基于区块链的分布式文件系统发展道路研究

刘轩铭，3180106071

（浙江大学，计算机科学与技术学院，软件工程专业）

前　言：分布式文件系统是指文件系统管理的物理存储资源不一定直接连接在本地节点，而是通过网络进行分布式文件存储和转发的系统整体。P2P 作为最早的解决方案被应用于分布式文件系统领域，但是它存在着缺乏激励机制，安全性能较差等问题。作为新兴的技术领域，区块链被应用于分布式文件系统的研究中。典型的区块链分布式文件系统如 IPFS、Swarm 在结构上分为七个层级，其具有的激励层和共识层创新性地为分布式文件系统发展赋能。以 IPFS 为代表的分布式文件系统应用已经落地并发展得很好，但该技术仍存在着隐私和访问控制技术上的不足。

一、　最初的尝试：点对点分布式文件系统介绍

分布式文件系统（Distributed File System，DFS）是指文件系统管理的物理存储资源不一定直接连接在本地节点上，而是通过计算机网络与节点（可简单的理解为一台计算机）相连，由此组成的完整而有层次的文件系统。在分布式文件系统中，存储资源和系统客户端分散在网络中。每个用户既是系统中存储数据的创建者又是消费者。

到目前为止，最大的分布式文件系统是超文本传输协议（HTTP），它是用于上传数据的 Web 服务器。然后，其他对等方可以访问特定数据。但是，该协议却存在许多问题。例如，为了确保 Web 服务器中的数据可访问性，需要支付维护费用。随着数据受欢迎程度的增加，这种维护成本也随之增加。另一个问题是客户的信息传播负担比较重。此外，中心服务器（群）的稳定性和鲁棒性也需要经受非常苛刻的考验。

目前，世界上已经拥有了许多致力于构建更好的分布式文件系统的尝试。其中，点对点（P2P）服务的惊人普及和研究，使分布式文件系统的实现令人兴奋且充满希望。

例如，作为最成功的 P2P 分布式文件系统之一，BitTorrent 支持超过 1 亿在线用户。与传统的 HTTP 服务不同，在 BitTorrent 中，分配器或文件的持有者将文件发送给其中一名用户，再由这名用户转发给其它用户，用户之间相互转发自己所拥有的文件部分，直到每个用户的下载都全部完成。这种方法可以使下载服务器同时处理多个大体积文件的下载请求，而无须占用大量带宽。它可以大规模部署，有成千上万个节点链接和互动。

但是，传统的 P2P 形式分布式文件系统存在一些弊端：

1) 下载不稳定，这限制了 BitTorrent 在特定场合下的广泛使用。

2) 无法验证文件发布者，并且很难保证下载内容的可信性。

3) 没有激励机制，种子节点不因共享其带宽和存储资源而获得奖励。

二、　基于区块链的分布式文件系统结构介绍

近年来，区块链已成为行业和学术界的流行语，区块链和分布式文件系统的结合正成为一种有前途的解决方案。其原因是，区块链有着对数据的加密和不可篡改数据的特点，这使得系统有着较高的安全性；此外，其共识机制所需要的“挖矿”等过程为系统提供了激励层。

当前，流行的基于区块链的分布式文件系统包括 IPFS，Swarm 等。IPFS 是一个对等分布式文件系统，用于存储和访问文件，网站，应用程序和数据。Swarm 是一个基于以太坊的分布式存储平台和内容分发服务工具。

一般而言，流行的基于区块链的分布式文件系

统都分为如下几个主要的层次：身份层，数据层，数据交换层，网络层，路由层，共识层和激励层。每一层都是分布式文件系统的关键模块。

1. 身份层

基于区块链的分布式文件系统本质上还是 P2P 的网络。在这样的网络中，首先我们需要识别网络中的每个节点和节点的内容——每个节点都必须由唯一的标识符标识从而保证身份不发生冲突。

这个问题一般用哈希算法和数字签名机制进行解决（这和传统区块链应用是类似的）。例如，在 IPFS 中，使用公钥的加密哈希值，即 NodeId 来标识每个节点。蜂拥而至

