



下载APP



## 10 | 技术迭代：美好背后的遗憾

2021-08-11 赵铭

《说透区块链》

课程介绍 >



讲述：赵铭

时长 15:02 大小 13.77M



你好，我是赵铭。

前面几讲我带着你深入了解了区块链技术的基本技术要点，不过我并没有与你提及关于每种基础技术的弊病，但你要知道，世界上并不存在十全十美的事物，技术也一样。

区块链技术是 4 种基础技术加智能合约的融合创新，讨论其弊病其实是很复杂的。既有区块链从基础技术那里继承的“遗传病”，也有因为融合多个技术而导致的冲突，想解决这些问题我们不得不引入新的技术修补、妥协。



技术的发展是一个渐进的平衡过程，并不存在从诞生就能一成不变的技术手段。区块链技术虽然才诞生十几年，但已经进行了多次迭代，**每一次的技术更新其实都是服务于实际问题的破解。**

这一讲，我会带你去探讨区块链发展过程中 3 个比较典型的共性问题。

## 存储冗余

还记得我讲的区块链技术特点么，第一个特点就是去中心化。而这个特点的负面作用就是导致了存储冗余，这是怎么回事儿呢？

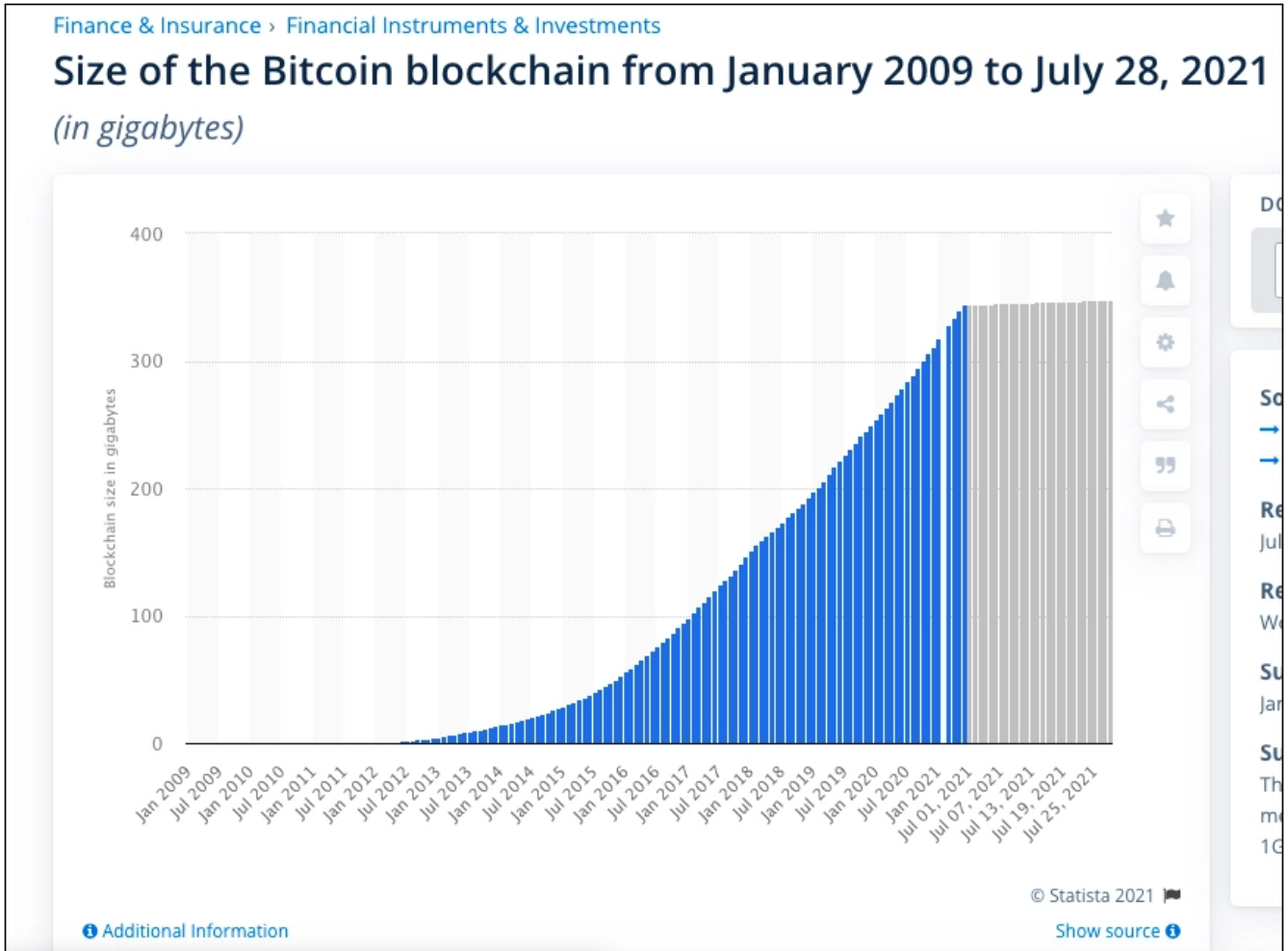
我们还是延续之前的思路，聚焦到区块链网络节点分析。

要想实现去中心化，就要让区块链网络中所有节点都是平等的，没有服务端客户端的区别，即便其他区块链节点都失效了，只要保留了一个节点，整个网络还是可以恢复的。所以在区块链网络中，是没有绝对集权存在的。

我们推演一下，要想实现只有一个节点也能恢复的目的，就意味着每个节点都要继承区块链网络里的全套“DNA”，也就是保存从创始区块到最新区块的完整数据。这样，随着区块链网络持续运行、交易量不断增加，单节点的存储负载自然越来越多。

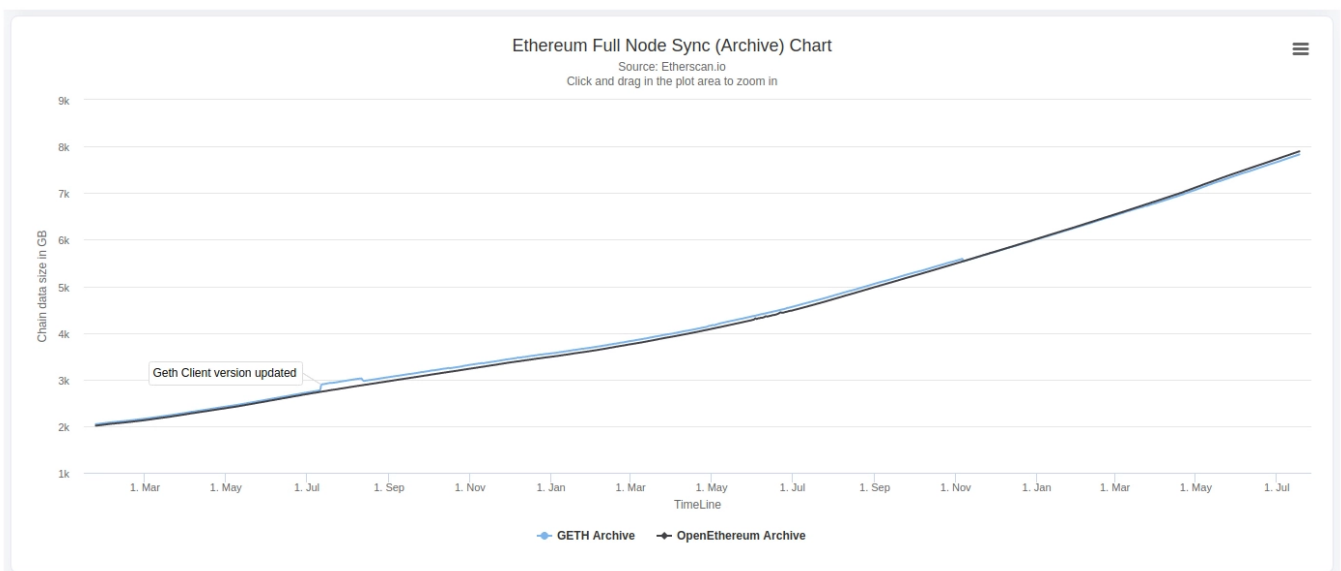
以上是从单个节点的角度进行的分析，如果我们将视角转换至整个区块链网络，就会发现，每增加一个节点，整个网络就多了一份数据的备份，整个系统的存储负载会随着区块链节点的增加而线性增长。

