

第10讲 | 深入区块链技术（二）：P2P网络

2018-04-16 陈浩

深入浅出区块链

[进入课程 >](#)



讲述：黄洲君

时长 11:19 大小 5.19M



在上一篇文章中，我大致讲解了一下区块链技术的几个核心要素。P2P 网络协议、分布式一致性算法（共识机制）、加密签名算法、账户与存储模型。今天我们就来看看区块链技术的第一个核心要素：P2P 网络。

如果我们简单来看 P2P 技术，它的应用领域已经非常广泛了，从流媒体到点对点通讯、从文件共享到协同处理，多种领域都有它的身影出现。

同样的，P2P 的网络协议也有很多，比较常见的有 BitTorrent、ED2K、Gnutella、Tor 等，也就是我们常说的 BT 工具和电驴。

比特币、以太坊等众多数字货币都实现了属于自己的 P2P 网络协议，但是这种模式并不同于以上讨论的 P2P 网络协议，所以我们今天讨论的重点主要是区块链技术的 P2P 技术，也

就是比特币和以太坊的 P2P 网络。

由于区块链的 P2P 网络技术知识繁多，我们主要提炼其中的四个内容进行讲解：区块链的网络连接与拓扑结构、节点发现、局域网穿透与节点交互协议。

希望读完本篇可以让你对目前成熟的区块链 P2P 网络的拓扑结构以及运行原理有个大体的认知。

网络连接与拓扑结构

1. 网络连接

除去少数支持 UDP 协议的区块链项目外，绝大部分的区块链项目所使用的底层网络协议依然是 TCP/IP 协议。

所以从网络协议的角度来看，区块链其实是基于 TCP/IP 网络协议的，这与 HTTP 协议、SMTP 协议是处在同一层，也就是应用层。

在“区块链的常见误区”这篇文章中，我们提到了“区块链是否会颠覆互联网”这一说法，如果要是认真分析的话，它颠覆的层面其实最多只到 HTTP 协议，不能再多了。

以 HTTP 协议为代表的、与服务端的交互模式在区块链上被彻底打破了，变更为完全的点对点拓扑结构，这也是以太坊提出的 Web3.0 的由来。

比特币的 P2P 网络是一个非常复杂的结构，考虑到矿池内部的挖矿交互协议与轻节点。我们仅仅讨论全节点这种场景下的 P2P 网络发现与路由。

