

## 42 | 互联网架构模板：“网络层”技术

2018-08-02 李运华 来自北京

《从0开始学架构》



除了复杂度，互联网业务发展的另外两个关键特点是“高性能”和“高可用”。通常情况下，我们在设计高可用和高性能系统的时候，主要关注点在系统本身的复杂度，然后通过各种手段来实现高可用和高性能的要求，例如我前面介绍的计算高性能架构模式、存储高可用架构模式等。但是当我们站在一个公司的的角度来思考架构的时候，单个系统的高可用和高性能并不等于整体业务的高可用和高性能，互联网业务的高性能和高可用需要从更高的角度去设计，这个高点就是“网络”，所以我将相关措施统一划归为“网络层”。注意这里的网络层和通常理解的如何搭建一个局域网这种概念不一样，这里强调的是站在网络层的角度整体设计架构，而不是某个具体网络的搭建。

接下来我将介绍[互联网架构模板](#)的“网络层”技术的几个关键架构设计点，部分内容专栏前面已经有深入阐述，今天作为概要的总结把它们归纳一下。

### 负载均衡

顾名思义，负载均衡就是将请求均衡地分配到多个系统上。使用负载均衡的原因也很简单：每个系统的处理能力是有限的，为了应对大容量的访问，必须使用多个系统。例如，一台 32 核 64GB 内存的机器，性能测试数据显示每秒处理 Hello World 的 HTTP 请求不超过 2 万，实际业务机器处理 HTTP 请求每秒可能才几百 QPS，而互联网业务并发超过 1 万是比较常见的，遇到双十一、过年发红包这些极端场景，每秒可以达到几十万的请求。

## 1.DNS

DNS 是最简单也是最常见的负载均衡方式，一般用来实现地理级别的均衡。例如，北方的用户访问北京的机房，南方的用户访问广州的机房。一般不会使用 DNS 来做机器级别的负载均衡，因为太耗费 IP 资源了。例如，百度搜索可能要 10000 台以上机器，不可能将这么多机器全部配置公网 IP，然后用 DNS 来做负载均衡。有兴趣的读者可以在 Linux 用 “dig baidu.com” 命令看看实际上用了几个 IP 地址。

DNS 负载均衡的优点是通用（全球通用）、成本低（申请域名，注册 DNS 即可），但缺点也比较明显，主要体现在：

DNS 缓存的时间比较长，即使将某台业务机器从 DNS 服务器上删除，由于缓存的原因，还是有很多用户会继续访问已经被删除的机器。

DNS 不够灵活。DNS 不能感知后端服务器的状态，只能根据配置策略进行负载均衡，无法做到更加灵活的负载均衡策略。比如说某台机器的配置比其他机器要好很多，理论上来说应该多分配一些请求给它，但 DNS 无法做到这一点。

所以对于时延和故障敏感的业务，有实力的公司可能会尝试实现 **HTTP-DNS** 的功能，即使用 HTTP 协议实现一个私有的 DNS 系统。HTTP-DNS 主要应用在通过 App 提供服务的业务上，因为在 App 端可以实现灵活的服务器访问策略，如果是 Web 业务，实现起来就比较麻烦一些，因为 URL 的解析是由浏览器来完成的，只有 Javascript 的访问可以像 App 那样实现比较灵活的控制。

HTTP-DNS 的优缺点有：

灵活：HTTP-DNS 可以根据业务需求灵活的设置各种策略。

可控：HTTP-DNS 是自己开发的系统，IP 更新、策略更新等无需依赖外部服务商。

及时：HTTP-DNS 不受传统 DNS 缓存的影响，可以非常快地更新数据、隔离故障。

开发成本高：没有通用的解决方案，需要自己开发。

侵入性：需要 App 基于 HTTP-DNS 进行改造。

## 2.Nginx 、LVS 、F5

DNS 用于实现地理级别的负载均衡，而 Nginx、LVS、F5 用于同一地点内机器级别的负载均衡。其中 Nginx 是软件的 7 层负载均衡，LVS 是内核的 4 层负载均衡，F5 是硬件的 4 层负载均衡。

