The Evolution of Bitcoin Hardware

Michael Bedford Taylor, University of Washington

Since its deployment in 2009, Bitcoin has achieved remarkable success and spawned hundreds of other cryptocurrencies. The author traces the evolution of the hardware underlying the system, from early GPU-based homebrew machines to today's datacenters powered by application-specific integrated circuits. These ASIC clouds provide a glimpse into planet-scale computing's future.

Bitcoin, since its January 2009 deployment,[1] has experienced exponential growth. As of July 2017, there are about 16.5 million Bitcoins (BTCs) in circulation; given the current (as of this writing) BTC-to-USD exchange rate of $2,500, Bitcoin's market capitalization therefore exceeds $41 billion, making it the most successful of the nearly 1,000 cryptocurrencies in use today (coinmarketcap.com).

Underpinning Bitcoin's success is a series of technological innovations that span from algorithms to distributed software to hardware. Amazingly, these innovations were not initiated by corporations or governments but rather emerged through a grass-roots collaboration of enthusiasts.

In this article, I discuss the hardware that maintains the integrity of the Bitcoin system, which evolved from CPUs to GPUs to field-programmable gate arrays (FPGAs) to application-specific integrated circuits (ASICs).[2] As Bitcoin's value grew, the industry rapidly matured and the system attained extraordinary scale, equivalent to 3.2 billion high-end GPUs. The latest round of Bitcoin hardware—dedicated ASICs—has co-evolved with datacenter design, and now most computation is performed in specialized ASIC datacenters that collectively form an *ASIC cloud*.[3,4]

HOW THE BITCOIN SYSTEM WORKS

The Bitcoin system maintains a global, distributed cryptographic ledger of transactions, or blockchain, through a consensus algorithm running on hardware scattered across the world. These machines perform a computationally intense proof-of-work function called mining, which integrates BTC transactions into the blockchain. Each transaction debiting a sender's account and crediting a receiver's account is aggregated with other pending transactions into a block by a single machine and posted to the blockchain's head. A block also contains a hash of the previous head block, creating a total order. Upon receiving notice of a block's posting, other nodes in the system will verify that the transaction is in order—for instance, not improperly creating, moving, or destroying BTCs—and then use the new block as the head block for future blockchain updates.

0018-9162/17/$33.00 © 2017 IEEE

Bitcoin mining

Bitcoin mining is the heart of the distributed consensus algorithm that enforces the consistency of BTC transactions. The earliest Bitcoin mining hardware was developed by a wide spectrum of enthusiasts from students to tech hobbyists to aspiring entrepreneurs.[5] Over the years, Bitcoin mining has consolidated and today is performed largely in custom corporate-owned datacenters, using ASICs in the latest (16 nm) technology nodes that aggressively optimize energy efficiency to unprecedented levels. The system has become increasingly vertically integrated, with single companies owning one or more datacenters, designing the chips, and maintaining the hardware. Bitcoin datacenters have migrated to regions with the lowest datacenter-related costs, including land, construction, power, taxation, and regulation.

Mining incentives

What incentivizes Bitcoin miners to perform the mining operation that is integral to BTC transaction verification? For each block they add to the blockchain, miners receive two rewards:

- *Block reward*. This was originally 50 BTCs in 2009, but the reward halves every 210,000 blocks, which occurs about every four years. As of July 2017, more than 475,000 blocks have been generated and the block reward is 12.5 BTCs (www.bitcoinblockhalf.com). Due to this halving, the number of BTCs will never exceed 21 million; 78 percent of BTCs have been mined, and 99 percent of all BTCs will be mined by 2032.
- *Transaction fees*. These fees, attached to the transactions in the block, are paid by Bitcoin users as a kind of tip to motivate the miner to incorporate their transaction into the block.

After earning BTCs, a miner can sell or trade them on an exchange like Coinbase, Bitfinex, OKCoin, or BTCC, or hold them for appreciation.

Block generation

Bitcoin mining is a key technical component of ensuring that the Internet has sufficient time to attain consensus on new blockchain updates. Miners must find a nonce value that makes a double SHA-256 hash of the block's header be less than (65535 << 208)/*difficulty*. Because SHA-256 is designed to be noninvertable, the primary approach is to use brute force. If the difficulty value is twice as large, then it takes twice as many brute-force tries on average to find the corresponding nonce.

The difficulty is scaled every 2,016 blocks using the world's collective hash rate, the *network hash rate*, in the preceding period to target an average block-creation time of 10 minutes. In practice, the time to generate blocks is somewhat random, with some blocks taking seconds and others hours.

The Bitcoin system is always searching for a new equilibrium. In the typical situation where network mining capacity increases because more machines or better hardware has been deployed, groups of 2,016 blocks will be mined more quickly than the targeted two weeks, and difficulty will be adjusted upwards. Each machine, or rig, that mines gets a correspondingly smaller fraction of the current 24 × 6 × 12.5 = 1,800 BTCs bounty that is available per day.

