Blockchain and Internet of Things Data Provider for Smart Applications

Bogdan Cristian FLOREA
Department of Applied Electronics and Information Engineering
Politehnica University of Bucharest
Bucharest, Romania
bogdan.florea@upb.ro

Abstract— **This paper describes the use of blockchain technology as a data provider in Internet of Things (IoT) applications. Blockchain is a novel technology, which has gained a lot of attention in the last years, mainly due to its use as a backbone for cryptocurrencies. The main purpose of blockchain technology is to provide anonymous transactions between participants, over a peer-to-peer network, using a decentralized distributed ledger. The goal of this novel approach is to eliminate any 3rd party validation and replace the trust of a central authority for transaction validation with cryptographic proof. While most applications of the blockchain revolve around cryptocurrencies, the blockchain can be used in many other fields, such as finance, distributed data storage, health and medicine, automation, etc. By creating an open, decentralized network, the blockchain can be used to develop decentralized applications and enable data access and sharing on a much higher level than the common implementations of client-server architectures which are in use today. In this paper, we will present a proof of concept method for field devices to store and share data using a distributed ledger built on the IOTA tangle, as well as provide means of access to the data which can be used in IoT and decentralized applications.**

Keywords- blockchain; internet of things; decentralized; distributed; data provider; tangle

I. INTRODUCTION

Blockchain technology was introduced in 2008, as a platform for secure, anonymous transactions, using a decentralized network of computers or devices. The first application of blockchain technology was the cryptocurrency Bitcoin, which promises anonymity, by allowing users to transfer tokens over a peer-to-peer (P2P) network without the regulation of a central authority [1]. The blockchain network is similar to distributed ledger technologies (DLT), which are distributed datasets over multiple locations, using peer-to-peer networks, where every change in the ledger is reflected in all copies over the network [2]. As such, a consensus mechanism is required, so the replicated information can't be altered by a user or a group of users. The blockchain achieves consensus using cryptographic functions with increasing difficulty.

The main idea of the blockchain network is to substitute trust (provided by a 3rd party) with proof [1]. To achieve this, a cryptographic hash function is used, which is validated by the network before a new block is created. The security is provided by using the resulting hash of a previous block as the starting point for the next block (Fig. 1). To achieve consensus, the network nodes, or "miners", will validate the resulting hash, thus confirming the previous transaction and start work on "discovering" the next block, by finding the next suitable hash of the block data. The transaction is bundled in blocks by using a Merkle tree, which is a type of tree structure where each leaf node contains the data and each non-leaf node contains the cryptographic hash of the descending nodes [3]. Using this method, only the Merkle root hash is added in the block hash function, allowing for old blocks to be compacted.

This consensus method is known as Proof-of-Work (POW), and the nodes are rewarded for the work done on the network, using a token as a reward. In most blockchain implementation, the tokens are "mined" (rewarded) for each new block added to the chain and they can be further traded on the network, or, in most cases, exchanged for fiat currency. This reward system motivates users to use the network and gain some value in exchange for the computational power they provide.

One important feature of blockchain technology is data immutability [4]. Once a block has been validated by the network, its resulting hash is used in the next block. If a participant would try to change an existing block and broadcast it to the network, it would not be accepted, as it would change its computed hash and all subsequent blocks hashes. The network would reject such a block, as long as the computing power of the network remains neutral [1]. This novel approach ensures that no one entity can have control over the data.

With the increased usage of the blockchain network, some limitations or drawbacks have emerged, such as increased difficulty of the hashing function, which results in high transaction times and very high electrical power usage [5][6].

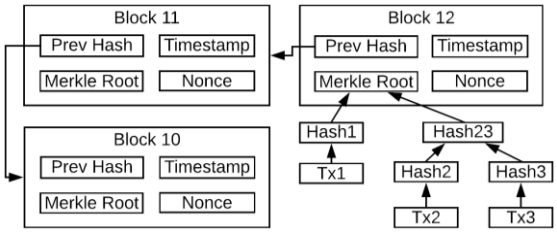

Figure 1. Typical blockchain structure

This work has been funded by University Politehnica of Bucharest, through the "Excellence Research Grants" Program, UPB – GEX 2017. Identifier: UPB-GEX 2017, Ctr. No. 39/2017".

978-1-5386-5683-9/18/$31.00 ©2018 IEEE

This has led the development of alternative blockchains, with different approaches for block validation. One such alternative is to use a Proof-of-Stake (POS) mechanism, in order to ensure the consensus [7]. In this approach, the creator of a new block is chosen deterministically, based on his stake of tokens. The rewards will be awarded from transaction fees, eliminating the reward for miners who generate the hash function. The idea is that users lock a certain number of tokens, which provide interest over time. This approach would reduce the energy footprint and the necessary hardware required to run the network and validate the transactions.