软件和硬件的区别就在于性能，硬件远远高于软件，Nginx 的性能是万级，一般的 Linux 服务器上装个 Nginx 大概能到 5 万 / 秒；LVS 的性能是十万级，没有具体测试过，据说可达到 80 万 / 秒；F5 性能是百万级，从 200 万 / 秒到 800 万 / 秒都有。硬件虽然性能高，但是单台硬件的成本也很高，一台最便宜的 F5 都是几十万，但是如果按照同等请求量级来计算成本的话，实际上硬件负载均衡设备可能会更便宜，例如假设每秒处理 100 万请求，用一台 F5 就够了，但用 Nginx，可能要 20 台，这样折算下来用 F5 的成本反而低。因此通常情况下，如果性能要求不高，可以用软件负载均衡；如果性能要求很高，推荐用硬件负载均衡。

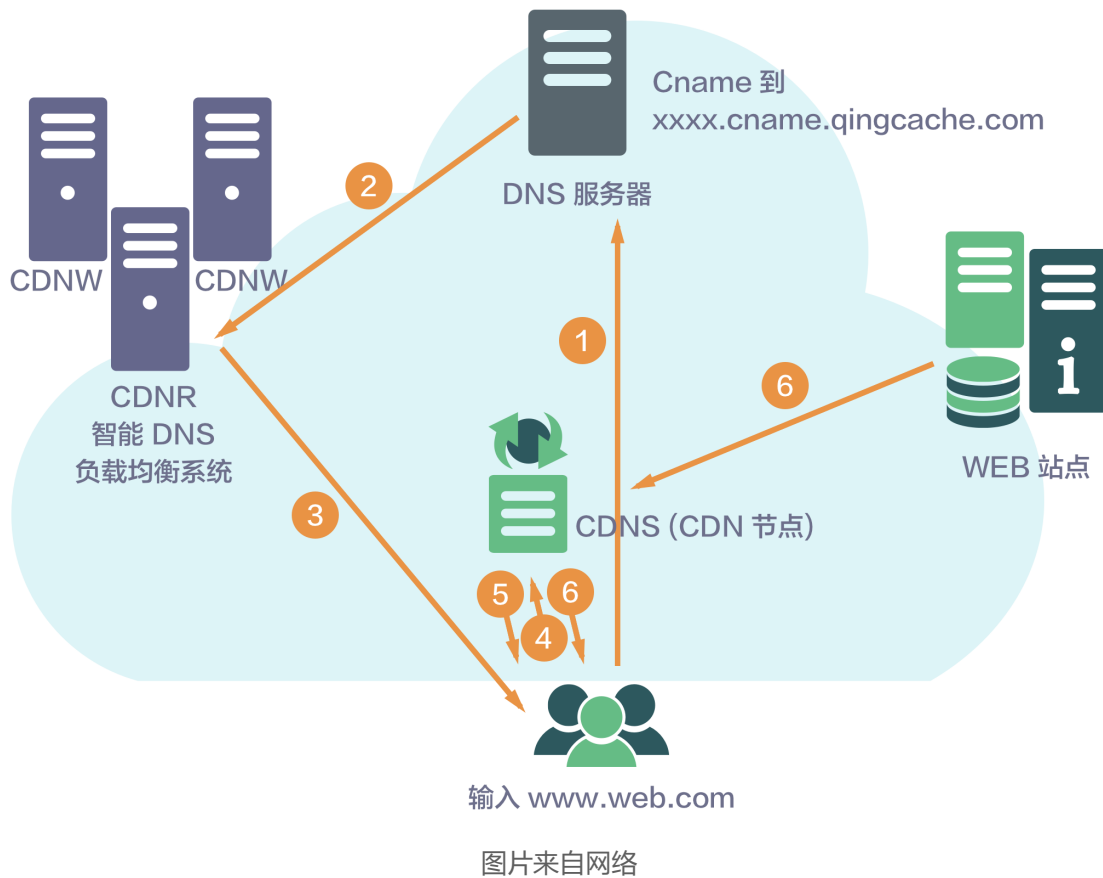
4 层和 7 层的区别就在于协议和灵活性。Nginx 支持 HTTP、E-mail 协议，而 LVS 和 F5 是 4 层负载均衡，和协议无关，几乎所有应用都可以做，例如聊天、数据库等。

