14-HTTP有哪些优点？又有哪些缺点？

上一讲我介绍了HTTP的五个基本特点，这一讲要说的则是它的优点和缺点。其实这些也应该算是HTTP的特点，但这一讲会更侧重于评价它们的优劣和好坏。

上一讲我也留了两道课下作业，不知道你有没有认真思考过，今天可以一起来看看你的答案与我的观点想法是否相符，共同探讨。

不过在正式开讲之前我还要提醒你一下，今天的讨论范围仅限于HTTP/1.1，所说的优点和缺点也仅针对HTTP/1.1。实际上，专栏后续要讲的HTTPS和HTTP/2都是对HTTP/1.1优点的发挥和缺点的完善。

简单、灵活、易于扩展

首先，HTTP最重要也是最突出的优点是“**简单、灵活、易于扩展**”。

初次接触HTTP的人都会认为，HTTP协议是很“**简单**”的，基本的报文格式就是“header+body”，头部信息也是简单的文本格式，用的也都是常见的英文单词，即使不去看RFC文档，只靠猜也能猜出个“八九不离十”。

可不要小看了“简单”这个优点，它不仅降低了学习和使用的门槛，能够让更多的人研究和开发HTTP应用，而且我在第1讲时就说过，“简单”蕴含了进化和扩展的可能性，所谓“少即是多”，“把简单的系统变复杂”，要比“把复杂的系统变简单”容易得多。

所以，在“简单”这个最基本的设计理念之下，HTTP协议又多出了“**灵活和易于扩展**”的优点。

“灵活和易于扩展”实际上是一体的，它们互为表里、相互促进，因为“灵活”所以才会“易于扩展”，而“易于扩展”又反过来让HTTP更加灵活，拥有更强的表现能力。

HTTP协议里的请求方法、URI、状态码、原因短语、头字段等每一个核心组成要素都没有被“写死”，允许开发者任意定制、扩充或解释，给予了浏览器和服务器最大程度的信任和自由，也正好符合了互联网“自由与平等”的精神——缺什么功能自己加个字段或者错误码什么的补上就是了。

“请勿跟踪”所使用的头字段 DNT（Do Not Track）就是一个很好的例子。它最早由Mozilla提出，用来保护用户隐私，防止网站监测追踪用户的偏好。不过可惜的是DNT从推出至今有差不多七八年的历史，但很多网站仍然选择“无视”DNT。虽然DNT基本失败了，但这也正说明HTTP协议是“灵活自由的”，不会受单方面势力的压制。

“灵活、易于扩展”的特性还表现在HTTP对“可靠传输”的定义上，它不限制具体的下层协议，不仅可以使用TCP、UNIX Domain Socket，还可以使用SSL/TLS，甚至是基于UDP的QUIC，下层可以随意变化，而上层的语义则始终保持稳定。

应用广泛、环境成熟

HTTP协议的另一大优点是“**应用广泛**”，软硬件环境都非常成熟。

随着互联网特别是移动互联网的普及，HTTP的触角已经延伸到了世界的每一个角落：从简单的Web页面到复杂的JSON、XML数据，从台式机上的浏览器到手机上的各种APP，从看新闻、泡论坛到购物、理财、“吃

鸡”，你很难找到一个没有使用HTTP的地方。

不仅在应用领域，在开发领域HTTP协议也得到了广泛的支持。它并不限定某种编程语言或者操作系统，所以天然具有“**跨语言、跨平台**”的优越性。而且，因为本身的简单特性很容易实现，所以几乎所有的编程语言都有HTTP调用库和外围的开发测试工具，这一点我觉得就不用再举例了吧，你可能比我更熟悉。

HTTP广泛应用的背后还有许多硬件基础设施支持，各个互联网公司和传统行业公司都不遗余力地“触网”，购买服务器开办网站，建设数据中心、CDN和高速光纤，持续地优化上网体验，让HTTP运行的越来越顺畅。