A QUANTITATIVE HISTORY OF BITCOIN MINING

Bitcoin was invented by a programmer or group of programmers self-identifying as Satoshi Nakamoto in a white paper[1] posted on a cryptography mailing list on 31 October 2008. Refining previous ideas about digital currency, which went back a decade, Nakamato described it as "a new electronic cash system that's fully peer-to-peer, with no trusted third party" (www.metzdowd.com/pipermail/cryptography/2008-October/014810.html). The system went live in January 2009; use initially grew slowly, then exponentially. Nakamato maintained the code base in collaboration with others online until April 2011, when he handed off responsibility and disappeared.

BTC price trends

Figure 1a shows the BTC-to-USD exchange rate over time. BTC valuations took off in mid-2010, rising from $0.08 in July 2010 to $1 in April 2011. Since then, the price has risen steadily but has also been highly volatile. There have been four bubbles: BTC prices peaked at $32 on 8 July 2011, $266 on 11 April 2013, $1,242 on 29 November 2013, and $3,000—its all-time-high—on 12 June 2017 (en.wikipedia.org/wiki/History_of_bitcoin#Prices_and_value_history). As of July 2017, the BTC price is around $2,500.

Mining difficulty trends

Figure 1b shows mining difficulty over time. The initial difficulty value of 1 corresponded to four to eight general-purpose cores running the nonce-search algorithm, trying out about 7 million double-SHA hashes per second; in July, the collective network hash rate reached 850 billion times

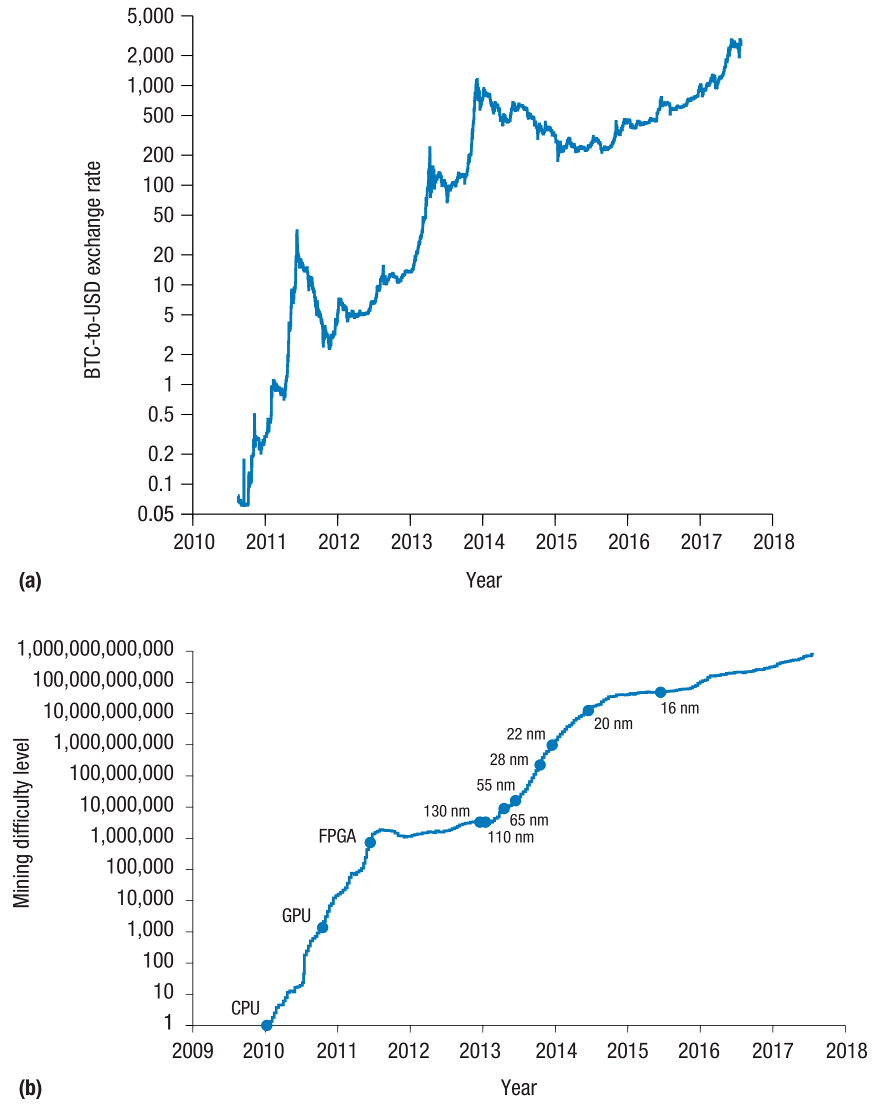

FIGURE 1. Bitcoin price and mining difficulty trends. (a) The price of Bitcoins (BTCs) took off in mid-2010, a year and a half after the system went live, and has since risen steadily but with periods of considerable volatility. (Source: bitcoincharts.com.) (b) Finding a block header hash value below the target threshold—the algorithm underlying Bitcoin's blockchain—is 850 billion times more difficult than it was originally. The approximate introduction dates of new mining technologies are indicated: CPUs, GPUs, field-programmable gate arrays (FPGAs), and application-specific integrated circuits (ASICs) in different VLSI nodes. (Data from blockchain.info.)

that (6 exahashes per second). Earning one block corresponds to about 2^{71} double SHA-256 hashes, an impressive amount of computation since each double hash is a few thousand operations itself.

