# Chapter 5: Bitcoin Mining

This chapter is all about mining. We've already seen quite a bit about miners and how Bitcoin relies on them — they validate every transaction, they build and store all the blocks, and they reach a consensus on which blocks to include in the block chain. We also have already seen that miners earn some reward for doing this, but we still have left many questions unanswered. Who are the miners? How did they get into this? How do they operate? What's the business model like for miners? What impact do they have on the environment? In this chapter, we will answer all of these questions.

## 5.1 The task of Bitcoin miners

Do you want to get into Bitcoin mining? If you do, we're not going to completely discourage you, but beware that Bitcoin mining bears many similarities to gold rushes. Historical gold rushes are full of stories of young people rushing off to find fortune and inevitably many of them lose everything they have. A few strike it rich, but even those that do generally endure lots of hardship along the way. We'll see in this section why Bitcoin mining shares many of the same challenges and risks as traditional gold rushes and other get-rich-quick schemes.

But first, let's look at the technical details. To be a Bitcoin miner, you have to join the Bitcoin network and connect to other nodes. Once you're connected, there are six tasks to perform:

1. *Listen for transactions.* First, you listen for transactions on the network and validate them by checking that signatures are correct and that the outputs being spent haven't been spent before.
2. *Maintain block chain and listen for new blocks.* You must maintain the block chain. You start by asking other nodes to give you all of the historical blocks that are already part of the block chain before you joined the network. You then listen for new blocks that are being broadcast to the network. You must validate each block that you receive — by validating each transaction in the block and checking that the block contains a valid nonce. We'll return to the details of nonce checking later in this section.
3. *Assemble a candidate block.* Once you have an up-to-date copy of the block chain, you can begin building your own blocks. To do this, you group transactions that you heard about into a new block that extends the latest block you know about. You must make sure that each transaction included in your block is valid.
4. *Find a nonce that makes your block valid.* This step requires the most work and it's where all the real difficulty happens for miners. We will see this in detail shortly.
5. *Hope your block is accepted.* Even if you find a block, there's no guarantee that your block will become part of the consensus chain. There's bit of luck here; you have to hope that other miners accept your block and start mining on top of it, instead of some competitor's block.
6. *Profit.* If all other miners do accept your block, then you profit! At the time of this writing in early 2015, the block reward is 25 bitcoins which is currently worth over $6,000. In addition, if

any of the transactions in the block contained transaction fees, the miner collects those too. So far transaction fees have been a modest source of additional income, only about 1% of block rewards.

We can classify the steps that a miner must take into two categories. Some tasks — validating transactions and blocks — help the Bitcoin network and are fundamental to its existence. These tasks are the reason that the Bitcoin protocol requires miners in the first place. Other tasks — the race to find blocks and profit —- aren't necessary for the Bitcoin network itself but are intended to incentivize miners to perform the essential steps. Of course, both of these are necessary for Bitcoin to function as a currency, since miners need an incentive to perform the critical steps.

***Finding a valid block.*** Let's return to the question of finding a nonce that makes your block valid. In Chapter 3 we saw that there are two main hash-based structures. There's the block chain where each block header points to the previous block header in the chain, and then within each block there's a Merkle tree of all of the transactions included in that block.

The first thing that you do as a miner is to compile a set of valid transactions that you have from your pending transaction pool into a Merkle tree. Of course, you may choose how many transactions to include up to the limit on the total size of the block. You then create a block with a header that points to the previous block. In the block header, there's a 32 bit nonce field, and you keep trying different nonces looking for one that causes the block's hash to be under the target — roughly, to begin with the required number of zeros. A miner may begin with a nonce of 0 and successively increment it by one in search of a nonce that makes the block valid. See Figure 5.1.

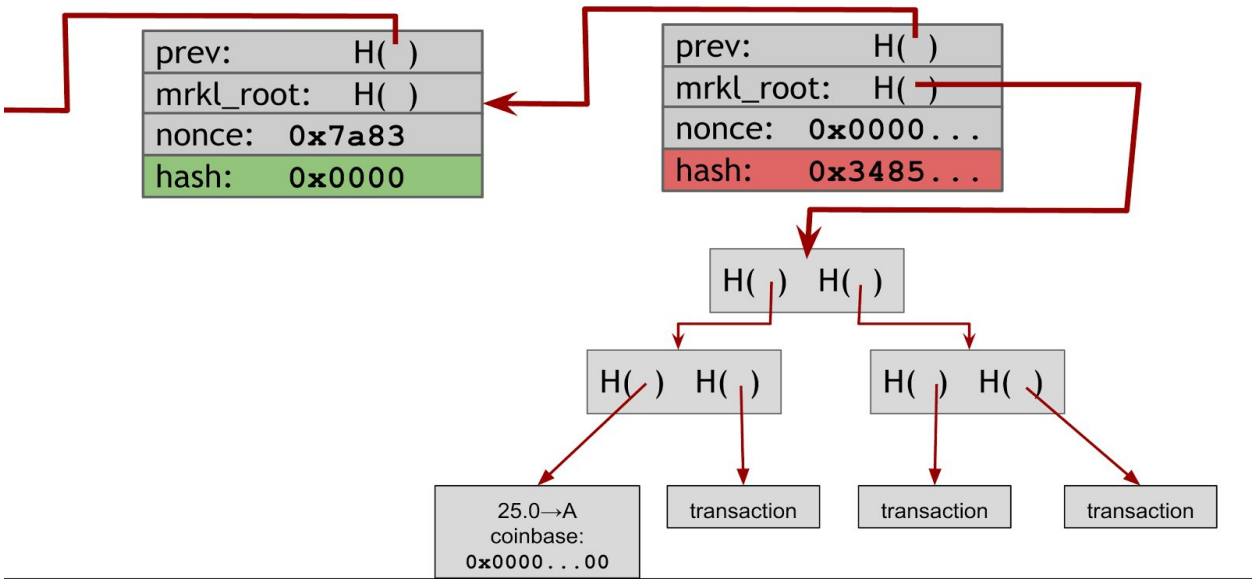



**Figure 5.1: Finding a valid block.** In this example, the miner tries a nonce of all 0s. It does not produce a valid hash output, so the miner would then proceed to try a different nonce.

In most cases you'll try every single possible 32-bit value for the nonce and none of them will produce a valid hash. At this point you're going to have to make further changes. Notice in Figure 5.1 that there's an additional nonce in the coinbase transaction that you can change as well. After you've exhausted all possible nonces for the block header, you'll change the extra nonce in the coinbase transaction — say by incrementing it by one — and then you'll start searching nonces in the block header once again.

When you change the nonce parameter in the coinbase transaction, the entire Merkle tree of transactions has to change (See Figure 5.2). Since the change of the coinbase nonce will propagate all the way up the tree, changing the extra nonce in the coinbase transaction is much more expensive operation than changing the nonce in the block header. For this reason, miners spend most of their time changing the nonce in the block header and only change the coinbase nonce when they have exhausted all of the $2^{32}$ possible nonces in the block header without finding a valid block.

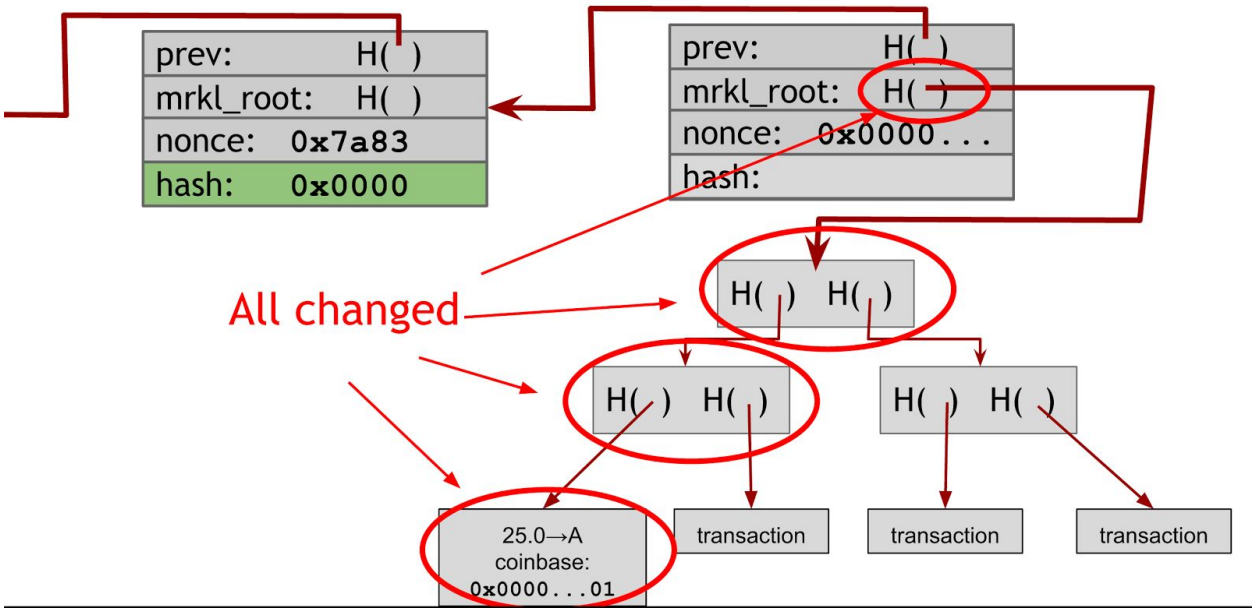



**Figure 5.2:** Changing a nonce in the coinbase transaction propagates all the way up the Merkle tree.

The vast, vast majority of nonces that you try aren't going to work, but if you stay at it long enough you'll eventually find the right combination of the extra nonce in the coinbase transaction and the nonce in the block header that produce a block with a hash under the target. When you find this, you want to announce it as quickly as you can and hope that you can profit from it.

