**上海大学“区块链”课程论文考试**

**Filecoin项目分析**

姓名：**王思楠** 学号： **21726048**

**字数：5081**

**摘 要：作为分布式存储项目赛道的领头羊，Filecoin从提出之时就备受关注。在2017年ICO时，该项目就收到了红杉资本等八大知名投资，可谓风光无限。本文想通过对Filecoin项目理念、代币机制、交易过程、共识机制的解读，以及对其当前发展状况的分析，发现项目的投资价值。**

**关键字：共识机制；交易流程；未来前景**

**论文评语：**

成绩： 任课教师：吴建刚

1. 经济模型分析
   1. 项目、创始人及团队简介

胡安·贝内特（Juan Benet）1988年3月出生于墨西哥，2006年，胡安以优异的成绩进入斯坦福大学学习计算机科学。从入学的第一年开始，他就先后在Anticancer、Verve Wireless等企业实习，这为他后来的创业打下了一定的基础。本科毕业后，胡安又继续在斯坦福攻读计算机硕士学位。与此同时，他还与其他几位斯坦福大学计算机科学的毕业生一起创办了一家手机AR游戏公司（Loki Studios），并出任CTO一职，后来该公司被雅虎收购。2012年，胡安从斯坦福研究生毕业后又再次创业，创办了Athena。由于Athena发展得并不顺利，一年之后，胡安选择了离开。在此期间，胡安第一次接触到了区块链技术，这为他后续的区块链创业奠定了基础。[1]

2014年，胡安与几个好友一起创办了协议实验室。和胡安一样，团队中的大多数成员都来自国际知名名校，其中斯坦福大学的居多。作为一个开源组织，协议实验室也十分乐于其他有志者参与到项目之中。2015年1月，协议实验室发布了IPFS项目，旨在建立一个分布式存储和共享文件的网络传输协议。2017年，他们又以IPFS为基础，构建了去中心化存储网络Filecoin。2020年10月15日，在经历多次公开测试和两次太空竞赛后，Filecoin主网正式上线。可以说，Filecoin项目的成功离不开协议实验室团队和整个开源社区的努力。

* 1. 项目想要实现什么

作为一种新型的网络传输协议，IPFS的提出弥补了原先HTTP协议的不足。传统的HTTP协议在使用时，浏览器会根据我们输入的网址，找到相应的服务器，然后从中心化的服务中读取数据显示在页面上。Dias D, Benet J(2016)指出在IPFS协议下，文件的内容分散成不同的碎片存储在多个节点当中，每一个碎片生成独一无二的哈希值，当用户需要访问时，只需要提供文件的哈希值，就能找到存储该文件的节点进行读取。[2]由于数据是分布式存储的，不仅黑客无法进行精准攻击，同时回应用户的访问申请速度也会加快，在大幅提高数据存储的安全性的同时也节约了带宽，降低用户的成本。

Juan Benet（2014）年提出Filecoin是针对IPFS提出的一项激励机制，目标是构建一个去中心化、高效率的存储市场。在IPFS网络中，人们不仅可以提供自己闲置的硬盘，也可以购买所需的存储空间。无论是提供硬盘，还是硬盘存储到数据后，都会获得相应的奖励。[3]这些奖励就是FileCoin，即FIL币。当Filecoin的矿工越多，对IPFS网络的支持也就越大，同样地，IPFS网络使用的越多，对 Filecoin的需求就越大。协议实验室正是通过这种方式来促进IPFS的发展。

* 1. 为什么需要分布式记账

Filecoin之所以采用分布式记账，最主要的原因就是安全。

随着互联网的快速发展，越来越多的数据被生产出来，而且伴随着物联网以及5G技术的发展，数据的生产速度也将不断加快。大量的数据给本地存储带来了巨大的压力，也增加了用户存储数据的成本。于是便产生了百度云、阿里云、腾讯云、华为云等云端存储，尽管这种存储方式在一定程度降低了用户的成本，但本质上依旧是存储在中心化的数据中心，数据安全隐患依旧存在。Filecoin是通过将数据上传至本地节点后，在经由矿工存储在相应的节点中。这些节点由不同的矿工提供，外界无法实现精准的攻击，有效地保证了数据的安全性。此外，用户存储的数据也可以通过预先进行加密来保证数据的隐私性。