While the systems described above have huge potential, there has been very little use of them outside the cryptocurrency world. In the next section, we present the concept of IOTA, a novel approach of blockchain technology for distributed applications, and demonstrate the use of the network for IoT applications and data marketplaces.

II. THE IOTA TANGLE

The main disadvantage on most blockchain networks is that they require a certain number of tokens (value) to be transferred between network participants. Most blockchain systems have transaction fees which are used to generate block rewards as incentives, in order to keep the network alive. Due to this fact, blockchain networks haven't gained much traction outside the cryptocurrency world, such as using the system for transfer of data, instead of tokens.

IOTA is a new type of DLT, which aims to mitigate two of the most important issues of the current blockchain solutions: high transaction fees and high processing time.

The IOTA network was designed with IoT applications in mind (labeling itself as "The backbone of IoT"), by enabling zero-fee and zero-value transactions [8]. These two features are unique to the IOTA network, making it usable without actually requiring the network tokens (value) to be transferred between participants and thus allowing only data messages to be stored.

IOTA has a different approach to blockchain technology, by replacing the sequential blockchain used in other systems (Fig. 1) with a type of directed acyclic graph (DAG), which is referred to as "The Tangle" (Fig. 2).

Instead of chaining transactions in a linear fashion, on the IOTA network they are stored in a graph-like manner, each new transaction referencing two previous transactions, which it has to validate before attaching to the tangle [8]. Because of this novel mechanism, the Proof-of-Work is only performed when generating a new transaction.

The main advantage of this approach is network scalability. While in other blockchain systems, scalability is becoming an issue, as the network speed decreases with the increase of transfers, on the tangle it is quite the opposite, the more transactions on the network, the higher the speed.

In IOTA, newly attached transactions are called tips. As each new transaction T (Fig. 3) references two previous transactions, the network will pick two unconfirmed tips, to

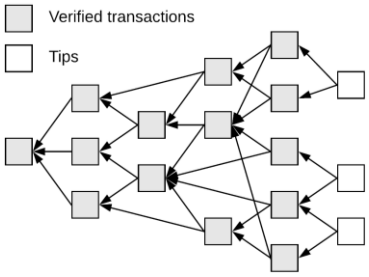

Figure 2. The IOTA tangle

which the new transaction will attach itself. The two transactions are validated by their signatures [8]. If the tips are validated, the new transaction is attached to the tangle, but it also confirms all other transactions linked by the two tips, generating a validation path, which is highlighted in Fig. 3. This increases the trust of all the transactions in the validation path. The more transactions are added, the higher the certainty levels of confirmation in the transactions path.

To understand how the network works, we can consider that, given the total number of tips at a given time $L(t)$, the tangle can be modelled as a Poisson process of rate λ [9], which is assumed constant over time [8]. Assuming h to be the average time required by a device to perform the necessary computations for issuing a transaction, at any given moment in time, a node does not view the actual state of the tangle, but the state from exactly h time units ago (when the selection of tips was made). The current transaction will become visible to the network at the moment $t + h$.

In the ideal network state, the number of tips is stationary in time. We will denote this value $L_0 > 0$. Given the previous assumptions, the number of tips L_0 can be expressed as a function of h and λ [8].

At any given moment, there are λh hidden tips (which were attached in the last h time units, so are not yet visible to the network) [8]. If r is the number of known tips at time $t - h$, and considering the stationarity assumption, we have:

$$L_0 = r + \lambda h. \tag{1}$$

Using the same logic, we can deduce that at time t, there are λh transactions which were tips at time $t - h$ but are not tips anymore.

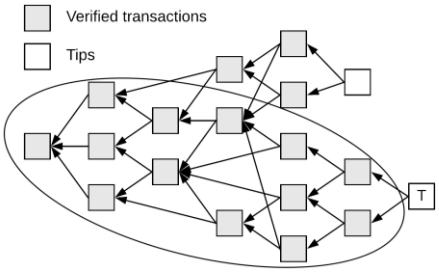

Figure 3. Transaction validation path

The probability for a new transaction of selecting a tip is:

$$p = r/(r + \lambda h). \quad (2)$$

Since every new transaction selects two previous tips, the probability becomes $2p$. Since in the stationarity assumption the number of tips should remain constant, then new tips should not change the overall number L_0. This results in $2p = 1$. Solving this with respect to r will result in $r = \lambda h$ and so:

$$L_0 = 2\lambda h. \quad (3)$$

From (3), we have the expected time t_a for a transaction to be approved for the first time is:

$$t_a = h + L_0/2\lambda = 2h, \quad (4)$$

since for the first h time units it cannot be approved.