**Is everyone solving the same puzzle?** You may be wondering: if every miner just increments the nonces as we described, aren't all miners solving the exact same puzzle? Won't the fastest miner always win? The answer is no! Firstly, it's unlikely that miners will be working on the exact same block as each miner will likely include a somewhat different set of transactions and in a different order. But more importantly, even if two different miners were working on a block with identical transactions, the blocks would still differ. Recall that in the coinbase transaction, miners specify their own address as the owner of the newly minted coins. This address by itself will cause changes which propagate up to the root of the Merkle tree, ensuring that no two miners are working on exactly the same puzzle unless they share a public key. This would only happen if the two miners are part of the same mining pool (which we'll discuss shortly), in which case they'll communicate to ensure they include a distinct nonce in the coinbase transaction to avoid duplicating work.

***Difficulty.*** Exactly how difficult is it to find a valid block? As of March 2015, the mining difficulty target (in hexadecimal) is:

0000000000000000172EC0000000000000000000000000000000000000000000

so the hash of any valid block has to be below this value. In other words only one in about $2^{67}$ nonces that you try will work, which is a really huge number. One approximation is that it's greater than the human population of Earth squared. So, if every person on Earth was themselves their own planet Earth with seven billion people on it, the total number of people would be close to $2^{67}$.

***Determining the difficulty.*** The mining difficulty changes every 2016 blocks, which are found about once every 2 weeks. It is adjusted based on how efficient the miners were over the period of the previous 2016 blocks according to this formula:

**next_difficulty = (previous_difficulty * 2016 * 10 minutes) / (time to mine last 2016 blocks)**

Note that 2016*10 minutes is exactly two weeks, so 2016 blocks would take two weeks to mine 2016 blocks if a block were created exactly every 10 minutes. So the effect of this formula is to scale the difficulty to maintain the property that blocks should be found by the network on average about once every ten minutes. There's nothing special about 2 weeks, but it's a good trade-off. If the period were much shorter, the difficulty might fluctuate due to random variations in the number of blocks found in each period. If the period were much higher, the network's hash power might get too far out of balance with the difficulty.

Each Bitcoin miner independently computes the difficulty and will only accept blocks that meet the difficulty that they computed. Miners who are on different branches might not compute the same difficulty value, but any two miners mining on top of the same block will agree on what the difficulty should be. This allows consensus to be reached.

You can see in Figure 5.3 that over time the mining difficulty keeps increasing. It's not necessarily a steady linear increase or an exponential increase, but it depends on activity in the market. Mining difficulty is affected by factors like how many new miners are joining, which in turn may be affected by the current exchange rate of Bitcoin. Generally, as more miners come online and mining hardware gets more efficient, blocks are found faster and the difficulty is increased so that it always takes about ten minutes to find a block.

In Figure 5.3 you can see that in the red line on the graph there's a step function of difficulty even though the overall network hash rate is growing smoothly. The discrete step results from the fact that the difficulty is only adjusted every 2016 blocks.

Another way to view the network's growth rate is to consider how long it takes to find a block on average. Figure 5.4 (a) shows how many seconds elapse between consecutive blocks in the block chain. You can see that this gradually goes down, jumps up and then gradually goes down again. Of course what's happening is that every 2016 blocks the difficulty resets and the average block time goes back up to about ten minutes. Over the next period the difficulty stays unchanged, but more and more miners come online. Since the hash power has increased but the difficulty has not, blocks are found more quickly until the difficulty is again adjusted after 2016 blocks, or about two weeks.

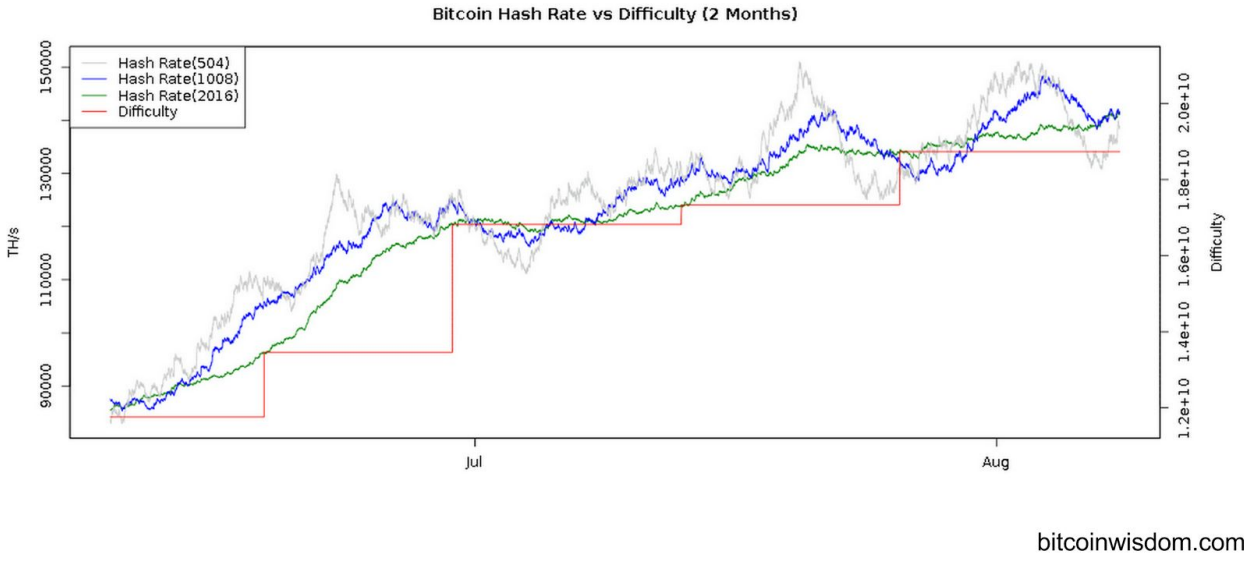



**Figure 5.3: Mining difficulty over time (mid-2014).** Note that the y-axis begins at 80,000 TH/s.

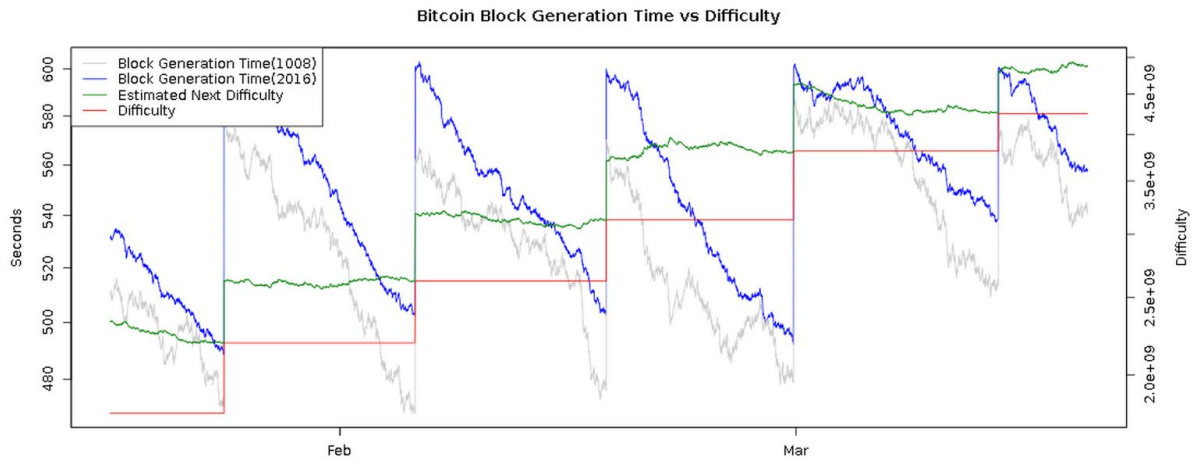



**Figure 5.4 (a) : Time to find a block (early 2014).** Note that the y-axis begins at 460 seconds. Due to continued rapid growth in mining power during this time, the time to find a block decreased steadily within each two-week window. Source: bitcoinwisdom.com