目前很多云服务商都已经提供了负载均衡的产品，例如阿里云的 SLB、UCloud 的 ULB 等，中小公司直接购买即可。

## CDN

CDN 是为了解决用户网络访问时的“最后一公里”效应，本质上是一种“以空间换时间”的加速策略，即将内容缓存在离用户最近的地方，用户访问的是缓存的内容，而不是站点实时的内容。

下面是简单的 CDN 请求流程示意图：



CDN 经过多年的发展，已经变成了一个很庞大的体系：分布式存储、全局负载均衡、网络重定向、流量控制等都属于 CDN 的范畴，尤其是在视频、直播等领域，如果没有 CDN，用户是不可能实现流畅观看内容的。

幸运的是，大部分程序员和架构师都不太需要深入理解 CDN 的细节，因为 CDN 作为网络的基础服务，独立搭建的成本巨大，很少有公司自己设计和搭建 CDN 系统，从 CDN 服务商购买 CDN 服务即可，目前有专门的 CDN 服务商，例如网宿和蓝汛；也有云计算厂家提供 CDN 服务，例如阿里云和腾讯云都提供 CDN 的服务。

## 多机房

从架构上来说，单机房就是一个全局的网络单点，在发生比较大的故障或者灾害时，单机房难以保证业务的高可用。例如，停电、机房网络中断、地震、水灾等都有可能导致一个机房完全瘫痪。

多机房设计最核心的因素就是如何处理时延带来的影响，常见的策略有：

## 1. 同城多机房

同一个城市多个机房，距离不会太远，可以投入重金，搭建私有的高速网络，基本上能够做到和同机房一样的效果。

这种方式对业务影响很小，但投入较大，如果不是大公司，一般是承受不起的；而且遇到极端的地震、水灾等自然灾害，同城多机房也是有很大风险的。

## 2. 跨城多机房

在不同的城市搭建多个机房，机房间通过网络进行数据复制（例如，MySQL 主备复制），但由于跨城网络时延的问题，业务上需要做一定的妥协和兼容，比如不需要数据的实时强一致性，只是保证最终一致性。

例如，微博类产品，B 用户关注了 A 用户，A 用户在北京机房发布了一条微博，B 在广州机房不需要立刻看到 A 用户发的微博，等 10 分钟看到也可以。

这种方式实现简单，但和业务有很强的相关性，微博可以这样做，支付宝的转账业务就不能这样做，因为用户余额是强一致性的。

## 3. 跨国多机房

和跨城多机房类似，只是地理上分布更远，时延更大。由于时延太大和用户跨国访问实在太慢，跨国多机房一般仅用于备份和服务本国用户。

## 多中心

多中心必须以多机房为前提，但从设计的角度来看，多中心相比多机房是本质上的飞越，难度也高出一个等级。

简单来说，多机房的主要目标是灾备，当机房故障时，可以比较快速地将业务切换到另外一个机房，这种切换操作允许一定时间的中断（例如，10 分钟、1 个小时），而且业务也可能有损失（例如，某些未同步的数据不能马上恢复，或者要等几天才恢复，甚至永远都不能恢复

了)。因此相比多机房来说，多中心的要求就高多了，要求每个中心都同时对外提供服务，且业务能够自动在多中心之间切换，故障后不需人工干预或者很少的人工干预就能自动恢复。

