第18讲 | DNS协议：网络世界的地址簿

2018-06-27 刘超

第18讲 | DNS协议：网络世界的地址簿
朗读人：刘超　11′55″ | 5.49M

前面我们讲了平时常见的看新闻、支付、直播、下载等场景，现在网站的数目非常多，常用的网站就有二三十个，如果全部用 IP 地址进行访问，恐怕很难记住。于是，就需要一个地址簿，根据名称，就可以查看具体的地址。

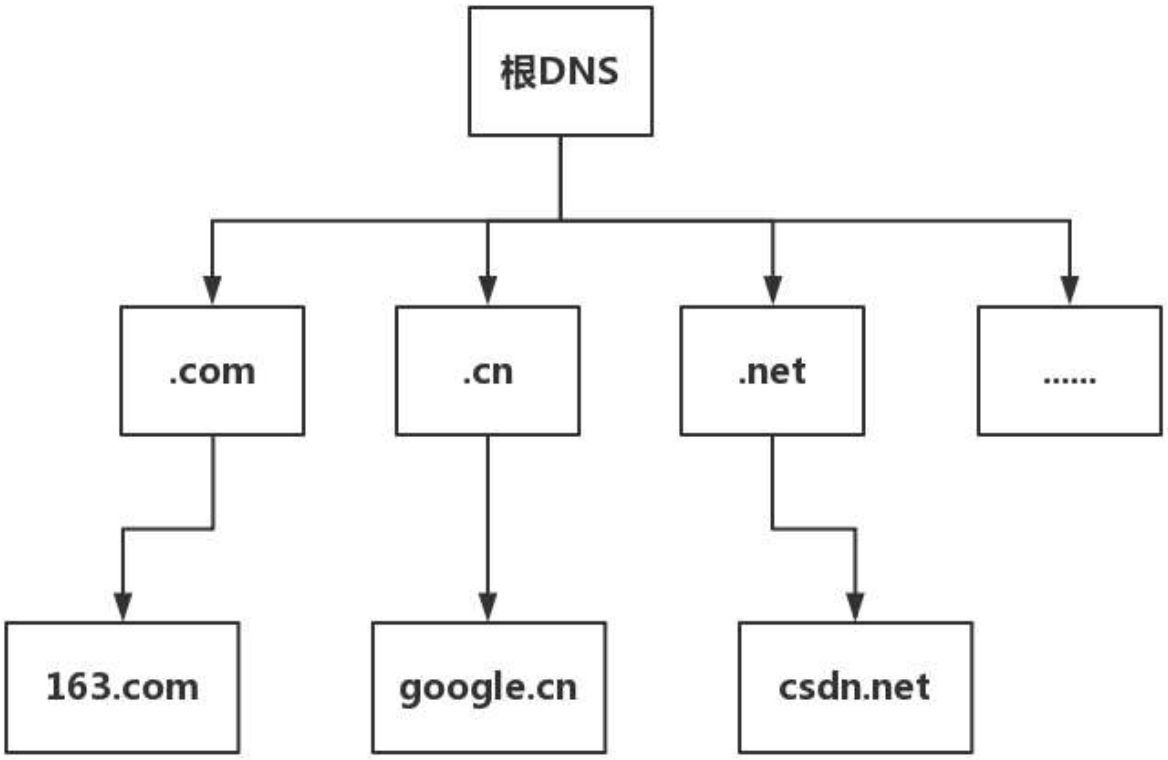
例如，我要去西湖边的“外婆家”，这就是名称，然后通过地址簿，查看到底是哪条路多少号。

DNS 服务器

在网络世界，也是这样的。你肯定记得住网站的名称，但是很难记住网站的 IP 地址，因而也需要一个地址簿，就是DNS 服务器。

由此可见，DNS 在日常生活中多么重要。每个人上网，都需要访问它，但是同时，这对它来讲也是非常大的挑战。一旦它出了故障，整个互联网都将瘫痪。另外，上网的人分布在全世界各地，如果大家都去同一个地方访问某一台服务器，时延将会非常大。因而，DNS 服务器，一定要设置成高可用、高并发和分布式的。

