原文链接：<https://cloud.tencent.com/developer/news/292539>

**深入浅出区块链之详解P2P网络**

文章来源：企鹅号 - 区块链兄弟

**公众号回复“1”，拉你进区块链技术讨论微信群**

作者：纳兰少

本文约3800字+，阅读（观看）需要22分钟

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。

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。

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。

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。

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可以控制的，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。

**文章发布只为分享区块链技术内容，版权归原作者所有，观点仅代表作者本人，绝不代表区块链兄弟赞同其观点或证实其描述。**

* 发表于: 2018-08-04
* 原文链接：https://kuaibao.qq.com/s/20180804B1GRZ200?refer=cp\_1026
* 腾讯「云+社区」是腾讯内容开放平台帐号（企鹅号）传播渠道之一，根据[《腾讯内容开放平台服务协议》](https://om.qq.com/notice/a/20160429/047194.htm)转载发布内容。
* 如有侵权，请联系 yunjia\_community@tencent.com 删除。