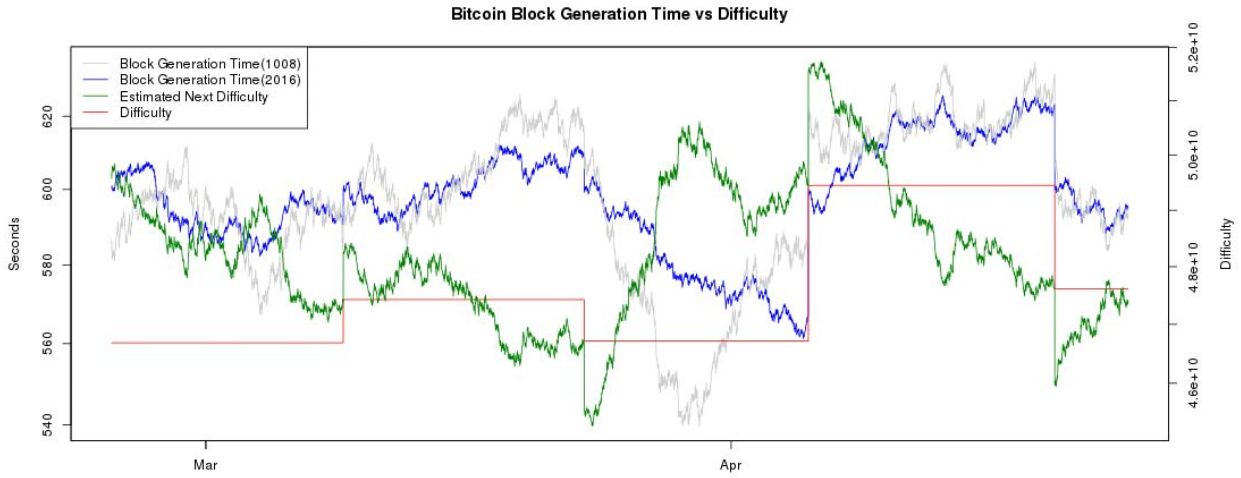



**Figure 5.4 (b) : Time to find a block (early 2015).** Note that the y-axis begins at 540 seconds. As the growth of the network has slowed, the time to find each block is much closer to 10 minutes and is occasionally over during periods where the network's hash power actually shrinks. Source: bitcoinwisdom.com

Even though the goal was for a block to be found every ten minutes on average, for most of 2013 and 2014 it was closer to about nine minutes on average and would approach 8 minutes at the end of each two week cycle. Quick calculations show that this requires an astonishing 25% growth rate every two weeks, or several hundred fold per year.

Unsurprisingly, this was not sustainable forever and in 2015 the growth rate has been much slower (and occasionally negative). In Figure 5.4(b), we can see that as the mining power is closer to a

steady-state, the period to find each block stays much closer to 10 minutes. It can even take longer than 10 minutes, in which case there will be a difficulty *decrease*. Once considered unthinkable, this has happened fairly regularly in 2015.

While there have been no catastrophic declines of the network's mining power so far, there's no inherent reason why that cannot happen. One proposed scenario for Bitcoin's collapse is a "death spiral" in which a dropping exchange rate makes mining unprofitable for some miners, causing an exodus, in turn causing the price to drop further.

## 5.2 Mining Hardware

We've mentioned that the computation that miners have to do is very difficult. In this section, we'll discuss why it is so computationally difficult and take a look at the hardware that miners use to perform this computation.

The core of the difficult computation miners are working on is the SHA-256 hash function. We discussed hash functions abstractly in Chapter 1. SHA-256 is a general purpose cryptographic hash function that's part of a bigger family of functions that was standardized in 2001 (SHA stands for Secure Hash Algorithm). SHA-256 was a reasonable choice as this was strongest cryptographic hash function available at the time when Bitcoin was designed. It is possible that it will become less secure over the lifetime of Bitcoin, but for now it remains secure. Its design did come from the NSA (US National Security Agency), which has led to some conspiracy theories, but it's generally believed to be a very strong hash function.

***A closer look at SHA-256.*** Figure 5.5 shows more detail about what actually goes on in a SHA-256 computation. While we don't need to know all of the details to understand how Bitcoin works, it's good to have a general idea of the task that miners are solving.

SHA-256 maintains 256 bits of state. The state is split into eight 32-bit words which makes it highly optimized for 32-bit hardware. In each round a number of words in the state are taken — some with small bitwise tweaks applied — and added together mod 32. The entire state is then shifted over with the result of the addition becoming the new left-most word of the state. The design is loosely inspired by simpler bitwise Linear Feedback Shift Registers (LFSRs).

**Sidebar: the SHA family.** The "256" in SHA-256 comes from its 256-bit state and output. Technically SHA-256 is one of several closely-related functions in the SHA-2 family, including SHA-512 (which has a larger state and is therefore more secure). There is also SHA-1, an earlier generation with 160-bit output which is now considered insecure but is still implemented in Bitcoin script.

Although the SHA-2 family, including SHA-256, are still considered to be cryptographically secure, the next generation SHA-3 family has now been picked by a contest. SHA-3 is in the final stages of standardization today, but it wasn't available at the time Bitcoin was designed.

Figure 5.5 shows just one round of the SHA-256 compression function. A complete computation of SHA-256 does this for 64 iterations. During each round, there are slightly different constants applied so that no iteration is exactly the same.

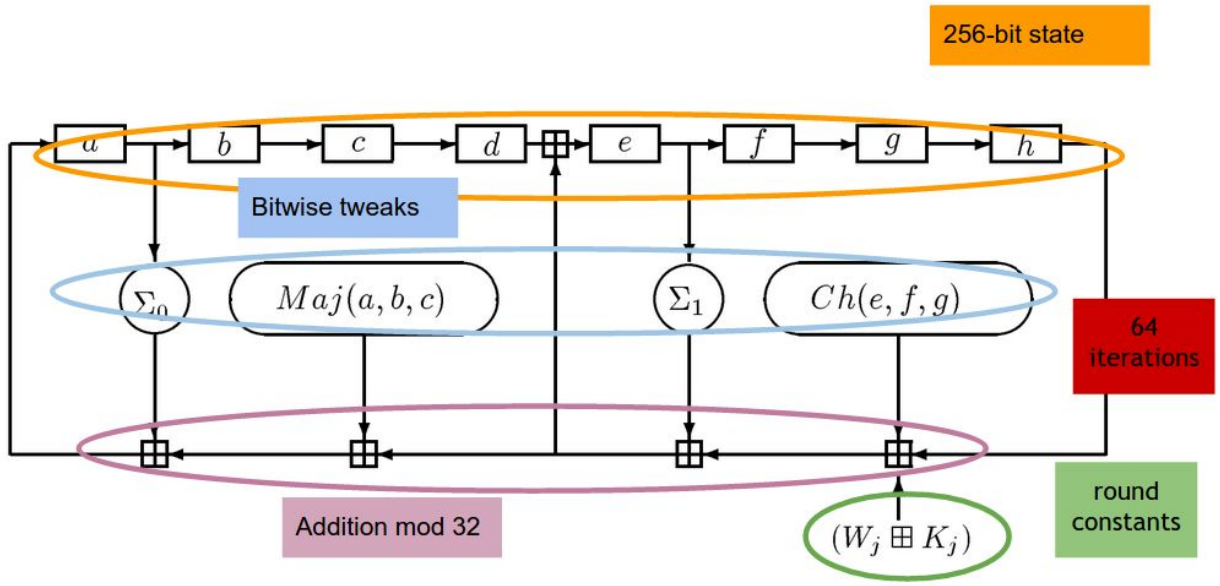



**Figure 5.5 : The structure of SHA-256.** This is one round of the compression function.

The task for miners is to compute this function as quickly as possible. Remember that miners are racing each other so the faster they do this, the more they earn. To do this, they need to be able to manipulate 32-bit words, do 32-bit modular addition and also do some bitwise logic.

As we will see shortly, Bitcoin actually requires SHA-256 to be applied twice to a block in order to get the hash that is used by the nodes. This is a quirk of Bitcoin. The reasons for the double computation are not fully specified, but at this point, it's just something that miners have to deal with.

***CPU mining.*** The first generation of mining was all done on general purpose computers — that is general purpose central processing units (CPUs). In fact, CPU mining was as simple as running the code shown in Figure 5.6. That is, miners simply searched over nonces in a linear fashion, computed SHA 256 in software and checked if the result was a valid block. Also, notice in the code that as we mentioned, SHA-256 is applied twice.

```
TARGET = (65535 << 208) / DIFFICULTY;
coinbase_nonce = 0;
while (1) {
      header = makeBlockHeader(transactions, coinbase_nonce);
      for (header_nonce = 0; header_nonce < (1 << 32); header_nonce++){
            if (SHA256(SHA256(makeBlock(header, header_nonce))) <
      TARGET)
                  break; //block found!
      }
      coinbase_nonce++;
}
```

**Figure 5.6 : CPU mining pseudocode.**

How fast will this run on a general purpose computer? On a high-end desktop PC you might expect to compute about 20 million hashes per second (MH/s). At that speed, it would take you several hundred thousand years on average at the early-2015 difficulty level ($2^{67}$) to find a valid block. We weren't kidding when we said mining was going to be a difficult slog!

If you're mining on a general purpose PC today, CPU mining is no longer profitable with the current difficulty. For the last few years, anyone trying to mine on a CPU probably doesn't understand how Bitcoin works and was probably pretty disappointed that they never made any money doing it.

***GPU mining.*** The second generation began when people started to get frustrated with how slow their CPUs were and instead used their graphics card, or graphics processing unit (GPU).

Almost every modern PC has a GPU built-in to support high performance graphics. They're designed to have high throughput and also high parallelism, both of which are very useful for Bitcoin mining. Bitcoin mining can be parallelized easily because you can try computing multiple hashes at the same time with different nonces. In 2010, a language called OpenCL was released. OpenCL is a general purpose language to do things other than graphics on a GPU. It's a high level-language and over time people have used it to run many types of computation more quickly on graphics cards. This paved the way for Bitcoin mining on GPUs.

