HKUST President's Cup Final Report

Vivian: Decentralized Global Naming and Storage System on Tangle

Written by:

TIAN Xiangan

Student ID: 20583620

Year of Study: 3

Supervised by:

Prof. TSOI, Yau Chat

Department of CSE

Department of Computer Science and Engineering

School of Engineering

Hong Kong University of Science and Technology

Vivian: Decentralized Global Naming and Storage System on Tangle

TIAN Xiangan
xtianae@connect.ust.hk

Supervisor: Prof. TSOI, Yau Chat
desmond@cse.ust.hk

***Abstract*—With the booming of distributed ledger technology (DLT) such as blockchain, many previous IT architectures can have alternative decentralized approaches for more secure, transparent, and immutable data storage. In this paper, we present the design and implementation of Vivian, a new decentralized global naming and storage system based on IOTA Tangle for re-decentralizing the current Internet service and building decentralized applications. Unlike the traditional Domain name System (DNS), trust points like DNS root servers are removed and critical data bindings are secured by the distributed ledger. All the nodes in the system form a peer-to-peer (P2P) network for data sharing. The P2P network is established through Kademlia DHT, mDNS peer discovery and eventually consistency of data is ensured by Gossip protocol. In this system, users can own their application data directly rather than relying on the central authorities. The system has no single point failure and the nodes in the network do not need to trust each other. By using IOTA Tangle, a directed-acyclic-graph (DAG) structure distributed ledger, the system inherits its scalable, lightweight, and feeless characteristics and enables the possibility of application in Internet-of-Thing (IoT) services.**

I. Introduction

A distributed ledger is a database which tolerates nodes with malicious intentions in a distributed manner. And distributed ledger technology (DLT) enables the realization and operation of distributed ledgers, which allows almost all the nodes in the network, to agree on an almost immutable record of transactions with Byzantine failure tolerance (BFT) and eventual consistency via a predefined consensus mechanism [1]. Blockchain is one of the most well-known DLTs which was first implemented on Bitcoin. It proposed a simple but robust way for transaction data storage without relying on trust of third parties [2]. Blockchain also ensures improved security and anonymity of Bitcoin transactions compared with traditional electronic transactions. Since the introduction of Bitcoin in 2009, DLT based cryptocurrencies have made a great impact on financial sectors. Later on people also discovered that the usefulness of DLTs is beyond exchange of currencies and significant adoption of DLTs were made in many other industries for other different services. Namecoin is the first altcoin[1] for being the first to create its own blockchain separate from Bitcoin's [3]. And its functionalities are not limited to financial transactions. The creation of Namecoin was inspired by the idea of BitDNS [4] and for establishing a decentralized domain name looking up system.

[1]Altcoin: any cryptocurrencies that are not Bitcoin.

The Internet today is a widespread information infrastructure and its history can date back to 1970s, when ARPANET[2] was developed. For most of the current Internet applications, data is stored in a centralized manner and users don't really own data by themselves. As shown in figure 1, if users want to do actions like checking their emails or browse the content of a website, they need to connect to the web servers via the Internet with web browsers, then the web servers retrieves the data from the database and then send it back to users. Usually users' data is hidden behind service providers' application code. This kind of arrangement has been very successful as it is easy to implement. However, it is not ideal since:

- Users must use requested web user interface if they want to access their data.
- The websites control the rules and access right of the data.
- The websites may snoop your data and sell users' information to others.
- Illegal use of data by websites' employees for personal purpose.

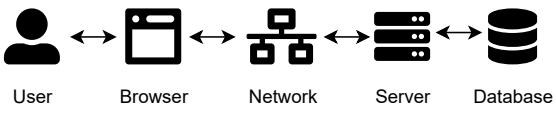


Fig. 1: User data arrangement traditional (centralized) Internet application

Since the early Internet, hosts in the network were assigned names for more convenient use and memorization by humans. With the growth of the network, it became impossible to store all the hosts in a single table. And Domain Name System (DNS) invented by Paul Mockapetris of USC/ISI permitted a scalable distributed mechanism for resolving hierarchical host names into Internet addresses [5]. The coordination and management of DNS Root, IP Addressing, and other Internet protocols are in the charge of IANA[3] [6]. These DNS root servers are central nodes of trust and failure, and cyber attack such as DDoS[4] may leads to the whole system taken down.

[2]ARPANET: Advanced Research Projects Agency Network. The first wide-area packet-switching network with distributed control originally established by United States Department of Defense.

