多么有名，不是说书讲得多么高深，让人看不懂（比如，小时候觉得余秋雨的书多么牛逼，晦涩的啥都看不懂），而是说读者看完书后能产生多少共鸣，多少为读者所用。这本书就是一本对我产生价值的书。作者说，软件架构是指有关软件整体结构与组件的抽象描述，用于指导大型软件系统各个方面的设计。书中先介绍了一个架构的简单发展史，然后分别讲了性能、可用性、伸缩性、扩展性、安全性等五个核心要素。在架构发展史中，我们看到，不是所有网站上来都追求庞大的架构，而多是从LAMP开始，通过架构的分层、业务的分层进行扩展。从这点来看，公司的很多线上事故确实是由于架构不合理造成的。由于工作原因，我重点关注了可用性部分。一个高可用网站的软件生产工艺，起码要关注以下几个方面：发布策略、自动化测试、预发布验证、代码的管理、自动化发布、灰度发布等。这些点我司一个都还没有做好，软件质量差也就不足为奇了。再加上实时监控、数据的冗余灾备，基本保证了网站的高可用。后面有一个案例是讲淘宝技术架构发展的，可以结合《淘宝技术这十年》一起看。

13、内容太单薄了，饼很大但内容深度有限，基本上都是网络上的各大型网站的综合，现在很多就是打着这种幌子来谈架构，个人以为架构按需定制越简单越好，此书的内容只是把网络流传的大型网站架构杂烩了一下，当然也有一个人的经验与心得，但只是点到为止，真正的干货还是比较少，架构还是需要大家结合自己项目情况，从工作实践，而不是堆砌框架

14、不愧是架构师写的书，脉络清晰，条理分明。非技术岗的我只是把他作为科普书籍来读，细节深入的部分也可以带过，并不影响了解全局。以下读书建议也是针对我这类读者来说——全书四大部分：第一章概述可以重点阅读，即是概览亦是总结，对网站架构的搭建，形态演化，基础模式和核心要素有很全面清晰的阐述。读完可以了解整个网站架构的基本情况。第二章针对网站架构的5大核心要素展开深入的说明，解读了每一个核心要素的检验标准和常见优化策略，不过具体到有些专业词汇还需要查阅其他资料，有些难懂。但是作者对基础知识的解读已经足够让我这种技术小白理解。第三章结合各大公司实例进行网站架构优化的分析，如果有实例工作经验的阅读起来可能会更有体会。最后一章在说架构师，不管是什么职业岗位，有些道理总是相通的。整本书读下来除了个别难懂的专业点，很顺畅，没有多余的套话，有些地方作者会插入自己的感性领悟，也能得到一些启发。作为小白觉得最好的读此书方法是认真读第一章，然后通读第二章先理解纲领，再填充阅读，希望深入的部分

继续补充阅读，最后两章可以轻松泛读。

15、这本书对大型网站的演化发展的过程，做了一个比较详尽的概述，虽然比较宽泛，但对于对架构敢兴趣的同学，非常值得一读，网站的演化过程，架构模式，性能，高可用性，伸缩性，扩展性，安全性等等都有涉及，看完这本书，让我对“架构”这两个字，有了更进一步的了解了，现在公司项目的架构很有问题，在加上本人技术菜鸟级别，还是要多学习学习~~~

16、书中大多数知识点比较熟悉所以读起来比较快，整体看来基本网站所用的技术都有提及，或多或少都有介绍。作为一个网站架构的入门书籍还是胜任的，理论梳理得比较清晰，技术介绍也比较流畅。不太满意的是一本200页的书定价要59元，这个苦了荷包紧张的读书人了。

17、这本书总结了大型网站技术的点点滴滴，让我一方面渐渐梳理了一些架构上的内容。比如何时何地用什么架构和框架，比如怎么进行有效的横向和纵向扩展。虽然有一点毕竟宽泛，不怎么实战化。有时候觉得这本书是个理论狂魔，但是理论建立在实践之上了。值得通读一遍，然后到时候有用到其中的某一块，比如分布式缓存，可以详细了解。我认为是这样的，花了两周时间看的。

18、这本书重点在最前面。“目前全球大型网站没有一个一开始规模大的，都是一台服务器，一台电脑开始做出来慢慢发展出来的。”我觉得跟“淘宝技术10年”这本书差不多，但是2位作者的写作方式不一样。对刚开始学习网站的人来说这本书还是很有帮助的，怎样架构，怎么解决问题等等，

19、内容还是很不错，但是自己去搜集也是很容易找到的。定价有点偏高，买了之后还是有点亏的感觉，网站架构理论对于从业稍微久一点的人基本都清楚。所以感觉干货不是很多，没事看看加深印象还是可以的。书还没有看完，至少从前面的几章来看，只能给70分，后面看完再说吧。

