White Paper

Quantum Hardware and Decentralized Components

anaxa.tech

Anaxa

Introduction

The Internet currently is a distributed network that consists of billions of different nodes. Though the Internet itself is not centralized, most of the data available online is stored and controlled by a few technology companies that are able to build these massive data centers that can handle a vast amount of information. The convenience and affordability of the Internet has provided many the opportunity of leveraging different platforms across the network that allows for global communication and interaction. There is a continually increasing number of devices within the Internet of Things (IoT) in both the individual and industry levels, with a total of 9.5 billion devices connected to the Internet as of 2019 [1]. The industry is only increasing, and is expected to have 24 billion devices by 2030 at a compound annual growth rate of 11% per year [2].

However, as a cost of this provided by the Internet, with the growing number of devices adding to our network, rising security hacks and latency delays are both major concerns. As of 2018, there are 80,000 cyberattacks being made per day or over 30 million attacks per year [3]. Ransomware attacks are growing more than 350% annually. In the healthcare industry itself, over 40 million records get stolen or lost each year, resulting in the industry losing over $6.2 billion per year.

Each data breach can account for more than $3.7 million.The use of decentralized storage and quantum key distribution hardware has risen as an answer to the challenge of providing secure, private and trusted cloud storage. With decentralized storage, the process of protecting data makes data breaches more difficult than current methods used by data centers. The technological use of both decentralized cloud storage and quantum key distribution are both rapidly increasing but are met with many constraints and parameters including speed, capacity, cost, bandwidth and latency.

We propose a framework that is able to scale data storage across the globe with a system that is able to encrypt and distribute data to different hosting nodes around the world. The component of quantum hardware is also added to the system in order to make the transfer of information seemingly instantaneous. With these components added together, data, specifically geared towards health data, can be stored and served in a manner purposefully designed to prevent any breaches.

Quantum Key Distribution

Quantum Key Distribution (QKD) is a technology that relies on the fundamental aspect of quantum physics (physics of particles like photons, electrons and other molecules). They rely on the fact that the behavior of quantum particles in most situations is random, and if used to generate a bit stream of 0's and 1's, it results in a set of bits that are completely random.

These random bits are ideal for the use as a cryptographic key in order to protect data. According to the No Cloning Theorem, quantum particles are also unable to be duplicated. In a standard QKD system, single particles of light (also known as photons) are generated in random states and are able to be transmitted between two separate locations.

However, due to the extremely fragile nature of quantum particles, as well as the inefficiencies of the transmission medium (fiber), only a small fraction of the photons are able to be received at the far end.

This connection that these two sides have is known as “entanglement” where data in the two common ends of the fiber are able to be traded publicly.

They share a common set of data containing a stream of random bits that are suitable for use as an encryption key. If a third party attempts to intercept the photons traveling in the QKD system, the eavesdropper must measure some property of the photon, which automatically destroys the photon in the process.

Now, the eavesdropper must try to recreate the photon with the same property as what was measured by guessing which one to measure and recreate based on the two ends of the system that used different properties to encode the information between them. Since the third party can only measure a single property because the photon is destroyed, the third party would automatically guess wrong and recreate the wrong property.

The endpoints can then automatically determine an error since the transmitted and received values for the number of bits are different. Once the third party is detected, the bits from both sides are automatically discarded. In addition, because of this “entanglement” property, which forms the connection between the two endpoints, the sending and receiving of data would be instantaneous with zero latency in between.

Decentralized Storage based on Blockchain Framework

With the component of decentralized storage, files are stored on multiple computers (nodes on a decentralized network. Similar to today’s conventional cloud storage, when a file is needed, the user is able to request and receive the file by downloading different fragments of that files from participants in the network until the user has the full file. The system automatically encrypts files and only the specific user holds the encryption key guaranteeing that the specific user’s files can only be read by them. No single person holding the files has the entirety of it, which automatically adds an extra layer of security and protection.

Our framework will enable users to complete the following actions:

.

Store data - the data stored within the network is encrypted and broken up into multiple, smaller pieces that are then distributed to peers across the network. Metadata is generated, containing the information on where to retrieve the data
Retrieve data - the user is able to reference the metadata in order to identify the specific locations of the stored pieces. The pieces are retrieved and the original data will be reassembled locally.
Pay for usage - the unit of value is sent in exchange for the services Within the decentralized storage network lies storage nodes. The storage node's role is to store and return data. Users within the system are able to choose to "host" ie. their device becomes a storage node based on criteria such as throughput bandwidth caps, sufficient disk space, geographic location and uptime.