我们拿比特币 / 以太坊网络举例。截止到今年 7 月份，比特币网络已经累积超过 350G 的历史数据。这也就意味着，如果你想要在自己的电脑同步全量的比特币历史数据，需要占用至少 350G 的存储空间，而且时间越长，空间使用量就越大。



比特币空间使用量示意图

再来看以太坊的空间使用量，那就更加恐怖了。以太坊在比特币的基础上增加了智能合约，所有的使用者都可以通过智能合约存入自定义数据，所以网络越活跃，产生的数据量也就越多。目前为止，从 [Etherscan 网站](#) 得到的数据，按最全的节点数据统计，单节点已经有超过 7800G 的历史数据了，而且有加快增长的趋势。



以太坊空间使用量示意图

接下来，我们再分析一下数据量增长会带来怎样的影响。虽然随着技术的发展，磁盘的成本在逐渐降低，区块数据的增长不会线性的增加存储成本。站在参与节点的角度来看，每个节点只保留一份区块链数据，成本是分摊给了网络的每一个参与者，并不会因为存储而增加额外的成本。

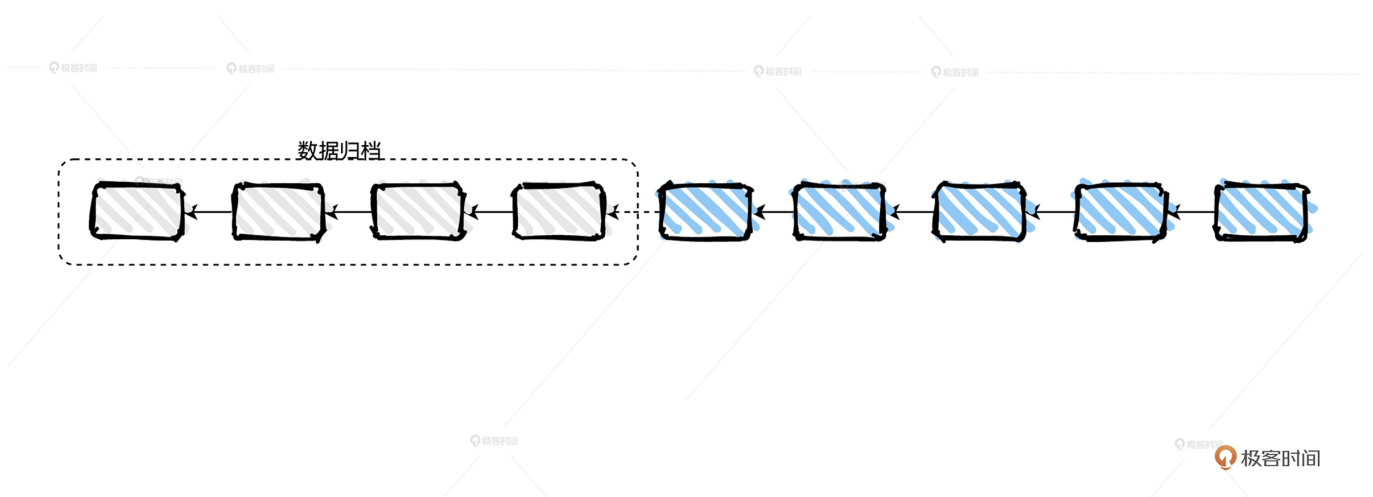
但终究会有一个时间临界点，会导致独立节点的存储使用量超过现阶段个人电脑的存储上限，而这也就会失去大量的小节点，从而导致去中心化程度的降低，不利于区块链网络的持续长久发展。而且，随着数据量的增多，通过网络同步区块数据也变得越来越困难，有些节点同步区块的速度还跟不上整个网络区块产生的速度，如果这样的情况越来越普遍，同步将变得没有意义。