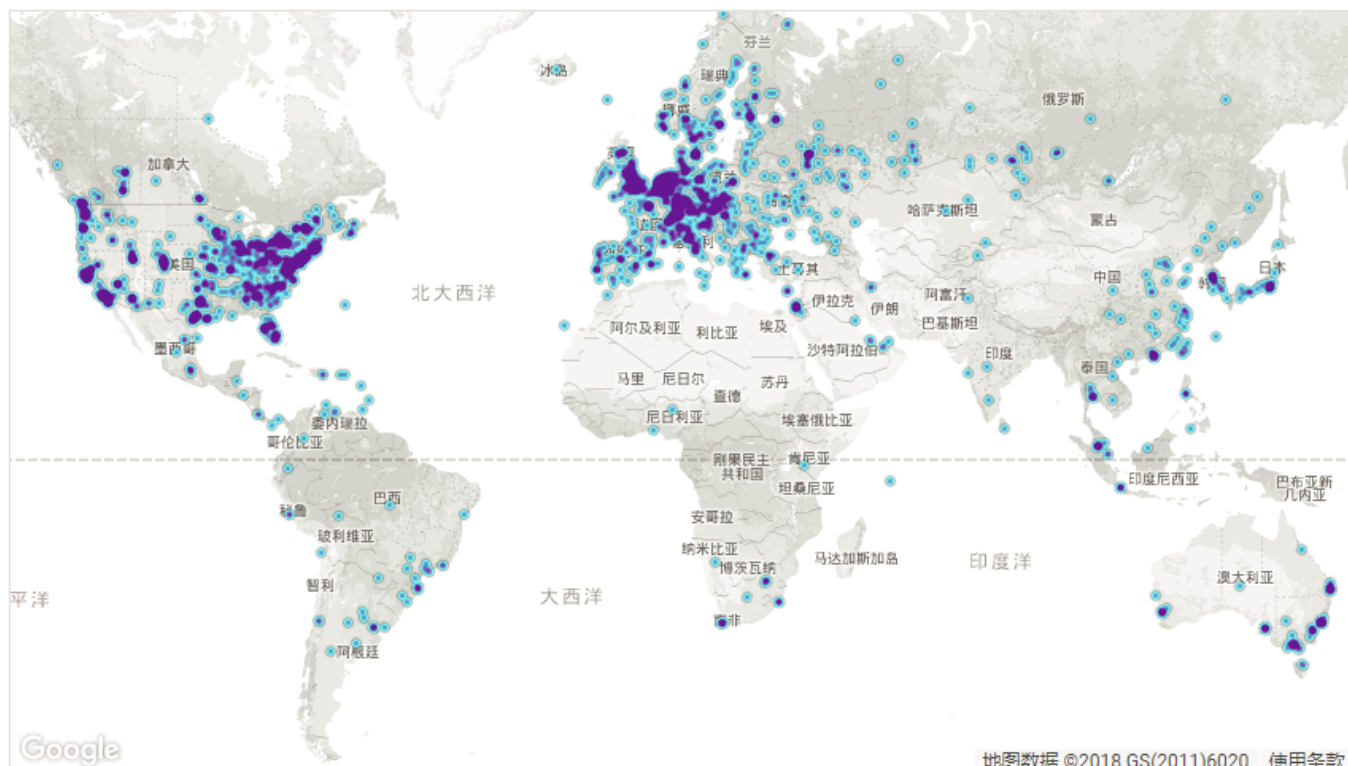
比特币的 P2P 网络基于 TCP 构建，主网默认通信端口为 8333。

以太坊的 P2P 网络则与比特币不太相同，以太坊 P2P 网络是一个完全加密的网络，提供 UDP 和 TCP 两种连接方式，主网默认 TCP 通信端口是 30303，推荐的 UDP 发现端口为 30301。

2. 拓扑结构

P2P 网络拓扑结构有很多种，有些是中心化拓扑，有些是半中心化拓扑，有些是全分布式拓扑结构。

□比特币全节点组成的网络是一种全分布式的拓扑结构，节点与节点之间的传输过程更接近“泛洪算法”，即：交易从某个节点产生，接着广播到临近节点，临近节点一传十十传百，直至传播到全网。



Map shows concentration of reachable Bitcoin nodes found in countries around the world.

LIVE MAP

(比特币全球节点图，图来自网络)

全节点与 SPV 简化支付验证客户端之间的交互模式，更接近半中心化的拓扑结构，也就是 SPV 节点可以随机选择一个全节点进行连接，这个全节点会成为 SPV 节点的代理，帮助 SPV 节点广播交易。

节点发现

节点发现是任何区块链节点接入区块链 P2P 网络的第一步。这与你孤身一人去陌生地方旅游一样，如果没有地图和导航，那你只能拽附近的人问路，“拽附近的人问路”的这个动作就可以理解成节点发现。

节点发现可分为初始节点发现，和启动后节点发现。初始节点发现就是说你的全节点是刚下载的，第一次运行，什么节点数据都没有。启动后发现表示正在运行的钱包已经能跟随网络动态维护可用节点。


1. 初始节点发现

在比特币网络中，初始节点发现一共有两种方式。

第一种叫做 DNS-seed，又称 DNS 种子节点，DNS 就是中心化域名查询服务，比特币的社区维护者会维护一些域名。

比如 seed.bitcoin.sipa.be 这个域名就是由比特币的核心开发者 Sipa 维护的，如果我们通过 nslookup 会发现大约二十多个 A 纪录的 IPv4 主机地址。