[3]IANA: Internet Assigned Numbers Authority. Website: https://www.iana.org

[4]DDoS: Distributed Denial-of-service Attack. Usually the attack attempts to disrupt the normal traffic of the victim's server by a large amount of requests made by attacker devices.

It is reported that 13 root servers were under DDoS attack on March 21st, 2002. Fortunately the attack only lasted for one hour and didn't cause severe damage [7]. These central points may also be exploited and misleading users into connecting to malicious attack like the incident of Turkish fake site certs [8].

Internet-of-Things (IoT) refers to *"physical or virtual objects which connects to the Internet and has the ability to communicate with human users or other objects"* [9]. These devices such as smart webcam and wearable health monitors are widely used in our daily life. It is estimated that there will be approximately 30.9 billion active IoT device connections installed worldwide by 2021 [10]. Due to heterogeneity and complexity of IoT devices, their security and privacy issues are becoming more and more severe [9]. And with the increasing number of devices connected to the network, load of centralized servers for handling the connection will become much higher. DLT supported IoT has been created for addressing the challenges like security, data integrity and reliability, and secured P2P sharing. It is a new decentralized and distributed solution to IoT services and enables the opportunity for developing new and creative applications and business models in vertical domains, e.g., from healthcare to supply chain, energy industry, and smart manufacturing [11].

Motivation. Many data management issues like security, integrity, access control has been exposed from centralized data model of traditional Internet. When accessing the web services, user data control are maintained by service vendors rather than users themselves. Domain Name System containing central nodes like DNS root servers are vulnerable to cyber attacks such DDoS. Distributed ledger technology such as blockchain can enhance security and data integrity of IoT services. However, many of the current DLTs are based on Proof-of-Work (PoW) consensus mechanism [12], which requires very strong computing power and large energy consumption for solving hash computational puzzles. These mechanisms are not suitable for IoT devices that have poor computational power and strict energy consumption limitation. Transaction fees paid to the miners in the network caused an extra cost for the service. DLTs like Bitcoin blockchain are also facing problems like low TPS[5], bad scalability, etc. They are not suitable for IoT service scenarios like sending large amount of micro-transactions in a short period of time. We wish to re-decentralize the current Internet service via distributed ledger technology for better security and data integrity. We also require the DLT used should be feeless, lightweight, and scalable which can support the use cases of IoT services.

Contribution. We introduce the design and implementation of Vivian, a global naming and storage system secured by IOTA Tangle distributed ledger [13]. It is a new decentralized Public Key infrastructure (PKI) system that enables users to register human-readable and unique domain name with binding information. By using IOTA Tangle DLT, no central trust points are needed and users can control their own data. Peer-to-peer network based on Kademlia DHT and mDNS peer discovery, and Gossip protocol ensures secure data sharing among nodes in the network. IOTA Tangle is a directed-acyclic-graph (DAG)[6] based distributed ledger which performs better scalability than traditional blockchains. There is no miner involved in the network so no transaction fee is needed. The whole system is lightweight and enables the possibility of building decentralized IoT applications.

[5]TPS: Transaction Per Second. The approximate average TPS of Bitcoin blockchain is around 5.

II. Background

A. Blockchain and Tangle (DAG)

Blockchain DLT become widely known in 2009 with the launch of Bitcoin network. Participants in the network can validate and verify the transactions independently, without relying on central trust parties. Blockchain is usually maintained and managed by a distributed group of participants independently. This along with its cryptographic mechanisms ensures the data recorded on the ledger immutable [14]. The structure of a traditional blockchain can be simplified as a singly linked list[7], you can traverse from the latest block to the Genesis block[8] (as shown in figure 2). Transactions are hashed in a Merkle Tree [15] for saving storage space and simplifying transaction validation.

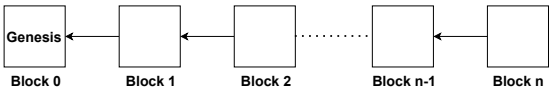


Fig. 2: Blockchain structure

In the original Bitcoin whitepaper [2], Satoshi Nakamoto has listed the procedures for handling the transactions and record them on the blockchain ledger. All the nodes listen to the transactions broadcast to the network and each node collects new transactions for generating a new block. Then each node will do proof-of-work and broadcast the block to the network once it finds the solution for other nodes' validation. However, in real case, almost all the transaction wrapping and PoW are finished by specific miners in the network. Miners also require transaction fees for the reward of doing these. The increasing of PoW complexity makes it nearly impossible to gain profit from mining Bitcoin with general PC hardware. Instead, dedicated miners have to use equipments like ASIC[9] devoted specifically for the mining algorithm. It is a potential hazard that the blockchain network

[6]Directed-acyclic-graph (DAG): a directed graph with no directed cycles. It is a directed graph, which means each edge has an orientation, from one vertex to another. However, no path in the graph forms a circle.