于是，就有了这样树状的层次结构。



- 根 DNS 服务器 ：返回顶级域 DNS 服务器的 IP 地址
- 顶级域 DNS 服务器：返回权威 DNS 服务器的 IP 地址
- 权威 DNS 服务器 ：返回相应主机的 IP 地址

DNS 解析流程

为了提高 DNS 的解析性能，很多网络都会就近部署 DNS 缓存服务器。于是，就有了以下的 DNS 解析流程。

1. 电脑客户端会发出一个 DNS 请求，问 www.163.com 的 IP 是啥啊，并发给本地域名服务器 (本地 DNS)。那本地域名服务器 (本地 DNS) 是什么呢？如果是通过 DHCP 配置，本地 DNS 由你的网络服务商（ISP），如电信、移动等自动分配，它通常就在你网络服务商的某个机房。
2. 本地 DNS 收到来自客户端的请求。你可以想象这台服务器上缓存了一张域名与之对应 IP 地址的大表格。如果能找到 www.163.com，它直接就返回 IP 地址。如果没有，本地 DNS 会去问它的根域名服务器：“老大，能告诉我 www.163.com 的 IP 地址吗？”根域名服务器是最高层次的，全球共有 13 套。它不直接用于域名解析，但能指明一条道路。

3. 根 DNS 收到来自本地 DNS 的请求，发现后缀是 .com，说：“哦，www.163.com 啊，这个域名是由.com 区域管理，我给你它的顶级域名服务器的地址，你去问问它吧。”

4. 本地 DNS 转向问顶级域名服务器：“老二，你能告诉我 www.163.com 的 IP 地址吗？”顶级域名服务器就是大名鼎鼎的比如 .com、.net、 .org 这些一级域名，它负责管理二级域名，比如 163.com，所以它能提供一条更清晰的方向。

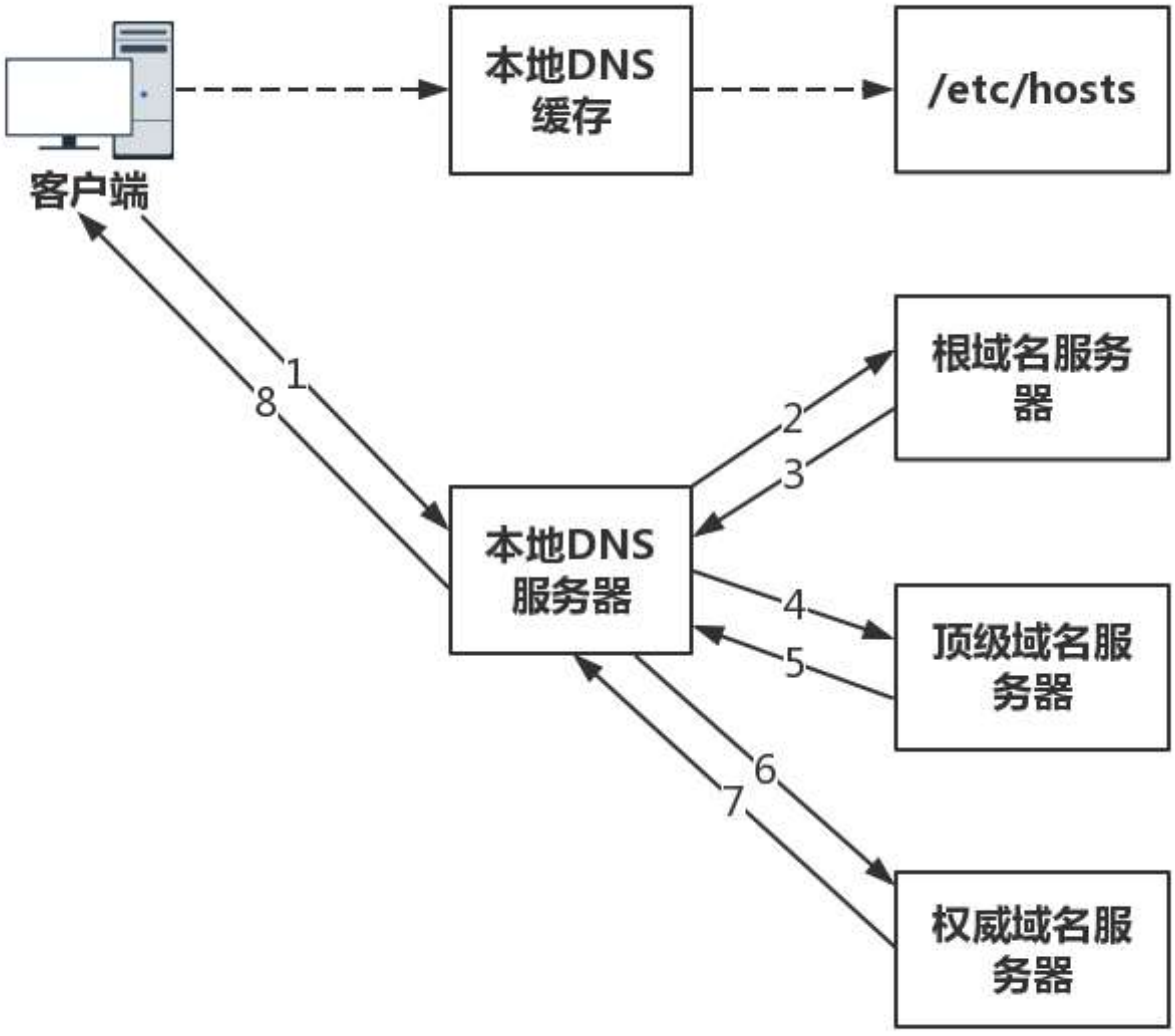
5. 顶级域名服务器说：“我给你负责 www.163.com 区域的权威 DNS 服务器的地址，你去问它应该能问到。”