20、这书比较薄，定价59比较反感的是这书的开头拉了一堆所谓的有title的架构师来撑场面，吹捧气息十足。在知乎上看过好几次这作者自荐这书的，真的拿上手看了遍，定价和内容质量严重不匹配，感觉这书有点像早几年的一本叫做 开源技术选型手册 的码农实践版 借来翻看足矣，不值得购买。

21、我在读这本书时，翻得非常快，呵呵。内容其实涵盖还是挺全的，大型网站的方方面面，面广但讲得不够深，或者说可当通俗的技术小说一样来读，作者了解的东西挺多，但里面的很多东西可能不一定是自己做的，只是拿过来一些同事的经验来说事。总体来说，不要期望太高，但作为一个想了解淘宝网等大型电子商务网站技术架构的人来说，也可以读一读，长长见识。

22、作者在阿里有一线的架构经验，但是本书中谈得并不深入，老生常谈的一些东西，实例部分更是点到即止。不过作者作了一定的归纳，可以看作是一般的方法论入门。下面是部分摘抄内容：4.3.4 代码优化(P54)1.多线程（1）使用多线程的原因：IO阻塞与多CPU；（2）启动线程= $[\text{任务执行时间}/(\text{任务执行时间}-\text{IO等待时间})] * \text{CPU内核数}$ （3）线程安全的主要手段：将对象设计为无状态、使用局部对象、并发访问资源时使用锁；2.资源复用：单例、线程池；3.数据结构；4.垃圾回收。5.3.2 应用服务器集群的Session管理1.Session复制；2.Session绑定；3.利用cookie记录Session；4.Session服务器；7.1 1.伸缩性不需要改变网站的软硬件设计，仅仅通过改变部署服务器数量就可以扩大或者缩小网站的服务处理能力。或者说，系统能够通过增加（减少）自身资源规模的方式增强（减少）自己计算处理事务的能力。2.扩展性对现在系统影响最小的情况下，系统功能可持续扩展或者提升的能力。表现在系统基础设施稳定不需要经常变更，应用之间较少依赖和耦合，对需求变更可以敏捷响应，它是系统架构设计层面的开闭原则。8.4.规则引擎规则引擎是一种将业务规则和规则处理相分离的技术，业务规则将文件由运营人员通过管理界面编辑，当需要修改规则时，无须更改代码发布程序，即可实时使用新规则，而规则处理逻辑则调用处理输入的数据。

23、从大型网站的高性能架构、高可用架构、伸缩性架构、可扩展性架构和安全架构阐述了大型网站的技术架构。但是具体如何实现，如集群如何配置没有描叙。阅读本书，可以大致了解大型网站架构的全貌。普及了一些网站技术架构的基础知识。如何真正构建一个大型的网站。还是得通过持续不断的运营和发展，逐步从小网站演进到大网站，不断由业务驱动技术架构的发展。总的来说本书对于了解互联网网站架构还是值得一读。

24、架构上的方方面面都谈了些，也因此对各个方面一般都是浅浅带过。200来页的书，就别指望包罗万象了。可以作为一本闲书，翻翻回顾回顾，整理一下自己的技术思路。不过对于刚接触互联网架构的人来说，还是有些价值的。秒杀的章节上讲的技巧还可以，值得一学。

25、作者的背景可谓是强力背书，而且从书中也可以看出作者本身的工程经验非常丰富，对架构技术的方方面面都应该涉猎不少。本书的优点是对大型网站的技术架构有一个系统性的梳理，让人看起来很有条理；缺点是讲得太浅，不过要一个人一本书讲透架构所有的方方面面也显然是不太现实的。所

《大型网站技术架构》

以总得来说，感觉入门是极好的选择，能让你对大部分的架构技术有一个宏观了解，继续深入则是后续的工作了。

26、网站架构入门科普。刚刚看完，写得挺好的，通俗易懂，开篇就明义：“大型网站是演化出来的，不是设计出来的。”，通过对这句经典的“任何计算机科学中的问题，都可以加一层抽象层来解决。”的阐述，衍生出了，缓存为王，分层，解耦，模块化等各个网站架构中应该遵循的原则。总的来说，作者认为纵向分层，横向分层，可线性伸缩是大型网站面对复杂业务和海量访问的致胜法宝！书中负载均衡那一节，基本上是参考的《HTTP权威指南》负载均衡的内容。也推荐大家可以看看《HTTP权威指南》这本书。

27、其实书里前几章内容对于对网站架构没有什么概念的人（譬如我）来说还是挺有用的（虽然还是有大把东西都已经知道并且在用了）。但给人的感觉始终是博客文章的范。内容讲的挺实在，各个部分的组件需要的相关手段都有提及。但是，也就这样了。整本书概念提得挺多，解释挺少，感觉起来也就是几篇博客就能搞定的事儿。为了出成一本书，添加了许多不必要的东西，外加一堆的重复和废话。就这样还只有218页，可见内容是多么的少了。总结：这书网上搜搜前三四章相关的博客快速看看差不多了，买书实在没必要。