Two factors increase mining difficulty. First, due to rising exchange rates, mining can cover the cost of more rigs. Second, mining software and hardware have both continually improved. Dips in difficulty often align with BTC price bubble bursts; in these cases, BTC value did not justify operating costs for the more inefficient miners, and their operators pulled them offline.

A Cambrian explosion of mining technology

The dots in Figure 1b indicate when new Bitcoin mining technology was introduced. The first publicly available CUDA-based GPU miner appeared in September 2010, followed a month later by the first OpenCL miner. Shortly afterward, in November 2010, *pooled mining* emerged, allowing parties to mine together and split the rewards pro rata.[6] These mining pools rapidly scaled to thousands of members, giving users small, frequent payouts instead of large 50- or 25-BTC payouts every several months. By this time, mining a block was equivalent to several months of computation for a single high-end consumer GPU.

Developers released the first open source FPGA miner code in June 2011. The first ASIC miner debuted in January 2013 in 130-nm VLSI technology, and more advanced ASICs rapidly followed, racing to the most advanced 16-nm node by mid-2015.

Performance and energy-efficiency advances

High-end, overclocked six-core CPUs like the Intel Core i7-990x eventually reached 33 megahashes per second (MH/s) when using SIMD (single instruction, multiple data) extensions. Top-tier consumer-grade Nvidia GPUs like the GTX 570 reached 155 MH/s, while $450 AMD GPUs like the 7970 performed even better, reaching 0.675 gigahashes per second (GH/s).

The next evolutionary step was FPGA-based miners, which emerged in June 2011. Open source versions used four Xilinx Spartan-6s, which were less cost-effective in terms of hash search time than AMD GPUs but operated on 60 W instead of 200 W. A commercial company, Butterfly Labs (BFL), began to market and sell a range of FPGA miners. These would have supplanted GPU miners due to energy costs, but the appearance of ASICs provided orders of magnitude cost

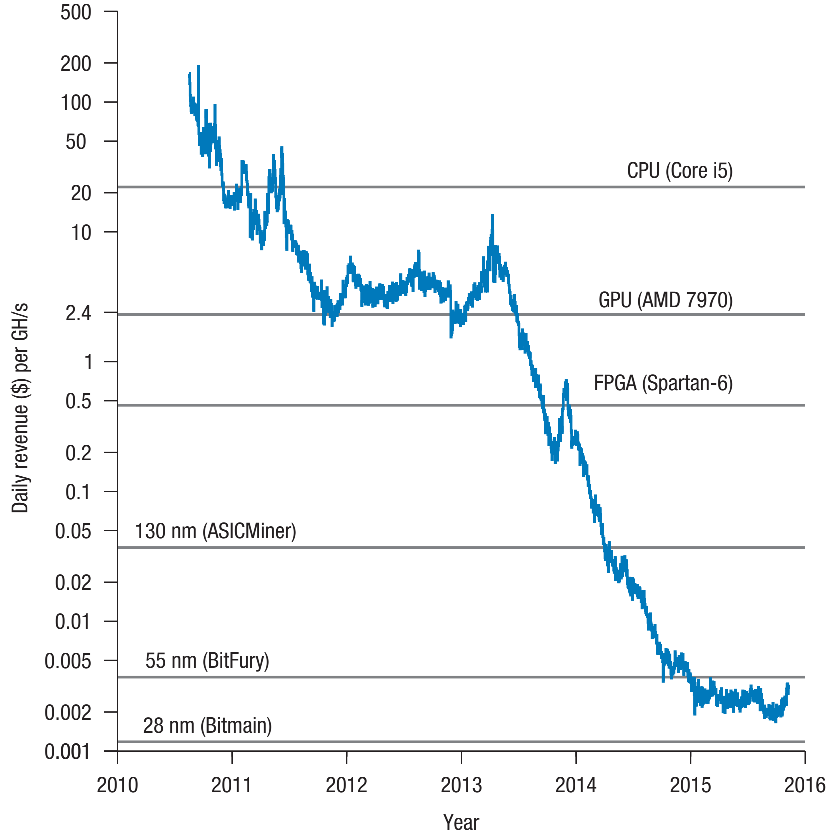

FIGURE 2. Daily Bitcoin revenue in dollars, per gigahash per second (GH/s) of mining performance, over time. The horizontal lines show the daily energy cost, at 20 cents/kWh, per GH/s of different hardware implementations as technology evolved. When mining revenue per GH/s drops below these costs, profits turn negative and the rig should be unplugged. After a GPU plateau, the system experienced a large-scale buildout of ASIC capacity, which dropped revenue per GH/s below the FPGA line and ultimately past all but the latest ASIC nodes.

reduction, driving up network hash rates and inexorably turning GPU and then FPGA profits negative. Thereafter, each more advanced generation of ASIC miners obsoleted the prior generation. Bitmain's Antminer S9 costs $2,100 and does 13.5 terahashes per second (TH/s) on 1,323 W, using 189 16-nm ASICs packed into a shoebox-size machine.