* 1. 代币机制

FIL币总发行总量为20亿枚，它采用了线性递减释放的机制，每6年便减半一次，42年后FIL币将接近被挖完。根据协议实验室的设置，FileCoin的分配方案由6个部分组成：

55%作为矿工挖矿的奖励，通过区块奖励的方式线性释放，用于维护区块链，运行合约；15%作为矿工储备，用于给检索矿工和维修矿工提供激励；10.50%作为协议实验室团队的研发及运营费用，按6年线性释放；4.50%分配给协议实验室团队和其他主要贡献者；10%分配给参与私募与公募的投资者，按6-36个月线性释放；剩余的5%则留给基金会，作为长期社区建设，网络管理费用等，按6年线性释放。[4]

在Filecoin中目前有两类矿工，存储矿工和检索矿工（如图1）。

存储矿工主要是在网络上提供数据存储。在交易开始后，矿工需要用一定数量的FIL币作为存储数据的抵押。在存储过程中，矿工需要不断生成时空证明并提交给区块链来证明自己一直在存储数据，如果证明无效或者丢失，将被扣除一部分FIL币作为惩罚。[4]

存储矿工的奖励主要分为以下三个部分：交易费、区块奖励、网络交易加速费。交易费和网络交易加速费是客户在发起交易时已经确认好的，一旦找到意向的矿工，这笔费用便会在钱包中锁定，在确认交易后将根据存储矿工不断提供的时空证明递增支付。一般来说，交易费用是由存储矿工根据文件大小以每千兆字节确定的。而区块奖励则是系统生成的，如果通过时空证明验证矿工存储的数据为有效数据，就会获得区块奖励。[5]

检索矿工主要负责为网络提供数据检索服务。与存储矿工不同，他们不需要进行抵押，也不用承诺存储数据或是提供存储证明。由于需要快速进行数据传输，检索矿工对带宽要求也较高。检索矿工的奖励也十分单一，只能获得检索的交易费。交易的费用同样是由矿工进行报价，包括每字节价格、解封价格等。这笔交易费用根据用户接收到的数据逐步支付的。每拿到一份数据，用户就会在链下通过支付通道支付相应的费用给矿工，有效保障了用户的权益。[5]

