链塔智库
Block Data

星际文件系统IPFS
生态研究报告

2018年6月6日

前言

近期币圈出现两大热点，EOS和IPFS。随着超级节点竞选拉开帷幕，EOS博尽眼球，引领市场行情。而IPFS相对而言就低调了很多，但IPFS引领的全民挖矿热情日益高涨。

IPFS的中文名是星际文件系统，由Juan Benet在2014年5月份发起。2015年，IPFS在YCombinator孵化竞赛中拿到了巨额的投资，同时成立了协议实验室。实验室团队由14位核心开发者以及社区中上百位代码贡献者组成。

IPFS在2017年8月份仅仅出售了10%的代币，就募集到2.5亿美金，这意味着IPFS还没正式上线，市值已经达到了25亿美金。

IPFS本质上是一种内容可寻址、版本化，点对点超媒体的分布式存储及传输协议，有望补充甚至取代过去20年里使用的超文本媒体传输协议（HTTP），效率提升显著。

在技术上有创新性的IPFS前景虽然看好，但是要完成替代HTTP的道路还比较长。

目录

PART.1
IPFS综述

用户每天上网使用APP和浏览网页都是凭借着HTTP协议，它基于TCP / IP的计算机应用层面，从服务器传输超文本数据到本地浏览器，本地浏览器或APP，经过渲染再呈现给用户。基于这样的网络环境构成了CS或BS架构，最后提供给大型网络供应商。

HTTP模式主要分为两种：

第一个模式属于集中化，部分企业的互联网由于业务不能分散去做，只有一个中心服务群，所有流量直接搭载在这个服务群上，承载的压力极大，极易造成系统崩溃。

第二个模式属于分散集群，各个网站需要建立不同区域下的服务群，它们背后的IDC机房会让同样的服务在一个局域内分散，这就减轻了中心服务器的压力。

两种模式容易产生一些弊端：在第一种模式里，服务高度依赖中心网络，大公司或者创业公司无法承受宕机，大公司需要雇一批运维专家或专业人士去保障系统的稳定性；在第二种模式里，中心化数据库一旦遭受人为因素或不可抗力损害，所有数据将全部丢失。

同时，这两种模式的带宽成本都比较高，皆会造成一定带宽资源的浪费。

而IPFS想打造一个点对点的网络拓扑，相当于颠覆HTTP所代表的分布关系，它具有内容可寻址的特点，通过文件内容生成唯一的哈希标识，一定程度上节约了空间开销的成本以及运维成本。

IPFS中文件通常不会都存在一个节点，而是分片存在其他一些子节点上。提取文件时，IPFS把这些节点列表全部并行抓取，最后在本地拼成完整的文件。并行的速度远远大于直接下载完整文件的速度，用户很快就能在本地获得文件，还可以继续分享给其他人。分布式的储存方式能有效解决数据丢失的风险，同时减轻个体数据库的存储压力。

PART.2
IPFS架构分析

2.1IPFS架构简述

IPFS架构分为八层子协议栈，从上至下为身份、网络、路由、交换、对象、文件、命名、应用，每个协议栈各司其职，又互相搭配。

IPFS八层协议栈	
身份层	S/Kademlia生成 对等节点身份信息生成
网络层	任意传输层协议 ICE NET&NAT穿透
路由层	分布式松散哈希表（DSHT） 定位对等点和存储对象需要的信息
交换层	BitTorrent&BitSwap 管理区块如何分布
对象层	Merkle-DAG 内容可寻址的不可篡改、去冗余的对象链接
文件层	类似Git 版本控制的文件系统：blob、list、tree、commit
命名层	具有SFS（Self-Certified Filesystems） IPNS：DAG对象命名可变
应用层	在IPFS上运行的应用程序利用最近节点提供服务 提升效率、降低成本