THE ECONOMICS OF BITCOIN MINING

Bitcoin entrepreneurs must weigh the costs of buying mining hardware against buying BTCs on an exchange, especially as rig maintenance requires round-the-clock monitoring and considerable energy consumption. A simple solution is to compare the purchase price and operating expenses, converted into BTCs, to the net mining returns in BTCs at the end of the machine's life.

With Bitcoin's exponential increase in hashing difficulty, a rig's ability to generate BTCs drops exponentially over time. At the lifetime average of 1.137× difficulty growth per 14-day period (see Figure 1b), more than 56.7 percent of a rig's lifetime BTC earnings comes in Q1, 24.6 percent in Q2, 10.6 percent in Q3, and 8.1 percent in Q4–Q∞. Lifetime BTC earnings top out at about 8.4 times the first two weeks' earnings. Practically speaking, a rig will be unplugged in two cases: when the earnings in dollars are less than operating costs (power, rent, and so on) and to clear space for newly purchased, quickly depreciating replacement hardware.

A rig should cost no more than the sum of these exponentially declining expected payments, minus operating costs and plus the resale value of the hardware at end of life. Custom hardware such as FPGA boards and especially ASICs have much more significant risks centered on delivery date. Receiving a new generation of hardware after other customers forfeits the early, most valuable, profits of the technology. For these reasons, large Bitcoin operations negotiate receipt of the first batches of machines, leapfrogging other customers.

Figure 2 plots daily revenue, in USD, per GH/s of mining performance paid out by the Bitcoin system since 2010, combining hashing difficulty data with the BTC-to-USD exchange rate. The horizontal lines show the daily energy cost per GH/s of CPUs (Intel Core i5), GPUs (AMD Radeon HD 7970), FPGAs (BitForce SHA256), and 130-nm through 28-nm ASICs at 20 cents/kWh energy cost. When mining revenue per GH/s drops below these costs, profits turn negative and the rig should be unplugged. After a GPU plateau, Bitcoin experienced a large-scale buildout of ASIC capacity, which dropped revenue per GH/s below the FPGA line and ultimately past all but the latest ASIC nodes. Downward voltage scaling provides a few extra months of life. Because difficulty largely increases exponentially, flat or upward regions in daily revenue per GH/s are typically the result of appreciation of BTCs relative to dollars.

EARLY BITCOIN MINING HARDWARE: THE FIRST THREE GENERATIONS

In the rest of this article, I examine some notable challenges and developments in the evolution of hardware

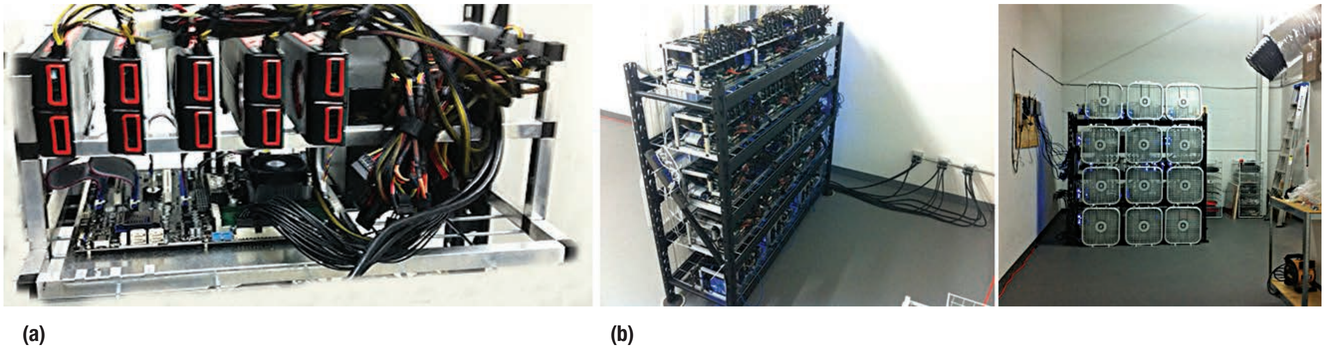

FIGURE 3. GPU Bitcoin miners. (a) Open-air rig with five GPUs suspended above the motherboards and connected via PCI Express extender cables and a single high-wattage power supply. (b) Homebrew 69-GPU mining datacenter. Note the ample power cabling (left) and cooling system, consisting of box fans and an air duct (right). Photos by James Gibson (gigavps).

customized for Bitcoin mining. Those interested in details on the first four generations should consult my paper from the 2013 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems (CASES).[2] Much of the information in this paper was drawn from an analysis of bitcointalk.org's mining hardware forum (bitcointalk.org/index.php?board=76.0), which as of July 2017 had more than 525,000 posts.