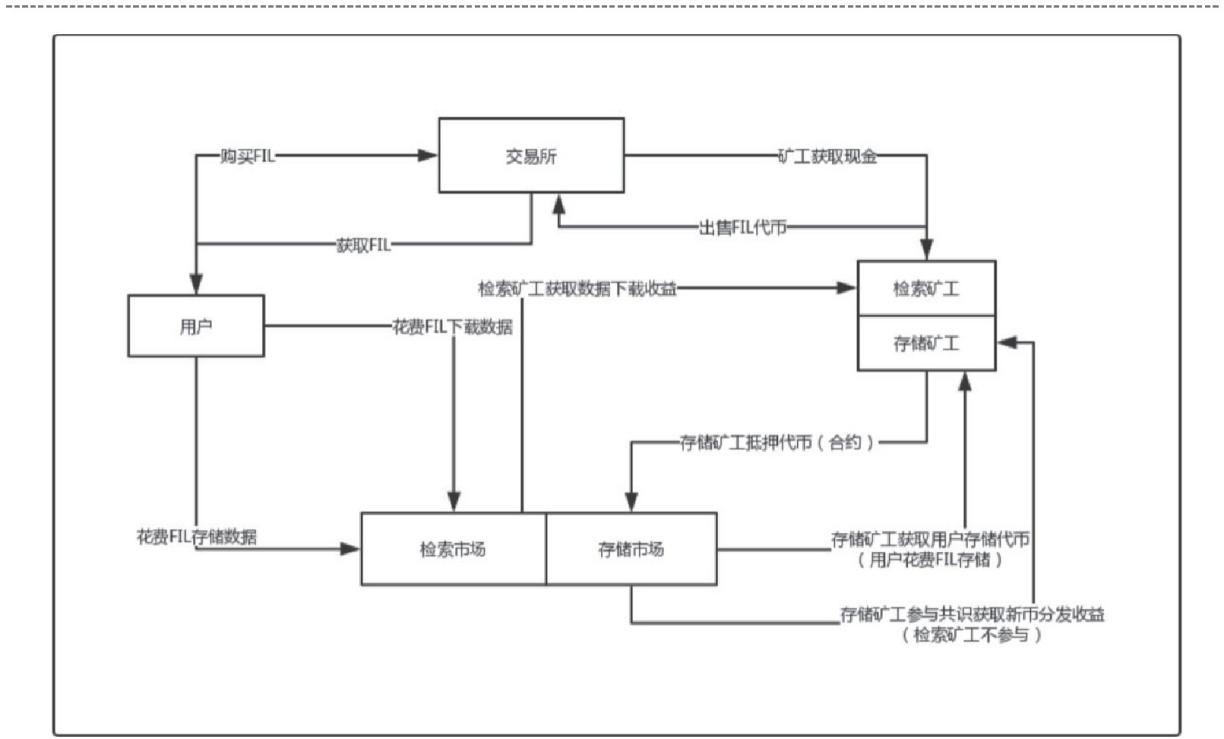


图1 关于Filecoin交易机制的图[6]

1. 技术模型分析
   1. 共识机制

与其他区块链项目不同，Filecion创新性的采用了混合共识机制：即以预期共识为主，复制证明与时空证明为辅。

预期共识机制，就是在每一轮选举出一名或者多名矿工来创建新的区块，每个矿工每轮只有1次机会，矿工赢得选举的可能性跟他们当前的有效存储能力成比例。尽管被称为选举，但实际上这个过程是由节点自己进行计算的，矿工之间并不产生交互。选举出的矿工只是得到了出块权，随后系统会给出一个随机数，如果矿工有效存储的占比大于该随机数，就能够进行出块。满足条件的矿工必须在30秒内完成所有消息的打包并在区块链上进行记账，同时提交自己的在获得出块权时的时空证明，才算成功出块。

在Filecoin中，每一轮的选举都被看做是一个纪元。由于满足出块条件的矿工数量不定，因此在每个纪元中就可能出现空块或者多个区块的情况。随着新一轮选举的开始，新的区块又在前一纪元中生成的区块上不断增加，这就生成了多条链条（如图2）。为了避免有矿工为了获得更多奖励而同时参与多条链，Filecoin赋予了每个区块权重，并结合抵押机制来促使矿工选择一条最有效的链，对同时挖多个链的矿工进行惩罚，从而保证资源的集中。[7]

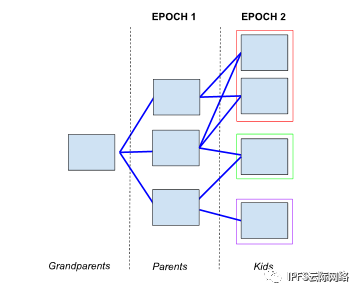


图2 关于Filecoin区块产出过程的图[7]

而复制证明和时空证明则主要是用来证明矿工的存储能力的。

复制证明是矿工将存储数据的区域转换为唯一的副本，并将验证秘钥提交给区块链的过程，用来证明自己确实存储了用户的数据。

时空证明则是要求矿工在固定时间间隔内随机提交文件的某个部分的证明给区块链，证明其在一定时间内一直存储着该数据。即使验证者不在线，也能够在未来去验证矿工在该段时间内生成了证明链。如果矿工没有通过证明，那么之前提交的抵押便会被扣除一部分，有效地避免了数据的丢失。