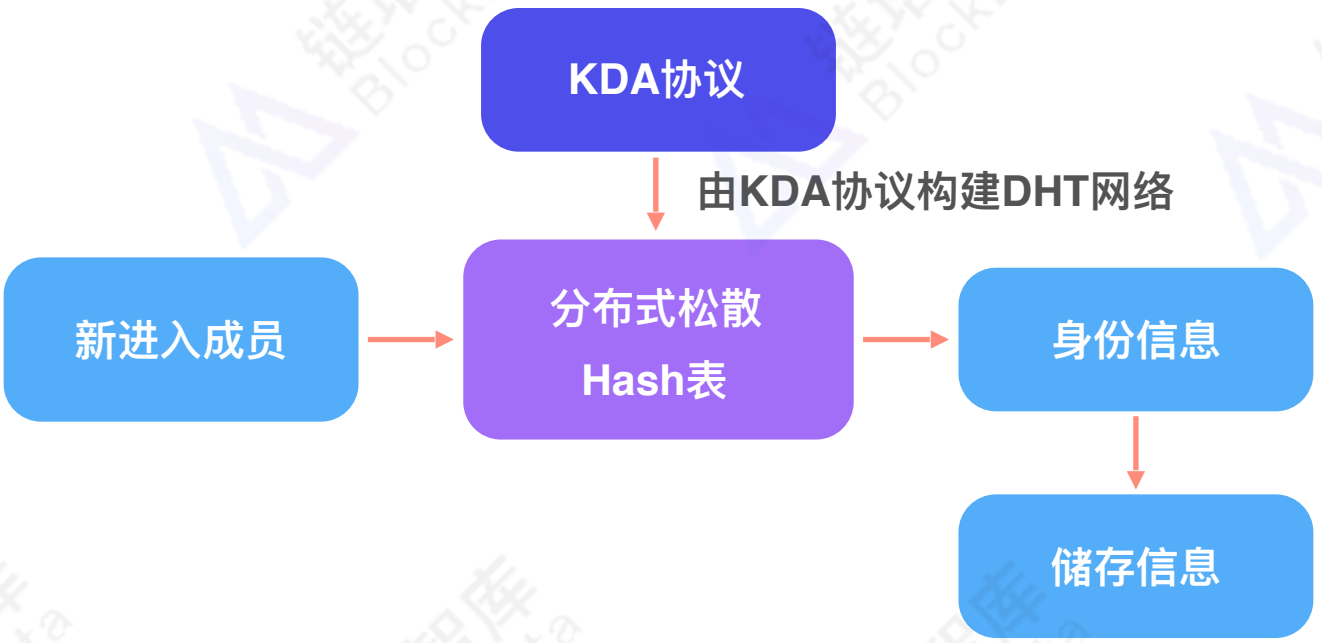
链塔智库研究绘制

2.2 IPFS架构解析

一、身份层及路由层

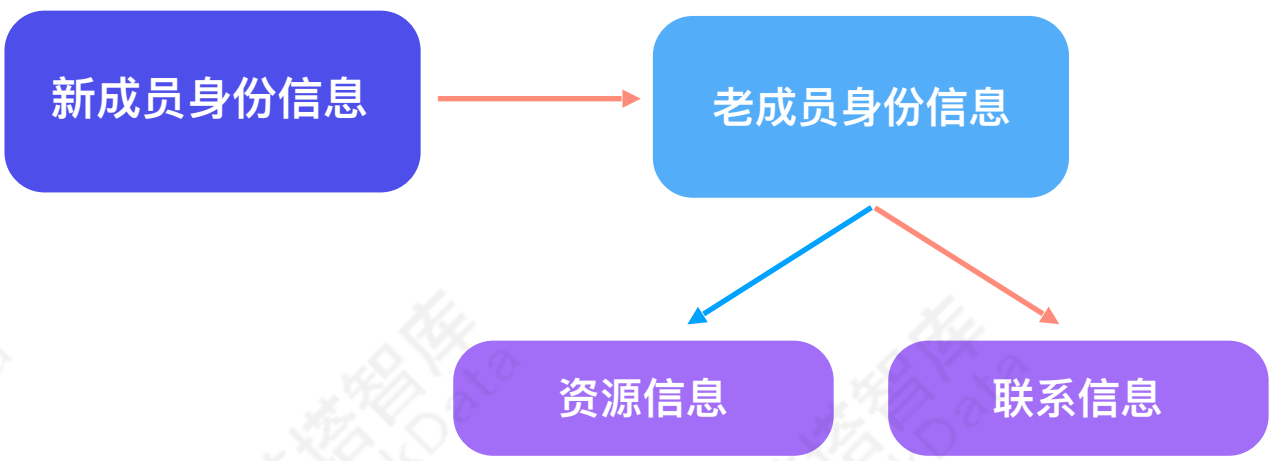
身份层和路由层属于捆绑性质。对等节点身份信息的生成以及路由规则是通过Kademlia协议生成制定，KAD协议实质是构建了一个分布式松散Hash表，简称DHT，每个加入这个DHT网络的人都要生成自己的身份信息，然后才能通过这个身份信息去负责存储这个网络里的资源信息和其他成员的联系信息。如果新成员需要寻找一位老成员A的联系信息，而他没有这位老成员A的联系方式，那么他可以通过联系任意一位存储老成员A联系信息的成员来获取这位老成员A的联系信息，同理在IPFS中获取资源信息也是一样的道理。

