**区块链的笔记**

# 数据存放的两类模式

　　如果说合约是内容，当内容完成数字化，区块链就成为了信息存放的载体。因此，结合梅万妮•斯万意见，数据的存放就形成了两类模式。一类是数据就存放在块中，另一类是将数据压缩形成数字串列，也就是降维哈希过程[ 参见梅兰妮•斯万《区块链新经济蓝图及导读》第21页]。对于前者容易实现，但一致性问题突出。后者只同步数字串列，通过减少同步量解决一致性问题，但哈希过程的降维成为关键。

# **一致性问题**

　　一致性问题是分布式存储系统一个比较难解决的问题，通常这一个环节起容灾备份功能，同步操作大多放在“闲时”，是一种后台操作。区块链采用“前置”的同步操作，当一个块生成，块链的更新需要所有副本完成内容一致的复制。如果内容存储在块内，那么更新的数据量为更新的内容乘以副本的数量，再加上副本间的传输距离和带宽，整个操作是耗时的，整体的效率体现在系统中最长的同步时间。

因此采用这种方式，将影响对高频应用的响应，例如诸多的交易场景。所以解决一致性问题的办法一是控制更新内容的大小，二是限定副本的数量。比特币系统的实现是采用前者，私有链是应用了后者。

# **降维哈希问题**

　　一致性问题的确是区块链技术未来真正应用的关键；如果同步的数据量仅仅是64位或者是128位的数字串列，那么一致性起码是可控的。在此模式下，区块链服务的任何数字资产(例如文件、图片、视频录像、监控记录、软件等)数据最终不需要存放在区块链上，更新的内容只是对应的数列。因此如何降维是这一模式下的问题所在，共识的内容是自然语义。当自然语义过渡到程序识别的数字表达，多维的自然语义和元数据需要完成一对一的关系映射，这一过程需要对信息进行降维，其中会用到哈希、编码等技术。例如一个9G的基因组文件，就需要降维为1维的二进制数列来平衡区块链同步前置的负外部性，最终使智能合约的执行算法能够以此甄别对象的“真伪”。需要提醒的一点是这种方式哈希过程不是地址(地址是数字资产数据的存储地址，例如指基因组文件的存放地址。要区别比特币中的哈希地址，其实质是明确比特币的所有权，类似于交易的甲乙方，是交易的内容)映射，如果区块链当中保存的是地址信息：

一是违背了内容共识的机制，二是大大降低了系统的安全性。一旦非法入侵者获得任何一个副本，就可以找到合约内容的存储空间进行篡改。降维哈希这一过程类似于“压缩”，需要基于数据内容、文件特征、时间戳等特定信息进行压缩。这样智能合约的执行算法才能够以此判定指定内容的“真伪”。所以这个过程是复杂而艰巨的，工程量浩大。

# **IPFS：基于区块链的去中心化存储网络**

IPFS是做基于区块链的去中心化存储网络。

怎么进行内容寻址？

1，给所有发布的内容进行哈希值计算；  
 2，构建哈希值索引地址（所有文件都有唯一哈希值，不产生冗余内容）  
 3，用户访问某个文件，会广播哈希请求，找到存储该文件的节点，传输给用户  
 这里面，有一个关键点，矿工奖励机制，奖励矿工贡献存储，贡献数据检索服务，让IPFS网络最终成为真正可用的网络。  
 也就是说，用户通过购买代币Filecoin获得数据存储和检索的服务，矿工则通过提供数据存储和检索服务获得代币的奖励。

**访问或下载：**

传统的网络下，访问网络或服务器文件都离不开http协议。当你输入网址，点击网页连接，就会向中心服务器地址寻找文件。如果有很多人同时访问，就容易造成网络堵塞，速度很慢。  
 在互联网上太多重复的文件，比如音乐、电影等。如果这些文件可以共享，就不用存储在很多的服务器上，造成存储和计算资源的浪费。IPFS通过内容寻址技术实现定位公有资源。它将多重文件通过哈希计算分散存储在不同的计算机上，用户访问一个文件，则按照哈希地址访问即可。另外，它有点对点的对等特性，它将文件碎片化存储在距离用户最近的计算机或服务器众，加载速度大幅提升，如果你所在的城市或者邻居有节点，那速度就更快了。甚至可以达到访问本地文件的速度。如果你想下热播剧，这就是最好的体验。  
 当然这也取决于IPFS网络节点本身的效应。这也是为什么绝大多数的代币都会奖励给矿工的原因。