多中心设计的关键就在于“数据一致性”和“数据事务性”如何保证，这两个难点都和业务紧密相关，目前没有很成熟的且通用的解决方案，需要基于业务的特性进行详细的分析和设计。以淘宝为例，淘宝对外宣称自己是多中心的，但是在实际设计过程中，商品浏览的多中心方案、订单的多中心方案、支付的多中心方案都需要独立设计和实现。

正因为多中心设计的复杂性，不一定所有业务都能实现多中心，目前国内的银行、支付宝这类系统就没有完全实现多中心，不然也不会出现挖掘机一铲子下去，支付宝中断 4 小时的故障。

有关多中心设计更详细的内容，你可以查看专栏第 [28](#)、[29](#)、[30](#) 期的内容。

## 小结

今天我为你讲了互联网架构模板中的“网络层”技术，希望对你有所帮助。

这就是今天的全部内容，留一道思考题给你吧，为什么可以购买负载均衡和 CDN 服务，但却不能购买多机房和多中心服务？

欢迎你把答案写到留言区，和我一起讨论。相信经过深度思考的回答，也会让你对知识的理解更加深刻。（编辑乱入：精彩的留言有机会获得丰厚福利哦！）

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

## 精选留言 (34)



narry

2018-08-02

负载均衡和cdn基本是和业务无关，具有通用性，而每个业务对数据的一致性和事务要求都不一样，需要单独设计，所以无法将多机房和多中心作为基础服务对外提供

作者回复: 赞同



👍 59



J

2018-09-04

老师你好，请教一个问题，像车联网这种业务，拥有上百万的车载终端，车端会实时上传当前数据，同时服务器会需要控制单独某台车（下行消息），这种场景一般采用哪种架构呢？如果采用MQ的话，一般的消息中间件只能支持几千上万的队列，满足不了要求，请老师指点，谢谢🙏

作者回复: 通常的做法是将连接管理和消息管理分开，消息队列存储需要下发的消息，连接管理从消息队列取消息然后下发，消息中标识要发给哪台车

共 4 条评论 >

👍 25



许鑫

2018-08-07

想问下，类似支付宝的场景如何保证异地金额的强一致性

作者回复: 我不清楚技术细节，从理论上来说，余额和库存异地强一致性会导致性能大幅下降，不适合互联网场景。支付宝有个TCC事务模型，你可以查一下资料。

即使是银行，对并发性能要求没那么高，余额也难以做到异地强一致性，我了解一般都是做备份，异地强一致性太复杂

共 3 条评论 >

👍 15



SMTCode

2018-08-02

CDN和LB是基于通用技术和协议，实现请求的调度与转发等功能。而多机房、多中心，是与业务强相关，对业务有很大侵入性的，只能是服务的实现者根据自己的侧重点，选择合适的技术，有针对性的实现服务的三高指标。

作者回复: 正确👍



👍 15





钱  
2019-09-04

## 课后思考及问题

1: 为什么可以购买负载均衡和 CDN 服务, 但却不能购买多机房和多中心服务?

这个问题, 我第一感觉是个伪问题, 商业社会没什么不能买的, 土地、房屋、武器、公司都能买卖, 多机房和多中心服务更如是。

当然也理解老师的意思, 主要是站在架构师的角度来看, 负载均衡和CDN与业务关联性小, 花钱也相对少, 买来直接就能用, 需要的也多一些, 所以, 购买使用比较顺当自然。

多机房和多中心服务, 实现上和业务关联性大, 除非订制化否则难以买来就用, 如果自己的公司都发展到必须使用多机房和多中心了, 招兵买马自己搞一套才是自然而然的想法。

作者回复: 思考和回答很到位👍👍



👍 13



kyq叶鑫

2018-12-05

老师你好, 你在评论或者专栏中多次说道金融类业务无法做到跨城强一致性, 并以支付宝举例, 我一直有一个疑问: 那么如果一个支付宝用户走出其数据所在大区, 支付宝是如何处理这个用户的数据, 保证业务不出错的呢?