身份生成流程示意图



链塔智库研究绘制

信息查询流程示意图

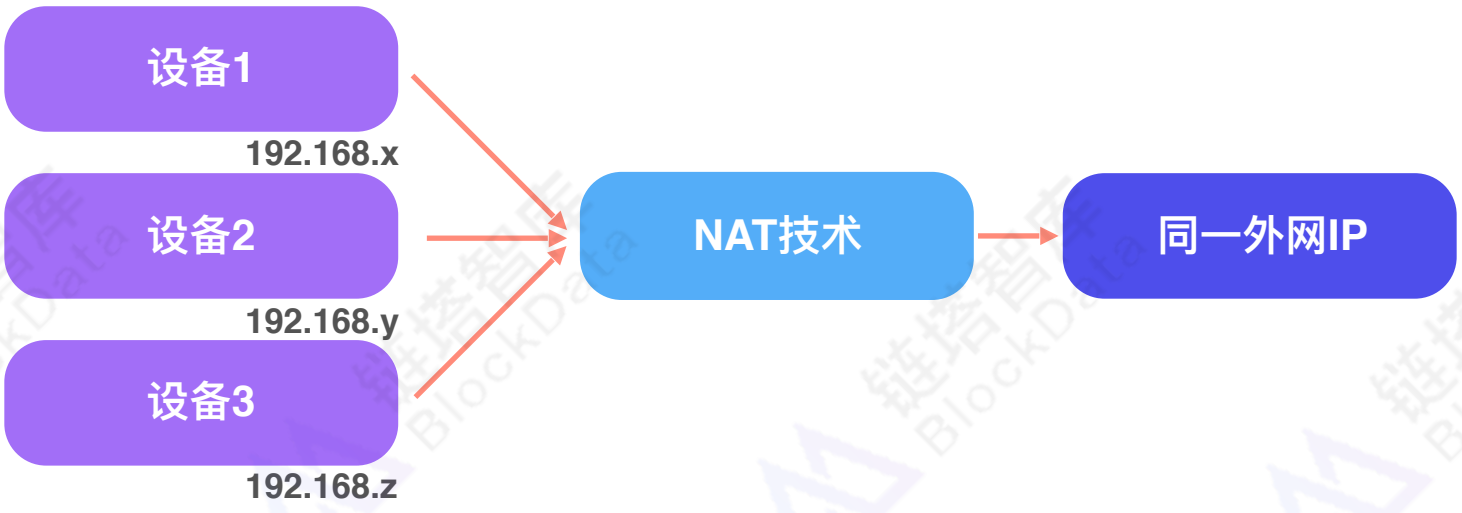


链塔智库研究绘制

二、网络层

网络层属于IPFS架构中的核心之一，使用的LibP2P可以支持任意传输层协议。NAT技术能让内网中的设备共用同一个外网IP，家庭路由器使用的就是这个原理。

IPFS IP转换流程示意图

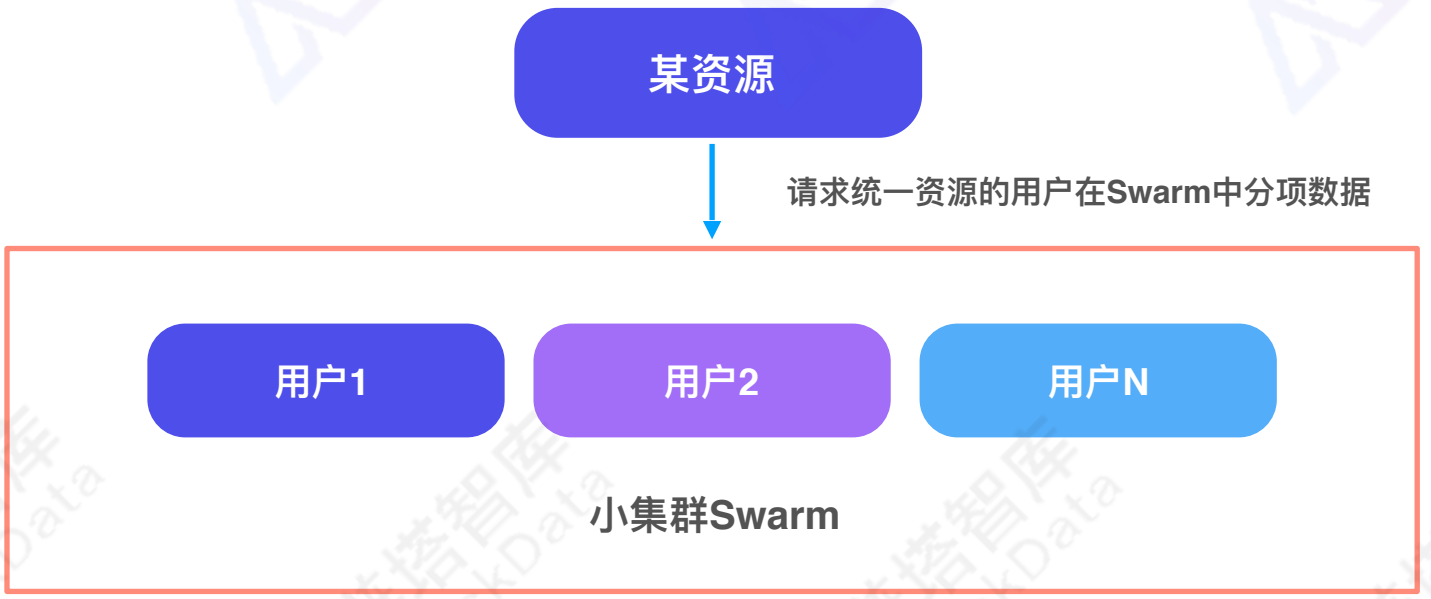


链塔智库研究绘制

三、交换层

交换层模拟了P2P网络，并创建中心服务器，当服务器登记用户请求资源时，让请求同样资源的用户形成一个小集群Swarm，在这里分享数据。在中心化的处理方式中这种方式有弊端，因为服务器是由中心化的服务提供商统一维护，如果出现了故障、宕机时，下载操作无法进行。

IPFS 资源请求流程示意图

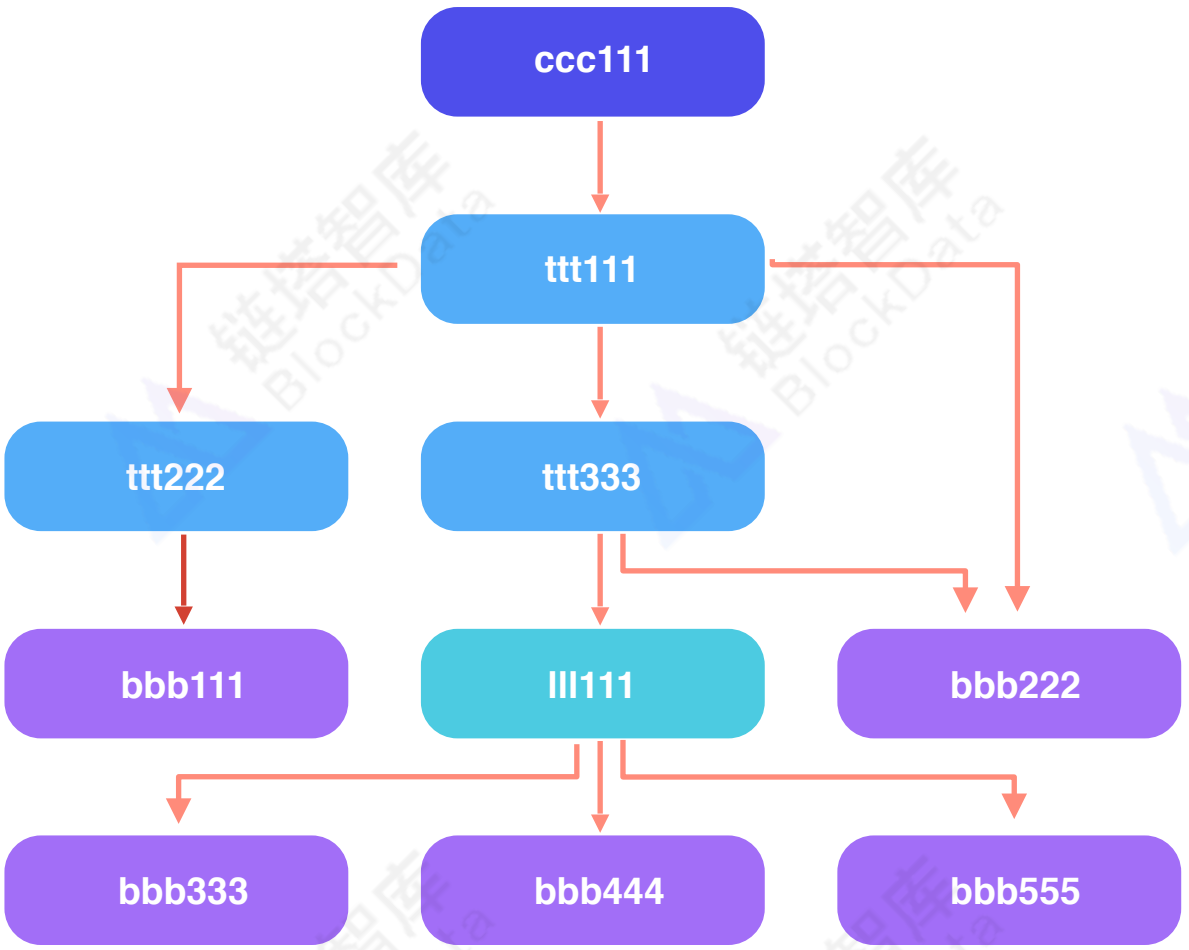


链塔智库研究绘制

三、对象层及文件层

对象层和文件层需要结合来谈，它们管理的是IPFS上80%的数据结构，大部分数据对象都是以MerkleDag[Merkle directed acyclic graph（默克有向无环图）]的结构存在，这为内容寻址和去重提供了便利。文件层是一个新的数据结构，和DAG并列，采用Git一样的数据结构来支持版本快照。