Mining with graphics cards had several attractive properties at the time. For one thing, they're easily available and easy for amateurs to set up. You can order graphics cards online or buy them at most big consumer electronics stores. They're the most accessible high-end hardware that's available to the general public. They also have some properties that make them specifically good for Bitcoin mining. They're designed for parallelism so they have many Arithmetic Logic Units (ALUs) that can be used for simultaneous SHA-256 computations. Some GPUs also have specific instructions to do bitwise operations that are quite useful for SHA-256.

Most graphics cards can also be ***overclocked***, meaning you can run them faster than they're actually designed for if you want to take on the risk that they might overheat or malfunction. This is a property

gamers have demanded for years. With Bitcoin mining, it might be profitable to run the chip much faster than it was designed for even if you induce a few errors by doing so.

For example, say you can run your graphics card 50 percent faster but doing so will cause errors in the SHA-256 computation to 30 percent of the time. If an invalid solution is erroneously declared valid by the graphics card — something that would happen rarely — you can always double-check it on your CPU. On the other hand, if a valid solution is erroneously missed, you'd never know. But if your speed increase from overclocking can overcome the decrease in output due to errors, you'd still come out ahead. In the above example, the throughput is 1.5x compared to not overclocking, whereas the success rate is 0.7x. The product is 1.05, which means overclocking increases your expected profits by 5%. People have spent considerable time optimizing exactly how much they should overclock a given chip to maximize profits.

Finally, you can drive many graphics cards from one motherboard and CPU. So you can take your computer, which will be running your actual Bitcoin node which gathers transactions from the network and assembles blocks, and attach multiple graphics cards to it to try to find the right nonces to make the SHA-256 of the block valid. Many people created some really interesting home-brewed setups like this one shown in Figure 5.7 to drive many, many GPUs from a single CPU. This was still in the early days of Bitcoin when miners were still mostly hobbyists without much experience running servers, but they came up with some quite ingenious designs for how to pack many graphics cards into a small place and keep them cool enough to operate.

**Figure 5.7: A home-built rack of GPUs used for Bitcoin mining.** You can also see the fans that they used to build a primitive cooling system. Source: LeonardH, cryptocurrenciestalk.com.

***Disadvantages of GPU mining.*** GPU mining has some disadvantages. GPUs have a lot of hardware built into them for doing video processing that can't be utilized for mining. Specifically, they have a large number of floating point units that aren't used at all in SHA-256. GPUs also don't have the greatest cooling characteristics when you put a lot of them next to one another. They're not designed to run side by side as they are in the picture; they're designed to be in a single box doing graphics for one computer.

**Miners vs. Gamers.** According to folklore, by 2011 Bitcoin miners were purchasing enough GPUs to upset the normal market. This caused friction with the gaming community who found it increasingly difficult to find certain popular GPUs in local electronics stores. Interestingly, however, it may have increased interest in Bitcoin mining as many of these frustrated gamers learned about the currency to understand where all the GPUs were going, with some of gamers becoming miners themselves!

GPUs can also draw a fairly large amount of power, so a lot of electricity is used relative to a computer. Another disadvantage initially was that you had to either build your own board or buy expensive boards to house multiple graphics cards.

On a really high-end graphics card with aggressive tuning you might get as high as 200 MH/s, or 200 million hashes per second, an order of magnitude better than you would be doing with a CPU. But even with that improved performance, and even if you're really enterprising and used one hundred GPUs together, it would still take you over 300 years on average to find a block at the early-2015 difficulty level. As a result, GPU mining is basically dead for Bitcoin today, though it still shows up sometimes in early-stage altcoins.

***FPGA mining.*** Around 2011 some miners started switching from GPUs to FPGAs, or Field Programmable Gate Arrays, after the first implementation of Bitcoin mining came out in Verilog, a hardware design language that's used to program FPGAs. The general rationale behind FPGAs is to try to get close as possible to the performance of custom hardware while also allowing the owner of the card to customize it or reconfigure it "in the field." By contrast, custom hardware chips are designed in a factory and do the same thing forever.

FPGAs offer better performance than graphics cards, particularly on "bit fiddling" operations which are trivial to specify on an FPGA. Cooling is also easier with FPGAs and, unlike GPUs, you can theoretically use nearly all of the transistors on the card for mining. Like with GPUs, you can pack many FPGAs together and drive them from one central unit, which is exactly what people began to do (see Figure 5.8). Overall, it was possible to build a big array of FPGAs more neatly and cleanly than you could with graphics cards.

Using an FPGA with a careful implementation, you might get up to a GH/s, or one billion hashes per second. This is certainly a large performance gain over CPUs and GPUs, but even if you had a hundred

boards together, each with a 1 GH/s throughput, it would still take you about 50 years on average to find a Bitcoin block at the early-2015 difficulty level.

**Figure 5.8: A home-built rack of FPGAs.** Although you don't see the cooling setup pictured here, a rack like this would need a cooling system.

Despite the performance gain, the days of FPGA mining were quite limited. Firstly, they were being driven harder for Bitcoin mining — by being on all the time and overclocked — than consumer grade FPGAs were really designed for. Because of this, many people saw errors and malfunctions in their FPGAs as they were mining. It also turned out to be difficult to optimize the 32-bit addition step which is critical in doing SHA-256. FPGAs are also less accessible-you can't buy them at most stores and there are fewer people who know how to program and set up an FPGA than a GPU.

Most importantly though, even though FPGAs improved performance the cost-per-performance was only marginally improved over GPUs. This made FPGA mining was a rather short-lived phenomenon. Whereas GPU mining dominated for about a year or so, the days of FPGA mining were far more limited — lasting only a few months before custom ASICs arrived.

***ASIC mining.*** Mining today is dominated by Bitcoin *ASICs*, or ***application-specific integrated circuits***. These are chips that were designed, built, and optimized for the sole purpose of mining Bitcoins. There are a few big vendors that sell these to consumers with a good deal of variety: you can choose between slightly bigger and more expensive models, more compact models, as well as models with varying performance and energy consumption claims.

Designing ASICs requires considerable expertise and their lead-time is also quite long. Nevertheless, Bitcoin ASICs were designed and produced surprisingly quickly. In fact, analysts have said that this

may be the fastest turnaround time in the history of integrated circuits from specifying a problem and to have a working chip in people's hands. Partially as a result of this, the first few generations of Bitcoin ASICs were quite buggy and most of them didn't quite deliver the promised performance numbers. Bitcoin ASICs have since matured and there are now fairly reliable ASICs available.

Up until 2014, the lifetime of ASICs was quite short due to the rapidly increasing network hash rate, with most boards in the early ASIC era growing obsolete in about six months. Within this time, the bulk of the profits are made up front. Often, miners will make half of the expected profits for the lifetime of the ASIC during just the first six weeks. This meant shipping speed can become a crucial factor in making a profit. Due to the immaturity of the industry though, consumers often experienced shipping delays with boards often nearly obsolete by the time they arrived. As the growth rate of Bitcoin's hash power has stabilized, mining equipment has a longer life time, but the early era saw many frustrated customers and accusations of fraud by vendors.

For much of Bitcoin's history, the economics of mining haven't been favorable to the small miner who wants to go online, order mining equipment, and start making money. In fact, in most cases people who have placed orders for mining hardware would have lost money based on the calculation that they made at the time. Until 2013 though, the exchange rate of Bitcoin rose enough to bail most customers out from losing money outright. In effect, mining has been an expensive way to simply bet that the price of Bitcoin would rise, and many miners — even though they've made money mining Bitcoins — would have been better off if they had just taken the money that they were going to spend on mining equipment, invested it in bitcoins, and eventually sold them at a profit.

You can still order Bitcoin mining equipment today and we wouldn't want to discourage that as a way to learn about Bitcoin and cryptocurrencies. However, we'll note again that this is not an advisable way to make money. Most ASICs sold commercially today are unlikely to pay for themselves in mining rewards once you factor in the price of electricity and cooling.

***Today : Professional mining.*** Today mining has mostly moved away from individuals and toward professional mining centers. Exact details about how these centers operate are not very well known because companies want to protect their setups to maintain a competitive advantage. Presumably, these operations maintain profitability by buying slightly newer and more efficient ASICs than are available for general sale at a bulk discount. In Figure 5.9, we see a picture of a professional mining center in the Republic of Georgia.

**Figure 5.9: BitFury mining center,** a professional mining center in the republic of Georgia.

When determining where to set up a mining center, the three biggest considerations are: climate, cost of electricity, and network speed. In particular, you want a cold climate to keep cooling costs low. Cooling is particularly challenging with Bitcoin mining, which is estimated to use an order of magnitude more electricity per square foot than traditional data centers (and hence give off an order of magnitude more heat). You obviously want cheap electricity. You also want a fast network connection to be well connected to other nodes in the Bitcoin peer-to-peer network so that you can hear about new blocks as quickly as possible after they've been announced. Georgia and Iceland have reportedly been popular destinations for Bitcoin mining data centers.

***Similarities to gold mining.*** While 'mining' may seem to be just a cute name, if we step back and think about the evolution of mining, we can see interesting parallels between Bitcoin mining and gold mining. For starters, both saw a similar gold rush mentality with many young, amateur individuals eager to get into the business as soon as possible.

Whereas with Bitcoin mining we've seen a slow evolution from CPUs to GPUs to FPGAs, to now ASICs, gold mining saw an evolution from individuals with gold pans to small groups of people with sluice boxes, to placer mining — consisting of large mining groups blasting away hillsides with water — to modern gold mining which often utilizes gigantic open pit mines to extract tons of raw material from the earth (See Figure 5.10). Both with Bitcoin and with gold, the friendliness and accessibility to individuals has gone down over time and large companies have eventually consolidated most of the operations (and profits). Another pattern that has emerged in both places is that most of the profits have been earned by those selling equipment, whether gold pans or mining ASICs, at the expense of individuals hoping to strike it rich.