望解惑。

作者回复: 有全局路由会把用户路由到其数据归属的大区



👍 11



LB

2018-08-06

仔仔老师你好, 有两个问题请帮忙解答:

1. 请问银行异地存取款的数据强一致性是怎么做到的? 还可以在国外消费做到强一致性?

2. 银行常用的两地三中心是真正的多中心吗? 如何判断?

谢谢。

作者回复: 1. 异地取款其实只是用户异地操作, 系统还是会回到用户开户行判断余额; 异地存款不需要校验余额, 只需要校验账号, 因此做异地存款是可以的, 因为银行能收到物理上真正的钱, 不用担心数据丢失的问题

2. 两地三中心, 是指同城两个机房是双活, 异地机房是备份, 当同城两个机房都挂掉, 异地机房不能接管业务, 只能用来备份恢复





7



呵呵

2018-08-27

大神问两个问题，1.假如有很多台nginx做负载均衡，那么请求是如何分发到这么多nginx服务上的？2. 对于支付宝的转账业务，需要数据强一致性，那是不是就无法做到跨城市区域存储相关的数据？

作者回复: 1. 硬件负载均衡设备，F5了解一下

2. 是的，跨城不行，但同城多机房可以基于paxos算法设计分布式一致性存储系统，例如oceanbase

共 2 条评论 >

6



凡凡

2018-08-05

多机房多中心，业务相关性太强，很难抽象绝对合适的模型，通用差。一般都是企业按照自己的特定业务，适当的做多中心，或者做到数据异地灾备，或者做到多中心服务。

或许，现在的云服务厂商，会逐渐面向行业，抽象出多种多中心方案，也是可能得。

毕竟云服务厂商在全球范围，已经有了很多机房，如果能抽象出来一定会做尝试的，不论用户多少，这是一种能力。

作者回复: 是的，现在有金融云，游戏云，证券云等解决方案了



6



LB

2018-08-05

仔仔老师，服务器中间件的连接数有限，例如WAS的连接池一共有几千。而咱们说的一台服务器都并发好几万，请问是怎么搞的？谢谢。

作者回复: 一般说并发可以是并发连接，也可以是并发请求，nginx这类基于事件驱动架构的服务器，连接数和请求数都可以上万，以前的apache基于ppc模式，连接数受C10K问题限制，达不到上万



3

敬艺



2018-08-05

即使机房接入CDN，如果CDN支持的网络服务商不够的话，用户访问时不时出现超时。如CDN只接入电信和联通（一时很难覆盖所有网络服务商），用户使用移动网络，访问服务会出现超时。请问大致从几方面优化？

作者回复：接CDN一般都是三线都接的



空档滑行

2018-08-02

负载均衡和cdn虽然复杂，但是局限于技术层面。多中心和多机房则是技术和业务都要涉及到改造，比如文中说的淘宝多中心业务，每个业务模块都要定制，所以提供通用服务是不现实的

作者回复：正确👍



正是那朵玫瑰

2018-08-02

华仔老师：

- 1、DNS负载均衡不一定是地埋级别的吧，如consul就实现了dns服务器，内部服务注册到consul，消费者调用服务可以通过dns服务器做负载均衡，服务故障下线也能及时更新！
- 2、cdn和负载均衡应该是通用的技术吧，多机房和多中心跟业务强关联，没法抽象通用的东西来，就算能，使用起来也不灵活吧！

