Deloitte.

Perspective

Quantum computers and the Bitcoin blockchain

An analysis of the impact quantum computers might have on the Bitcoin blockchain

One of the most well-known applications of quantum computers is breaking the mathematical difficulty underlying most of currently used cryptography. Since Google announced that it achieved quantum supremacy there has been an increasing number of articles on the web predicting the demise of currently used cryptography in general, and Bitcoin in particular. The goal of this article is to present a balanced view regarding the risks that quantum computers pose to Bitcoin.

Authors: Itan Barmes, Bram Bosch and Olaf Haalstra

The main focus of this article will be to answer the following questions:

1. How many Bitcoins could be stolen now if a sufficiently large quantum computer was available?
2. What can one do to mitigate the risk of Bitcoins being stolen by an adversary with a quantum computer?
3. Is the Bitcoin blockchain inherently resilient to quantum attacks now and in the future?

Quantum computers and cryptography

A great amount of digital ink has been spilled on the topic of how quantum computers pose an existential threat to currently used asymmetric

cryptography. We will therefore not discuss this in detail, but only explain the aspects that are relevant for the analysis in this article.

In asymmetric cryptography, a private-public key pair is generated in such a manner that the two keys have a mathematical relation between them. As the name suggests, the private key is kept as secret, while the public key is made publicly available. This allows individuals to produce a digital signature (using their private key) that can be verified by anyone who has the corresponding public key. This scheme is very common in the financial industry to prove authenticity and integrity of transactions.

The security of asymmetric cryptography is based on a mathematical principle called a “one-way function”. This principle dictates that the public key can be easily derived from the private key but not the other way around. All known (classical) algorithms to derive the private key from the public key require an astronomical amount of time to perform such a computation and are therefore not practical. However, in 1994, the mathematician Peter Shor published a quantum algorithm that can break the security assumption of the most common algorithms of asymmetric cryptography. This means that anyone with a sufficiently large quantum computer could use this algorithm to derive a private key from its corresponding public key, and thus, falsify any digital signature.

Bitcoin 101

To understand the impact of quantum computers on Bitcoin, we will start with a brief summary about how Bitcoin transactions work. Bitcoin is a decentralized system for transferring value. Unlike the banking system where it is the responsibility of a bank to provide customers with a bank account, a Bitcoin user is responsible for generating his own (random) address. By means of a simple procedure, the user's computer calculates a random Bitcoin address (related to the public key) as well as a secret (private key) that is required in order to perform transactions from this address.

Moving Bitcoins from one address to another is called a transaction. Such a transaction is similar to sending money from one bank account to another. In Bitcoin, the sender must authorize their transaction by providing a digital signature that proves they own the address where the funds are stored. Remember: someone with an operational quantum computer who has your

public key could falsify this signature, and therefore potentially spend anyone's Bitcoins!

In the Bitcoin network, the decision of which transactions are accepted into the network is ultimately left to the so called miners. Miners compete in a race to process the next batch of transactions, also called a block. Whoever wins the race, is allowed to construct the next block, awarding them new coins as they do so. Bitcoin blocks are linked to each other in a sequential manner. Together, they form a chain of blocks, also called the "blockchain".

The victorious miner who creates a new block, is free to include whichever transaction they wish. Other miners express their agreement by building on top of blocks they agree with. In case of a disagreement, they will build on the most recently accepted block. In other words, if a rogue miner attempts to construct an invalid block, honest miners will ignore the invalid block and build on top of the most recent valid block instead.

Address types

Bitcoin transactions allow for a custom logic to be implemented, enabling a myriad of financial transaction types such as escrow and shared ownership. However, for the purpose of this article, we restrict ourselves to simple person-to-person payments. These can be divided into 2 categories, each affected differently by a quantum computer.