Note that in the previous assumptions, the tips are chosen at random. This may not be the best strategy, as it can encourage new transactions to reference a fixed set of older transactions (reconfirming them) and leaving out more recent tips. This type of transaction is called a lazy tip. In order to prevent this type of behavior, the network uses a Markov chain Monte Carlo (MCMC) method of selecting the next tips to be validated [8].

As we have shown, the IOTA tangle acts different than other blockchain networks for linking together transactions on the tangle. For each new transaction, POW must be done in order to approve two past transactions and increase the confidence level of older transactions in the validation path (Fig. 3). One very important thing to note, which sets IOTA apart from other blockchain-type systems, is that the computations should be done locally, on the device which initiates the transaction (client), but they can also be delegated to the node to which the device is connected. Although most public nodes don't allow this function, one can always deploy a full node containing the full tangle and API, and leverage this functionality, allowing simpler devices, such as single board computers or microcontrollers to push data to the network.

For the actual validation, IOTA uses the Curl cryptographic function, which is based on the SHA-3 family of functions, with the addition of converting the input/output to trinary. Trinary alphabet consists of digits –1, 0 and 1 (trits). For actual representation of messages, IOTA uses trytes, which are made out of 3 trits, representing a total of 27 possible values. The IOTA alphabet contains the characters A-Z (uppercase) and 9.

III. The Data Provider System

In order to demonstrate the use of the IOTA tangle as a data provider for IoT applications, we have deployed a full node which is connected to the tangle. This node will receive the transaction data, do the required POW (validate two previous transactions) and attach the new transactions to the tangle.

The data is collected from a sensor station which consists in a DS18B20 temperature sensor, connected to a Raspberry Pi board. The sensor has a measuring range from –55°C to +125°C, with best accuracy between –10°C and +85°C.

The Raspberry Pi collects data at a fixed time interval of 90 minutes. The data is formed into JSON (JavaScript Object Notation) messages, which are then tryte-encoded, using the IOTA API functions. Each message is assigned a tag, which allows the end users to search the tangle for the required data. A sample message is shown in Table 1.

After the message and tag are encoded, a transaction is created by preparing the transfers and creating a new bundle which is attached to the tangle. A bundle can contain a set of transfers, with different values and messages and different recipient addresses. The resulting bundle from the previous transfer in Table 1 is attached to the tangle by referencing two previous tips.

Ideally, the POW should be done on the client-side (in this case, the Raspberry Pi station) and attach the newly formed transaction to the tangle. The current implementation of the Curl SHA-3 algorithm, which is used to sign and verify the transactions, requires OpenCL support, which is not yet available on the Raspberry Pi board. Since IOTA is a work in progress, other implementations will become available, which may allow for smaller embedded devices to do local POW, thus allowing them to connect to one of the public nodes which do not allow the computations of the actual POW.

Fortunately, since the POW can be delegated to a running node, it is possible for even the simplest devices to collect data and attach it to the tangle, by simply providing a node which allows the Proof-of-Work. In this way, multiple devices, such as microcontrollers, can be connected to the same node which can handle the transaction operations.

Another advantage of running a full node is that one can control how and when the transactions are generated, by implementing a middleware which can collect data from a group of sensors and group it into one or more transactions.

TABLE I. JSON AND TRYTE-ENCODED MESSAGES

	Message	Tag

JSON	{
	"name": "IOTA Station"
	"message": [{
	"variable": "temperature"
	"value": 10.062
	"unit": "°C"
	"timestamp": "1519786031"
	}]
	}
		ist_180228[a]

Trytes	ODGABDPCADTCGADBGASBYBCCKBEA BCHDPCHDXCCDBDGAQAGAADTCGDG DPCVCTCGADBJCODGAJDPCFDXCPCQC 9DTCGADBGAHDTCADDDTCFDPCHDIDF DTCGAQAGAJDPC9DIDTCGADBVAUASA UA9BWAQAGAIDBDXCHDGADBGANFM BGAQAGAHDXCADTCGDHDPCADDDGA DBGAVAZAVACBABBB9BUAXAVAGAQ DLCQD	GAXCGD HDNCVA BBUAWA WABBGA

a. The tag contains a station identifier (ist) and the date formatted as YYMMDD

To secure the data, a new method is proposed by the IOTA network: Masked Authenticated Messaging (MAM). The method uses a Merkle tree signature scheme to encrypt the messages, allowing the implementation of provider/subscriber systems, with controlled access and forward secrecy. In MAM, each message contains only the root of the next Merkle tree, which means that at any given time, a subscriber can gain access to the data feed, without being able to decrypt the data of previous roots of the tree.

IV. RESULTS

In order to run the system, we have deployed a full node, which runs on a regular desktop computer with a dual core 3GHz processor and 4GB RAM. The system can process an average of 15 transactions/minute, which is sufficient for providing the necessary Proof-of-Work for the data provider.