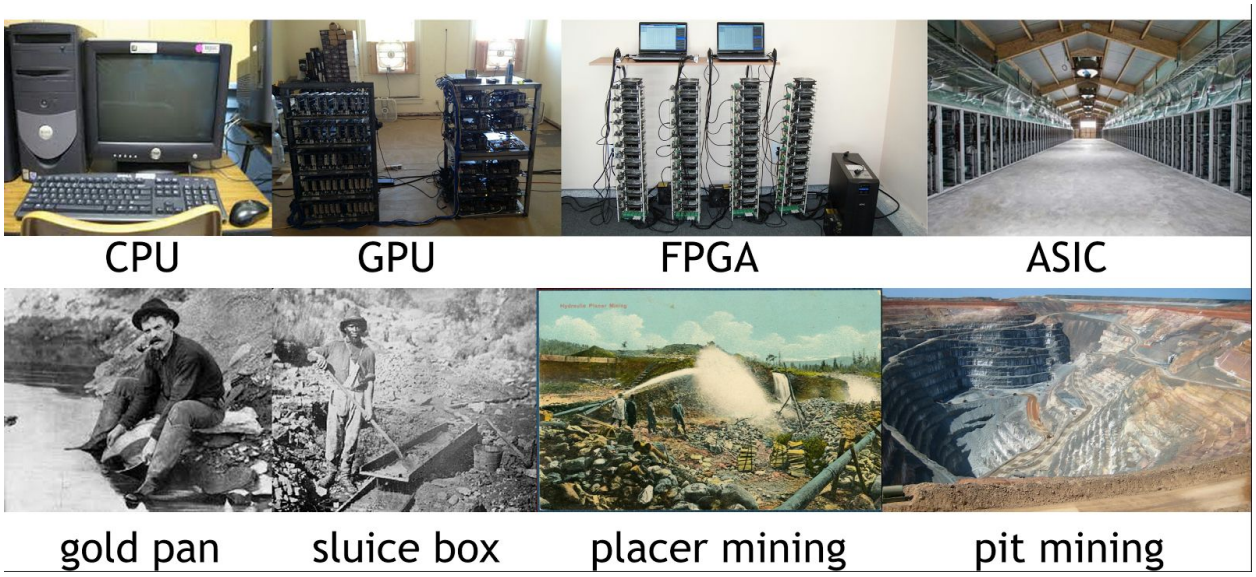



**Figure 5.10: Evolution of mining.** We can see a clear parallel between the evolution of Bitcoin mining and the evolution of gold mining. Both were initially friendly to individuals and over time became massive operations controlled by large companies.

***The future.*** Currently ASIC mining is the only realistic means to be profitable in Bitcoin and it's not very friendly to small miners. This raises a few questions about what will happen going forward. Are small miners out of Bitcoin mining forever, or is there a way to re-incorporate them? Moreover, does ASIC mining and the development of professional mining centers violate the original vision of Bitcoin which was to have a completely decentralized system in which every individual in the network mined on his or her own computer?

Furthermore, if this is indeed a violation of Satoshi Nakamoto's original vision for Bitcoin, would we be better off with a system in which the only way to mine was with CPUs? In Chapter 8, we'll consider these questions and look at ideas for alternative forms that might be less friendly to ASICs.

***The cycle repeats itself.*** It's also worth noting here that several smaller altcoins have indeed used a different puzzle than SHA-256, but have seen a similar trajectory in mining as Bitcoin. We'll discuss these altcoins more in Chapter 9 but recall that for ASICs there is still a long lead time between designing a chip and shipping it, so if a new altcoin uses an new puzzle (even just a modified version of SHA-256), this will buy some time in which ASICs are not yet available. Typically, mining will proceed just at Bitcoin did from CPUs to GPUs and/or FPGAs to ASICs (if the altcoin is very successful, like LiteCoin).

Thus, one strategy for smaller miners may be to try to pioneer new altcoins which aren't yet valuable enough for large mining groups to invest in-just like small gold miners who have been driven out of

proven goldfields might try prospecting unproven new areas. Of course, this means the pioneers are facing a significant risk that the altcoin will never succeed.

## 5.3 Energy consumption and ecology

We saw how large professional mining data centers have taken over the business of Bitcoin mining, and how this parallels the movement to pit mining in gold mining. You may be aware that pit mines have been a major source of concern over the years due to the damage they cause to the environment. Bitcoin is not quite at that level yet, but it is starting to use a significant amount of energy which has become a topic of discussion. In this section we'll see how much energy Bitcoin mining is using and what the implications are for both the currency and for our planet.

***Thermodynamic limits.*** There's a physical law known as ***Landauer's principle*** developed by Ralph Landauer in the 1960s that states that any non-reversible computation must use a minimum amount of energy. Logically irreversible computations can be thought of as those which lose information. Specifically, the principle states that erasing any bit must consume a minimum of ($kT$ ln 2) joules, where $k$ is the Boltzmann constant (approximately $1.38\times10^{-23}$ J/K), $T$ is the temperature of the circuit in kelvins, and ln 2 is the natural logarithm of 2, roughly 0.69. This a tiny amount of energy per bit, but this does provide a hard lower bound on energy usage from basic physics.

We're not going to go through the derivation here, but the high-level idea is that every time you flip one bit in a non-reversible way there's a minimum number of joules that you have to use. Energy is never destroyed; it's converted from one form into another. In the case of computation the energy is mostly transformed from electricity, which is useful, high-grade energy, into heat which is dissipated into the environment.

As a cryptographic hash function, SHA-256 is not a reversible computation. We can recall from Chapter 1 that this is a basic requirement of cryptographic hash functions. So, since non-reversible computation has to use some energy and SHA-256 — the basis of Bitcoin mining — is not reversible, energy consumption is an inevitable result of Bitcoin mining. That said, the limits placed by Landauer's principle are far, far below the amount of electricity that is being used today. We're nowhere close to the theoretical optimal consumption of computing, but even if we did get to the theoretical optimum we would still be using energy to perform Bitcoin mining.

How does Bitcoin mining use energy? There are three steps in the process that requires energy, some of which may not be so obvious:

***1. Embodied energy.*** First, Bitcoin mining equipment needs to be manufactured. This requires physical mining of raw materials as well as turning these raw materials into a Bitcoin mining ASIC, both of which require energy. This is the embodied energy. As soon as you receive a Bitcoin mining ASIC in the mail, you've already consumed a lot of energy — including the shipping energy, of course — before you've even powered it on!

Hopefully, over time the embodied energy will go down as less and less new capacity comes online. As fewer people are going out to buy new mining ASICs, they're going to be obsoleted less quickly, and the embodied energy will be amortized over years and years of mining.

***2. Electricity.*** When your ASIC is powered on and mining, it consumes electricity. This is the step that we know has to consume energy due to Landauer's principle. As mining rigs get more efficient, the electrical energy cost will go down. But because of Landauer's principle, we know that it will never disappear; electrical energy consumption will be a fact of life for Bitcoin miners forever.

***3. Cooling.*** A third important component of mining that consumes energy is cooling off your equipment to make sure that it doesn't malfunction. If you're operating at a small scale in a very cold climate your cooling cost might be trivial, but even in cold climates once you get enough ASICs in a small space you're going to have to pay extra to cool off your equipment from all of the waste heat that it is generating. Generally, the energy used to cool off mining equipment will also be in the form of electricity.

***Mining at scale.*** Both embodied energy and electricity decrease (per unit of mining work completed) when operating at a large scale. It's cheaper to build chips that are designed to run in a large data center, and you can deliver the power more efficiently as you don't need as many power supplies.

When it comes to cooling, however, the opposite is usually true: cooling costs tend to increase the larger your scale is. If you want to run a very large operation and have a lot of Bitcoin mining equipment all in one place, there's less air for the heat to dissipate into in the area surrounding your equipment. Your cooling budget will therefore increase at scale (per unit of mining work completed) unless you scale your physical area along with the number of chips you have in use.

***Estimating energy usage.*** How much energy is the entire Bitcoin system using? Of course, we can't compute this precisely because it's a decentralized network with miners operating all over the place without documenting exactly what they're doing. But there are two basics approaches to estimating how much energy Bitcoin miners are using collectively. We'll do some back-of-the-envelope calculations here based on early 2015 values. We must emphasize that these figures are very rough, both because some of the parameters are hard to estimate and because they change quickly. At best they should be treated as order-of-magnitude estimates.

***Top-down approach.*** The first approach is a top-down approach. We start with the simple fact that every time a block is found today 25 bitcoins, worth about 6,500 US dollars, are given to the miners. That's about 11 dollars every second, being created out of thin air in the Bitcoin economy and given to the miners.

Now let's ask this question: if the miners are turning all of those 11 dollar per second into electricity, how much can they buy? Of course miners aren't actually spending all of the revenue on electricity, but this will provide an upper bound on the electricity being used. Electricity prices vary greatly, but

we can use as an estimate that electricity costs around 10 cents per kilowatt-hour (kWh) at an industrial rate in the US, or equivalently 3 cents per megajoule (MJ). If Bitcoin miners were spending all 11 dollars per second of earnings buying electricity, they could purchase 367 megajoules per second, consuming a steady 367 megawatts (MW).