[7]Singly linked list: linked list which is unidirectional and can only be traversed on one direction.

[8]Genesis block: the first block of a blockchain. It is a special case as it does not reference a previous block.

[9]ASIC: application-specific integrated circuit. It is the integrated circuit designed for a specific use case. Bitcoin ASIC chip can only handle computing tasks for Bitcoin mining, and cannot be used for any other tasks.

will be centralized and controlled by the parties that owns most of the "mining rigs". In addition, blockchains relying on PoW are consuming massive energy [16]. According to CBECI[10], Bitcoin network consumes approximately 130.51 TWh electricity per year, which is far more than the annual electricity consumption of some countries like Ukraine and Argentina. Carbon dioxide emission caused by these PoW blockchains will cause environmental issues like global warming.

Transaction throughput and transaction confirmation latency are two most critical performance issues about blockchain technology [17]. Transactions can only be recorded on the blockchain in sequence due to its linear structure. Also the limitation of size of each block makes it struggling to handle the enormous volume of transactions nowadays. Blockchains including Bitcoin and Ethereum are facing problems like low TPS and bad scalability which results in transaction backlog and high transaction fees. In order to tackle these issues, people have put forward many alternative solutions such as side-chain, cross-chain, improved consensus, sharding, DAG, etc. IOTA Tangle is a new type of DLT addressing solving the problems above for IoT services. It has following advantages comparing with traditional blockchain technologies:

1) **High scalability**. Directed-acyclic-graph structured Tangle ledger enables its high scalability. Serguei Popov has analyze the performances of the system under two different regimes: low load and high load in the article *The Tangle* [13]. In high load regime when more new tips[11] are attached to Tangle, the typical time of a tip being approved is reduced. So the larger the scale of the network, the more efficient it will be.
2) **Feeless & Environmental Friendly**. In DLT network like Bitcoin, we need to pay transaction fees to the miners for rewarding them wrapping our transactions to the block and conducting PoW computations. Transaction fees are considered part of the incentive for nodes to support the network. In IOTA Tangle, however, PoW consensus and miners are removed [18]. So it is more economical and environmental Friendly to use Tangle for sending transactions.
3) **Quantum Computation Resistance**. IOTA Tangle uses post-quantum cryptography for securing data on the ledger [19]. For instance, IOTA uses Winternitz One-Time Signature (WOTS), which is promising to be resistant to quantum computers [20]. as a signature scheme protocol.
4) **Lightweight**. IOTA node applications like GoHornet[12] is lightweight and can be easily installed and run on low-end devices such as Raspberry Pi 4.

These characteristics are very crucial for IoT applications and favorable for building this project.

[10]CBECI: Cambridge Bitcoin Electricity Consumption Index. Website: https://cbeci.org/

[11]Tip: every new (unconfirmed) transaction is known as a tip.

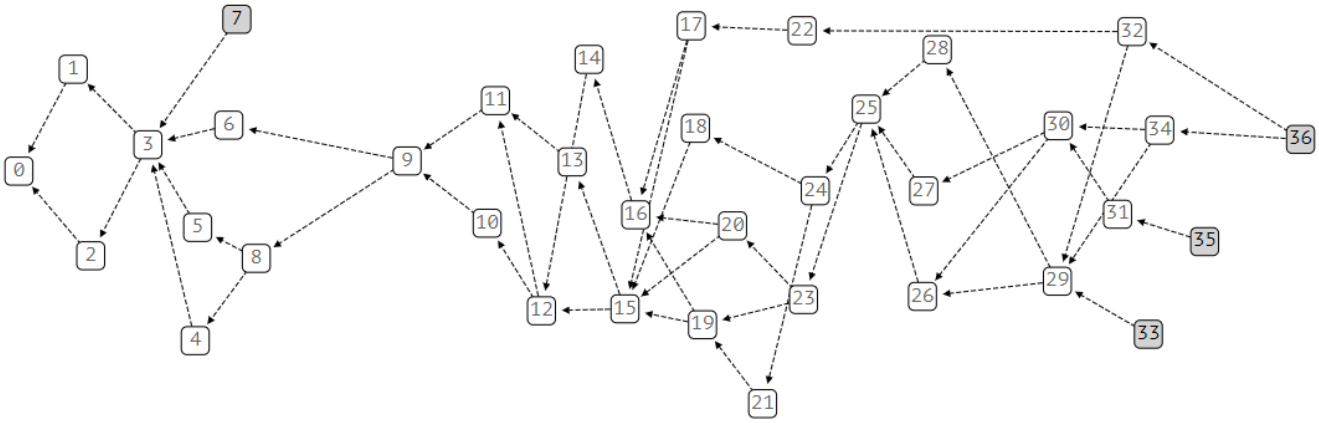
[12]GoHornet: IOTA full node software built in Go. Github source code: https://github.com/gohornet/hornet