6. 本地 DNS 转向问权威 DNS 服务器：“您好，www.163.com 对应的 IP 是啥呀？”163.com 的权威 DNS 服务器，它是域名解析结果的原出处。为啥叫权威呢？就是我的域名我做主。

7. 权限 DNS 服务器查询后将对应的 IP 地址 X.X.X.X 告诉本地 DNS。

8. 本地 DNS 再将 IP 地址返回客户端，客户端和目标建立连接。

至此，我们完成了 DNS 的解析过程。现在总结一下，整个过程我画成了一个图。



负载均衡

站在客户端角度，这是一次DNS 递归查询过程。因为本地 DNS 全权为它效劳，它只要坐等结果即可。在这个过程中，DNS 除了可以通过名称映射为 IP 地址，它还可以做另外一件事，就是负载均衡。

还是以访问“外婆家”为例，还是我们开头的“外婆家”，但是，它可能有很多地址，因为它在杭州可以有很多家。所以，如果一个人想去吃“外婆家”，他可以就近找一家店，而不用大家都去同一家，这就是负载均衡。

DNS 首先可以做内部负载均衡。

例如，一个应用要访问数据库，在这个应用里面应该配置这个数据库的 IP 地址，还是应该配置这个数据库的域名呢？显然应该配置域名，因为一旦这个数据库，因为某种原因，换到了另外一台机器上，而如果有多个应用都配置了这台数据库的话，一换 IP 地址，就需要将这些应用全部修改一遍。但是如果配置了域名，则只要在 DNS 服务器里，将域名映射为新的 IP 地址，这个工作就完成了，大大简化了运维。

在这个基础上，我们可以再进一步。例如，某个应用要访问另外一个应用，如果配置另外一个应用的 IP 地址，那么这个访问就是一对一的。但是当被访问的应用撑不住的时候，我们其实可以部署多个。但是，访问它的应用，如何在多个之间进行负载均衡？只要配置成为域名就可以了。在域名解析的时候，我们只要配置策略，这次返回第一个 IP，下次返回第二个 IP，就可以实现负载均衡了。