2. 路由层

其次，需要解决的问题是节点和节点之间的路由问题。

通常，分布式文件系统的路由层的功能包括：

1）维护对等连接拓扑图，以便可以定位特定的对等体和数据对象（这和我们在计算机网络课程中学习到的内容是类似的）；

2）响应来自本地和远程对等体的查询；

3）通过分布式哈希表进行通信。

IPFS 采用了基于分布式哈希表（DSHT）来解决这个问题。位于对等节点中的 DSHT 可以帮助一个特定的节点找到对等方的网络地址。

而 Swarm 使用分布式原像存档（DPA）技术实现其路由层。

3. 网络层

这一层的作用是保证对等节点之间的连接不中断。从而数据和文件能够在网络中稳定的传输。

其使用的方法和传统的 P2P 网络是类似的。

4. 数据层

由于文件是分布式的，所以我们不能把一个很大的文件直接进行分发和存储。事实上，分布式文件系统会将文件进行分块，然后分别储存，之后查找数据时，需要将文件块重新合并。

IPFS 中有多个级别的数据模型：

•块：任意大小的数据。

•列表：块或其他列表的集合。

•树：块，列表或其他树的集合。

这样的数据模型和常见的版本控制工具 Git 是类似的。基于此数据模型，IPFS 系统使用默克尔树存储数据。默克尔树是区块链中常见的技术应用。它有着防篡改，便于验证数据真实性等特点。

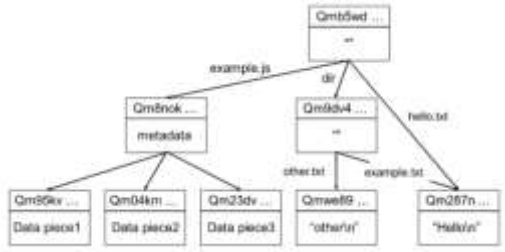

图 1 IPFS 中的默克尔树示意图

为了将文件分成独立的块，IPFS 还利用了许多算法，例如 rsync 滚动校验和算法等。

5. 激励层

之前提到，传统的 P2P 网络缺乏激励机制，无法促使节点之间互相进行通信，也无法鼓励节点保存其他节点的数据和文件——谁都不会在无利可图的情况下去做这些事情。

而区块链中存在的共识机制，促使矿工们在利益的驱使下去“挖矿”，让网络达成共识，使网络不断发展。这正好可以作为激励机制，让分布式文件系统不断发展。

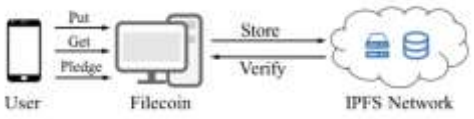

图 2 IPFS 中的激励层示意图

为了达成激励的效果，IPFS 的设计者专门发放了 Filecoin 这样一种虚拟数字货币。它们向参与系统，维护系统和存储文件的节点发放这种货币，

从而对这些节点起到了奖励和激励的作用。

具体而言，节点被分为两类：存储客户端是需要文件存储服务的客户端，它们位于存储市场的需求方，不参与激励层；而矿工是使用其可用磁盘空间为分布式文件系统提供存储的节点。它们位于存储市场的供应方。在 Filecoin 中，存储客户端首先将数据存贮的订单提交给矿工们。之后存储矿工对接受请求并完成储存，然后获得一定量的数字货币奖励。

可以说，正是这样的机制，使得基于区块链的分布式文件系统具有了生命力，从而能够不断的延续和发展。

6. 数据交换层

基于区块链的分布式文件系统中，数据交换层的作用类似于常见网络系统中的数据交换层，主要是对数据进行封装和传递。

7. 共识层

共识机制对于每个区块链系统都是至关重要的。在大型分布式网络中，多个对等点通常通过异步通信形成网络群集。如果每个节点认知的网络不一样，那么网络可能会拥塞，从而导致错误消息在整个系统中传播。

因此，如果对等点无法通过共识网络与其他人通信，系统就不可能有效的存在，这就是共识的作用——让每个节点都认知到一个相同的网络

与以太坊等常见区块链网络不同，Filecoin 不仅包含一个主链，而且还包含一个存储市场。Filecoin 中的用户与存储市场进行交互，来完成对时空状态和用户行为的证明，从而让整个网络达成一致。用户的这些交互存储在主链的分类帐中。以下是在 Filecoin 共识过程中起重要作用的三个证明机制：