B. Decentralized Naming System

Before distributed ledger technology came into being,

C. P2P Network

Fig. 3: Tangle structure



REFERENCES

[1] A. Sunyaev, *Distributed Ledger Technology*, pp. 265–299. Cham: Springer International Publishing, 2020.

[2] S. Nakamoto and A. Bitcoin, "A peer-to-peer electronic cash system," *Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf*, vol. 4, 2008.

[3] H. A. Kalodner, M. Carlsten, P. Ellenbogen, J. Bonneau, and A. Narayanan, "An empirical study of namecoin and lessons for decentralized namespace design.," in *WEIS*, Citeseer, 2015.

[4] F2b, "BitDNS and Generalizing Bitcoin." https://bitcointalk.org/index.php?topic=1790.0/, 2010. [Online; accessed 25-Feb-2021].

[5] B. M. Leiner, V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. Wolff, "A brief history of the internet," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 5, pp. 22–31, 2009.

[6] J. Postel, "Domain name system structure and delegation," *RFC*, vol. 1591, 1994.

[7] D. McGuire and B. Krebs, "Attack on internet called largest ever," *The Washington Post*, vol. 22, 2002.

[8] S. Rosenblatt, "Fake turkish site certs create threat of bogus google sites." https://www.cnet.com/news/fake-turkish-site-certs-create-threat-of-bogus-google-sites/, Jan 2013. [Online; accessed 27-Feb-2021].

[9] Z. Zhang, M. C. Y. Cho, C. Wang, C. Hsu, C. Chen, and S. Shieh, "Iot security: Ongoing challenges and research opportunities," in *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, pp. 230–234, 2014.

[15] R. C. Merkle, "Protocols for public key cryptosystems," in *1980 IEEE Symposium on Security and Privacy*, pp. 122–122, IEEE, 1980.

[10] S. R. Department, "Internet of things - active connections worldwide 2015-2025." https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/, Jan 2021. [Online; accessed 28-Feb-2021].

[11] B. Farahani, F. Firouzi, and M. Lücking, "The convergence of iot and distributed ledger technologies (dlt): Opportunities, challenges, and solutions," *Journal of Network and Computer Applications*, p. 102936, 2020.

[12] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, "On the security and performance of proof of work blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, (New York, NY, USA), p. 3–16, Association for Computing Machinery, 2016.

[13] S. Popov, "The tangle," *White paper*, vol. 1, p. 3, 2018.

[14] D. Yaga, P. Mell, N. Roby, and K. Scarfone, "Blockchain technology overview," *arXiv: Cryptography and Security*, 2018.

[16] J. Sedlmeir, H. U. Buhl, G. Fridgen, and R. Keller, "The energy consumption of blockchain technology: beyond myth," *Business & Information Systems Engineering*, vol. 62, no. 6, pp. 599–608, 2020.

[17] Q. Zhou, H. Huang, Z. Zheng, and J. Bian, "Solutions to scalability of blockchain: A survey," *IEEE Access*, vol. 8, pp. 16440–16455, 2020.

[18] S. Popov and Q. Lu, "Iota: feeless and free," *IEEE Blockchain Technical Briefs*, 2019.

[19] L. Tennant, "Improving the anonymity of the iota cryptocurrency," 2017.

[20] J. Buchmann and J. Ding, *Post-Quantum Cryptography: Second International Workshop, PQCrypto 2008 Cincinnati, OH, USA October 17-19, 2008 Proceedings*, vol. 5299. Springer Science & Business Media, 2008.