我们通过 nc 命令尝试连接域名下的某个主机的 8333 端口会发现连接成功，运行结构如下。

 复制代码

```
1
2 X chenhao@chenhaodeMacBook-Pro ~ nc -nv 149.202.179.35 8333
3 found 0 associations
4 found 1 connections:
5     1: flags=82<CONNECTED,PREFERRED>
6     outif en0
7     src 192.168.1.104 port 62125
8     dst 149.202.179.35 port 8333
9     rank info not available
10    TCP aux info available
11 Connection to 149.202.179.35 port 8333 [tcp/*] succeeded!
```

好的，到目前为止我们已经手动做了一遍初始节点发现的工作，这些操作是由比特币的代码完成的。

第二种方式就是，代码中硬编码（hard-code）了一些地址，这些地址我们称之为种子节点（seed-node），当所有的种子节点全部失效时，全节点会尝试连接这些种子节点。

用在以太坊中，思路也大致相同，也是在代码中硬编码（hard-code）了一些种子节点做类似的工作。

2. 启动后节点发现

在 Bitcoin 的网络中，一个节点可以将自己维护的对等节点列表 (peer list) 发送给临近节点，所以在初始节点发现之后，你的节点要做的第一件事情就是向对方要列表：“快把你的节点列表给我复制一份。”

所以在每次需要发送协议消息的时候，它会花费固定的时间尝试和已存的节点列表中的节点建立链接，如果有任何一个节点在超时之前可以连接上，就不用去 DNS seed 获取地址，一般来说，这种可能性很小，尤其是全节点数目非常多的情况下。

而在以太坊网络中，也会维护类似的一个节点列表 (NodeTable)，但是这个节点列表与比特币的简单维护不同，它采用了 P2P 网络协议中一个成熟的算法，叫做 Kademlia 网络，简称 KAD 网络。

它使用了 DHT 来定位资源，全称 Distributed Hash Table，中文名为分布式哈希表。KAD 网络会维护一个路由表，用于快速定位目标节点。由于 KAD 网络基于 UDP 通信协议，所以以太坊节点的节点发现是基于 UDP 的，如果找到节点以后，数据交互又会切换到 TCP 协议上。

3. 黑名单与长连接

公有区块链面临的网络环境是非常开放的，任何人只要下载好钱包，打开运行就进入了这个 P2P 网络，这也会带来被攻击的可能。

所以在比特币的代码中，会有一段去控制逻辑，你可以手动将你认为可疑的节点移除并加入禁止列表，同时去配置可信的节点。当然，以上并不属于客户端的标准协议的一部分，任何人都可以实现属于自己的 P2P 网络层。

以太坊上有针对账户进行的黑名单处理，但是这属于业务层。我没有找到很详尽的资料，所以你有兴趣的话，可以自己尝试一下。

不过总的来说，黑名单我们也可以通过操作系统的防火墙去处理，这并不算一个特别棘手的问题。

局域网穿透

前面我们说到了区块链的 P2P 网络结构是一种全分布式的拓扑结构。但是，如今我们的网络环境是由局域网和互联网组成的。也就是说，当你在局域网运行一个区块链节点，在公网

是发现不了的，公网上的节点只能被动接受连接，并不能主动发起连接。

如果这个局域网是你控制的，那么很好说，咱们只需要在 VPC 网络中配置路由，将公网 IP 和端口映射到局域网中你的 IP 和端口即可。

这个条件是非常苛刻的，那么到底有没有一种方案可以自行建立映射呢？答案是：有，就是 NAT 技术和 UPnP 协议。

NAT 技术非常常见，这里使用的是源 NAT，简而言之就是替换 TCP 报文中的源地址并映射到内网地址。

UPnP 是通用即插即用（Universal Plug and Play）的缩写，它主要用于设备的智能互联互通，所有在网络上的设备马上就能知道有新设备加入。

这些设备彼此之间能互相通信，更能直接使用或者控制它，一切都不需要人工设置。有关 UPnP 的资料比较多，这里就不赘述了，你可以自行搜索相关的信息。

比特币和以太坊均使用了 UPnP 协议作为局域网穿透工具，只要局域网中的路由设备支持 NAT 网关功能、支持 UPnP 协议，即可将你的区块链节点自动映射到公网上。