这三种机制相辅相成，保证了Filecoin有效存储的概率，同时也证明了其公平性、保密性与公开可验证性。

* 1. 帐户机制及智能合约

Filecoin的每个账户都有自己的地址、余额、状态，用户可以随时查看自己的账户信息。当用户发布交易后，便会在用户的账户中将该笔交易费用进行锁定，并根据交易进展逐步转出至矿工账户。交易完成后，用户的账户余额状态也会得到更新。由于奖励的获得是按照180天内线性释放的，因此矿工账户的余额状态并不会立刻发生变化。这种方式在一定程度上避免了参与者的短期投机行为，保证了Filecoin项目的稳定发展。

Filecoin的智能合约计划通过FVM虚拟机来实现，目前仍处于开发阶段。根据协议实验室发布的路线图，在今年的第2季度，虚拟机第一阶段可以正式上线，而智能合约的实现预计到9月才能正式发布。

* 1. 数据库、隐私保护、身份验证等

作为一个存储类项目，Filecoin本身就是一个非关系型的数据库，为其他区块链项目提供底层的存储。在Filecoin中，用户的数据就存放在各个不同的节点中，这些数据不仅可以长久的保存下来，而且通过验证秘钥，网络上的所有人都可以随时检验数据的存储状态。此外用户也可以通过存放多份数据、预先加密数据、或者将数据分散成多份的方式进行存储，有效的保护了用户数据的隐私安全。而且Filecoin独有的惩罚机制也能够保证用户数据被完整的保存在节点中，不用担心数据的恶意丢失。目前在Filecoin中，用户在交易时并不需要进行身份的验证。每一个用户都有独一无二的节点ID与账户ID，有存储需求的用户通过提供账户ID和存储需求给矿工，就可以进行交易。[8]而要提取数据的用户则通过提供自己上传数据时生成的内容ID来在市场上进行检索，找到对应的矿工并获取数据。

* 1. 现有技术文档及代码评价

Lotus是Filecoin分布式存储网络的一个简单实现，采用Go语言进行编写，实现了挖矿、存储、数据检索这些主要功能。Filecoin整个的代码结构是模块化的，不同程序功能的实现放在不同的模块里。例如：不同类型的账户尽管程序中包含的内容不同，但都是归属在账户这一模块下。存储过程的协议中包含了从发起交易、到执行交易再到检验交易等整个过程，而检索过程的协议则只包含了简单的检索请求。这种代码模块化的形式完全符合Filecoin项目开放的特征，开发者可以单独对某个模块进行调整，即使同时协作，大家相互之间也不干扰。而且整个代码的结构十分清晰，便于后续的维护。

1. 项目未来
   1. 未来规划及激励机制

Filecoin未来的发展主要分为两部分：

第一是技术方面，目前Filecoin正在进行虚拟机的开发，并希望通过虚拟机来实现可编程的智能合约。协议实验室在自己编写的同时，也邀请到更多的开发者参与其中。与此同时，Filecoin通过黑客松发布了Filecoin VM Bug赏金计划，邀请外部的开发人员来帮助发现FVM M1的各种潜在漏洞，为之后加入智能合约奠定好的基础。届时Filecoin的机制会更加完善，甚至实现跨链操作的可能性。

此外，提升数据读取速度也是Filecoin所关注的一个点。由于用户从存储数据到取出数据的过程中步骤较多，并且打包数据与检索数据都需要一定的时间，因此目前Filecoin中存储的数据大多为不经常访问的冷数据。对于偶尔需要修改的温数据，Filecoin目前采取重新写入的方式来进行更新，这在一定程度上也花费了不少时间。因此未来协议实验室将在这一方面做出调整，降低用户读取数据的成本，甚至实现热数据的存储与读写。

第二是项目生态方面。近年来游戏行业快速发展，加上VR技术的实现与元宇宙概念的提出，游戏行业预期还会有较大的增长。而该行业的特点就是数据存储量大，并且对数据安全性要求较高，因此Filecoin 也将其列为重点发展的方向。未来Filecoin将向更多行业领域进军，提高自己的影响力。