另外一个更加重要的是，DNS 还可以做全局负载均衡。

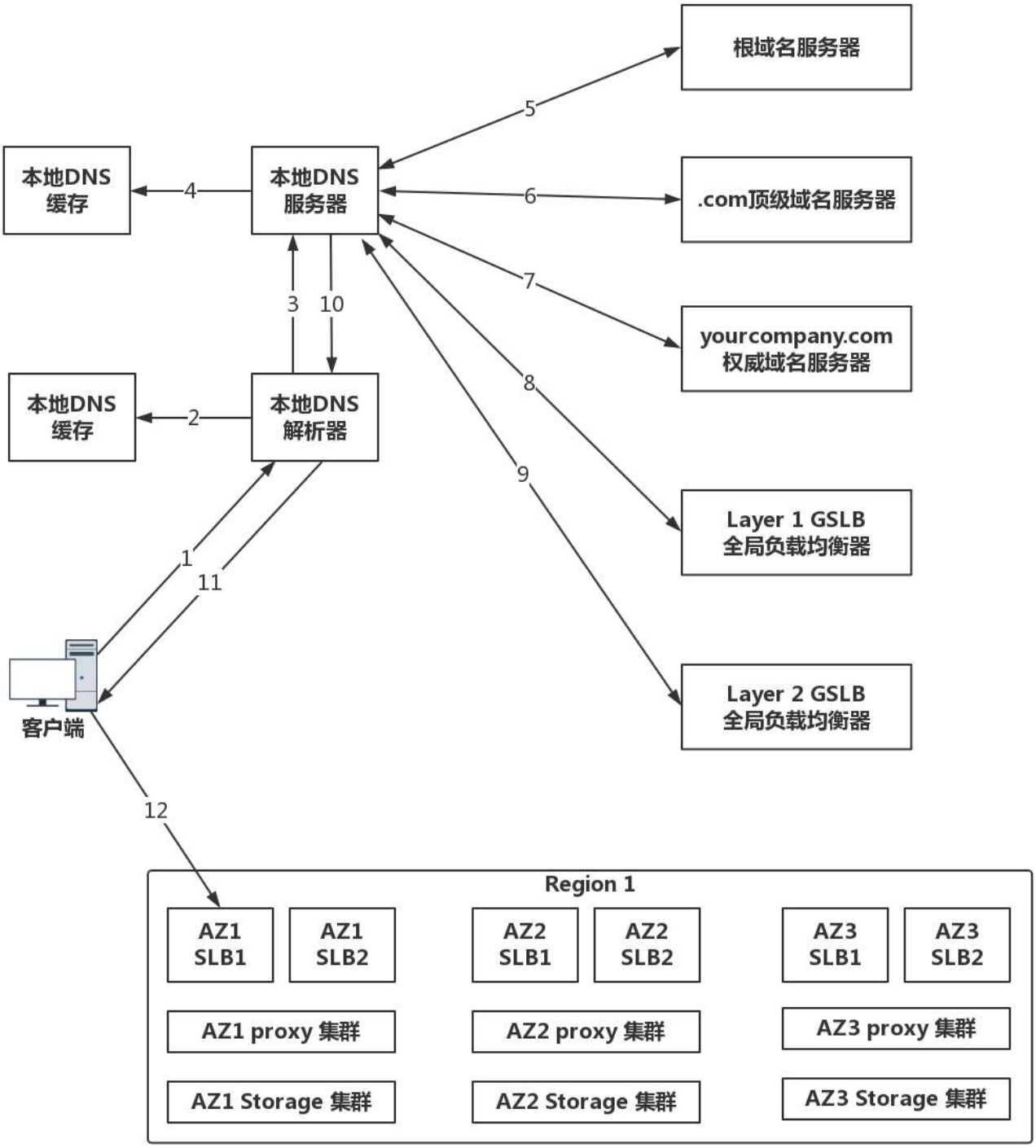
为了保证我们的应用高可用，往往会部署在多个机房，每个地方都会有自己的 IP 地址。当用户访问某个域名的时候，这个 IP 地址可以轮询访问多个数据中心。如果一个数据中心因为某种原因挂了，只要在 DNS 服务器里面，将这个数据中心对应的 IP 地址删除，就可以实现一定的高可用。

另外，我们肯定希望北京的用户访问北京的数据中心，上海的用户访问上海的数据中心，这样，客户体验就会非常好，访问速度就会超快。这就是全局负载均衡的概念。

示例：DNS 访问数据中心中对象存储上的静态资源

我们通过 DNS 访问数据中心中对象存储上的静态资源为例，看一看整个过程。

假设全国有多个数据中心，托管在多个运营商，每个数据中心三个可用区（Available Zone）。对象存储通过跨可用区部署，实现高可用性。在每个数据中心中，都至少部署两个内部负载均衡器，内部负载均衡器后面对接多个对象存储的前置服务器（Proxy-server）。