**Units of energy and power.** In the International System of Units (SI), energy is measured in ***joules***. A ***watt*** is a unit of power, where one ***watt*** is defined as one joule per second.

***Bottom-up approach.*** A second way to estimate the cost is to use a bottom-up approach. In this approach, we look at the number of hashes the miners are actually computing, which we know by observing the difficulty of each block. If we then assume that all miners are using the most efficient hardware, we can derive a lower bound on the electricity consumption.

Currently, the best claimed efficiency figure amongst commercially available mining rigs is about 3 GH/s/W. That is, the most cutting-edge ASICs claim to perform three billion hashes per second while consuming 1 watt of power. The total network hashrate is about 350,000,000 GH/s, or equivalently 350 petahashes per second (PH/s). Multiplying these two together, we see that it takes about 117 MW to produce that many hashes per second at that efficiency. Of course this figure excludes all of the cooling energy and all of the embodied energy that's in those chips, but we're doing an optimal calculation and deriving a lower bound so that's okay.

Combining the top down and bottom up approaches, we can derive a ballpark estimate of the amount of power being used for Bitcoin miners is probably on the order of a few hundred MW.

How much is a megawatt? To build up intuition, we can see how much large power plants produce. One of the largest power plants in the world, the Three Gorges Dam in China is a 10,000 MW power plant. A typical large hydroelectric power plant produces around 1,000 MW. The largest nuclear power plant in the world, Kashiwazaki-Kariwa in Japan, is a 7,000 MW plant, whereas the average nuclear power plant is about 4,000 MW. A major coal-fired plant produces about 2,000 MW.

According to our estimates then, the whole Bitcoin network is consuming perhaps 10% of a large power plant's worth of electricity. Although this is a significant amount of power, it's still small compared to all the other things that people are using electricity for on the planet.

***Is Bitcoin mining wasteful?*** It's often said Bitcoin "wastes" energy because the energy expended on SHA-256 computations which don't serve any other useful purpose. It's important to recognize, however that any payment system requires energy and electricity. With traditional currency, considerable energy is consumed printing currency and running ATM machines, coin sorting machines, cash registers, and payment processing services, as well as transporting money and gold bullion in armored cars. You could equally argue that all of this energy is "wasted" in that it doesn't serve any purpose besides maintaining the currency system. So, if we value Bitcoin as a useful currency system, then the energy required to support it is not really being wasted.

Still, if we could replace Bitcoin mining with a less energy-intensive puzzle and still have a secure currency, this would be a positive change. We'll see in Chapter 8, however, that we don't know if that's actually possible

***Repurposing energy.*** Another idea to make Bitcoin more eco-friendly is to capture the heat generated from Bitcoin mining do something useful with it instead of just heating up the atmosphere. This model of capturing waste heat from computation is called the approach is called the *data furnace* approach. The concept is that instead of buying a traditional electric heater to heat your home, or to heat water in your home, you could buy a heater which doubled as a Bitcoin mining rig, mining bitcoins and heating up your home as a byproduct of that computation. It turns out that the efficiency of doing this isn't much worse than buying an electric heater, and perhaps this would be no more complicated for a home consumer than plugging their heater into their Internet connection as well as their electricity outlet.

There are a few drawbacks to this approach. Although it's about as efficient as using an electric heater, electric heaters are themselves much less efficient than gas heaters. Besides, what happens when everybody turns off their Bitcoin mining rig during the summer (or at least everybody in the Northern Hemisphere)? Mining hash power might go down seasonally based on how much heat people need. It might even go down on days that happen to be warmer than average! This would caused many interesting effects for Bitcoin consensus if the data furnace model actually caught on.

The question of ownership is also not clear. If you buy a Bitcoin data furnace, do you own the Bitcoin mining rewards that you get, or does the company that sold them to you? Most people don't have any interest in Bitcoin mining — and probably never will — so it might make more sense to buy it as an appliance and have the company that sold it to you keep the rewards. This might mean the heater is sold at a slight loss then, in which case some enterprising users might want to buy them and modify them to keep the mining rewards for themselves, leading to a potentially ugly DRM (Digital Rights Management) battle.

***Turning electricity in cash.*** Another long-term question posed by Bitcoin is that it might provide the most efficient means of turning electricity into cash. Imagine a world in which Bitcoin mining ASICs are a readily-available commodity and the dominant cost of mining is electricity. In effect, this would mean providing free or low-cost electricity is open to new forms of abuse.

In many countries around the world, governments subsidize electricity, particularly industrial electricity. Among other reasons, they often do so to encourage industry to be located in their country. But Bitcoin provides a good way to turn electricity into cash, which might cause governments to rethink that model if their subsidized electricity is converted en masse to bitcoins. Electricity subsidies are intended to attract businesses that will contribute to the country's economy and labor market and subsidizing Bitcoin mining may not have the intended effect.

An even bigger problem is the billions of freely available electrical outlets around the world in people's homes, universities, hotels, airports, office buildings and so on. People might try to plug in mining equipment so that they can profit while someone else is paying the electricity bill. In fact, they might use outdated hardware and not bother to upgrade, considering that they will not be paying the electricity bill. It is quite daunting to consider the possibility of monitoring every power outlet in the world of for potential unauthorised used a source of electricity for Bitcoin mining.

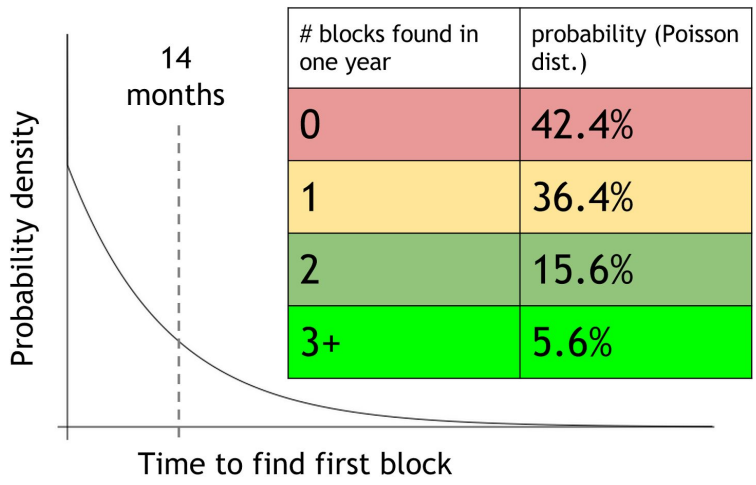
## 5.4 Mining pools

Consider the economics of being a small miner. Suppose you're an individual who spent $6,000 of your hard-earned money to buy a nice, shiny, new Bitcoin mining rig. Say that the performance is such that you expect to find a block every 14 months (and remember that a block is worth about 6,500 dollars as of early 2015).

Amortized, the expected revenue of your miner is perhaps $400 per month once you factor in electricity and other operating costs. If you actually got a check in the mail every month for $400, it would make a lot of sense to buy the mining rig. But remember that mining is a random process. You don't know when you're going to find the next block, and until that happens you won't earn anything.

***High variance.*** If we look at the distribution of how many blocks you're likely to find in the first year, the variance is pretty high and the expected number is quite low. Because you find blocks at a fixed, low rate which is independent of the time since the last block you found, your expected number of blocks is very well approximated by a ***Poisson distribution***. A Poisson distribution arises if you have *N* independent trials each with a chance $\lambda/N$ of success as *N* approaches infinity. With Bitcoin mining, each individual nonce attempted is in fact a random trial with a small chance of success, so *N* is indeed very large even for small miners and the approximation is very good.

If you expect to find about 1 block per 14 months (a Poisson distribution with $\lambda$=6/7 blocks/year), there's a greater than 40% chance that you won't find any blocks within the first year. For an individual miner, this could be devastating. You spent thousands of dollars on the miner, paid lots in electricity to run it, and received nothing in return. There's a roughly 36% chance that you'll find one block within the first year which means maybe you're barely scraping by, provided your electricity costs weren't too high. Finally, there's a smaller chance that you'll find two or more blocks, in which case you might make out with a nice profit.




| # blocks found in one year | probability (Poisson dist.) |
|---|---|
| 0 | 42.4% |
| 1 | 36.4% |
| 2 | 15.6% |
| 3+ | 5.6% |

**Figure 5.11: Illustration of uncertainty in mining.** Assuming that the global hash rate is constant and the mean time to find a block is 14 months, the variance for a small miner is quite high.

These numbers are only approximate, but the main point here is that even though on expectation you might be doing okay — that is, earning enough to make a return on your investment — the variance is sufficiently high that there's a big chance that you'll make nothing at all. For a small miner, this means mining is a major gamble.

***Mining pools.*** Historically, when small business people faced a lot of risk, they formed mutual insurance companies to lower that risk. Farmers, for example, would get together and agree that if any individual farmer's barn burned down the others would share their profits with that farmer. Could we have a mutual insurance model that works for small Bitcoin miners?

A mining pool is exactly that — mutual insurance for Bitcoin miners. A group of miners will form a pool and all attempt to mine a block with a designated coinbase recipient. That recipient is called the pool manager. So, no matter who actually finds the block, the pool manager will receive the rewards. The pool manager will take that revenue and distribute it to all the participants in the pool based on how much work each participant actually performed. Of course, the pool manager will also probably take some kind of cut for their service of managing the pool.