作者回复：1. 可以做你说的功能，但这样做不常见，出问题可能处理比较麻烦

2. 正解



Monday

2020-10-16

dig baidu.com

结果：

```
; <<<>> DiG 9.11.3-1ubuntu1.13-Ubuntu <<<>> baidu.com
```

```
:: global options: +cmd
:: Got answer:
:: ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 47856
:: flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 1

:: OPT PSEUDOSECTION:
; EDNS: version: 0, flags;; udp: 4000
:: QUESTION SECTION:
;baidu.com.                IN      A

:: ANSWER SECTION:
baidu.com.                300     IN      A      220.181.38.148
baidu.com.                300     IN      A      39.156.69.79

;; Query time: 8 msec
;; SERVER: 172.16.237.4#53(172.16.237.4)
;; WHEN: Fri Oct 16 10:23:55 CST 2020
;; MSG SIZE rcvd: 70
```

作者回复: 你加个vpn再看看, 结果就不同了



👍 2



**LouisLimTJ**

2018-08-06

其实谷歌有些产品号称可以做到同城和跨城高可用了, 比如cloud Spanner, 当然我没有在这些产品针对高可用的问题进行针对性的测试。

回到问题上来, 主要第一. 想用的公司本来就少, 方案贵; 第二. 既然要用也是趋向自搭, 可以针对业务进行定制化处理; 第三. 个人看法, 所提的问题层次最多是PaaS, 而既然同城跨城高可用跟业务有关, 如果云平台技术实力强, 不见得不可以考虑实现SaaS级别的跨城高可用

作者回复: 希望云平台能解决, 目前很多公司都有多活的需求, 但技术上自己又实现不了



👍 3



**Geek\_b04b12**

2018-08-07

比较好奇: 像支付宝银行这些实现异地多中心的情况是如何达到的? 专门部署自己的传输路

线？包括蚂蚁金融、以前用mysql实现异地多数据中心，发现延迟有点大，稍微不对的操作就不同步了、

作者回复: 1. 有钱的公司买专线，非常有钱的公司自己搭建网络

2. 有的数据不能实现多中心，但可以实现数据分区，例如分为南方，北方，西南，西北四个数据大区



👍 2



**Geek\_b04b12**

2018-08-07

前者业务性不大，后者一般都自己搞，请技术顾问，毕竟自己的东西，数据什么的都是自己的核心，还是自己折腾比较好、烧钱、还是自己烧比较好、

今天用了亚马逊的CDN（router 53），发现对全球支持的比较好变成熟，看了阿里云，折腾了半天发现要达到一定的等级、感叹：国内和国外的技术还是有差距的、果断选择了亚马逊！

作者回复: 如果你们的业务聚焦国内，我还是推荐阿里云 😊

共 2 条评论 >

👍 1



**flykyle**

2018-08-02

华仔老师，本地服务（没有固定IP）和云端平台（在阿里云服务器上）之间的通信如何保证高可用？

作者回复: 通信保证高可用比较难，因为线路不是你控制的，是运营商控制的，一般我们是电信联通移动三条线路一起上



👍 1



**宇飞**

2022-06-29

CDN和LB是基于通用技术和协议，实现请求的调度与转发等功能。而多机房、多中心，是与业务强相关，对业务有很大侵入性的，只能是服务的实现者根据自己的侧重点，选择合适的技术，有针对性的实现服务的三高指标。

作者回复: 正解





**William Ning**

2022-05-02

→ xx-api git:(dev) dig baidu.com

```
; <<>> DiG 9.10.6 <<>> baidu.com
```

```
:: global options: +cmd
```

```
:: Got answer:
```

```
:: ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 48292
```

```
:: flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 0
```

```
:: QUESTION SECTION:
```

```
;baidu.com.                IN      A
```

```
:: ANSWER SECTION:
```

```
baidu.com.                427 IN    A    220.181.38.148
```

```
baidu.com.                427 IN    A    220.181.38.251
```

```
:: Query time: 19 msec
```

```
:: SERVER: 192.168.31.1#53(192.168.31.1)
```

```
:: WHEN: Mon May 02 23:53:14 CST 2022
```

```
:: MSG SIZE rcvd: 59
```