默克有向无环图示意图

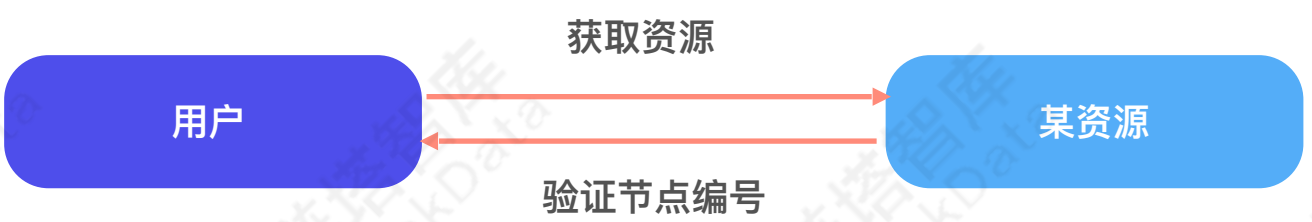


链塔智库研究绘制 www.blockdata.club

四、命名层

命名层具有自我验证的特性【当其他用户获取该对象时，使用指纹公钥进行验签，即验证所用的公钥是否与NodeId(节点编号）匹配，这验证了用户发布对象的真实性，同时也获取到了可变状态】，并且加入了IPFS这个设计来使得加密后的DAG对象名可定义，增强可阅读性。

资源获取流程示意图



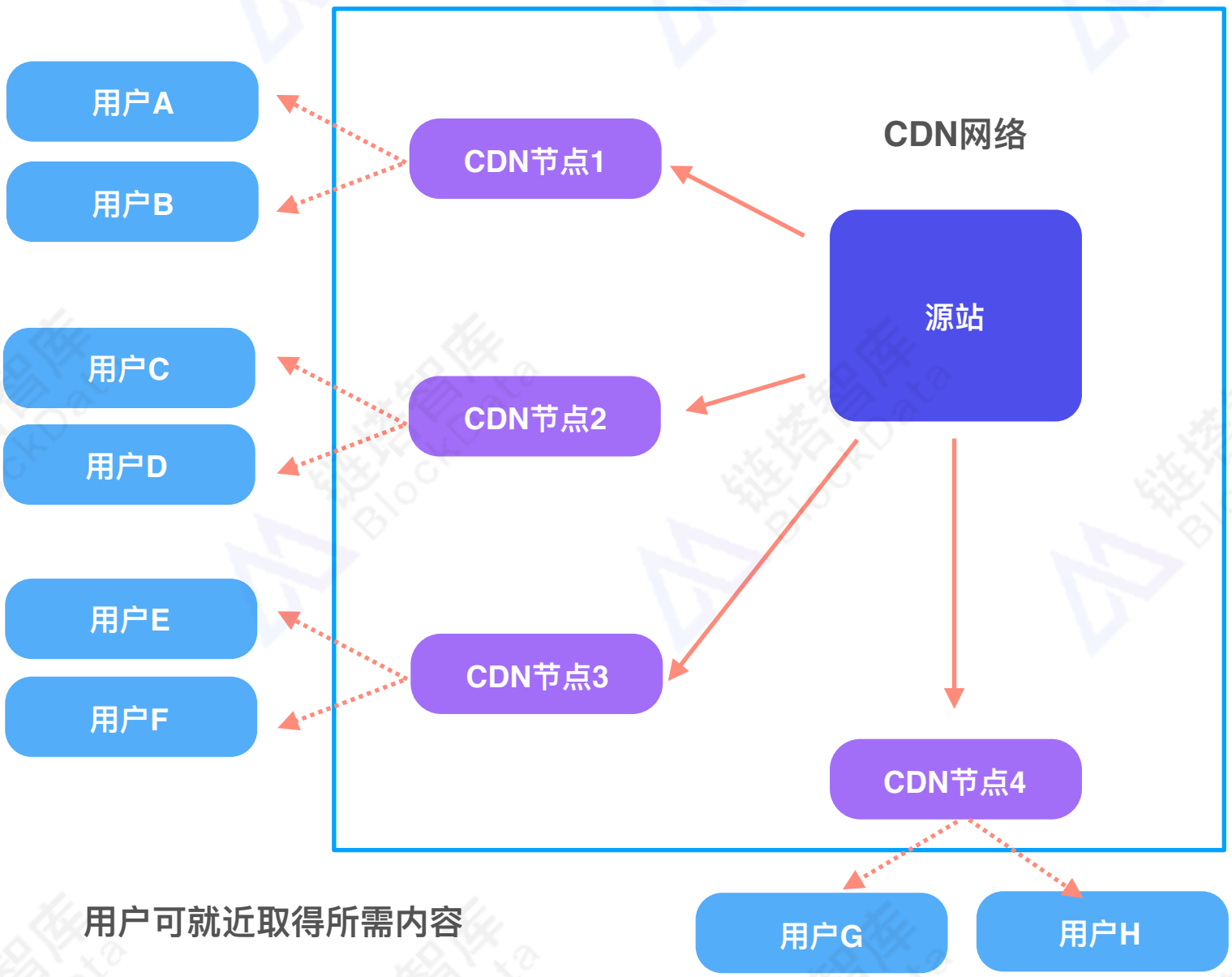
链塔智库研究绘制 www.blockdata.club

五、应用层

应用层，IPFS核心价值就在于其上运行的应用程序，可以利用它类似CDN的功能，在成本很低的带宽下，去获得想要的数据，从而提升整个应用程序的效率。

CDN的全称是Content Delivery Network，即内容分发网络。其基本思路是尽可能避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，使内容传输的更快、更稳定。通过在网络各处放置节点服务器所构成的在现有的互联网基础之上的一层智能虚拟网络，CDN系统能够实时地根据网络流量和各节点的连接、负载状况以及到用户的距离和响应时间等综合信息将用户的请求重新导向离用户最近的服务节点上。其目的是使用户可就近取得所需内容，解决Internet网络拥挤的状况，提高用户访问网站的响应速度。