Assuming everybody trusts the pool manager, this works great for lowering miners' variance. But how does a pool manager know how much work each member of the pool is actually performing? How can the pool manager divide the revenue commensurate with the amount of work each miner is doing? Obviously the pool manager doesn't want to just take everyone's word for it because people might claim that they've done more than they actually did.

***Mining shares.*** There's an elegant solution to this problem. Miners can prove probabilistically how much work they're doing by outputting *shares*, or near-valid blocks. Say the target is a number beginning with 67 zeros. A block's hash must be lower than the target for the block to be valid. In the process of searching for such a block, miners will find some blocks with hashes beginning with a lot of

zeros, but not quite 67. Miners can show these nearly valid blocks to prove that they are indeed working. A share might require say 40 or 50 zeros, depending on the type of miners the pool is geared for.

```
4AA087F0A52ED2093FA816E53B9B6317F9B8C1227A61F9481AFED67301F2E3FB
D3E51477DCAB108750A5BC9093F6510759CC880BB171A5B77FB4A34ACA27DEDD
00000000008534FF68B98935D090DF5669E3403BD16F1CDFD41CF17D6B474255
BB34ECA3DBB52EFF4B104EBBC0974841EF2F3A59EBBC4474A12F9F595EB81F4B
00000000002F891C1E232F687E41515637F7699EA0F462C2564233FE082BB0AF
0090488133779E7E98177AF1C765CF02D01AB4848DF555533B6C4CFCA201CBA1
460BEFA43B7083E502D36D9D08D64AFB99A100B3B80D4EA4F7B38E18174A0BFB
0000000000000000078FB7E1F7E2E4854B8BC71412197EB1448911FA77BAE808A
652F374601D149AC47E01E7776138456181FA4F9D0EEDD8C4FDE3BEF6B1B7ECE
785526402143A291CFD60DA09CC80DD066BC723FD5FD20F9B50D614313529AF3
000000000041EE593434686000AF77F54CDE839A6CE30957B14EDEC10B15C9E5
9C20B06B01A0136F192BD48E0F372A4B9E6BA6ABC36F02FCED22FD9780026A8F
```

**Figure 5.12: Mining Shares.** Miners continually try to find blocks with a hash below the target. In the process, they'll find other blocks whose hashes contain fewer zeros — but are still rare enough to prove that they have been working hard. In this figure, the dull green hashes are shares, while the bright green hash is from a valid block (which is also a valid share).

The pool manager will also run a Bitcoin node on behalf of participants, collecting transactions and assemble them into a block. The manager will include their own address in the coinbase transaction and send the block to all of the participants in the pool. All pool participants work on this block, and they prove that they've been working on it by sending in shares.

When a member of the pool finds a valid block, they sends it to the pool manager who distributes the reward in proportion to the amount of work done. The miner who actually finds the block is not awarded a special bonus, so if another miner did more work than, that other miner will be paid more even though they weren't the one who ended up finding a valid block. See Figure 5.13.

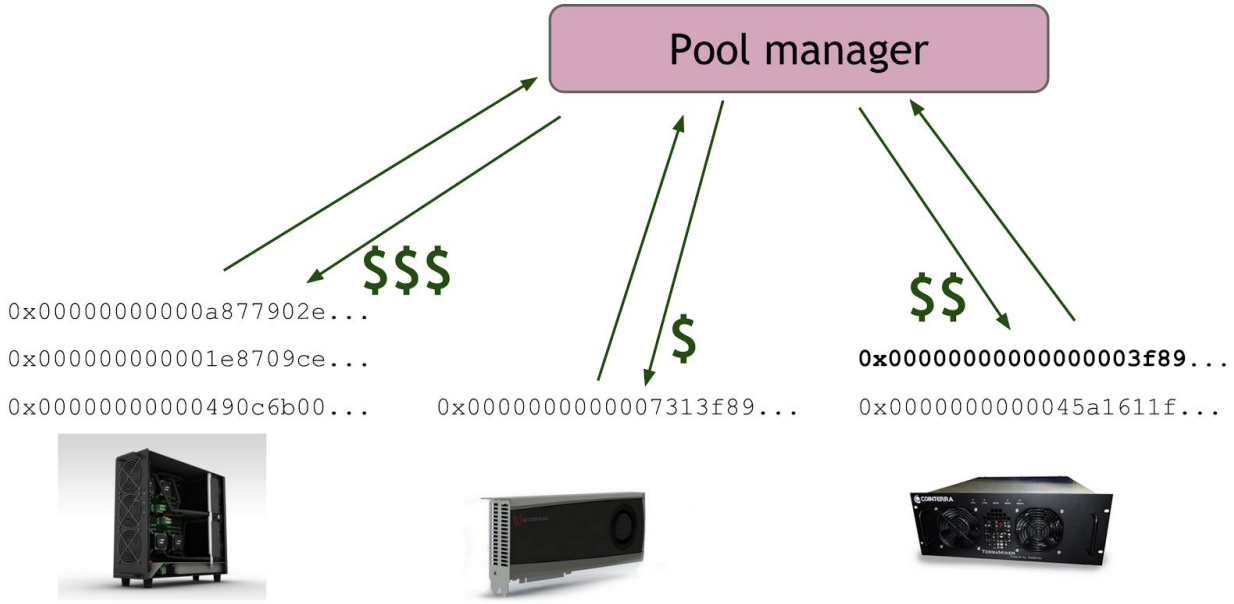



**Figure 5.13: Mining rewards.** Three participants pictured here are all working on the same block. They are awarded commensurate with the amount of work done. Even though the miner on the right was the one to find the valid block, the miner on the left is paid more since this miner did more work. There is (typically) no bonus paid to the miner who actually finds the block.

There are a few options for exactly how exactly the pool manager calculates how much to pay each miner based on the shares they submit. We'll look at two of the common, simpler ones. There are many other schemes that are also used, but these will illustrate the trade-offs between reward schemes.

***Pay-per-share.*** In the pay per share model, the pool manager pays a flat fee for every share above a certain difficulty for the block that the pool is working on. In this model, miners can send their shares to the pool manager right away and get paid without waiting for the pool to find a block.

In some ways, the pay-per-share model is the best for miners. They are guaranteed a certain amount of money every time they find a share. The pool manager essentially absorbs all of the risk since they pay rewards even if a block is not found. Of course, as a result of the increased risk, in the pay-per-share model, the pool manager will probably charge higher fees as compared with other models.

One problem with the pay-per-share model is that miners don't actually have any incentive to send valid blocks to the pool manager. That is, they can discard valid blocks but still be paid the same rewards, which will cause a big loss to the pool. A malicious pool manager might attack a competing pool in this fashion to try to drive them out of business.

***Proportional.*** In the proportional model, instead of paying a flat fee per share, the amount of payment depends on whether or not the pool actually found a valid block. Every time a valid block is found the rewards from that block are distributed to the members proportional to how much work they actually did.

In the proportional model, the miners still bear some risk proportional to the risk of the pool in general. But if the pool is large enough, the variance of how often the pool finds blocks will be fairly low. Proportional payouts provide lower risk for the pool manager because they only pay out when valid blocks are found. This also gets around the problem that we mentioned with the pay-per-share model, as miners are incentivized to send in the valid blocks that they find because that triggers revenue coming back to them.

The proportional model requires a little more work on behalf of the pool managers to verify, calculate, and distribute rewards as compared to the flat pay-per-share model.

***Pool hopping.*** Even with just these two types of pools, we can see that miners might be incentivized to switch between the pools at different times. To see this, consider that a purely proportional pool will effectively pay out a larger amount per share if a block is found quickly, as it always pays one block reward no matter how long it has been since the last block was found.

A clever miner might try mining in a proportional pool early in the cycle (just after the previous block was found) while the rewards per share are relatively high, only to switch ("hop") to a pay-per-share pool later in the cycle, when the expected rewards from mining in the proportional pool are relatively low. As a result of this, proportional pools aren't really practical. More complicated schemes, such as "pay per last *N* shares submitted" are more common, but even these are subject to subtle pool hopping behavior. It remains open how to design a mining pool reward scheme that is not vulnerable to this kind of manipulation.

***History and standardization.*** Mining pools first started around 2010 in the GPU era of Bitcoin mining. They instantly became very popular for the obvious reason that they lowered the variance for the participating miners. They've become quite advanced now. There are many protocols for how to run mining pools and it has even been suggested that these mining pool protocols should be standardized as part of Bitcoin itself. Just like there's a Bitcoin protocol for running the peer-to-peer network, mining pool protocols provide a communication API for the pool manager to send all of the members the details of the block to work on and for the miners to send back to the pool manager the shares that they're finding. getblocktemplate (GBT) is officially standardised as a Bitcoin Improvement Proposal (BIP). A competing protocol, Stratum, is currently more popular in practice and is a proposed BIP. Unlike the Bitcoin protocol itself, it is only a minor inconvenience to have multiple incompatible mining pool protocols. Each pool can simply pick whichever protocol they like and the market can decide.

Some mining hardware even supports these protocols at the hardware level, which will ultimately limit their development flexibility somewhat. However, this makes it very simple to buy a piece of mining hardware and join a pool. You just plug it into the wall — both the electricity and your network connection — choose a pool, and then it will start immediately getting instructions from the pool, mining and converting your electricity into money.