In the first type, a public key directly serves as the Bitcoin address of the recipient. A transaction to such an address is called 'pay to public key' (p2pk) for obvious reasons. In the early days of Bitcoin, in 2009, this was the dominant address type. Many of the original coins mined by Satoshi Nakamoto himself are still stored in such addresses. One of the issues with these addresses is the lack of a mechanism to detect mistyping of addresses (for example a last checksum digit which is used, for example, in credit card numbers). An additional problem is that these addresses are very long, which results in a larger transaction file and therefore longer processing time. Regarding the threat from a quantum computer, the public key is directly obtainable from the address. Since all transactions in Bitcoin are public, anyone can obtain the public key from any p2pk address. A quantum computer running Shor's algorithm could then be used to derive the private key from this address. This would allow an adversary who has a quantum computer to spend the coins that the address had.

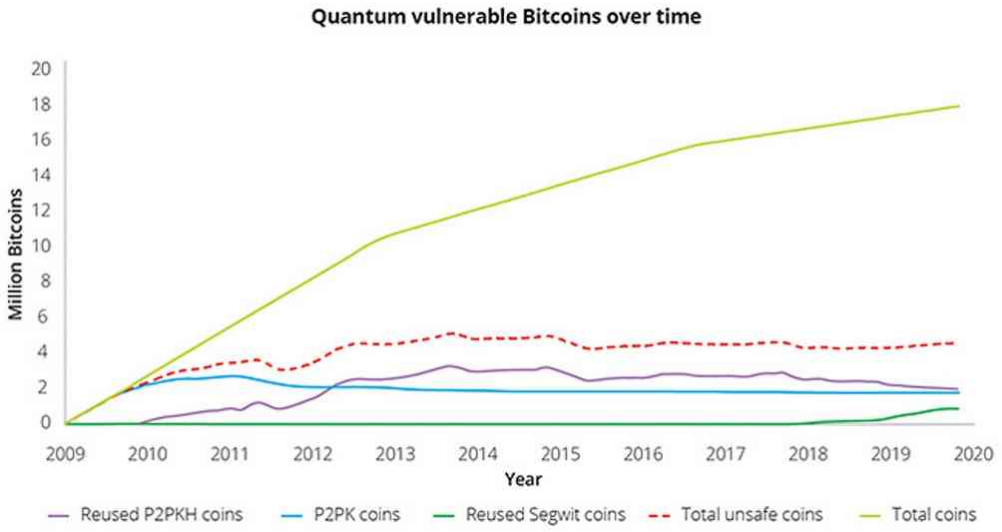
In the second type of transaction, the address of the recipient is composed of a hash of the public key. As a hash is a one-way cryptographic function, the public key is not directly revealed by the address. The first and most popular implementation of this is called ‘pay to public key hash’ (p2pkh) and was designed to solve the two issues described above (checksum and address length, for a more elaborate explanation we refer to this page. As was mentioned above, the public key cannot be retrieved from the address. The public key is only revealed at the moment when the owner wishes to initiate a transaction. This means that as long as funds have never been transferred from a p2pkh address, the public key is not known and the private key cannot be derived using a quantum computer. There is a ‘but’ though! If funds are ever transferred from a specific p2pkh address (no matter what amount), the public key is revealed. From that moment on, this address is marked "used" and should ideally not be used again to receive new coins. In fact, many wallets are programmed to avoid address reuse as best they can. Avoiding the reuse of addresses is considered best practice for Bitcoin users, but you would be surprised how many people do not take this advice to heart. More on that in the following chapter.

How many Bitcoins could be stolen now if sufficiently large quantum computers were available?

Imagine that someone manages to build a quantum computer today and is therefore able to derive private keys. How many Bitcoins will be in danger?

To answer this question, we analyzed the entire Bitcoin blockchain to identify which coins are vulnerable to an attack from a quantum computer. As explained in the previous section, all coins in p2pk addresses and reused p2pkh addresses are vulnerable to a quantum attack. The result of our analysis is presented in the figure below. It shows the distribution of Bitcoins in the various address types over time. As can clearly be seen in the graph, p2pk addresses dominated the Bitcoin blockchain in the first year of its existence. Interestingly, the number of coins in p2pk addresses has stayed practically constant (circa 2M Bitcoins). A reasonable assumption is that these coins were generated through mining and have never been moved from their original address.