CPUs: first-generation miners

The Bitcoin miner source code (github.com/bitcoin/bitcoin/blob/master/src/miner.cpp) is surprisingly simple. The basic computation

```
while (1)
 HDR[kNoncePos]++;
  IF  (SHA256(SHA256(HDR)) < (65535
       << 208)/ DIFFICULTY) return;
```

leverages existing high-performance SHA-256 hashing libraries. One simple optimization employs a midstate buffer, which hashes the block header's beginning portion that precedes the nonce and has a constant intermediate hash value. More optimizations are discussed elsewhere.[7]

The SHA-256 computation takes in 512-bit blocks and performs 64 rounds of a basic encryption operation involving several long chains of 32-bit additions and rotations, as well as bit-wise XOR, majority, and mux functions. An array of 64 32-bit constants is also used. Each round depends on the last, creating a chain of dependencies between operations. Successive SHA-256 rounds cannot be parallelized, but each nonce trial is parallel in a classic Eureka-style computation, making this amenable to parallelization. Furthermore, some operations inside a round are parallelizable. However, typical out-of-order multicore machines have extra hardware optimized for less regular computations, resulting in wasted performance and energy efficiency.

GPUs: second-generation miners

In October 2010, Bitcoin mining software for GPUs was released on the web, and it was rapidly optimized and adapted for use in several open source efforts. Typically, this software would implement the Bitcoin protocol and GPU voltage/temperature/error control in a language such as Java or Python, and the core nonce-search algorithm as a single OpenCL file (see, for example, github.com/Diablo-D3/DiabloMiner/blob/master/src/main/resources/DiabloMiner.cl) that was compiled down by installed runtimes into the GPU's hidden native instruction-set architecture.

GPUs proved much more accessible than FPGAs for Bitcoin enthusiasts, requiring PC-building skills but no formal training in parallel programming or FPGA tools. After investing resources in a GPU-based mining rig that was literally minting cash, the natural inclination was to scale up.

Efforts to scale hash rates through GPUs pushed the limits of consumer computing in novel ways. A crowdsourced standard evolved,[2] wherein five GPUs were suspended over an inexpensive AMD motherboard with minimum DRAM, connected via five PCI Express extender cables to reduce motherboard costs, and using a large high-efficiency power supply to drive all GPUs. The system was open-air to maximize airflow, as Figure 3a shows. These approaches enabled the mining hardware to be amortized across five GPUs, improving capital efficiency.

After optimizing per-GPU overhead, the next scaling challenge was meeting the prodigious power and cooling requirements of multiple GPUs. With each GPU consuming 300 W, the power density exceeded that supported by both high-density datacenters and residential electric grids. Most successful Bitcoin mining operations typically relocated to warehouse spaces with a large air volume for cooling and cheap industrial power rates. Figure 3b shows a homebrew datacenter consisting of a 69-GPU rack cooled by an array of 12 box fans and an airduct.

FPGAs: third-generation miners

June 2011 brought the first open source FPGA Bitcoin miner implementations. FPGAs are inherently good at processing SHA-256's rotate-by-constant and bit-level operations, but not its 32-bit add operations.

The typical FPGA miner replicated multiple SHA-256 hash functions and unrolled them. With full unrolling, the module created different hardware for the 64 hash rounds, each of which was separated by pipeline registers. These registers contained the running hash digest as well as the 512-bit block being hashed. The state for a given nonce trial would proceed down the pipeline, one stage per cycle, allowing for a throughput of one nonce trial (hash) per cycle.

Hackers developed custom boards that minimized unnecessary costs due to RAM and I/O and focused on providing sufficient power and cooling; these boards attained 215 MH/s rates with Spartan XC6SLX150 parts. Quad-chip boards were developed to reduce board fabrication, assembly, and bill-of-materials costs, reaching 860 MH/s at 216 MHz and 39 W, and costing $1,060. Kansas-based BFL offered a non–open source version for $599 with similar 830 MH/s performance. BFL was by all accounts the most successful commercial FPGA miner vendor.

FPGAs had trouble competing on cost per GH/s with high-volume GPUs that were on more advanced process nodes and sold on retail sites like Newegg. However, FPGAs were up to five times more energy efficient than GPUs, breaking even on total cost of ownership (TCO) after a year or two. Nevertheless, the reign of FPGA miners was brief because ASICs arrived soon after, providing orders of magnitude cost and energy-efficiency improvements.

THE ASIC RACE: FOURTH-GENERATION BITCOIN MINERS

Three companies came to market with ASIC Bitcoin miners in close succession. The designs were based loosely on FPGA miners. Because ASICs brought enormous benefits over prior devices,[3,4] the emphasis was on getting a working, not necessarily optimal, design out as quickly as possible.

Butterfly Labs

Fresh off the success of its FGPA miners, BFL was the first to announce an ASIC product line. The company took preorders in June 2012 for three types of machines; $149 Jalapenos rated at 4.5 GH/s, $1,299 Singles rated at 60 GH/s, and $30,000 Mini Rigs rated at 1,500 GH/s. At these prices, the machines could generate 20 to 50 times more BTCs per dollar invested versus GPUs. The preorder revenue, which exceeded $250,000 on day one, presumably covered the $500,000 nonrecurring engineering (NRE) mask costs[4] for BFL's 65-nm GlobalFoundries process.