数据存储冗余的问题急需解决。你可以想想如果是你来解决这个问题，你会如何处理呢？

最直观也是最简单的解决方案就是删除历史数据，完成数据的归档。首先我们可以先思考一个问题，就是区块链为什么要保留区块历史数据？

有了前面学习的基础（不熟悉可以回看 [第 3 讲](#)），这个问题并不难，区块链是由区块哈希前后关联形成的链式结构，主要目的就是保持数据的持续完整性，让区块链在保持不可篡改特性的同时获得可追溯特性。

而数据归档是将区块链从某历史区块开始，将之前的数据清理掉，但随后的区块还保持原有的顺序不变，因此**数据归档的本质是在保持不可篡改的前提下牺牲部分可追溯性**。但如果我们能在将区块历史数据清理以后，还提供一种方式将被归档的区块再次读取出来，那数据归档其实是一个比较优秀的解决方案。



数据归档示意图

不过，如果稍微查查资料你就会发现，包括比特币 / 以太坊在内的公有链，几乎都不采用数据归档这种方法，这又是为什么呢？

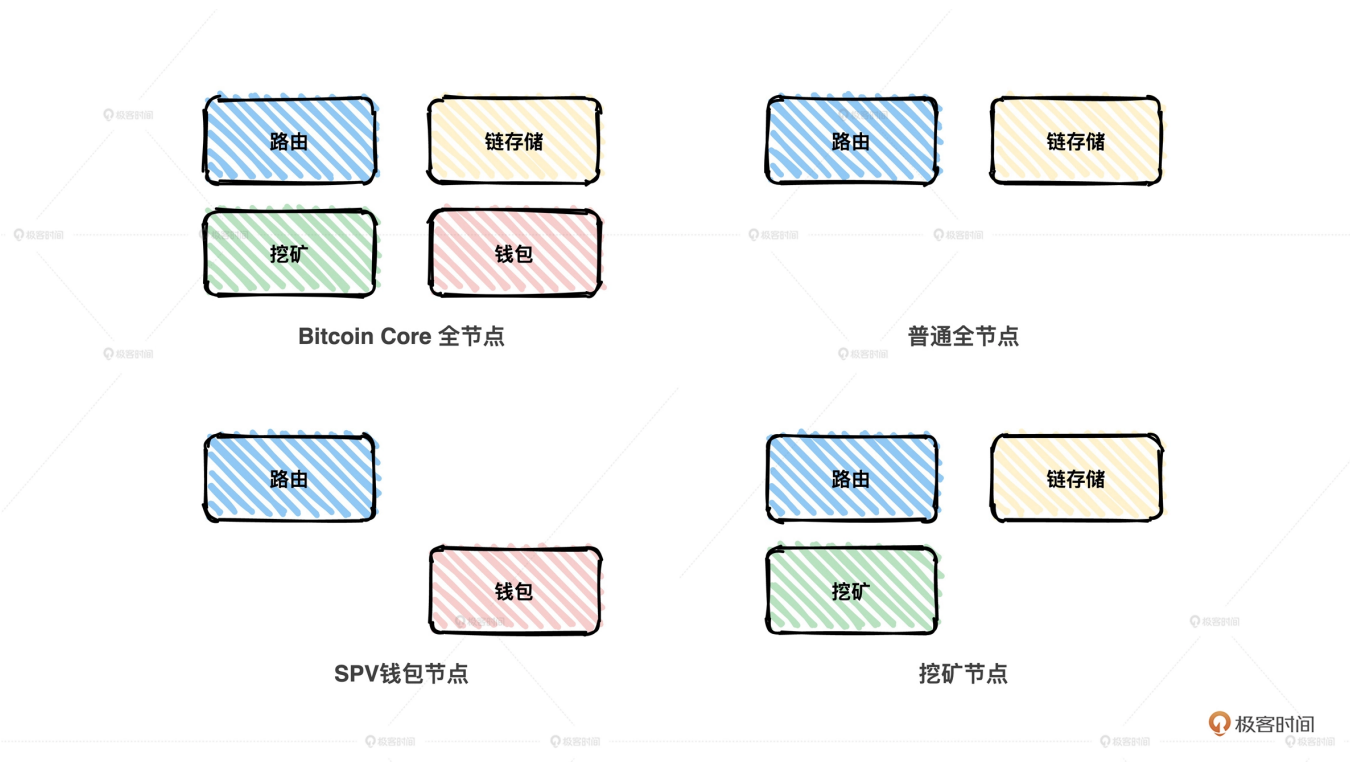
这主要是因为公有链是公开的，所以你不能预设所有用户的使用场景，有些应用可能赖以生存的根源就在于可追溯性，社区不能剥夺这部分用户的权利。