节点交互协议

一旦节点建立连接以后，节点之间的交互是遵循一些特定的命令，这些命令写在消息的头部，消息体写的则是消息内容。

命令分为两种，一种是请求命令，一种是数据交互命令。

节点连接完成要做的第一件事情叫做握手操作。这一点在比特币和以太坊上的流程是差不多的，就是相互问候一下，提供一些简要信息。

比如先交换一下版本号，看看是否兼容。只是以太坊为握手过程提供了对称加密，而比特币没有。

握手完毕之后，无论交互什么信息，都是需要保持长连接的，在比特币上有 PING/PONG 这两种类型的消息，这很明显就是用于保持节点之间长连接的心跳而设计的；而在以太坊的设计中，将 PING/PONG 协议移到了节点发现的过程中。

请求命令一般分为发起者请求，比如比特币中的 `getaddr` 命令是为了获取对方的可用节点列表，`inv` 命令则提供了数据传输，消息体中会包含一个数据向量。

我们说区块链最重要的功能就是同步区块链，而同步区块恰巧是最考验 P2P 网络能力的。区块同步方式分为两种，第一种叫做 HeaderFirst，它提供了区块头先同步，同步完成以后再从其他节点获得区块体。

第二种叫做 BlockFirst，这种区块同步的方式比较简单粗暴，就是从其他节点获取区块必须是完整的。第一种方案提供了较好的交互过程，减轻了网络负担。这两种同步方式会直接体现在节点交互协议上，他们使用的命令逻辑完全不同。

总结

今天我与你分享了区块链的 P2P 网络结构与节点交互过程，一般 P2P 网络技术要解决两个主要问题，第一是资源定位，第二是资源获取，这一篇文章也是主要围绕这两点展开，其中节点发现和局域网穿透是属于资源定位问题，节点交互协议是属于资源获取问题。

在这一篇文章中，我仅以比特币和以太坊为例进行分享，虽然区块链项目比较多，但是他们要做的事情大多是类似的，比如以太坊是改进版的实现，而比特币使用了简单版实现。

P2P 网络模块作为所有区块链的最底层模块，直接决定了整个区块链网络的稳定性。区块链网络是一个网状分布式的结构，与互联网结构有点相似，那么，亲爱的读者，我们是不是可以设计一个节点爬虫，去爬全网节点呢？你可以给我留言，我们一起讨论。

感谢你的收听，我们下次再见。


参考链接：<https://bitnodes.earn.com/>

深入浅出区块链

你的区块链入门第一课

陈浩 元界 CTO



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 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。

上一篇 第9讲 | 深入区块链技术（一）：技术基础

下一篇 第11讲 | 深入区块链技术（三）：共识算法与分布式一致性算法

精选留言 (17)

 写留言



许星昊

2018-08-02

 5

第二种方式就是，代码中硬编码（hard-code）了一些地址，这些地址我们称之为种子节点（seed-node），当所有的种子节点全部失效时，全节点会尝试连接这些种子节点。

...

展开

作者回复：不好意思，应该是基于DNS的种子节点全部失效时，会尝试连接Hard code的种子节点。



阿痕

2018-04-21

👍 3

爬虫完全没问题，可以设计一个递归算法，从一个peer节点出发，找到它相连的N个节点，再从这N个节点出发，以此类推。理论上应该可以找到所有的节点。

展开 ▾



区块链先锋

2018-05-30

👍 2

节点和区块有什么区别

展开 ▾

作者回复: 节点是启动的MySQL服务，区块是里面的表

◀ ▶



guanhua

2018-05-12

👍 2

陈老师，请问区块同步方式1中，先同步区块头过程中就可以进行合法性验证了吗？之后再同步区块体就直接复制吗？

每个区块头里面不都包含对上一个整个区块的哈希吗？

展开 ▾

作者回复: 每个头都包含上个区块的整块哈希的。

◀ ▶



Aaron

2018-04-16

👍 2

根据节点的发现机制，完全可以爬取所有节点的信息

展开 ▾

作者回复: Sipa有一套简单的代码，叫bitcoin-seeder。可以参考

◀ ▶



呼啦斯卡

2018-09-03

👍

tcp/ip不是应用层协议

展开 ▾

作者回复: 是的, 与HTTP, SMTP协议在同一层, 应用层。



wahaha

2018-07-01



没有公网IP的两个节点不能用TCP直接互通吧? 用UDP可以打洞来直接互通, 不知有哪种区块链支持UDP打洞?

