UPS 中文白皮书 1.0

-基于 TrueChain 的 IPFS 侧链网络

摘要

UPS 能够在对等节点之间形成存储合约。合约是存储提供商与其客户之间的协议，定义将以何种价格存储数据。他们要求存储提供商定期证明他们仍在存储客户的数据。合约存储在区块链中，使其公开可审计。

介绍

UPS 是以 TrueChain 侧链的形式提供了去中心化的存储服务，倾向于丰富 TrueChain 主网的功能及扩展。 UPS 网络不是从集中供应商处租用存储，而是从彼此租用存储。 UPS 主网提供内置文件存储合约，定义其合约细节。通过签订合约，存储提供商（也称为主机）同意存储客户数据，并定期提交其持续存储的证明，直至合约到期。主机补偿他们提交的每一份证据，并因缺少证据而受到处罚。由于这些证据是公开可验证的（并且可以在区块链中公开获得），因此可以使用网络共识来自动执行文件存储合约。重要的是，这意味着客户不需要亲自验证存储证明;他们可以简单地上传文件，然后让网络完成剩下的工作。

将数据存储在单个不可信任的主机上几乎无法确保可用性，带宽或服务质量的一致性。UPS 在多个主机上冗余存储数据。特别是，使用纠删码（erasure codes）可以实现高可用性，而不会出现过多冗余。

UPS 主网将实施基于 TrueChain 区块链预编译合约，与 TrueChain 挂钩，本身不发行任何通证，其共识安全也由 truechain 来保证。

需求

UPS 对于去中心化存储的需求将从安全性，可靠性，性能等方面阐述去中心化存储的必要性：

安全性

安全性指的是用户存在去中心化存储系统里的数据是绝对安全的，有非常严密的隐私保护。

这是与现在中心化云存储最大的不同和优势。为了实现安全，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。这些节点可能分布在美国、日本、欧洲、南美等等。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。另外，黑客在进行攻击时，也无法得知哪些数据对应着哪位用户。即使黑客找到了这些数据，也只能望洋兴叹。

安全这个点，现在的用户似乎不在意，没有需求，但并不代表着未来也不需要。有的时候市场需要培养，做产品需要多一些耐心。最近 Facebook 用户数据泄露的事件，影响非常的大。这其实就是一个数据安全的典型的场景。用户的数据需要更好的管理，在去中心化存储之上，可以通过智能合约实现授权的机制。在用户授权的条件下，完成基于用户行为的模型训练。所以，去中心化存储，引入的并不只是区块链技术本身，更多的是激励生态体系。

可靠性

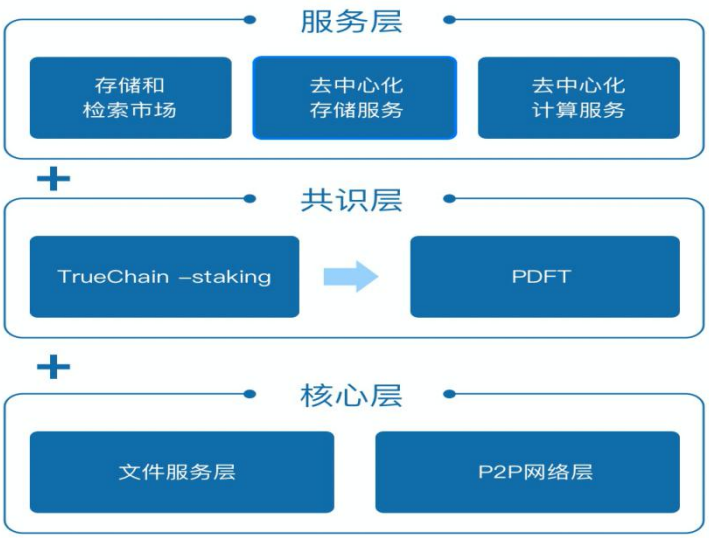
目前的云存储经过各大厂商的努力，已经做到很高的水平，非常的厉害。当然投入是巨大的，而且也有巨大的安全隐患。那去中心化存储的可靠性如何呢？从技术上客观的讲，这里有一个网络规模的问题。去中心化存储的网络越大，可靠性就越高，随着用户的认可、网络的规模越来越大，甚至可以媲美互联网的时候，数据的可靠性就是接近 100%的。而且隐私安全、控制权在用户的手里。这里有一个风险点，就是早期节点数量比较小的时候，需要有一些措施保证网络的稳定运行。

性能

这个大家联想一下 BT、电驴就好了。如果了解技术原理，差不多也是一样的。通过一种纠删码的技术，实现多路高速下载，计算适当的冗余完成性能和可靠性的权衡。这应该不用过多的解释。