The chip in all three products contained 16 double SHA-256 hash pipelines. The die was 7.5 × 7.5 mm and placed in a 10 × 10 mm BGA 144-lead package. BFL initially targeted a November 2012 ship date, but the schedule repeatedly slipped due to setbacks and delays from the ASIC foundry, packaging, and BFL itself. It took nearly a year to clear the order backlog. A major cause was that the chip consumed four to eight times more power than expected, requiring a redesign of all ASIC systems. For example, the Jalapenos, slated to use one chip, shipped with two chips to meet the 4.5 GH/s rate, and they typically operated at 30 W, close to 6 W per GH/s.

ASICMiner

ASICMiner was founded in early July 2012, after BFL had started taking preorders for their machines, by three Chinese nationals. A key motivation was to prevent BFL from being the sole Bitcoin ASIC purveyor and controlling the blockchain. ASICMiner's approach was quite different than BFL's; it initially intended not to sell hardware but to run an ASIC datacenter that mined BTCs on behalf of shareholders. This approach, arguably the first ASIC cloud, eliminated the need to ship hardware to customers and won the race to large-scale deployment.

Lacking BFL's name recognition, ASICMiner raised funding online through bitcointalk.org and some Chinese-language forums. The company carefully outlined its plan for developing an ASIC Bitcoin miner, and responded to hundreds of questions by the online community regarding its business model, technical decisions, and financial trustworthiness.[2]

In early August 2012, after completing an initial place-and-route, ASICMiner proceeded to raise funds through an IPO on the online stock exchange GLBSE, in which the securities were Bitcoin-related and further denominated in BTCs. The IPO closed 27 August, selling 163,962 shares—roughly equivalent to $160,000. By 22 September, ASICMiner had finalized the chip's specs, and a tapeout shortly followed. On 28 December, the company posted photos of its chip carrier—the first ASIC miner—on bitcointalk.org's forum. By 31 January 2013, ASICMiner had 64-chip boards in hand and aimed to deploy 800 of them, mounted in 10-board backplanes, the following month. By 14 February, it had 2-TH/s miners in the wild.

Over time, ASICMiner continued to deploy at capacity but had difficulty scaling its datacenter and started selling hardware. It first sold boards from its datacenter but later developed a USB miner stick with a single ASIC, the Block Erupter, which sold initially for 2 BTCs in large lots to be resold by

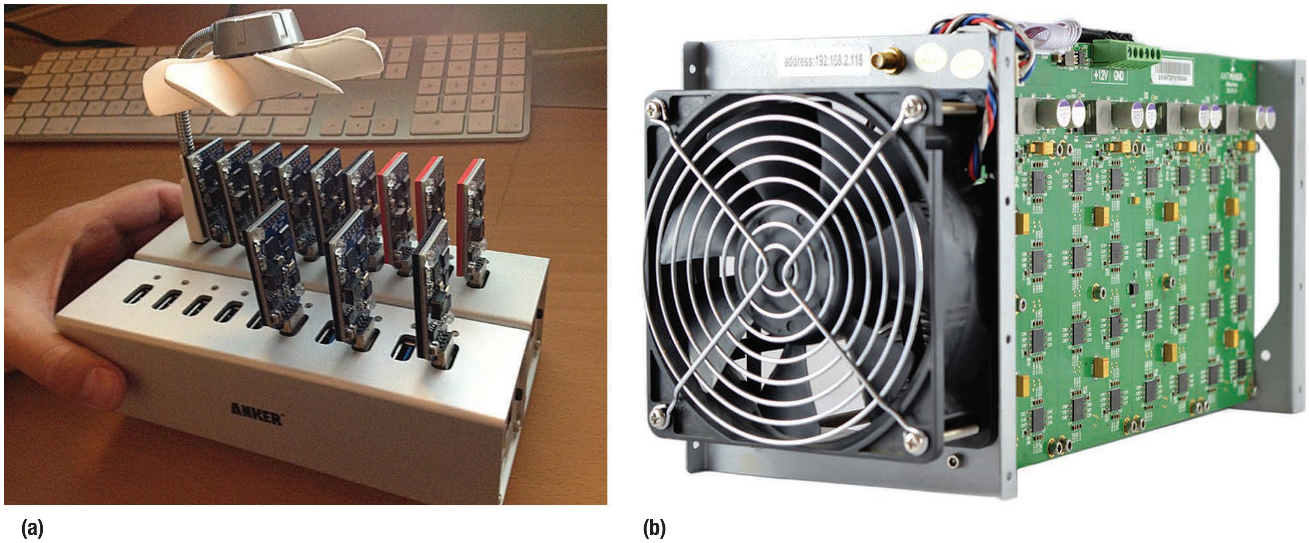

FIGURE 4. ASIC Bitcoin miners. (a) USB hub hosting an array of ASICMiner Block Erupter USB stick–style miners and a USB-powered cooling fan. Each USB stick's 130-nm ASIC hashes at 330 megahashes per second (MH/s), or about half the MH/s performance of a $450 28-nm AMD Radeon HD 7970 GPU. (b) Bitmain Antminer S1 machine with two parallel sea-of-ASICs printed circuit boards. Photos by DennisD7 and dogie of bitcointalk.org.