因此，数据归档方案一般都是联盟链网络在使用，和业务背景一结合就很容易理解了，因为联盟链的业务场景比较单一，而且可以提前预知。如果不太需要可追溯性，就可以将部分历史数据归档，但如果需要时，也可以通过技术手段恢复。

除了数据归档，解决数据冗余还有什么思路呢？如果你是个严谨的人，平时工作的时候通常有分类文档的好习惯。而这里我要分享的第二种解决思路也是这样，核心思想就是分类，这也符合大多数人的思维。

如果我只是一位比特币的持有者，我只关心我的余额以及历史交易记录，其他的信息我一概不关心。因此我并不需要因为这一小小的需求去同步 350G 的存量数据，这些数据对我来说并没有什么意义。

但对于某些大型的节点来说，它就可以同步历史数据，并以此数据作为生产资料，对外提供相关服务，比如挖矿的矿工节点以及区块链浏览器背后的节点。公链普遍采取的就是这种节点角色分类的方案。



比特币网络现在大概有 10000 个全节点，由它们维护完整的区块链数据，而持币者只需要在手机上运行一个钱包节点就可以进行交易。可以发现，**虽然节点角色区分是缓解存储冗余的一种有效手段，但这是以牺牲去中心化特性作为前提的**，从某种程度来说，这也是一种妥协。

解读了两种不同的存储冗余的解决方案，可以发现我们无法从根源上解决存储冗余的问题，只有通过不同的技术手段缓解数据的增长。联盟链以牺牲区块链可追溯性作为突破口，而公链以牺牲去中心化为前提。

另外，其实我们还可以有另外一种办法解决存储冗余，那就是减少数据量，减缓数据增长的趋势。比如在区块链上存储的是原始数据的哈希，而不是数据本身。当然，这只是一种业务层面的解决方案，并不是一个技术维度上的办法。

## 量子计算威胁

说完关于存储冗余的问题，我们再看看另一个大家津津乐道的问题，那就是量子计算对区块链的威胁。我们普遍觉得，量子计算机一旦成熟，以比特币为首的区块链系统就会面临崩溃，因为量子计算机的算力与经典计算机相比，不可同日而语。

但事实真的是这样吗？为了更好地帮你理解这个问题，我觉得有必要先为你科普一下什么叫量子计算。当然，我无法为你去解释叠加态 / 量子纠缠等复杂特性，因为太高深了，所以我只能通过一个从《科学声音公众号》借鉴过来的例子帮助你理解。