For peer-to-peer communication, all peers within the network must communicate via a standardized protocol provided by the quantum hardware (BB84 protocol) [4]. This specific protocol is a quantum cryptography protocol that relies on entanglement and QKD where information gain is not possible (due to the no cloning theorem) [5].

This protocol allows for 1) complete privacy in communication by default where the user and storage node are able to communicate without any risk of eavesdroppers 2) provides peer reachability even in the face of firewalls and NATs and 3) provides authentication with each user cryptographically proving the identity of the peer they are speaking to in avoiding man-in-the-middle attacks.

With the metadata is a result of the file being split up, nodes for where the new pieces would be stored are selected. Each user is able to choose a node based on location, performance characteristics and available space. Each user can also choose to host their own node for other users to store pieces of data on if they fit requirements based on bandwidth, disk space, etc. An explicit node selection scheme such as directory-based lookups [6] must be used. Users must also be able to choose an arbitrary key (similar to a path) in order to identify the mapping of data pieces of the node and to allow for storing/receiving of the data.

In terms of encryption, all data and metadata will be encrypted in order to provide total security and privacy. Data is encrypted before leaving the source computer, in the beginning stages of the data storage pipeline. The encryption uses a pluggable mechanism that enables users to choose their desired encryption scheme while also storing data about the specific encryption scheme that enables the user to recover their data using an appropriate decryption mechanism.

Each file is encrypted with a unique key, which would prevent the access of decryption keys in all files while being able to access certain selected files at the same time. The metadata and the information about its paths will be stored and encrypted by a deterministic hierarchical encryption scheme [7]. This scheme will allow for subtrees to be shared without sharing their parents and will allow some files to be shared without sharing other files.

In terms of data repair or damage, probabilistic per file audits or “proofs of retrievability” is used as a way to determine when and where to repair the files [8]. Audits confirm that a storage node is able to keep the data it claims and is not susceptible to any hardware failure by functioning as spot checks in order to calculate the usefulness of the given storage node [9]. Failed audits based on the storage node uptime and overall health result in the storage node being marked which will automatically redistribute the data to new nodes.

In the case of data loss due to storage node churn or software errors, the storage node that stops storing data is detected and the original data is recovered from an erasure code reconstruction [10]. This reconstruction process regenerates the missing pieces then stores them back in the network on new storage nodes.

References

[1] Lueth, Knud. “IoT 2019 in Review: The 10 Most Relevant IoT Developments of the Year.” IoT Analytics, 7 Jan. 2020, iot-analytics.com/iot-2019-in-review/.

[2] D'mello , Anasia. “Global IoT Market.” IoT Now - How to Run an IoT Enabled Business, 20 May 2020, www.iot-now.com/2020/05/20/102937-global-iot-market-to-grow-to-1-5trn-annual-revenue-by-2030/.

[3] PurpleSec LLC. “2019 Cyber Security Statistics Trends & Data.” PurpleSec, 19 Sept. 2020, purplesec.us/resources/cyber-security-statistics/.

[4] B. Archana and S. Krithika, "Implementation of BB84 quantum key distribution using OptSim," 2015 2nd International Conference on Electronics and Communication Systems (ICECS), Coimbatore, 2015, pp. 457-460.doi: 10.1109/ECS.2015.7124946

[5] Fan, Heng & Wang, Yi-Nan & Jing, Li & Yue, Jie-Dong & Shi, Han-Duo & Zhang, Yong-Liang & Mu, Liang-Zhu. (2013). Quantum Cloning Machines and the Applications. Physics Reports. 544. 10.1016/j.physrep.2014.06.004.

[6] J. Paiva and L. Rodrigues. Policies for Efficient Data Replication in P2P Systems. In 2013 International Conference on Parallel and Distributed Systems, pages 404–411, Dec 2013.

[7] Peter Wuille. BIP32: Hierarchical Deterministic Wallets. https://github.com/bitcoin/bips/blob/master/bip-0032.mediawiki, 2012.

[8] Ari Juels and Burton S. Kaliski, Jr. PORs: Proofs of Retrievability for Large Files. In Proceedings of the 14th ACM Conference on Computer and Communications Security, CCS ’07, pages 584–597, New York, NY, USA, 2007. ACM.

[9] Hovav Shacham and Brent Waters. Compact proofs of retrievability. Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology, ASIACRYPT ’08, pages 90–107, Berlin, Heidelberg, 2008. Springer-Verlag.

[10] Kevin D. Bowers, Ari Juels, and Alina Oprea. Proofs of retrievability: Theory and implementation. In Proceedings of the 2009 ACM Workshop on Cloud Computing Security, CCSW ’09, pages 43–54, New York, NY, USA, 2009. ACM.