“应用广泛”的这个优点也就决定了：无论是创业者还是求职者，无论是做网站服务器还是写应用客户端，HTTP协议都是必须要掌握的基本技能。

无状态

看过了两个优点，我们再来看看一把“双刃剑”，也就是上一讲中说到的“无状态”，它对于HTTP来说既是优点也是缺点。

“无状态”有什么好处呢？

因为服务器没有“记忆能力”，所以就不需要额外的资源来记录状态信息，不仅实现上会简单一些，而且还能减轻服务器的负担，能够把更多的CPU和内存用来对外提供服务。

而且，“无状态”也表示服务器都是相同的，没有“状态”的差异，所以可以很容易地组成集群，让负载均衡把请求转发到任意一台服务器，不会因为状态不一致导致处理出错，使用“堆机器”的“笨办法”轻松实现高并发高可用。

那么，“无状态”又有什么坏处呢？

既然服务器没有“记忆能力”，它就无法支持需要连续多个步骤的“事务”操作。例如电商购物，首先要登录，然后添加购物车，再下单、结算、支付，这一系列操作都需要知道用户的身份才行，但“无状态”服务器是不知道这些请求是相互关联的，每次都得问一遍身份信息，不仅麻烦，而且还增加了不必要的数据传输量。

所以，HTTP协议最好是既“无状态”又“有状态”，不过还真有“鱼和熊掌”两者兼得这样的好事，这就是“小甜饼”Cookie技术（第19讲）。

明文

HTTP协议里还有一把优缺点一体的“双刃剑”，就是**明文传输**。

“明文”意思就是协议里的报文（准确地说是header部分）不使用二进制数据，而是用简单可阅读的文本形式。

对比TCP、UDP这样的二进制协议，它的优点显而易见，不需要借助任何外部工具，用浏览器、Wireshark或者tcpdump抓包后，直接用肉眼就可以很容易地查看或者修改，为我们的开发调试工作带来极大的便利。

当然，明文的缺点也是一样显而易见，HTTP报文的所有信息都会暴露在“光天化日之下”，在漫长的传输链路的每一个环节上都毫无隐私可言，不怀好意的人只要侵入了这个链路里的某个设备，简单地“旁路”一下流量，就可以实现对通信的窥视。

你有没有听说过“免费WiFi陷阱”之类的新闻呢？

黑客就是利用了HTTP明文传输的缺点，在公共场所架设一个WiFi热点开始“钓鱼”，诱骗网民上网。一旦你连上了这个WiFi热点，所有的流量都会被截获保存，里面如果有银行卡号、网站密码等敏感信息的话那就危险了，黑客拿到了这些数据就可以冒充你为所欲为。

不安全

与“明文”缺点相关但不完全等同的另一个缺点是“不安全”。

安全有很多的方面，明文只是“机密”方面的一个缺点，在“身份认证”和“完整性校验”这两方面HTTP也是欠缺的。

“身份认证”简单来说就是“**怎么证明你就是你**”。在现实生活中比较好办，你可以拿出身份证、驾照或者护照，上面有照片和权威机构的盖章，能够证明你的身份。

但在虚拟的网络世界里这却是个麻烦事。HTTP没有提供有效的手段来确认通信双方的真实身份。虽然协议里有一个基本的认证机制，但因为刚才所说的明文传输缺点，这个机制几乎可以说是“纸糊的”，非常容易被攻破。如果仅使用HTTP协议，很可能你会连到一个页面一模一样但却是个假冒的网站，然后再被“钓”走各种私人信息。

HTTP协议也不支持“完整性校验”，数据在传输过程中容易被窜改而无法验证真伪。

比如，你收到了一条银行用HTTP发来的消息：“小明向你转账一百元”，你无法知道小明是否真的就只转了一百元，也许他转了一千元或者五十元，但被黑客窜改成了一百元，真实情况到底是什么样子HTTP协议没有办法给你答案。