•交易证明：

矿工和用户达成交易后，主链会锁定用户的令牌和矿工的存款。主链还记录有关交易的信息，包括矿工的硬盘扇区，存款明细，交易费用和存储期限等。

•复制证明（PoRep）：

为了防止矿工欺骗性行为的发生，Filecoin 要求每个矿工向主链提交复制证明。这样的复制证明可确保每个矿工真正独立地存储文件。

•时空证明（PoSt）：

为证明矿工在交易的有效时间内一直存储文件，每个矿工必须定期向主链提交时空证明，以证明文件不丢失。

三、 基于区块链的分布式文件系统应用和不足分析

1.应用场景

随着技术的成熟，IPFS 等分布式文件系统的生态系统中衍生出了较多的 Dapp、工具和项目。其中以 IPFS 的发展最为显著。下面，介绍几种 IPFS 文件系统的应用场景。

图 3 IPFS 生态系统中的应用

搜索引擎

该应用最先被 Firfox 火狐浏览器等搜索引擎所接纳。Mozilla 官网发布的浏览器扩展应用中包括了对 IPFS 分布式协议的支持，也就是说，使用“ipfs://”格式也可以进行内容的检索。

而 IPSE 是一款基于 IPFS 网络的搜索引擎，致力于打造下一代互联网的流量入口。在 IPSE 上可以搜索 IPFS 网络的文件。由于采用哈希标注技术，该引擎让内容的哈希地址转化为文字标题，可以实现快速访问。

内容平台

以 Netflix 为例，它将 IPFS 系统中的对等服务等技术整合到网飞的工具中，利用 IPFS 的技术加速云的构建、设计和测试。

为什么这个庞大的内容平台会和 IPFS 进行合作？其原因在于，Netflix 想要解决的容器分发挑战：如何在大规模，多区域环境中有效地提取容器图像。图像层通常位于不同的区域，利用 IPFS 作为点对点 CDN，可以使 Netflix 基础架构内的节点进行协作并将共同的种子播种到相邻节点，从而有助于更快地分发容器。

此外，DTube 是第一个加密分布式视频平台，建立在 STEEM 区块链和 IPFS 点对点网络之上，未来会支持 Filecoin 网络，它旨在成为 YouTube 的替代品，允许用户在 IPFS/Filecoin 基础上观看或上传视频。同时它承诺：在 DTube，内容是不受监管的，所有人都可以自由地发布自己的内容。

2.存在的问题

虽然 IPFS 已经落地，但是仍有不同的声音认为，该区块链的分布式文件系统技术是存在问题和缺陷的。

例如 DFS 的某些当前版本（例如 IPFS）不能容忍拜占庭式攻击。例如，只要每个对等方加入系统，它就可以访问 IPFS 上存储的每个文件。这种情况使隐私和安全问题成为 IPFS 系统的弱点。

此外，在 Swarm 和 IPFS 中，用户上载到分布式文件系统的数据分为几部分，然后存储在不同的对等点中。尽管上传的数据可以加密，但是存储在网络中的数据内容可供每个对等方访问。此外，根据 IPFS 和 Swarm 的设计，可以轻松收集记录对等方发展的事务。因此，存储在分布式文件系统后面的区块链中的交易是公开可见的。

如何更好地保护隐私，例如加入访问控制的方法，是今后 DFS 需要思考的主要问题。

四、 关于这一问题总结与思考

新一代基于区块链的分布式文件系统，例如 IPFS 和 Swarm，具有以下关键特征，显示出了巨大的潜力：激励，低延迟数据检索等新颖解决方案等。

通过分析我们可以看出，IPFS 确实可以作为一种较有创新的工具，用于分布式文件系统，服务于多个领域，例如对于大数据的存储和转发。

我相信，基于区块链的分布式文件系统可以成为下一代网站和数据共享平台非常有希望的解决方案。但是，在隐私和安全方面，IPFS 还有发展的潜力和空间，其实这也是分布式文件系统共同的弊端。期待分布式文件系统能够在这一方面有着更好的改进。