在激励机制方面，与比特币的指数衰减模型不同，Filecoin并没有采用以时间长短来确定铸币的数量的方法，而是根据网络中总存储能力的增长而增加的。如果存储能力没有达到预期值，则会延迟一部分的区块奖励。这种将奖励与存储能力挂钩的方式也使得参与者能够保持参与的热情，保证了项目的长久性。

* 1. 现有社区及应用评价

目前，在Filecoin上有1857882个账户，4083名活跃的存储提供者，9776名贡献代码的开发者,还有超过380个组织在Filecoin上建立，整个社区的活跃度较高。Messari研究报告（2022）中指出：从已使用的存储和网络能力来看，Filecoin占据了最大的市场份额，也受到许多组织与社区的支持（见图3、图4）。在这些组织的支持下，Filecoin生态也逐步发展起来。他们通过黑客松、加速器、资助计划等方式来帮助其他的项目尽快走向成熟，扩大了Filecoin的生态。[9]与此同时，应用在Filecoin上的项目数量也从2021年起步的40个项目，发展到超过1630个项目，并涵盖了从数据市场、公共数据、视频，再到网络出版、游戏等多个领域。目前，数字令牌存储和公共数据的存储是Filecoin主要的应用场景，NFT便是其中之一，Filecoin安全、永久性的存储正好满足了NFT上各种艺术品的存储需求。随着web3的到来，未来还会有更多项目将应用在Filecoin之上。

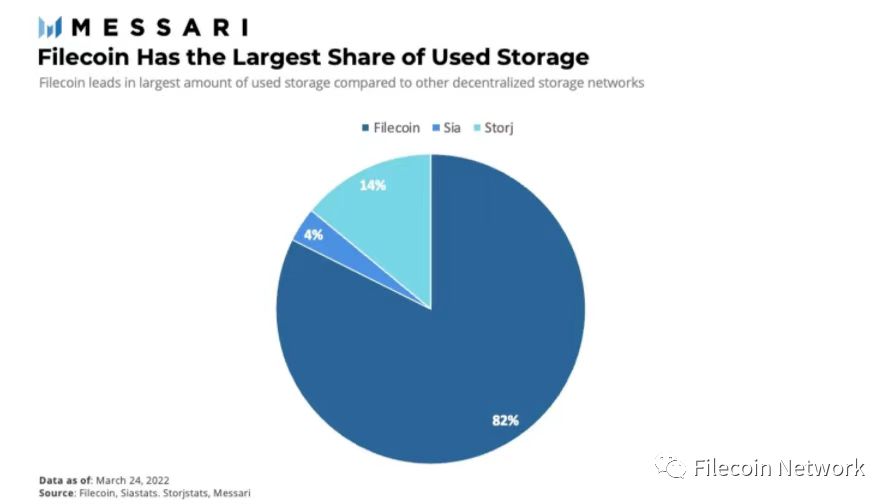


图3 关于Filecoin在存储市场占比的图



图4 关于Filecoin生态中主要组织的图

* 1. 项目投资分析

作为分布式存储项目的领军者，Filecoin从设计之初就一直致力于解决构建一个去中心化、高效率的网络，这与互联网技术发展所带来的困境不谋而合。而且项目的创业团队从建立之初就以一种开放的姿态面向大众，任何开发者都可以通过加入社区参与其中，大家集思广益，保证了项目的运转。社区中也会不定时组织各类活动，来提高开发者的积极性。此外，Filecoin也收到了来自红杉资本、斯坦福大学、文克莱沃斯兄弟基金、Y Combinator、DCG集团等八大知名机构的投资，这也是支持该项目继续发展的坚实后盾。

从协议实验室发布的白皮书和技术路线图来看，Filecoin目前还有许多功能并未实现，例如：智能合约、维修矿工、可扩展共识、高效检索等，都处于初步探索开发的过程。如果这些功能一旦全部实现，那么Filecoin将迎来质的飞跃。