想象一下，如果让你使用 1 只手，你可以同时表示出几个数字呢？毫无疑问，你只可能同时表示出 1-10 这 10 个数字中的一个，而这也正是我们使用经典计算机存储数据的基本规则，同一时间一个比特位只能存储一个二进制数，要么是 0，要么是 1。

而如果让你将你的手揣入口袋，请问如果让你将手从口袋里拿出来，那么你这只手有可能表达出多少数字呢？答案肯定是 10 种可能，而且在你没将手拿出来前，数字是不确定的。

例子说完了，从这个例子就可以看出经典计算机与量子计算机存储数据的差别，一个是**存储具体的值**，而另一个是**存储值的概率**。



什么是存储值的概率呢？其实就是把值的所有可能性叠加在一起存储。比如 5 比特位的经典计算机，同时只能表达一个数字，而同样是 5 比特位的量子计算机则可以同时表示  $2^5$  个数字，二者有着 32 倍性能的差距，而且存储的效率也会随着比特位的增多而成倍增长。

而这还仅仅是量子计算机的存储能力，只有强大的存储能力还不足以体现出量子计算机对经典计算机的碾压效应，更重要的是量子计算机的并行计算能力。怎么形容这种并行计算的可怕呢？我还是用例子说明。

假设现在你面前有  $64 \times 64$  根水管分成两列，其中只有一组水管可以联通，那么请问需要多少次尝试，我们才可以找到那唯一一组可以联通的水管呢？

如果按照经典计算机的思路，我们只有一组一组尝试，左侧列的第一个水管与右侧列的第一个水管联通，如果不行，保持左侧列的第一个水管不变，将右侧列水管换成第二个。

这样依次比较下去，直到找到可以联通的两根水管。极端情况下，我们需要尝试  $64 \times 64$  次才可以找到。而量子计算机则大不一样，它可以同时表示 64 根水管，因此一次尝试就可以找到那唯一匹配的结果。

通过以上对量子计算机在存储及计算两方面的解读，你是不是也对量子计算机时代的区块链产生了某种担心呢？不过要我说，这些担心仅仅是杞人忧天，你相信吗？这不仅是因为量子计算机还仅存在实验室环境中，离真正的商业化还很远，或许我们这一代人都不一定能够赶上。

另一方面，量子计算机在发展，区块链难道就会原地踏步吗？这几乎是不可能的，现阶段，我们主要担心量子计算机可以瞬间把以比特币为首的、使用 PoW 共识算法的区块链的奖励都挖完，从而让其他矿工无币可挖；或者使用量子计算机破解比特币的公私钥，盗取其他用户的比特币。

你可以停下来稍加思考，刚才说的担心真的有道理么？其实在我看来也是多余的。先不说现阶段的量子计算机，它能否完成哈希计算或者破解非对称加密算法，而且也这样推理忽略了区块链协议的作用。我们以比特币为例分析一下，你就懂了。

比特币协议中就有这样的规定：不管全网络的算力如何变化，区块的出块速度应保持在平均 10 分钟左右，如果出块间隔较短，则会加大挖矿难度，如果出块间隔较长，则会降低难

度。而如果量子计算机加入挖矿，出块间隔势必变短，因此挖矿的难度会增加。即便增加的难度不足以难倒量子计算机，也不过是后续所有的区块奖励都由量子计算机获得，瞬间将比特币挖光的情况并不可能出现。


再者，为什么比特币不能通过硬分叉的方式将现阶段的密码算法换成与量子计算机匹配的抗量子密码呢？量子计算在发展，区块链技术也不会原地踏步。我相信，车到山前必有路，终归会有解决之道。

## 智能合约安全

最后，我想特别强调的一点是智能合约安全方面的遗憾。相比前面两个问题，或多或少都是因为技术本身或者其他威胁造成的，而智能合约安全完全是人的问题。

智能合约是软件工程师基于以太坊等区块链平台自主编写的程序代码。而我们都清楚，天底下几乎不存在没有 BUG 的软件系统，智能合约当然也不例外。可以说几乎每天都存在因智能合约编写不规范而造成的攻击事件，不过是或大或小的差别罢了。