作者回复: 有的呀, uunp协议



张珩

2018-06-29



看起来像是gossip协议

展开 ▾

作者回复: 是的呢



Clancey

2018-06-26



陈老师, 您好, NDN+区块链是否能从ip层改变整个网络架构?

展开 ▾



K.o

2018-05-23



请问1、一个节点要与多少个节点保持长连接? 2、如果一个节点发起广播, 如何保证不被其他节点循环广播同一个消息

作者回复: 你好, 这个是算法可配置的, 一般是8个或更多。



慢摇哥哥

2018-05-20

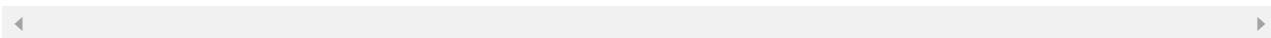


陈老师，两个问题请教：bitnodes.earn.com显示比特币网络上才1万出头的节点，会不会太少有安全问题；另外，有很多节点的NETWORK显示Hangzhou Alibaba Advertise Co., Ltd，这是因为这些矿工运行在阿里云的原因？

展开 ∨

作者回复: 一万个节点是全节点，属于核心节点，类似电信的骨干网络。算上轻节点，不开放的节点，可能在百万以上。

是的，很多开发和测试节点会在云上搭建，生产环境也首选在云上的。



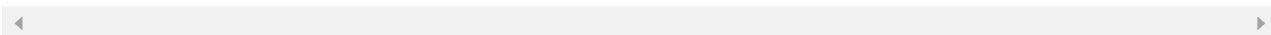
guanhua

2018-05-12



陈老师，请问，同步区块的第一种方法，先同步区块头时就进行合法性验证吗？之后再同步区块体就直接复制吗？

作者回复: 先同步头可以简单验证交易的存在和有效，完整同步依然要检验交易的有效性。



王由华

2018-05-01

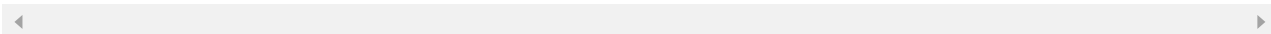


借前面读者的问题，"在挖出前，交易记录保存在普通节点还是核心服务器上?" 以及您的回答"保存在矿池的服务器中"。我想问:1) 交易记录只能保存在矿池服务器中吗？2) 普通节点与矿池服务器的区别是什么？3) 能详细描述下交易记录被发布到服务器的流程吗？

作者回复: 1-只要是全节点，都有的。

2-实际上全节点都一样，对等的，你也可以选择挖矿，只是中奖概率几乎为零。

3-就是全网广播交易，等待被打包的过程，如果被矿工打包，则会从内存中移除。



ytl

2018-04-22



区块链公链都可以做爬虫，获得经济数据。

展开 ▾

作者回复: 爬虫只能过去节点信息哦，账本本来就是公开的，直接解析分析即可。



刘诗晓osca...

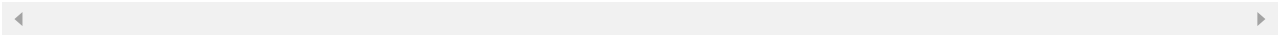
2018-04-21



您好，请问 矿工解出一个区块后获得记录近期交易的权利和收益，那么在挖出前，交易记录保存在普通节点还是核心服务器上?矿工成功打包交易信息的首次广播人是谁?

展开 ▾

作者回复: 保存在矿池的服务器上，叫做memory pool。首次广播就是挖出块的矿工自己



马蹄莲子

2018-04-16



应该可以做一个爬虫吧 😊

展开 ▾



oTo123

2018-04-16



挺好的

展开 ▾