Playing With Blocks: The Basics of Blockchain Databases (Part 1 - Blockchain for Everyone)

Author	josh
Date	2018-11-01 02:14:18

Overview

Blockchain is the latest and greatest buzzword in the information technology world. From open source, decentralized cryptocurrencies like Bitcoin to traditional financial institutions, it seems as though everyone is dying to create and release their own blockchain based applications. But what *is* blockchain? Why is it such a popular concept, and what is it actually good for? Let's discuss.

What and why: Blockchain simplified

What is blockchain?

So you've heard that blockchain is going to revolutionize everything, but what is it exactly? Let's cut through the hype and discuss the technical foundations of Blockchain. A blockchain is a *distributed*, *cryptographically secured* database that focuses on making historical data *immutable*. In a traditional database, information is often stored on one or a few machines, controlled by a central authority. Access is controlled by this authority (think IT administrator) and the data is kept secure by granting credentials to modify that data to a select few trusted parties. By contrast, a blockchain database is governed by what is called **distributed consensus**, using mechanisms such as proof-of-work. For more information on proof-of-work, you can read my series of articles on it.. The important thing to note is that (in general), *no one central person or authority decides what data is "verified"* in a blockchain, a community of network nodes and software does. If anyone can modify the data in a blockchain rather than a trusted party, then how is this consensus on what is correct achieved? Again, the secret lies in the science of cryptography. Through a mechanism like proof-of-work, a cryptographic puzzle is solved by software with some incentive to do so. In Bitcoin, the node that solves this puzzle is granted new currency. The real magic, however, is the fact that any other node in the network can verify that this answer is correct in a split second, so anyone can *independently* verify that a block meets the cryptographic standards set by that blockchain's protocol. You may be wondering how the cryptography in each block keeps the overall blockchain secure. This is the question of *immutability*, or how easy it is to modify the history stored in the blockchain. Blockchains solve this by cryptographically "linking" each block to the previous block, thereby making each individual block a critical part of the history stored by that "chain". Each block has a header full of useful metadata about that block - a timestamp, a "summary" of the included data or transactions, a difficulty target and nonce for mining (part of proof-of-work), and the hash of the *previous* block's header. Each block header is run through a one-way, cryptographically secure function called a "hash function" that creates a unique digital fingerprint for the data. Immutability is achieved when combining the proof-of-work consensus mechanism with this system of chaining each block together. In order to create each block, the cryptographic puzzle solved by the proof-of-work algorithm allows a unique block header hash to be generated. It is computationally difficult to get this value, but very easy to verify it is correct. Now, it's not that hard to re-solve that hard problem in a matter of minutes...it would be easy to create a fake block at the top of the chain. But what about 10 blocks back? Well, since each block contains a hash of the previous block header that is generated by solving this hard problem, you would have to now fake history for *ten* whole blocks! It is exponentially more difficult to do so the further back in the chain that you go. Unless you can truly do the work required to fake history in a blockchain, any independent network node could easily see that the rest of your history on forward is invalid. The immense difficulty of "faking" history in a blockchain gives it the most important property it has, its immutability.

Cool, so why is it useful then?

By far the most important aspect of blockchain, in my opinion, is its ability to *decentralize* applications. With a traditional database, a central authority has to be trusted, which can be a disadvantage in applications that are controversial or have high incentives for fraud. For example, previous attempts at digital money like DigiCash had central services for issuing currency and validating transactions. These were promptly shut down by governments that didn't like independent currencies very much. With blockchain, it is possible to have things like completely peer-to-peer money as with Bitcoin, Litecoin, and countless others because no central government or individual has to be trusted! The network is secured by math (cryptography) rather than trust thanks to the blockchain. You don't have to trust anyone to not defraud you of your money, because the math cannot lie about who owns what. The other critical function of blockchains beyond decentralization are the preservation of history. Because blockchains are immutable, they can be useful for keeping things like medical records, property transactions, court histories, and more secure from malicious tampering like a traditional database. This does rely on some degree of decentralization, but even within a single company a blockchain is far harder to tamper with than a traditional database.

Cool, now I want a blockchain!

Blockchains are a fascinating and novel way to handle problems with traditional databases in certain applications. Thanks to the decentralized and cryptographically secure nature of these databases, it's possible to create peer-to-peer applications that don't require trusting a third party - a key problem to solve for concepts like digital money. As well, their immutability makes them useful even beyond the first few money-centric applications that existed - they may be coming do a real-estate authority, doctor's office, or justice system near you!