others and rapidly dropped in price. Figure 4a shows a USB hub hosting an array of USB stick–style Bitcoin miners and a USB-powered cooling fan. Each USB stick has a 130-nm ASIC that hashes at 330 MH/s at 1.05 V and 2.5 W, reaching 392 MH/s at 1.15 V. The ASIC performs one hash per cycle, mirroring earlier FPGA designs. It is 40 times more energy efficient than the 28-nm AMD 7970 GPU and 4.4 times cheaper per GH/s.

ASICMiner shares reached 4 BTCs each in October 2013, signifying a 40× return to the initial investors. Of the three early ASIC mining companies, it was the most innovative in trying out new products and business models.

Avalon

Avalon also secured grass-roots funding through direct presales of units via an online store. A key founder, N.G. Zhang, established his reputation with the design of a top Bitcoin FPGA board, Icarus. Avalon focused on an 110-nm TSMC implementation of a double SH-256 pipeline, measuring 4 × 4 mm, and packaged 300 chips across three blades inside a 4U-ish machine. Like ASICMiner, Avalon was based in Shenzhen, China. The company preordered sales of 300 rigs, each priced at $1,299 or 108 BTCs at the time, and hashing at 66 GH/s at 600 W.

Avalon taped out slightly after ASICMiner, with a target date of 10 January 2013. On 30 January, Bitcoin developer Jeff Garzik became the first customer in history to receive an ASIC mining rig, which earned about 15 BTCs the first day. Avalon offered batches of 600 rigs for 75 BTCs on 2 February ($1,600) and 25 March ($5,500). They sold out almost immediately. Avalon followed up with direct chip sales, selling more than 100 batches of 10,000 chips for 780 BTCs per batch, or about $78,000, enabling others to design systems around the new chips.

THE ASIC WAR: FIFTH-GENERATION BITCOIN MINERS

The next generation of ASICs departed from the first in several ways. After first-generation ASICs had proven their value in Bitcoin mining, venture capitalists and other investors funded a swath of start-ups, many featuring industry veterans. Moreover, the competition was not easily beaten GPUs but rather other ASICs. New ASICs had to best the previous generation in cost/performance and energy efficiency to be competitive and stay ahead of ever-rising difficulty levels. These successive generations had two potential sources of innovation: better architectures and more advanced process nodes. To date, there have been more than 37 different ASIC efforts.

BitFury, with star chip designer Valery Nebesny, reached 55 nm first in mid-2013 with a best-of-class fully custom implementation in many ways superior to 28-nm designs, reaching 0.8 W per GH/s and 2.5 GH/s per chip. Sixteen chips were placed on a printed circuit board, and 16 PCBs went into a backplane. Unlike most other architectures that unrolled double SHA-256 hashes into long pipelines, BitFury's used "rolled" hashes that iterate in place. It also introduced support for string designs, with ASIC power pins connected serially like Christmas tree lights, eliminating the DC–DC converters that comprise 20–40 percent of Bitcoin server cost. BitFury's initial 40,000 chips went to a large datacenter provider that financed the NRE costs. Later, individual chips were sold, and interesting variants ranging from USB keys to blades were sold by third parties online, including on Amazon.com.

Sweden-based KnCMiner reached 28 nm by October 2013. Shortly afterward, San Francisco–based HashFast and Austin-based CoinTerra[8] also came out with 28-nm implementations. These ASIC miners were much more cost-efficient than the BitFury chips, but energy efficiency was actually worse: greater than 1.1 W per GH/s. The designs placed four dies on a shared substrate that reached several hundred watts and required water cooling. Because BitFury had several months to ramp up before these products came out, HashFast and CoinTerra were caught off guard by its deployment of massive quantities of highly efficient 55-nm chips, as well as concurrently shipping 28-nm chips. This limited the usefulness for HashFast and CoinTerra's machines and contributed to the companies going out of business.

BFL, Spondoolies, and Bitmain also implemented 28-nm miners, targeting energy efficiencies that matched or exceed BitFury's designs, at 0.7 W per GH/s. Figure 4b shows Bitmain's Antminer S1. There is evidence that 21 Inc reached 22 nm around December 2013, but the details are closely guarded secrets.

THE ASIC VICTORS: SIXTH-GENERATION BITCOIN MINERS

Current sixth-generation Bitcoin miners are the products of companies that survived the ASIC war and advanced to bleeding-edge nodes as they emerged (for example, 20 nm and 16 nm). The two main publicly known contenders are BitFury (bitfury.com) and Bitmain (www.bitmain.com), which have 16-nm chips. Both companies' implementations run at ultralow voltages; BitFury miners exceed 0.07 W per GH/s, which is 100 times more energy efficient than the first 130-nm ASIC miners and 8,000 times more energy efficient than GPU miners.