1. 当一个客户端要访问 object.yourcompany.com 的时候，需要将域名转换为 IP 地址进行访问，所以它要请求本地 DNS 解析器。
2. 本地 DNS 解析器先查看看本地的缓存是否有这个记录。如果有则直接使用，因为上面的过程太复杂了，如果每次都要递归解析，就太麻烦了。
3. 如果本地无缓存，则需要请求本地的 DNS 服务器。
4. 本地的 DNS 服务器一般部署在你的数据中心或者你所在的运营商的网络中，本地 DNS 服务器也需要看本地是否有缓存，如果有则返回，因为它也不想把上面的递归过程再走一遍。

5. 至 7. 如果本地没有，本地 DNS 才需要递归地从根 DNS 服务器，查到.com 的顶级域名服务器，最终查到 yourcompany.com 的权威 DNS 服务器，给本地 DNS 服务器，权威 DNS 服务器按说会返回真实要访问的 IP 地址。

对于不需要做全局负载均衡的简单应用来讲，yourcompany.com 的权威 DNS 服务器可以直接将 object.yourcompany.com 这个域名解析为一个或者多个 IP 地址，然后客户端可以通过多个 IP 地址，进行简单的轮询，实现简单的负载均衡。

但是对于复杂的应用，尤其是跨地域跨运营商的大型应用，则需要更加复杂的全局负载均衡机制，因而需要专门的设备或者服务器来做这件事情，这就是全局负载均衡器（GSLB，Global Server Load Balance）。

在 yourcompany.com 的 DNS 服务器中，一般是通过配置 CNAME 的方式，给 object.yourcompany.com 起一个别名，例如 object.vip.yourcomany.com，然后告诉本地 DNS 服务器，让它请求 GSLB 解析这个域名，GSLB 就可以在解析这个域名的过程中，通过自己的策略实现负载均衡。

图中画了两层的 GSLB，是因为分运营商和地域。我们希望不同运营商的客户，可以访问相同运营商机房中的资源，这样不跨运营商访问，有利于提高吞吐量，减少时延。

1. 第一层 GSLB，通过查看请求它的本地 DNS 服务器所在的运营商，就知道用户所在的运营商。假设是移动，通过 CNAME 的方式，通过另一个别名 object.yd.yourcompany.com，告诉本地 DNS 服务器去请求第二层的 GSLB。
2. 第二层 GSLB，通过查看请求它的本地 DNS 服务器所在的地址，就知道用户所在的地理位置，然后将距离用户位置比较近的 Region 里面，六个内部负载均衡（SLB，Server Load Balancer）的地址，返回给本地 DNS 服务器。
3. 本地 DNS 服务器将结果返回给本地 DNS 解析器。
4. 本地 DNS 解析器将结果缓存后，返回给客户端。
5. 客户端开始访问属于相同运营商的距离较近的 Region 1 中的对象存储，当然客户端得到了六个 IP 地址，它可以通过负载均衡的方式，随机或者轮询选择一个可用区进行访问。对象存储一般会有三个备份，从而可以实现对存储读写的负载均衡。

小结

好了，这节内容就到这里了，我们来总结一下：

- DNS 是网络世界的地址簿，可以通过域名查地址，因为域名服务器是按照树状结构组织的，因而域名查找是使用递归的方法，并通过缓存的方式增强性能；

- 在域名和 IP 的映射过程中，给了应用基于域名做负载均衡的机会，可以是简单的负载均衡，也可以根据地址和运营商做全局的负载均衡。

最后，给你留两个思考题：

1. 全局负载均衡为什么要分地址和运营商呢？
2. 全局负载均衡使用过程中，常常遇到失灵的情况，你知道具体有哪些情况吗？对应应该怎么来解决呢？

我们的专栏更新过半了，不知你掌握得如何？每节课后我留的思考题，你都有没有认真思考，并在留言区写下答案呢？我会从已发布的文章中选出一批认真留言的同学，赠送学习奖励礼券和我整理的独家网络协议知识图谱。