28、系统的讲述了网站架构，演进的各种技术和原理。可以让初学者对网站架构以及演进有一个很好的全貌，少走很多弯路，同时也能继续深入研究里面的一些技术点，如果能把第三篇和第四篇具体案例以及架构师写的再具体和深入点就更好了。

章节试读

1、《大型网站技术架构》的笔记-第1页

创新的而业务发展模式对网站架构逐步提出更高要求，才使得创新的网站架构得以发展成熟。是业务成就了技术，是事业成就了人，而不是相反。

一般来说，除了当前的系统需求外，软件架构还需要关注性能、可用性、伸缩性、扩展性和安全性这5个框架要素。

网站高可用的手段是冗余，应用部署在多台服务器上同时提供访问.....

所谓伸缩性是指通过不断向集群中加入服务器的手段来缓解不断上升的用户并发访问压力和不断增长的数据存储需求。

网站快速发展，功能不断扩展，如何设计网站的架构使其能够快速响应需求变化，是网站可扩展架构的主要目的。

网站可伸缩架构的主要手段是事件驱动架构和分布式服务。前者通常用消息队列实现，否则通常是将业务和可复用服务分流开来。

即使在技术层面，性能优化也需要全面考虑，综合权衡：性能提升一倍，但服务器数量也需要增加一倍；或者响应事件缩短，同时数据一致性也下降，这样的优化是否可以接受？这类问题的答案不是技术团队能回答的。归根结底，技术是为业务服务的，技术选型和架构决策依赖具体的业务规划乃至企业战略，离开业务发展的支撑和驱动，技术走不远，甚至还会迷路。

高可用服务有几个策略我很喜欢，分级管理。核心业务使用好的硬件，跟次业务进行分离等等；超时设置；异步调用；服务降级；幂等性设计，多次操作跟操作一次一样效果。

保证数据高可用的手段主要是利用数据备份和失效转移机制。

CAP,数据一致性（Consistency）、数据可用性（Availability）、分区耐受性（Partition Tolerance）。

数据一致性又可以细分成，数据强一致性，数据用户一致性、数据最终一致。

一个处理失败的理念，快速失败，尽早退出。

灰度发布，可以有效降低发布失败风险等。

数字签名的过程则相反，签名者用自己的私钥对信息进行加密，然后发送给对方，接收方用签名者的公钥对信息进行解密，获得原始明文信息，由于私钥只有签名者拥有，因此该信息是不可抵赖的，具有签名的性质。

HTTPS传输中浏览器使用的数字证书实质上是经过权威机构认证的非对称加密的公钥。

wikipedia通过约束业务获得更大的技术方案选择余地，很多时候业务后退一小步，技术就可以前进一大步。

《大型网站技术架构》

这些问题只要在大型网站技术一线呆上两三年，就一定会碰到。

遇到问题，经理了这个过程，技术才能升华，人和技术才能融为一体，才知道什么技术是真正有用的，什么技术是花拳绣腿。

培养一个网站架构师的成本不单要看付了他多少薪水，给了他多少股票，还要看为他引起的故障买了多少次单。

“软件设计有两种风格，一种是将软件设计得很复杂，以使其缺陷没那么明显；一种是将软件设计得很简单，以使其没有明显的缺陷。”

发掘人的优秀远比发掘优秀的人更有意义。

每个人都有自己成就的目标，而工作是达成自我成就的一种手段：通过工作的挑战，发掘自我的潜能，重新认知自我和世界。

最好的奖励是最终达成的目标本身，最大的惩罚就是这个美好的目标没有实现。

寻找一个值得共同奋斗的目标，营造一个让大家都能最大限度发挥自我价值的工作氛围。

所谓问题，就是体验-企望，当体验不能满足期望，就会觉得出了问题。消除问题有两种手段：改善体验或者降低期望。降低期望只是回避了问题，而如果直面期望和体验之间的差距，就会发现问题所在，找到突破点。

其实新员工首先要做的是融入团队，跟大家打成一片，只要能和团队一起共进退，你就不是一个人在战斗。等熟悉了情况，知道了水的深浅之后，再寻找突破口，择机而动。

2、《大型网站技术架构》的笔记-第103页

由于反向代理服务器转发请求在http协议层面，因此也叫应用层负载均衡。其优点是和反向代理服务器功能集成在一起，部署简单。缺点是反向代理服务器是所有请求和响应的中转站，其性能会成为瓶颈