内容分发网络示意图



用户可就近取得所需内容

链塔智库研究绘制

PART.3
IPFS生态系统

3.1 IPFS生态发展进程

2014年5月，Juan Benet发起IPFS项目

2015年1月，IPFS协议实验室成立，并发布IPFS

2016年，IPFS团队创建了libp2p、IPLD、multiformats、Orbit等模块

2017年6月，IPFS已储存50亿份文件

2017年7月，IPFS团队宣布成立Filecoin项目

Filecoin项目预计于2018年年底上线

3.2 IPFS生态系统分布

IPFS生态系统		
IPFS	应用数据	应用层
libp2p	传递数据	路由层、网络层、交换层
IPLD	定义数据	命名层、对象层、文件层
Multiformats	加密、描述数据	身份层
Filecoin	价值数据	用户激励

链塔智库研究绘制

IPFS的团队在开发时，采用高度模块集成化的方式。协议实验室团队2015年创立，到17年的这段时间里都在做IPLD、LibP2P、Multiformats这三个模块的开发，它们服务于IPFS底层。

Multiformats是一系列hash加密算法和自描述方式（从值上就可以知道值是如何生成）的集合，它具有SHA1\SHA256 \SHA512\Blake3B等6种主流的加密方式，用以加密和描述nodeID（节点编号）以及指纹数据的生成。

LibP2P是IPFS的核心，面对各式各样的传输层协议以及复杂的网络设备，它可以帮助开发者迅速建立一个可用P2P网络层，快速且节约成本。

IPLD是一个转换中间件，将现有的异构数据结构统一成一种格式，方便不同系统之间的数据交换和互操作。现在IPLD支持的数据结构，是比特币、以太坊的区块数据，也支持IPFS和IPLD。IPLD中间件可以把不同的区块结构统一成一个标准进行传递，为开发者提供了成功性比较高的标准，不用担心性能、稳定和bug。

Filecoin把以上这些应用的数据价值化，通过类似比特币的激励政策和经济模型，让更多的人去创建节点，提供服务，去让更多的人使用IPFS。

PART.4
IPFS挖矿机制

4.1 IPFS挖矿背景简述

IPFS系统下挖矿所得为FIL（ Filecoin），Filecoin是一个去中心化的存储网络，Filecoin有两个市场：存储市场和检索市场。这两个市场结构相同但设计不同。

存储市场上，客户付费给存储矿工，令其存储数据。检索市场上，客户向检索矿工付费后取回数据。

这两种情况下，客户和矿工都可以设置报价或接受报价。整个交易是由网络来运行—Filecoin中的所有节点构成了拟人化的网络。网络保证了矿工在提供服务时可以得到客户的奖励。

存储的需求和供给在两个Filecoin市场进行：存储市场和检索市场。这两个市场属于去中心化交易所，简而言之，客户和矿工们通过向各自的市场提交订单来为服务定价。交易所为客户和矿工们提供了匹配交易和牵线的方法。运行管理协议后，如果服务请求被成功提供，网络会确保矿工得到奖励，客户得到服务。

4.2 IPFS挖矿参与者

IPFS矿工分为存储矿工、检索矿工，客户在请求存储或检索数据时需要支付相应代币。

存储矿工为网络提供数据存储，存储矿工通过提供磁盘空间和响应客户请求来参与Filecoin运作。要想成为存储矿工，用户必须用与存储空间成比例的抵押品来抵押。（抵押品为扇区，扇区指存储矿工向网络提供的磁盘空间。矿工将客户的数据片段存储到扇区，并以此赚取代币。为了存储片段，矿工们必须向网络抵押他们的扇区。）

存储矿工通过在特定时间内存储数据，来响应用户的存储请求。存储矿工生成证明并提交到区块链网络，来证明他们在特定时间内存储了数据。如果数据失效或丢失，存储矿工将被罚没部分抵押品。存储矿工也可以挖掘新区块，如果挖到了新区块，矿工就能获得挖取新块的奖励和新区块中的交易费用。

检索矿工为网络提供数据检索服务，检索矿工通过提供用户检索请求所需要的数据来参与Filecoin运作。和存储矿工不同，他们不需要抵押品，不需要提交存储数据，也不需要提供存储证明。存储矿工同样可以担任检索矿工，检索矿工可以直接从客户或者从检索市场赚取收益。

4.3存储矿工工作流程

1、存储矿工在区块链中存放抵押品，来保证向网络提供稳定的存储。抵押品为了保证服务而存在，如果矿工为所存储的数据生成了存储量证明，抵押品就会被退回。如果没有成功生成存储量证明，矿工将失去抵押品。