The main advantage of this approach is that any type of data can be exchanged, with a high degree of trust and a relatively small time overhead (approximately 4 seconds/transaction with the current configuration).

The collected data is shown in Fig. 4 and 5 for a 1 day and 4 days interval. The results can be queried on the tangle using the tags attached to the transactions. It this way, we can retrieve data for any desired time interval or location. Data is provided as JSON strings, which can be used by any application.

For this implementation, the data is publicly available for anyone who has the tags or the address of the provider. Not all types of data should be publicly available however. For this purpose, the IOTA tangle can be used to implement a subscription model, where data can be available in exchange for network tokens or other forms of payment. This incentive can help run the data providers and allow users to subscribe only to the type of data they require, using MAM.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have implemented a data provider using blockchain technologies. The main purpose of the paper was to evaluate how these novel technologies can be used for real world applications other than cryptocurrencies.

The implementation shows that it is possible to create a data network on the IOTA tangle, which can be publicly available,

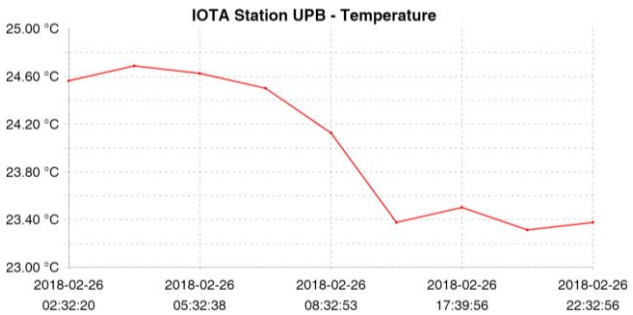

Figure 5. 1-day temperature

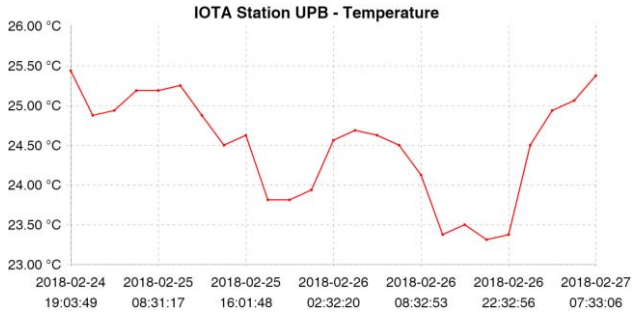

Figure 4. 4-days temperature

or accessed via permissions, using Masked Authenticated Messaging, and monetized using the network token.

The results show great potential to easily integrate remote sensor data, using a decentralized storage system, eliminating the need for complex databases and query models. The main limitation so far is that the field devices are not yet capable of doing POW, so running a system which requires a large number of data points can hit a bottleneck when the transactions are attached to the tangle. This issue can be addressed by different implementations of the signing algorithm, or by running a cluster of nodes with sufficient processing power to handle the required transaction volume.

In the future, we aim to expand the use of this technology and implement access controlled data providers, using different encryption techniques, such as Masked Authenticated Messaging.

REFERENCES

[1] Satoshi Nakamoto, “Bitcoin: A peer-to-Peer Electronic Cash”, https://bitcoin.org/bitcoin.pdf, 2009, [online], retrieved Feb. 2018

[2] L. D. Ibáñez, E. Simperl, F. Gandon and H. Story, "Redecentralizing the Web with Distributed Ledgers," in IEEE Intelligent Systems, vol. 32, no. 1, pp. 92-95, Jan.-Feb. 2017.

[3] R.C. Merkle, "Protocols for public key cryptosystems," In Proc. 1980 Symposium on Security and Privacy, IEEE Computer Society, pages 122-133, April 1980.

[4] R. Guo, H. Shi, Q. Zhao and D. Zheng, "Secure Attribute-Based Signature Scheme with Multiple Authorities for Blockchain in Electronic Health Records Systems," in IEEE Access, vol. PP, no. 99, pp. 1-1.

[5] K. J. O'Dwyer and D. Malone, "Bitcoin mining and its energy footprint," 25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014), Limerick, 2014, pp. 280-285.

[6] Bitcoin energy consumption index, http://digiconomist.net/bitcoin-energy-consumption [online], retrieved Feb. 2018.

[7] D. K. Tosh, S. Shetty, X. Liang, C. Kamhoua and L. Njilla, "Consensus protocols for blockchain-based data provenance: Challenges and opportunities," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York City, NY, 2017, pp. 469-474.

[8] Serguei Popov, “The Tangle”, https://iota.org/IOTA_Whitepaper.pdf, 2016, [online], retrieved Feb. 2018

[9] Ross, S. M., “Introduction to probability models”, 2010, Academic Press, ISBN: 978-0-12-407948-9