虽然银行可以用MD5、SHA1等算法给报文加上数字摘要，但还是因为“明文”这个致命缺点，黑客可以连同摘要一同修改，最终还是判断不出报文是否被窜改。

为了解决HTTP不安全的缺点，所以就出现了HTTPS，这个我们以后再说。

性能

最后我们来谈谈HTTP的性能，可以用六个字来概括：“**不算差，不够好**”。

HTTP协议基于TCP/IP，并且使用了“请求-应答”的通信模式，所以性能的关键就在这两点上。

必须要说的是，TCP的性能是不差的，否则也不会纵横互联网江湖四十余载了，而且它已经被研究的很透，集成在操作系统内核里经过了细致的优化，足以应付大多数的场景。

只可惜如今的江湖已经不是从前的江湖，现在互联网的特点是移动和高并发，不能保证稳定的连接质量，所以在TCP层面上HTTP协议有时候就会表现的不够好。

而“请求-应答”模式则加剧了HTTP的性能问题，这就是著名的“队头阻塞”（Head-of-line blocking），当顺序发送的请求序列中的一个请求因为某种原因被阻塞时，在后面排队的所有请求也一并被阻塞，会导致客户端迟迟收不到数据。

为了解决这个问题，就诞生出了一个专门的研究课题“Web性能优化”，HTTP官方标准里就有“缓存”一章（RFC7234），非官方的“花招”就更多了，例如切图、数据内嵌与合并，域名分片、JavaScript“黑科技”等等。

不过现在已经有了终极解决方案：HTTP/2和HTTP/3，后面我也会展开来讲。

小结

1. HTTP最大的优点是简单、灵活和易于扩展；
2. HTTP拥有成熟的软硬件环境，应用的非常广泛，是互联网的基础设施；
3. HTTP是无状态的，可以轻松实现集群化，扩展性能，但有时也需要用Cookie技术来实现“有状态”；
4. HTTP是明文传输，数据完全肉眼可见，能够方便地研究分析，但也容易被窃听；
5. HTTP是不安全的，无法验证通信双方的身份，也不能判断报文是否被窜改；
6. HTTP的性能不算差，但不完全适应现在的互联网，还有很大的提升空间。

虽然HTTP免不了这样那样的缺点，但你也不要怕，别忘了它有一个最重要的“灵活可扩展”的优点，所有的缺点都可以在这个基础上想办法解决，接下来的“进阶篇”和“安全篇”就会讲到。

课下作业

1. 你最喜欢的HTTP优点是哪个？最不喜欢的缺点又是哪个？为什么？
2. 你能够再进一步扩展或补充论述今天提到这些优点或缺点吗？
3. 你能试着针对这些缺点提出一些自己的解决方案吗？

欢迎你把自己的答案写在留言区，与我和其他同学一起讨论。如果你觉得有所收获，欢迎你把文章分享给你的朋友。

课外小贴士

01 “简洁至上”也是 Apple 公司前领导人乔布斯所崇尚的设计理念。

02 与 DNT 类似的还有 P3P（Platform for Privacy Preferences Project）字段，用来控制网站对用户的隐私访问，同样也失败了。

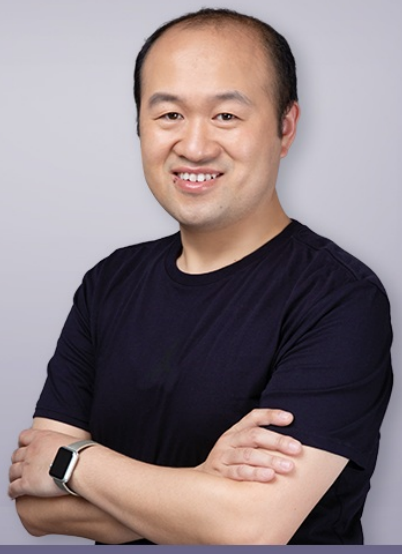
03 出于安全的原因，绝大多数网站都封禁了 80/8080 以外的端口号，只允许 HTTP 协议“穿透”，这也是造成 HTTP 流行的客观原因之一。

04 HTTP/1.1 以文本格式传输 header，有严重的数据冗余，也影响了它的性能。

极客时间

透视 HTTP 协议

深入理解 HTTP 协议本质与应用



罗剑锋

奇虎360技术专家

Nginx/OpenResty 开源项目贡献者