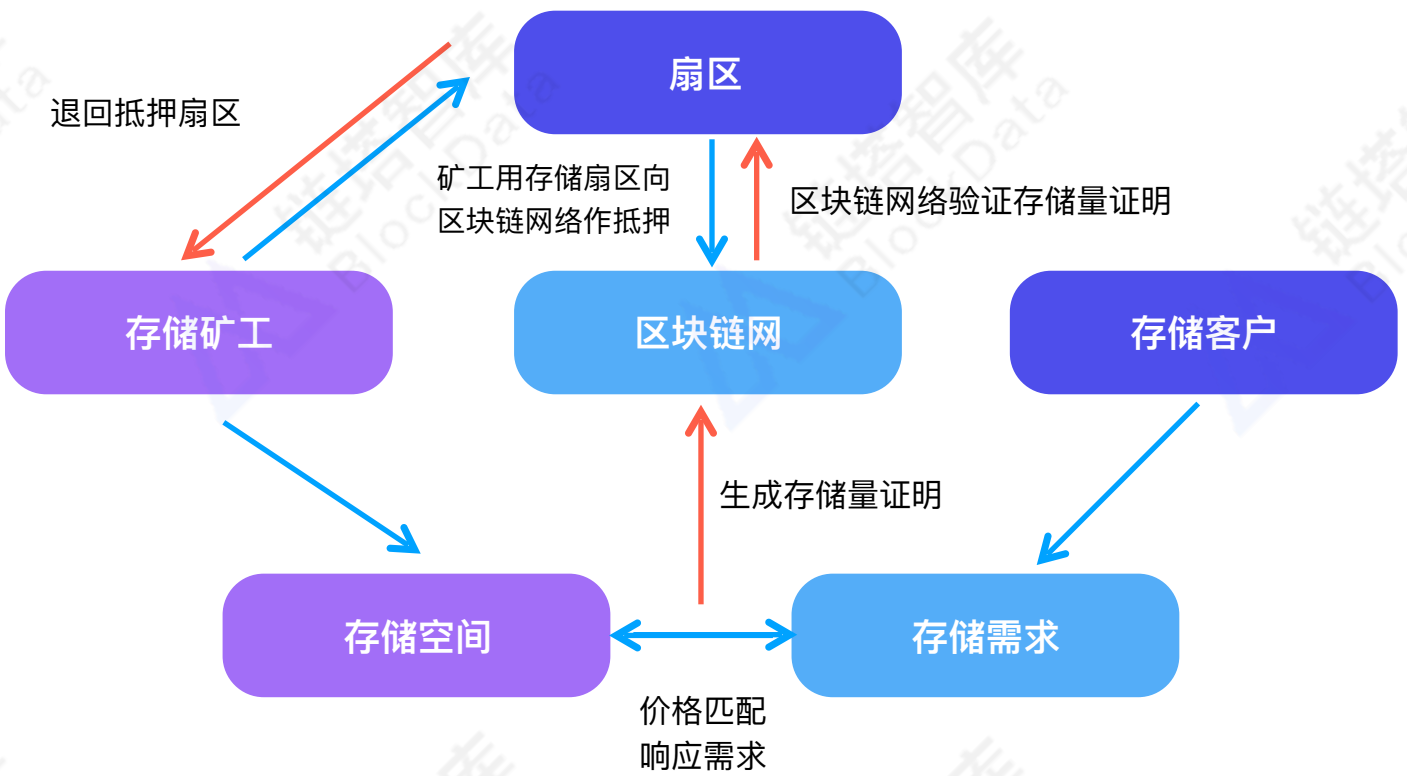
2、一旦抵押交易在区块链中出现，矿工就可以在存储市场中提供存服务。矿工们设置价格，并响应市场订单簿中的订单要求

3、一旦订单匹配，客户就将数据发给存储矿工。存储矿工数据接收完成后，矿工和客户签署交易订单并提交到区块链。

4、当存储矿工被分配了数据时，必须重复生成存储量证明来确保他们正在存储数据证明被发布在区块链中，并由网络来验证。

5、验证成功后，存储矿工将会获得相应的奖励

存储矿工工作流程示意图



链塔智库研究绘制

4.4 检索矿工工作流程

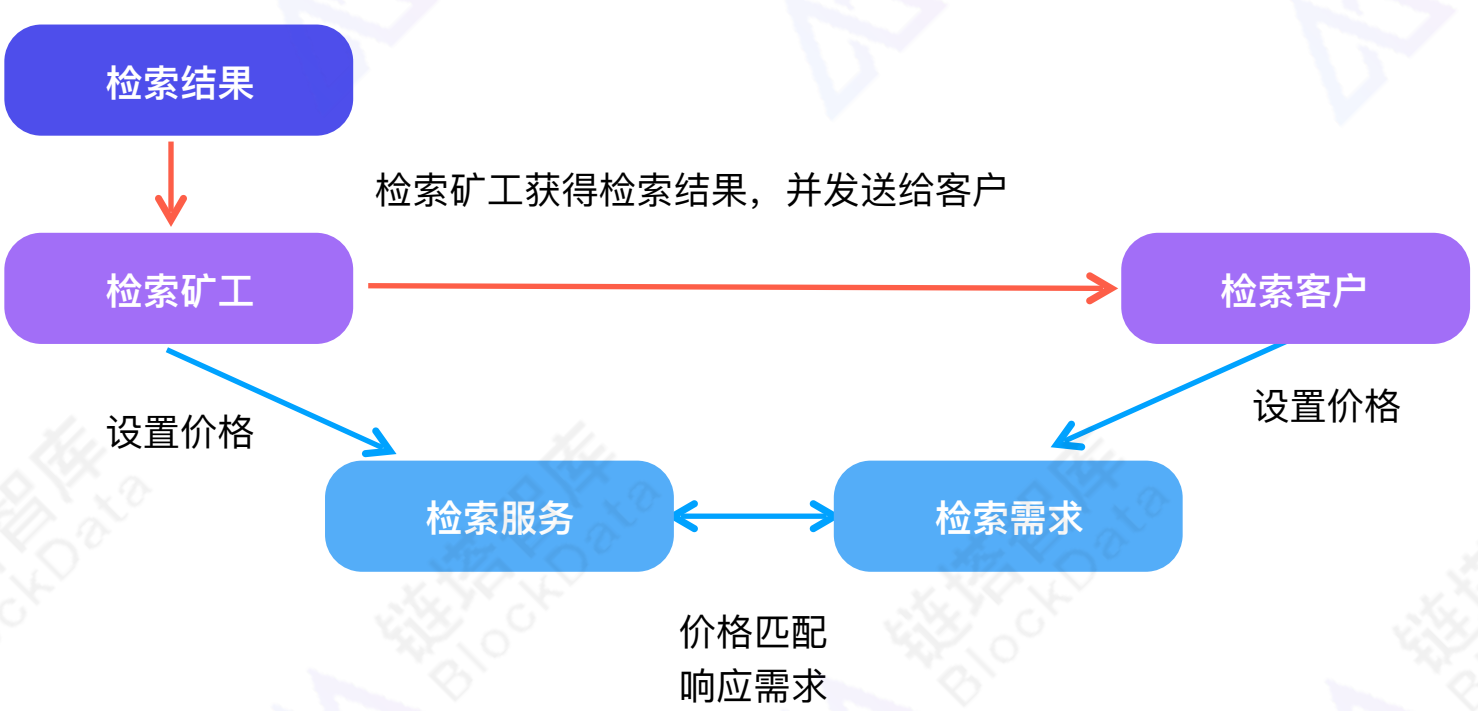
1、检索矿工设置价格并发送到市场订单簿，并通过向网络发送报价。

2、然后检索矿工检查是否与客户的订单报价匹配。

3、一旦订单匹配，检索矿工就将数据发送给客户。数据接收完成后，矿工和客户就签署交易订单提交到区块链。

4、当交易被验证后，检索矿工也将获得相应奖励

检索矿工工作流程示意图



链塔智库研究绘制

PART.5
IPFS工作机制

5.1 IPFS工作机制分析

IPFS为每一个文件分配一个独一无二的哈希值(文件指纹：根据文件的内容进行创建)，即使是两个文件内容只有1个比特的不相同，其哈希值也不相同。这个方式使得IPFS可以支持基于文件内容进行寻址。