***51% mining pools.*** As of early 2015, the vast majority of all miners are mining through pools with very few miners mining "solo" anymore. In June 2014, Ghash.io, the largest mining pool, got so big that it actually had over 50% of the entire capacity over the Bitcoin network. Essentially Ghash offered such a good deal to participating miners that the majority wanted to join.

This is something that people had feared for a long time and this led to a backlash against Ghash. By August, Ghash's market share had gone down by design as they stopped accepting new participants. Still, two mining pools controlled about half of the power in the network.

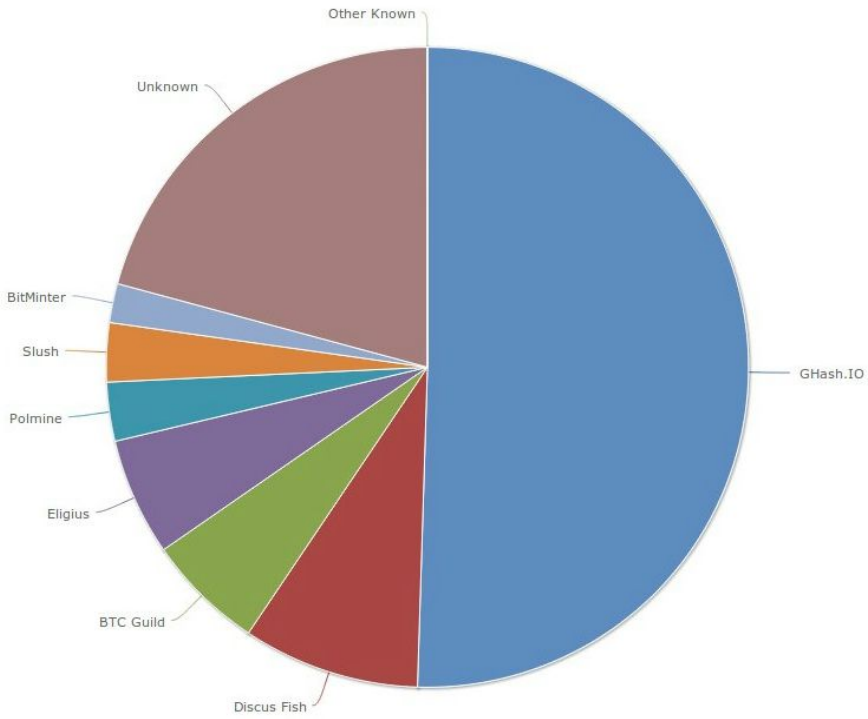



**Figure 5.14 (a) Hash power by mining pool, via blockchain.info (June 2014)**

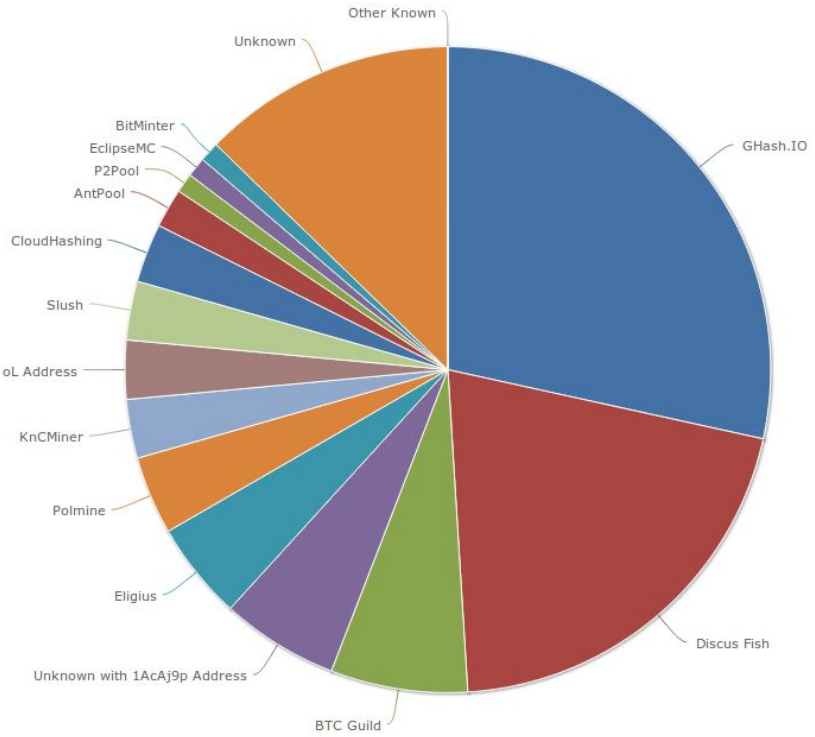



**Figure 5.14 (b) Hash power by mining pool, via blockchain.info (August 2014)**

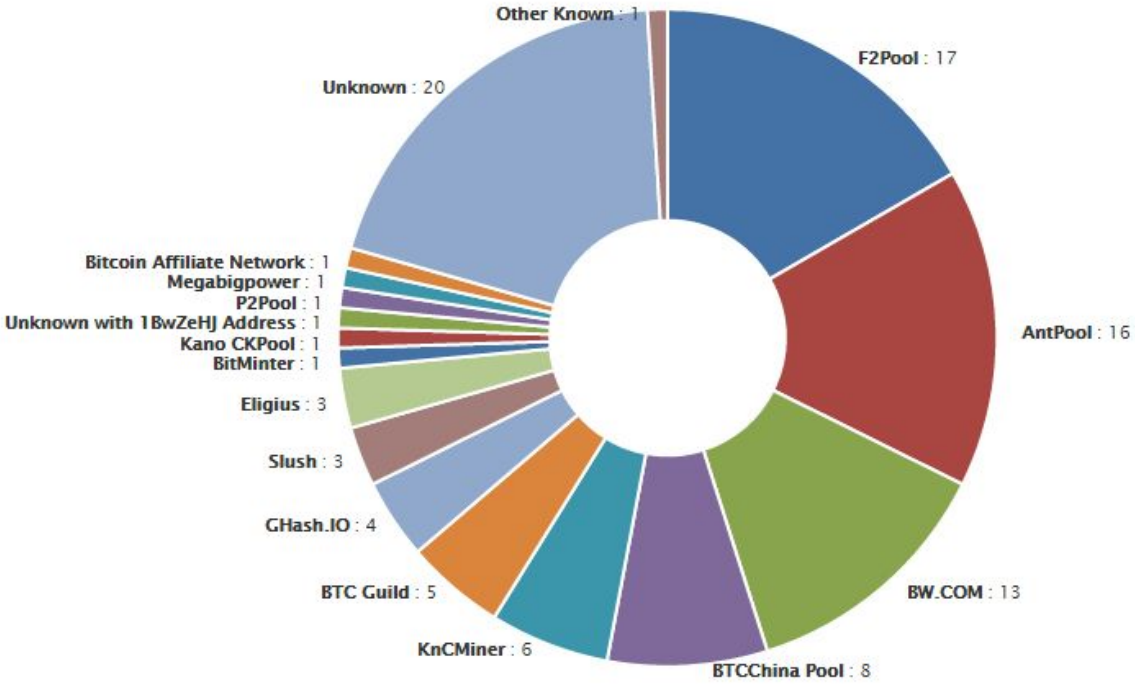



**Figure 5.14 (c) Hash power by mining pool, via blockchain.info (April 2015)**

By April 2015, the situation looks very different and less concentrated, at least on the surface. The possibility of a pool acquiring 51% is still a concern in the community, but the negative publicity GHash received has led pools to avoid becoming too large since then. As new miners and pools have entered the market and standardized protocols have increased the ease of switching between pools for miners, the market share of different pools has remained quite fluid. It remains to be seen how things will evolve in the long run.

However, it is worth noting that mining pools might be hiding actual concentration of mining power in the hands of a few large mining organizations which can participate in multiple mining pools simultaneously to hide their true size. This practice is called *laundering hashes.* It remains unknown how concentrated physical control of mining hardware actually is and mining pools make this quite difficult to determine from the outside.

***Are mining pools a good thing?*** The advantages of mining pools are that they make mining much more predictable for the participants and they make it easier for smaller miners to get involved in the game. Without mining pools, the variance would make mining infeasible for many small miners.

Another advantage of mining pools is that since there's one central pool manager who is sitting on the network and assembling blocks it makes it easier to upgrade the network. Upgrading the software that the mining pool manager is running that effectively updates the software that all of the pool members are running.

The main disadvantage of mining pools, of course, is that they are a form of centralization. It's an open question how much power the operators of a large mining pool actually have. In theory miners are free to leave a pool if it is perceived as too powerful, but it’s unclear how often miners do so in practice.

Another disadvantage of mining pools is that it lowers the population of people actually running a fully validating Bitcoin node. Previously all miners, no matter how small, had to run their own fully validating node. They all had to store the entire block chain and validate every transaction. Now, most miners offload that task to their pool manager. This is the main reason why, as we mentioned in Chapter 3, the number of fully validated nodes may actually be going down in the Bitcoin network.

If you're concerned about the level of centralization introduced by mining pools, you might ask: could we redesign the mining process so that we don't have any pools and everybody has to mine for themselves? We'll consider this question in Chapter 8.

## 5.5 Mining incentives and strategies

We've spent most of this chapter describing how the main challenge of being a miner is getting good hardware, finding cheap electricity, getting up and running as fast as you can and hoping for some