总的来说，我认为Filecoin项目在上线仅1年多的情况下就有数百个项目选择将数据存储存储在其网络之上，也表明其具有较好的成长性与较高的安全性。尽管存在着一些不确定的风险，但Filecoin仍然是一个值得投资的项目。

参考文献：

1. 胡安：22岁创业、29岁1小时融资2.57亿美元，区块链贡献度仅排在中本聪、V神之后[J/OL].币伙计.2020-12-03.
2. Dias D, Benet J. Distributed web applications with IPFS, tutorial[C]. International Conference on Web Engineering. Lugano,Switzerland,2016: 616-619.
3. Juan Benet. IPFS - Content Addressed, Versioned, P2P File System. 2014.
4. 董天一.解密基于区块链的分布式存储[J].软件和集成电路,2019,(11):23-27.
5. Filecoin是如何进行数据存储和检索交易的？[EB/OL].贝拉比.2021-03-23.
6. Protocol Labs. Filecoin: A Decentralized Storage Network[EB/OL].2017-07-19.
7. 区块.FILECOIN区块结构[EB/OL].简书.2020-10-22.
8. 猩动力.Filecoin如何创建账户钱包并获取FIL测试币[EB/OL].知乎.2020-04-08.
9. Messari研究报告：Filecoin生态概述[EB/OL].2022.

附1：

1. 请说明发起比特币交易到交易确认的整个过程。（说明中请包含：交易者如何产生地址、如何对交易签名、如何发出交易、矿工如何接受交易、如何选出记帐人、为何会使用矿池挖矿、矿池挖矿的过程、为何可能会有软分叉、为何限制区块大小、如何控制区块出块时间、区块中包含的内容，并讨论为什么使用UTXO设计）

（1）比特币交易过程：

①支付方打开钱包，复制交易方的地址；

②填上转账的金额以及交易费用；

③使用私钥进行签名并发送交易；

④交易被广播到网络上，附近节点接收并验证；

⑤由矿工将该笔交易以及上一个区块的Hash值打包进区块中；

⑥矿工之间竞争挖矿，通过PoW找出记账人；

⑦节点验证结果并检测区块；

⑧交易方确认交易。

（2）交易地址的产生过程：

①通过对一个长串随机字节，进行SHA256哈希算法，产生一个随机产生一个256位的二进制数字，也就是私钥。

②私钥通过椭圆加密曲线计算出公钥，并通过SHA256、RIPEMD160算法计算出公钥哈希。

③对公钥哈希增加版本号后进行连续两次SHA256运算，并将最后一次哈希结果的前4字节作为校验码附在版本号和公钥哈希之后，然后再经过BASE58编码计算得到钱包地址。

（3）如何对交易签名：

①查找该笔交易对应的UTXO；

②获得该UTXO对应的锁定脚本；

③复制该交易对象，并在复制副本中设置对应的锁定脚本；

④使用哈希计算得到交易的摘要；

⑤使用私钥对该哈希值进行签名。

（4）如何发出交易：

①创建用户，拿到自己和交易方的地址；

②将交易的金额以及支付的手续费填写好；

③使用自己的私钥对交易进行签名；

④签名后付款节点会将交易广播到全网。

（5）矿工如何接受交易：当交易信息被广播至整个网络时，附近的节点就会对该交易进行验证，矿工在交易池中根据每笔交易所出的手续费来挑选进行自己打包的交易进行打包。

（6）如何选出记账人：通过分发不同的题目给矿工，矿工计算出哈希值，然后通过两两计算得出根哈希值，即梅克尔数。然后对最后计算出的根哈希值加上某个随机数，如果该值在某个范围内，那么这个计算结果就是被认可的，计算出该值的矿工当选为记账人。