总体结构

UPS 侧链采用 DPOS 共识，基于 truechain 的预编译质押合约(UPS 质押合约)，通过 truechain 质押 True 代币，来选举 UPS 的委员会，UPS 质押合约规定了 UPS 委员会的任期及数量，UPS 通过跨链协议获取 TrueChain 上 UPS 质押合约的变动，委员会共识打包交易，并验证存储合约的有效性维护主链的安全性。如图:



UPS 支持从 TrueChain 定向兑换一定数量 True 代币作为 UPS 原始资金，用于 UPS 网络维护及出块奖励，ups 与 true 一比一兑换，即锁定 1 个 true，释放一个 ups，其中没有任何资产增发。ups 可以随时兑换回 true 代币；UPS 使用内置存储合约提供存储服务，并定期监控存储合约中主机的证据项（合约定义了主机必须提交存储证据的规则），并释放佣金给主机，UPS 提供提供文件访问合约用以支持用户获取存储文件。

主机定义了唯一的主机 ID，该 ID 定义了主机的标识，全网唯一且不重复，ID 绑定了主机的奖惩地址和其物理地址信息(ip;port;容量;已有文件数量,filelist 等)

基本组件

存储网络

UPS 网络聚集了由多个独立存储提供商(主机)提供的存储，并且能自我协调的提供存储数据和检索数据服务给客户。这种协调是通过智能合约执行实现的，是去中心化的、无需信任的；通过协议的协调与个体参与者能实施验证操作，系统可以获得安全性操作。

由主机提供协议元组用以客户端上传及获取文件，包括(Put, Get, Manage）：
Put(data） → key: 客户端执行 Put 协议以将数据存储在唯一的标识符秘钥下。
Get(key） → data: 客户端执行 Get 协议来检索当前使用秘钥存储的数据。
Manage(）：验证节点通过管理协议来协调：控制可用的存储，审核主机提供的服务并修复可能的故障、管理协议由存储提供商来运行，并且经常与客户或者审计网络结合（审计网络为验证人组成的委员会）。

存储网络必须保证数据的完整性和可恢复性，并且能够具有一定的容错性。存储故障表现为阻止了客户检索数据。例如存储矿工丢失了他们的数据，检索矿工停止了他们的服务。一个成功的 Put 操作的定义是(f, m），既是它的输入数据被存储在 m 个独立的存储提供商（总共有 n 个）中，并且它可以容忍最多 f 个拜占庭存储提供商。参数 f 和 m 取决于协议的实现。

协议设计者可以固定 f 和 m，或者留给用户自己选择。将 Put(data） 扩展为 Put(data, f, m）。如果有小于 f 个故障存储提供商，则对存储数据的 Get 操作是成功的。管理协议需要兼容这种错误，如果发现数据的有效副本在一段时间内小于临界值时需要主动执行复制操作增加数据的有效副本数量。

存储证明

在 UPS 协议中,主机必须让他们的客户相信，客户所付费的数据已经被他们存储。在实践中，主机将生成“存储证明”(POS）给验证人委员会（或客户自己）来验证。存储证明(POS）它允许一个将数据外包给服务器（既证明人 P）的用户（既验证者 V）可以反复检查服务器是否依然存储数据 D。用户可以用比下载数据还高效的方式来验证他外包给服务器的数据的完整性。服务器(委员会)通过对一组随机数据块进行采样和提交小量数据来生成拥有的默克尔证明作为给用户的响应协议。

存储证明方案只保证了证明人在响应的时候拥有某些数据。我们需要更强大的保障能阻止作恶矿工利用不提供存储却获得奖励的两种类型攻击：女巫攻击(Sybil attack）、外包攻击(outsourcing attacks）。