Several existing Bitcoin mining companies now develop their own ASICs and have created ASIC cloud datacenters in areas with low energy and cooling costs.[9] For example, BitFury optimizes its chips for use in new immersion-cooled datacenters in the Republic of Georgia, Iceland, and Finland.[10]

Merged ASIC development and datacenter operation have become prevalent in the industry for three reasons. First, the ASIC, enclosing machine, and datacenter can be codesigned. This eliminates the need to worry about varying customer environments (temperature, customs certification, 220-V/110-V compatibility, setup and tech support, shipping and returns, warranties, and so on) and enabling new cost, energy-efficiency, and performance optimizations. Second, the time to get an ASIC running is greatly shortened if the product does not have to be packaged, troubleshooted, and shipped to the customer, which means that the chips can start hashing earlier. This is particularly important when the network hash rate is increasing exponentially and the bulk of the profits are earned early in a machine's life. Third, tuning an ASIC chip to exactly meet promised energy-efficiency and performance specifications before shipping to a customer delays ASIC deployment and reduces ASIC lifetime.

Bitcoin mining is an example of the emerging class of planet-scale applications. Today, companies including Apple, Facebook, and Google are deploying planet-scale applications like Siri, Facebook Live, and Brain, respectively, for which computational demand scales with the number of users just like with Bitcoin. Ultimately, the TCO of the datacenters that run these computations becomes so large that it makes economic sense to build specialized ASICs to reduce hardware cost and power consumption. Following this trend, last year Google announced the creation of neural-network ASICs for their datacenter workloads.[11] Recent ASIC cloud research shows how the lessons from Bitcoin mining hardware apply to other workloads like YouTube's video transcoding.[12] The future of ASIC clouds is bright, in part due to the many pioneers who took financial, legal and, technical risks to accelerate Bitcoin development and design an entirely new class of hardware.

ABOUT THE AUTHOR

MICHAEL BEDFORD TAYLOR is a professor in the Computer Science and Engineering as well as the Electrical Engineering departments at the University of Washington. His research interests include ASIC clouds, Bitcoin mining hardware, dark silicon, tiled microprocessors, and open source hardware design. Taylor received a PhD in electrical engineering and computer science from MIT. Contact him at prof.taylor@gmail.com.

ACKNOWLEDGMENTS

This work was partially supported by NSF awards 1228992, 1563767, and 1565446, and by STARnet's Center for Future Architectures Research, a SRC program sponsored by MARCO and DARPA.

REFERENCES

1. S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008; bitcoin.org/bitcoin.pdf.
2. M.B. Taylor, "Bitcoin and the Age of Bespoke Silicon," *Proc. Int'l Conf. Compilers, Architectures, and Synthesis for Embedded Systems* (CASES 13), 2013, article no. 16.
3. I. Magaki et al., "ASIC Clouds: Specializing the Datacenter," *Proc. 43rd Int'l Symp. Computer Architecture* (ISCA 16), 2016, pp. 178–190.
4. M. Khazraee et al., "Moonwalk: NRE Optimization in ASIC Clouds," *Proc. 22nd Int'l Conf. Architectural Support for Programming Languages and Operating Systems* (ASPLOS 17), 2017, pp. 511–526.
5. J. Light, "For Virtual Prospectors, Life in the Bitcoin Mines Gets Real," *The Wall Street J.*, 19 Sept. 2013; www.wsj.com/articles/for-virtual-prospectors-life-in-the-bitcoin-mines-gets-real-1379644359.
6. M. Rosenfeld, "Analysis of Bitcoin Pooled Mining Reward Systems," 22 Dec. 2011; arxiv.orgpdf/1112.4980.pdf.
7. N.T. Courtois, M. Grajek, and R. Naik, "Optimizing SHA256 in Bitcoin Mining," *Proc. Int'l Conf. Cryptography and Security Systems* (CCSS 14), 2014, pp. 131–144.
8. J. Barkatullah and T. Hanke, "Goldstrike 1: CoinTerra's First-Generation Cryptocurrency Mining Processor for Bitcoin," *IEEE Micro*, vol. 35, no. 2, 2015, pp. 68–76.
9. N. Popper, "Into the Bitcoin Mines," *The New York Times*, 21 Dec. 2013; dealbook.nytimes.com/2013/12/21/into-the-bitcoin-mines.
10. A. Kampl, "Bitcoin 2-Phase Immersion Cooling and the Implications for High Performance Computing," *Electronics Cooling*, Mar. 2014, pp. 24–29.
11. C. Metz, "Google Built Its Very Own Chips to Power Its AI Bots," *Wired*, 18 May 2016; www.wired.com/2016/05/google-tpu-custom-chips.
12. M. Khazraee, et al, "Specializing a Planet's Computation: ASIC Clouds," *IEEE Micro*, vol. 37, no. 3, 2017, pp. 62–69.