就比如前面课程中提到的直接造成以太坊分裂的 The Dao 合约攻击事件，我们在这里可以列出该合约中有漏洞的代码，其实逻辑很简单。

 复制代码

```
1 function withdrawBalance() {  
2     amountToWithdraw = userBalances[msg.sender];  
3     if (!(msg.sender.call.value(amountToWithdraw)())) { throw; }  
4     userBalances[msg.sender] = 0;  
5 }
```

结合代码我来给你做个解读，这里要表达的意思是，如果用户需要提款，先给用户打款再将用户的余额清空。然而打款这一步骤会递归调用该方法，因此逻辑一直卡在打款那一步，而用户的余额一直没有被清空，悲剧就此发生。而其实解决方案也很简单，将两步动作调换顺序即可。

 复制代码

```
1 function withdrawBalance() {  
2     amountToWithdraw = userBalances[msg.sender];  
3     userBalances[msg.sender] = 0;  
4     if (!(msg.sender.call.value(amountToWithdraw)())) { throw; }  
5 }
```



```
5 }
```

所以，后续如果你有机会编写智能合约，就需要时刻谨记要搞懂你写的每一行代码，否则有漏洞的智能合约对黑客来说就是一份诚意满满的馈赠。

## 总结

这一讲我们通过三个例子，说明了区块链技术虽然已经发展了十多年，仍然存在不成熟的地方。

但问题的出现也仅代表着当前所处的困境，并不是无法改变，技术总是在不断的迭代向前，**每一次的技术更新其实都是服务于实际问题的破解**，当前的遗憾并不代表着后续无法美好。

当然，区块链技术存在的问题也不仅仅是这三点，比如区块链不可能三角还无法突破等。如果你感兴趣，后续我可以通过加餐的方式为你解析相关的知识点。

“

1. 解决存储冗余问题有两种思路，数据归档和节点角色区分。
2. 量子计算机离商业化还很遥远，无需过于担心。且区块链技术也是在不断的发展当中，当有一天量子时代到来，区块链也能够通过硬分叉等方式更新。
3. 时刻谨记智能合约Bug可能引起的后果，不要留给黑客可乘之机。



## 说透区块链

赵铭 | 区块链服务平台资深架构师

”



识别二维码  
免费试读

## 讨论

你认为区块链技术要想有更大的发展，还有哪些地方需要改进呢？

## 扩展阅读

关于 [比特币](#) 跟 [以太坊](#) 节点的分类，这两篇文章中做了详细解读。

关于量子密码对比特币的威胁，你可以参考科学声音的 [公众号文章](#)。

如果你对智能合约安全攻防感兴趣，可以参考蚂蚁安全实验室出品的 [智能合约安全系列](#)。

欢迎你在留言区跟我互动，主动思考、积极交流会让你更有收获。如果这一讲对你有帮助，也欢迎你分享给自己的朋友、同事。

分享给需要的人，Ta订阅后你可得 **20** 元现金奖励

👍 赞 1    💡 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇    09 | 智能合约：虚拟与现实的价值锚定载体

## 更多课程推荐

# 说透区块链

拨开迷雾，还原区块链真相

赵铭

区块链服务平台资深架构师



新版升级：点击「 请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

## 精选留言 (1)

[写留言](#)**宇宙全栈**

2021-08-11

观察目前的区块链行业，个人认为市场在关注以下几个方向

1. 更强大的性能：比特币转账成本越来越高，本质上是数据库的性能越来越差。
2. 真正解决更多实际问题：目前大家对区块链认知仅限于比特币，还没有另一个现象级的应用。
3. 更加易用：目前使用区块链应用是有学习成本的，小白根本无法使用。

展开 ▾