女巫攻击：作恶矿工可能通过创建多个女巫身份假装物理存储很多副本（从中获取奖励），但实际上只存储一次。

外包攻击：依赖于可以快速从其他存储提供商获取数据，作恶矿工可能承诺能存储比他们实际物理存储容量更大的数据。

算法

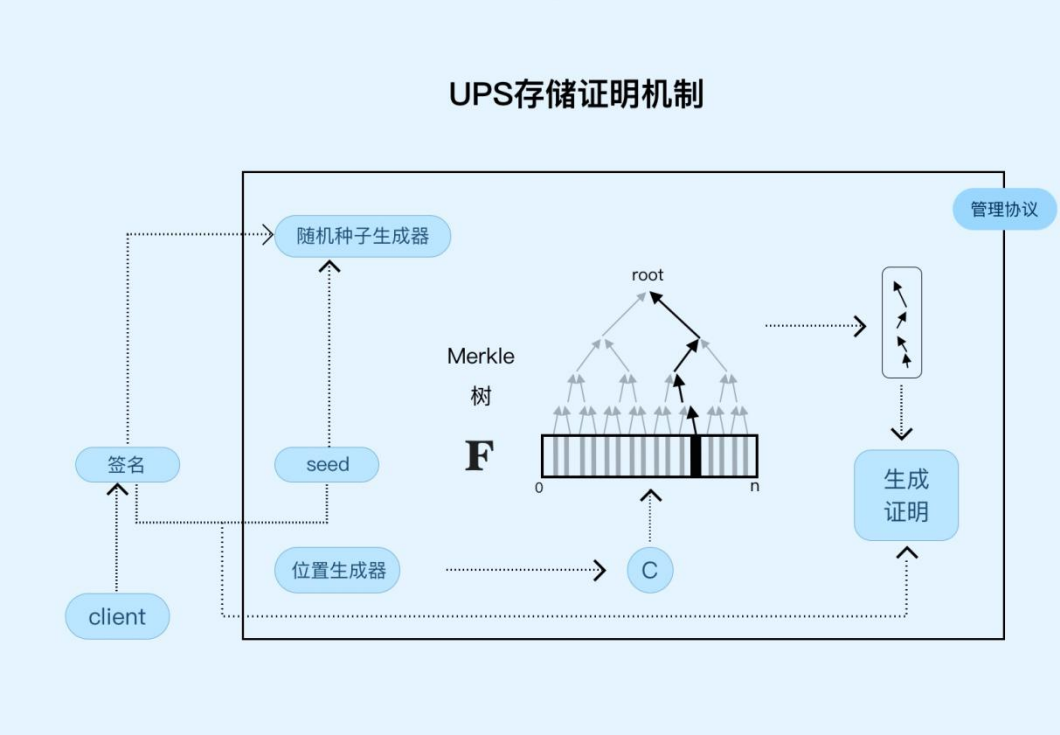
主机通过从文件的 Merkle 树中提供一段原始文件和一系列哈希来证明它们的存储。 这些信息足以证明该段来自原始文件。 由于证明提交给委员会并写入区块链，任何人都可以验证其有效性或无效性。 每个存储证明使用随机选择的段。方式：

H（H（B[Index-1]））

其中 B[i-1]是数据文件中原始文件的某一数据段。

如果主机始终能够证明拥有一个随机段，那么他们很可能会存储整个文件。 只存储 50％文件的主机将无法完成大约 50％的证明。

为了满足主机验证的随机性和防止'女巫攻击'，主机需要对来自委员会的随机种子进行签名，并使用该签名作为该文件数据段的索引，主机需要提供该段数据的原始数据并提交到委员会，并由委员会验证该证明。
IndexH = H(Sign(Random | fileid | dataid))



管理协议

文件管理合约是存储提供商与其客户之间的协议。文件管理合约的核心是文件的 Merkle 根哈希。为了构造这个散列，文件被分割成大小不变的段并散列到 Merkle 树中。根散列以及文件的总大小可用于验证存储证据。

文件管理合约还规定了校验频率，抽样方式和支付参数，包括有效证明的奖励，无效或缺少证据的奖励以及可以错过的最大证明数量。校验频率指定服务器在一定周期内必须提交存储证明次数；抽样方式规定了提交证明时，采样数据段索引的计算方式；如果证明有效将会在周期内获取一定的支付奖励(用户预先支付)；如果证明无效将受到一定的惩罚，系统会根据临界参数复制一定的数据副本。

文件服务

用户可以很方便的通过文件的散列码访问文件，通过文件服务合约请求下载文件并查找文件索引找到文件块的位置，将文件位置信息返回给用户，用户根据文件位置及序号，多路下载文件，文件下载完成后请求文件服务获取文件解密信息，用户发送文件解密交易到文件服务合约用于申请解密密钥，用户获取密钥用以解密文件获取完整文件信息。

共识机制

用户文件被加密后拆分成 N 份数据安全的存储在网络中主机服务器上，这样主机服务器即使作恶也无法知道文件的具体内容，这样保证了文件的隐私性，为了保证文件安全存储及及时响应用户的文件访问请求，UPS 采用强大的 DPOS 共识机制，基于 TrueChain 提供的质押选举合约，可以稳定获取高质量的验证人节点信息，验证人集合组成委员会，执行管理协议，验证文件索引服务及存储服务提供方，确保用户文件安全存储和快速读取，委员会随机验证抽取验证一批服务器上存储的文件，要求服务器在一定时间(高度)内，提交完整的文件存在证明，服务器将因此获取文件存储的收益。

总之, UPS 是以 TrueChain 侧链的形式提供了去中心化的存储服务，进一步丰富了 TrueChain 主网的功能及扩展。UPS 网络不是从中心化云供应商处租用存储，而是实现点对点的存储租用，同时 UPS 主网提供内置文件存储合约，定义交易细节。通过签订合约，存储提供商同意存储客户数据，并定期提交其持续存储的证明，直至合约到期。UPS 主网将实施基于 TrueChain 区块链预编译合约，与 TrueChain 挂钩，本身不发行任何通证，其共识安全也由 TrueChain 来保证。