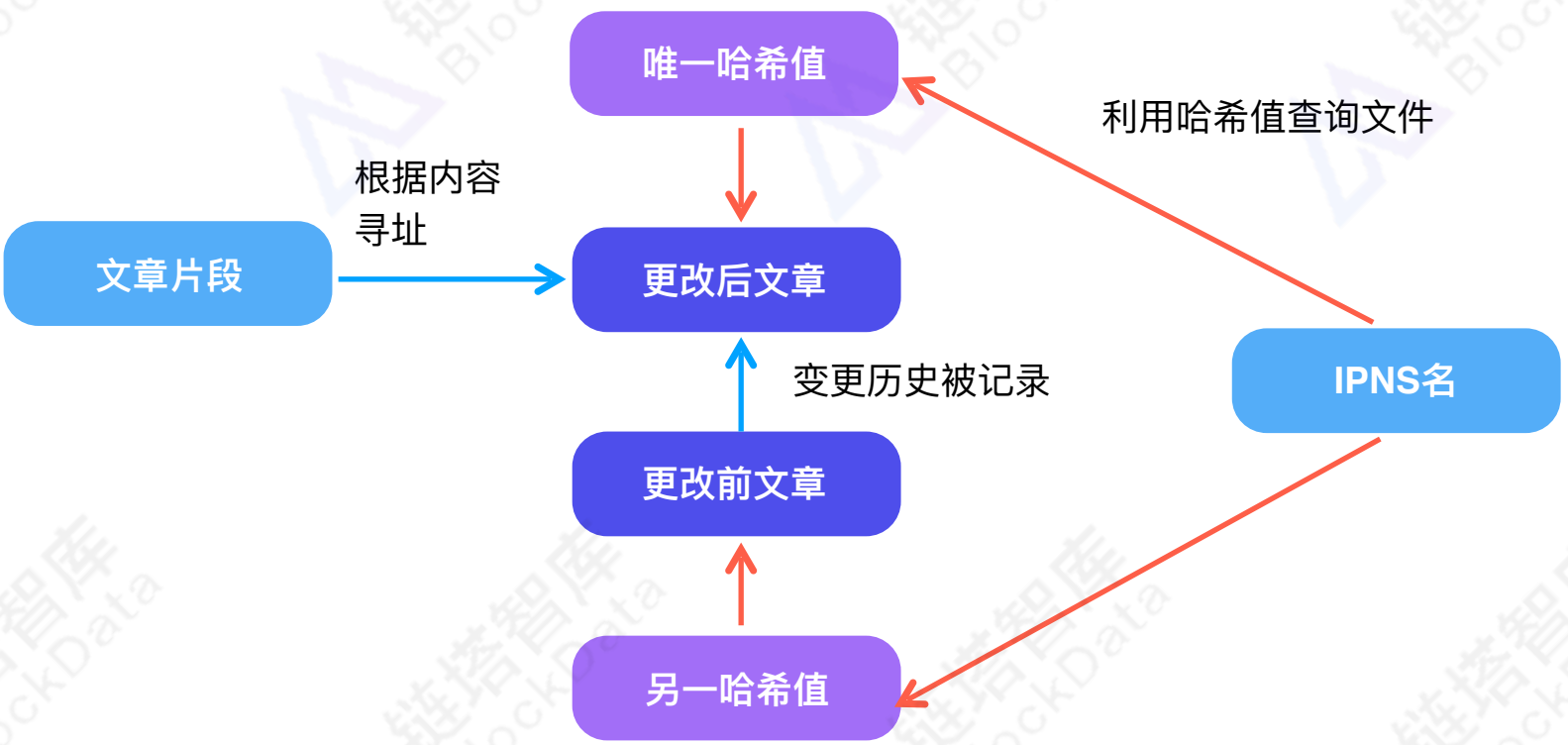
IPFS在整个网络范围内去掉重复的文件，并且为文件建立版本管理——每一个文件的变更历史都将被记录。版本管理是版本控制工具git，svn等的重要特性，依靠它系统可以很容易地查看文件的历史版本中的数据。

当查询文件的时候，IPFS网络根据文件的哈希值(全网唯一)进行查找。由于每个文件的哈希值全网唯一，查询的执行很容易，效率也比较高。

如果仅仅使用哈希值来区分文件的话，会给内容的传播造成困难，因为哈希值不容易记忆，这点和IP地址不容易记忆一样。同样，类似于域名，IPFS利用IPNS将哈希值映射为容易记的IPNS名字。

每个节点除了存储自己需要的数据，还存储了一张哈希表，用来记录文件存储所在的位置、进行文件的查询下载。

IPFS工作机制示意图



链塔智库研究绘制

5.2 IPFS工作机制与HTTP工作机制对比

1、下载速度

HTTP：低效，成本高。使用HTTP协议每次需要从中心化的服务器下载完整的文件(网页、视频、图片等)，速度慢，效率低。

IPFS：高效，低成本。整个IPFS系统是一个分布式的文件存储系统，在下载相关数据的时候，可以从多个节点同时下载。这点与P2P下载(比如: 迅雷、BitTorrent)一致。与HTTP从中心服务器下载相比较，P2P下载速度要快很多。

2、存储

HTTP：成本高，易丢失。基于HTTP的存储方式主要依赖于中心化的云存储，占用大量空间，且容易丢失。HTTP页面平均生存周期大约只有100天，Web文件经常被删除(由于存储成本太高，无法永久保存。

IPFS：低成本，不易丢失。由于IPFS使用的是区块链技术，一方面利用 FileCoin来激励矿工分享自己的硬盘，另一方面IPFS从全网去掉了冗余存储（见5.1），提升了存储使用效率，节省无效网络存储空间。FileCoin将采用供需对价的方式解决市场对接，存储通过挖矿获得收益，检索通过对价认可提供服务获得收益。作为IPFS生态的基础能力网络存储，因为持续的投放形成的供需平衡而非常便宜。

3、安全

HTTP：易受攻击，无法保护隐私。中心化服务器目前很难抵挡DDoS（分布式拒绝）攻击，当大量的访问请求涌来，中心化的服务器几乎会在一瞬间瘫痪。中心化云存储一旦遭到窃取，海量用户数据及文件将被泄露，无法保证用户隐私。