（7）为什么选择则矿池挖矿：

①集中了大量算力的矿池，挖到区块的概率会大大提升；

②挖矿得到的奖励会按照一定分配方式分给矿工，矿工可以获得更加稳定的收益；

③加入矿池后矿工只需要负责进行哈希运算，其他事情交由矿池处理，提高了挖矿效率。

矿池挖矿的过程：

①矿机发送请求接入矿池，矿池回应请求；

②矿机登陆矿池，并由矿池返回登陆结果；

③矿池将预先打包好的任务分配给矿机；

④矿机进行哈希运算，当找到符合条件的一组时，进行任务提交；

⑤矿池确认任务并将矿机提交的值与全网的难度条件进行验证，如果符合则成功爆块。

（8）为何可能会有软分叉：比特币每次代码升级都要获得比特币社区的一致认可，如果比特币社区无法达成一致，区块链就很可能形成分叉。当新的共识规则发布以后，没有完成升级的节点由于不了解新的共识规则，产生不合法的区块，从而产生临时性的分叉就是软分叉。

（9）为何限制区块大小：区块的大小会对每个块可以验证的事务数量产生限制。较大的区块需要更大的计算能力并且需要更长的时间来开采。打包一个大小为1M的区块，数据传输和验证的时间很短，全网同步账本的时间也很短。打包一个较大的区块耗时较久，在这个过程中容易受到外界攻击，影响比特币网络的安全性。

（10）控制出块时间的方法：通过让矿工计算数学问题得出哈希值来控制出块时间，只有计算出范围区间内的哈希值时，这个区块才能够产出。这个题目的计算难度一般需要10分钟才能完成。

（11）区块中包含的内容：版本号、本区块哈希值、本区块高度、上一个区块的哈希值、本区块交易数据打包数据的哈希、Unix时间戳、交易的数量统计、本区块大小、本区块总交易额、目标值（难度）、随机数，以及交易详情。

（12）为什么使用UTXO设计：

①UTXO设计能够证明用户拥有的比特币的真实性，易于确认比特币的所有权；

②UTXO设计状态不可更改，与区块链账本是完全融为一体的；

③UTXO设计特别的数据结构可以让双重花费更容易验证。

1. 请说明如果签名正确、余额足够的情况下，发出的交易迟迟没能出块，应该是什么原因，应该采取什么措施，分别说明不同措施的后果。

（1）原因：①手续费出价较低；②当前网络上提交的交易过多，造成网络拥堵。

（2）措施：针对手续费低的原因，一般有以下几种措施：

①耐心等待，直到有矿工将交易进行打包。这种方式一般会耗费大量的时间，如果该笔交易又急于完成，那么就会导致交易被对方口头取消。

②借助交易加速器来加速交易的打包。这种方式需要另行支付一笔费用，会加大交易的成本。

③选择网络畅通重新发起一笔交易，或者重新发起一笔高手续费的交易。使用这种方式同样也增加了交易的成本，而且前一次交易如果之后也被打包，就需要对方将收到的比特币再转回，过程比较繁琐。

1. 请说明以太坊项目的目标，它与比特币不同的技术设计及其原因。

（1）以太坊项目的目标是建立一个内置有成熟的图灵完备编程语言的区块链，使得任何人都能够创建合约。

（2）与比特币不同的技术设计及原因：

比特币使用未花输出模型，而以太坊使用的是账户模型。因为在比特币中并不存在账户，只有地址，而且每一笔比特币都源自于上一个交易，可以一直追溯到源头。这种模式下，系统中并没有显示的记录每个账户上有多少钱，如果想要清楚手中中的余额就需要一层层推算，十分不方便，于是以太坊便提出账户模型，使用户能够清楚的知道每个账户中到底拥有多少以太币，每个账户都有对应的用户，这也便于智能合约的实施。

4、请说明以太坊智能合约的运行原理以及合约的整个生命周期。

（1）智能合约的运行原理：智能合约就是在区块链上运行的自执行计算机程序，该程序规定和实施合约的协商和与执行。以太坊去中心化的性质，确保了合约发起的交易能够自动、如实地执行。

（2）合约的生命周期：

①使用智能合约编程语言编写智能合约；

②将合约代码编译成EVM字节码；

③将合约代码上传部署到以太坊区块链上；

④在网络上通过交易执行该合约；

⑤将合约进行销毁。