3、《大型网站技术架构》的笔记-网站的伸缩性架构

原文大意：许多计算机的问题都可以通过在上面增加一层抽象来解决。

以目前接触到的层次来说，确实是这样。搞不定了，加一层抽象就好搞一些。至于加抽象层是不是最优方案那就另说了。

4、《大型网站技术架构》的笔记-第23页

架构技术文摘，及每日图书推荐。

5、《大型网站技术架构》的笔记-第117页

在计算机数据存储领域，一直是关系数据库的天下，以至于传统企业应用领域，许多应用系统设计都是面向数据库设计——先设计数据库然后设计程序，从而导致关系模型绑架对象模型，并由此引申出旷日持久的业务对象贫血模型与充血模型之争。

6、《大型网站技术架构》的笔记-第37页

在网站产品设计初期，产品经理和运营人员就需要规划不同发展阶段的网站系统用户数，并以此为基础，根据产品特性和运营手段，推算在线用户数和并发用户数。测试程序通过多线程模拟并发用户的办法来测试系统的并发处理能力。为了真实模拟用户行为，在两次请求之间加入一个随机等待时间，这个时间被称为思考时间。

7、《大型网站技术架构》的笔记-第123页

软件架构师最大的价值不在于掌握多少先进的技术，而在于具有将一个系统切分成N个低耦合的子模块的能力，这些子模块包括横向的业务模块，也包含纵向的基础技术模块。这种能力一部分源于自专业技术和经验，还有一部分源自架构师对业务场景的理解、对人性的把握、甚至对世界的认知。

8、《大型网站技术架构》的笔记-第35页

开发人员可以利用的优化手段有，使用缓存加速数据读取，使用集群提高吞吐能力，使用异步消息加快请求响应及实现削峰，使用代码优化改善程序性能。

9、《大型网站技术架构》的笔记-第162页

而有些路，走过以后，再回头，一览众山小。

10、《大型网站技术架构》的笔记-第90页

由于火车发布模型是基于规则驱动的流程，所以这个流程可以自动化。采用火车发布模式的网站会开发一个自动化发布的工具实现发布过程的自动化。

11、《大型网站技术架构》的笔记-第194页

提醒自己是来做软件的，来实现客户价值的，不是来证明谁对谁错的，不是来给黑熊XX的。

12、《大型网站技术架构》的笔记-第191页

一群优秀的人做一件他们热爱的事，一定能取得成功。

最好的软件项目管理不是制定计划，组织资源，跟踪修正项目进展，对成员进行激励和惩罚，而是发掘项目组每个成员的优秀潜能，让大家理解并热爱软件产品最终的蓝图和愿景。每个人都是为了实现自我价值而努力，不是为了领工资而工作。

一旦做到这一点，项目组每个成员都会自我驱动，自觉合作，寻找达成目标的最终路径并坚韧不拔的持续前进。。。不需要拙劣胡萝卜和大棒，最好的奖励就是最终要达成的目标本身，最大的惩罚就是这个目标没有实现。

寻找一个值得共同奋斗的目标，营造一个让大家都能最大限度发挥自我价值的工作氛围。

13、《大型网站技术架构》的笔记-第85页

网站发布

网站需要保证7X24高可用支行，同时网站需要不断地发布新功能有为引用户以保证在激烈的市场竞争中获得成功。这相当于给飞行中的飞机换个引擎。但是网站发布毕竟是一次提前预知的服务器宕机，所以过程可以更柔和。

14、《大型网站技术架构》的笔记-第1页

```
测试一下
第一次写笔记
###这里有一个异常
###不知道是什么原因
```

15、《大型网站技术架构》的笔记-第105页

使用三角传输模式的链路层负载均衡是目前大型网站使用最广的一种负载均衡手段。在linux平台上最好的链路层负载均衡开源产品是LVS(linux virtual server)

16、《大型网站技术架构》的笔记-第56页

事实上,java开发常用的对象容器Spring默认构造的对象都是单例(需要注意的是Spring的单例是Spring容器管理的单例,而不是单例模式构造的单例).

这句话什么意思,不甚明了

17、《大型网站技术架构》的笔记-第15页

技术是用来解决业务问题的，而业务的问题，也可以通过业务的手段去解决。

18、《大型网站技术架构》的笔记-第144页

均衡负载/ dns/反向代理/ip层/链路层
MQ消息系统支撑
分布式缓存
主从数据库分离/读写
请求数优化
基于HDFS的HBase
NoSQL也出场了

需求层次分离/系统层次分离

19、《大型网站技术架构》的笔记-第37页

实践中通常采用的方法是重复请求，比如一个请求操作重复执行一万次，测试一万次执行需要的总响应时间之和，然后除以一万，得到单次请求的响应时间。

20、《大型网站技术架构》的笔记-第90页

火车发布模型：将每个应用的发布过程看作一次火车旅程，火车定点运行，期间有若干站点，每一站都进行一次例行检查，不通过的项目下车，剩下的项目继续坐着火车旅行，直到火车到达终点。

《大型网站技术架构》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu111.com