IPFS：不易攻击，保护隐私。IPFS拥有抵挡DDoS攻击的能力：所有的访问将会被分散到不同的节点，甚至攻击者自己也是节点之一。为了实现安全，IPFS中每一份用户的数据都进行了加密、分片，并且有多分冗余在全网的节点中。黑客在进行攻击时，无法得知哪些数据对应着哪位用户，即使黑客找到了相应数据，也无法解密数据。和数字货币一样，只有持有私钥的人才能够拿到数据，对数据进行解密，查看数据。

PART.6
IPFS应用实例

6.1 OpenBazaar

OpenBazaar是IPFS上的一个应用，中文名为开放集市，于年初获得比特大陆500万美元的投资。

应用IPFS前：利用ZeroMQ来实现P2P交易，一定程度也绕过了中心化的检查，把交易的手续费作为红利给用户。同时它整合了比特币作为支付渠道，用户数量在短时间内迅速提升。

应用IPFS后：官方加入审查机制，同时支持了比特币之外的BCH等数字货币，并且整合和重构IPFS，取代了之前的ZeroMQ。由于利用了IPFS，商户在OpenBazaar市场的产品可以通过其他用户进行上传，从而实现即使当商店不与互联网直接相连的情况下也能进行购买。以前必须同时登陆才可以交易，现在利用IPFS相当于实现了离线店铺。这同时意味着，访问你的店铺的人越多，店铺数据被复制越多，越有利于优质的店铺宣传和推广。

6.2 Mediachain

Mediachain则是一种基于IPFS的开放媒体库，支持多种区块链，包括比特币和以太坊。这个区块链应用将会帮助用户找到高质量的和正确归属的图片，可用于满足任何视觉成像需求——网站，博客和演讲——同时允许出版商和创造者保护他们的数字内容的所有权。 Mediachain同时推出Mediachain Attribute Engine（Mediachain内容归属引擎）。Mediachain项目目前的参与者包括现代艺术博物馆(MoMA)， 盖蒂图片社（Getty Images），美国数字公共图书馆 (DPLA) 。

Mediachain Attribute Engine的创建是基于Mediachain的开放媒体库项目。它允许发布者和内容创建者上传他们的内容和附加信息到他们的媒体，然后在区块链上加上时间戳并存储在IPFS。然后，这些信息可以通过感知搜索进行查看。

用户可以使用Attribution Engine来搜索Mediachain开放数据库中他们想要分享的图片。一旦他们找到喜欢的图片，他们可以复制粘贴链接的HTML码到他们的博客或者网站。这个图片的归属信息被嵌入到了html码本身，因此无需手动添加。

结语

- IPFS架构分为八层子协议栈，从上至下为身份、网络、路由、交换、对象、文件、命名、应用，每个协议栈各司其职，又互相搭配。IPFS生态系统分为五大模块，覆盖八个层级的子协议栈。Filecoin利用激励政策和经济模型将IPFS中的生态模块价值化，鼓励更多人去创建节点。
- IPFS矿工分为存储矿工和检索矿工，人人可成为矿工，使用内存挖矿并获取奖励。
- IPFS拥有去冗余机制，自动删除重复文件，保证区块链网络空间的合理分配。IPFS将文件分片存在子节点上，提取文件并行抓取，保证区块链网络的高效性。IPFS中所有的访问将会被分散到不同的节点，能有效防止DDoS攻击，保证区块链网络安全性。
- IPFS已拥有部分成熟应用及大量存储文件，生态布局较为广泛。
- IPFS特殊的挖矿机制容易导致类似于比特币中的大型矿池控制大量算力的情况，存在一定隐患。
- IPFS欲颠覆HTTP统治地位，但激励系统尚未完善，实际落地情况有待考证。

法律声明

知识产权声明

本报告为链塔智库BlockData制作，报告中所有数据、表格、图片均受有关商标和著作权法律保护，部分数据采集自公开信息，知识产权为原作者所有。我们相信数据的价值，我们同样相信分享也能创造价值，我们欢迎各组织和个人采用我们的报告和数据，在此之前告知我们即可。

免责条款

本报告中所载所有内容为链塔智库分析师通过访谈、市场调查、信息调研整理及其他方式方法获得，并结合链塔智库独有的数据和分析资源，建立相关预测模型估算而得，为区块链行业从业者提供基本参考，受研究方法和数据获取渠道所限，本报告只提供受众作为各类市场活动参考资料，不构成任何投资或交易买卖建议。如果访问者依据本报告信息进行投资或进行交易买卖而遭受损失，本公司对此不承担责任。

链塔智库BlockData

链 我们深刻认识到区块链数据的价值，专注用深度数据赋能区块链产业。

塔 我们关注每一个细分领域的头部项目，Top X只是我们展现的手段。

智 我们只与业内顶尖的合作伙伴、区块链专家、行业分析师为伴，提供专业的数据服务。

库 我们拥有全球最全的区块链项目库，时刻扫描和追踪全球区块链动态。

我们是链塔智库 推崇专注专心专业，坚持公开公正公平，“天赐时代 睿见未来”，预见更多可能。

全球首家区块链数据服务提供商

扫码关注公众号

扫码进入小程序

网址：www.blockdata.club

微信订阅号ID：liantazhiku

链塔智库合作伙伴

独家大数据支持平台：TalkingData

联合发布媒体（排名不分前后）：

BiaNews 三言财经 Sanyan Blockchain 区区 耳朵财经 金色财经 火星财经 猎云财经 荣格财经 Jung Media

媒体深度合作伙伴（排名不分前后）：

猛犸财经 链路财经 linkafia.com 蜂巢财经 www.bicecoc.com 链道学社 财链社 挖链网 区块链蓝海 千氪财经 www.kg.com 核财经 HE CAI JING 鸵鸟区块链 TuoniaoX.com 未来财经 weilaicaijing.com

链塔智库

Block Data

全球区块链数据服务提供商

1000+项目入库 /800+机构入驻 /100+专家学者观点

扫码关注公众号

ID：liantazhiku

扫码进入

小程序

『链塔智库BlockData』，全景式扫描和追踪全球区块链公司/项目，提供深度数据服务，专注于区块链行业研究、分析、项目评级。全球最全的区块链项目库1000+（数据每周都在更新）。