欢迎你留言和我讨论。趣谈网络协议，我们下期见！

精选留言

upstream 1

问题1，分成两级还是在于国内网络跨运营商访问慢
问题2，如果用户是上海电信，用户本地dns不是用的电信dns服务器，负载均衡调度会不准确。针对app的可以httpdns
针对网站的除非能支持edns协议才比较好。
另外当用户使用阿里云这类anycast 地址时，不清楚gslb调度策略如何的

2018-06-27

upstream 1

虽然有了域名，但对于大多数人而言，能记住的就那么一两个网站地址。于是，打开baidu.com 搜索想要打开的网站，也就不奇怪了。打开baidu.com 搜索谷歌，再去访问Google 也就成为很多人的默认行为。说白了，也还是记不住域名。
Baidu.com 搜索qq邮箱的大有人在

2018-06-27

羽毛 0

“我们希望不同运营商的客户，可以访问相同运营商机房中的资源”为什么不同运营商访问相同的运营商机房中的资源

2018-06-27

进化论 0

分运营商，是不是解决跨运营商访问慢的问题啊！

第二个是因为有缓存？

2018-06-27

灰灰 0

写得太好了，刘超老师，你还有其他技术专栏吗？或者有其他推荐的技术专栏吗？

2018-06-27

Jobs 0

1.防止单个地区或运营商的流量过爆而挂了。
2.服务器故障，网络故障如中间路由器故障，SLB设备故障，断电故障等。Solution:返回多重A记录？

2018-06-27

浪子恒心 0

还是没太想明白为什么分了两层全局负载，一层负载不能全解决吗？全局负载一般在用户配置了公共DNS服务时会失效吧，通过HTTPDNS能解决，但移动端需要配合做改造，难度比较大。

2018-06-27

LEON 0

超哥，目前我国DNS SEC可以使用了吗？最早听说我们运营商不支持。

2018-06-27

LEON 0

超哥，请教个问题，GSLB在判断站点故障后切换A记录到可用站点(GSLB上的A记录配置的TTL30秒)。但是部分用户并没有切换成功，是因为GSLB提供的A记录的TTL让部分运营商的local dns 改成了12个小时，导致用户本地TTL缓存长，必须清缓存换local dns。用户端的local d

ns配置不可控制。这种情况下除了不使用HTTP DNS改造外，有什么更好的办法？如果使用EDNS不知道目前支持情况怎么样？谢谢

2018-06-27

coldfire 0

从真名到别名怎么完成均衡负载，感觉没讲清楚，对于小白来说这就是一笔带过留了个影响

2018-06-27

龚极客 0

问题1，因为机房的网络对应的运营商速度不同

问题2，如果某一个节点不可达，需要弹性伸缩，数据中心master 和slave的延迟也是个问题。

2018-06-27

叮当 0

同求知识图谱阳光普照

2018-06-27

free 0

1 跨运营商慢的问题

2 gslb自身的高可用

2018-06-27

问题究竟系边度 0

分地址和运营商应该是使用适用的网络提高访问效率。

这种分配方式只是直接就近分配。并没有根据对应链路质量进行监测和调整。

是否应该部署两个gslb预防单点故障和网络攻击？

2018-06-27

问题究竟系边度 0

分地址和运营商应该是使用适用的网络提高访问效率。

这种分配方式只是直接就近分配。并没有根据对应链路质量进行监测和调整。

是否应该部署两个gslb预防单点故障和网络攻击？

2018-06-27

Kail 0

1.CDN?

2.DNS 劫持？HTTPDNS

2018-06-27

iiwoai 0

图谱应该都给赠送才对啊

2018-06-27

一箭中的 0

思考题2:当域名对应的ip地址更换了，那么dns缓存内容就失效了，此时会导致无法访问到该网址。所以dns缓存都会超时时间，超时后缓存就会失效。

2018-06-27

Jason 0

每篇都能学到新知识。谢超哥。

2018-06-27

LongXiaJun 0

😀😀😀现在一周三天都在期待专栏更新~

2018-06-27