As p2pkh was introduced 2010, it quickly became dominant. Most of the coins created since then are stored in this type of address. In the graph we see that the number of Bitcoins stored in reused p2pkh increases from 2010 to 2014, and since then is decreasing slowly to reach the current amount of 2.5M Bitcoins. This suggests that people are generally following the best practice of not using p2pk address as well as not reusing p2pkh addresses. Nevertheless, there are still over 4 million BTC (about 25% of all Bitcoins) which are potentially vulnerable to a quantum attack. At the current price this is over 40 billion USD!



What can one do to mitigate the risk of Bitcoins being stolen by an adversary with a quantum computer?

In the previous section we explained that p2pk and reused p2pkh addresses are vulnerable to quantum attacks. However, p2pkh addresses that have never been used to spend Bitcoins are safe, as their public keys are not yet public. This means that if you transfer your Bitcoins to a new p2pkh address, then they should not be vulnerable to a quantum attack.

The issue with this approach is that many owners of vulnerable Bitcoins have lost their private keys. These coins cannot be transferred and are waiting to be taken by the first person who manages to build a sufficiently large quantum computer. A way to address this issue is to come to a consensus within the Bitcoin community and provide an ultimatum for people to move their coins to a safe address. After a predefined period, coins in unsafe addresses would become unusable (technically, this means that miner will

ignore transactions coming from these addresses). Such a drastic step needs to be considered carefully before implemented, not to mention the complexity of achieving consensus about such a sensitive issue.

Is the Bitcoin blockchain inherently resilient to quantum attacks now and in the future?

Let's assume for a minute that all owners of vulnerable Bitcoins transfer their funds to safe addresses (everyone who lost their private key 'magically' finds them). Does that mean that the Bitcoin blockchain is no longer vulnerable to quantum attacks? The answer to this question is actually not that simple. The prerequisite of being "quantum safe" is that the public key associated with this address is not public. But as we explained above, the moment you want to transfer coins from such a "safe" address, you also reveal the public key, making the address vulnerable. From that moment until your transaction is "mined", an attacker who possesses a quantum computer gets a window of opportunity to steal your coins. In such an attack, the adversary will first derive your private key from the public key and then initiate a competing transaction to their own address. They will try to get priority over the original transaction by offering a higher mining fee.

In the Bitcoin blockchain it currently takes about 10 minutes for transactions to be mined (unless the network is congested which has happened frequently in the past). As long as it takes a quantum computer longer to derive the private key of a specific public key then the network should be safe against a quantum attack. Current scientific estimations predict that a quantum computer will take about 8 hours to break an RSA key, and some specific calculations predict that a Bitcoin signature could be hacked within 30 minutes. This means that Bitcoin should be, in principle, resistant to quantum attacks (as long as you do not reuse addresses). However, as the field of quantum computers is still in its infancy, it is unclear how fast such a quantum computer will become in the future. If a quantum computer will ever get closer to the 10 minutes mark to derive a private key from its public key, then the Bitcoin blockchain will be inherently broken.

Closing remarks

Quantum computers are posing a serious challenge to the security of the Bitcoin blockchain. Presently, about 25% of the Bitcoins in circulation are vulnerable to a quantum attack. If you have Bitcoins in a vulnerable address and believe that progress in quantum computing is more advanced than publicly known, then you should probably transfer your coins to a new p2pkh address (don't forget to make a secure backup of your private key).

In case your own Bitcoins are safe in a new p2pkh address, you might still be impacted if many people will not (or cannot) take the same protection measures. In a situation where a large number of Bitcoins is stolen, the price will most likely crash and the confidence in the technology will be lost.

Even if everyone takes the same protection measures, quantum computers might eventually become so fast that they will undermine the Bitcoin transaction process. In this case the security of the Bitcoin blockchain will be fundamentally broken. The only solution in this case is to transition to a new type of cryptography called 'post-quantum cryptography', which is considered to be inherently resistant to quantum attacks. These types of algorithms present other challenges to the usability of blockchains and are being investigated by cryptographers around the world. We anticipate that future research into post-quantum cryptography will eventually bring the necessary change to build robust and future-proof blockchain applications.

Get in touch

Marc Verdonk

Partner

✉ **mverdonk@deloitte.com**

✆ **+31652615027**