Streembit: decentralized P2P communication system for humans and machines

T.Z. Pardi
Email: tzpardi@streembit.com
Github: zsoltpardi

Contents

Abstract

Streembit is a decentralized, peer to peer (P2P), permissionless, real time communication system for humans and machines. The application aims to implement a system that securely manages humans to machine and machine to machine (M2M) communication without using a central server or client-server infrastructure. The actors of the system - both human users and Internet-of-Things devices – are the nodes of the Streembit peer-to-peer network. The system uses a distributed hash table (DHT) for contact discovery services. The system ensures data integrity using public-private key cryptography. Messages are signed with the contact's private key on the client side using ECDSA cryptography. The data is encrypted using 256-bit AES symmetric encryption. The symmetric keys are exchanged between contacts using ECDH key exchange to exchange. The video and audio communication between contacts uses the WebRTC protocol that allows end-to-end encrypted video and audio conversation in a true P2P manner. For all use cases the system end-to-end encrypts the communication between the peers without routing the conversation through any central server.

Background

Internet-of-Things refers to the network of uniquely identifiable embedded hardware devices accessed through the Internet infrastructure. Internet-of-Things devices are normally semi-autonomous, wireless devices participating in internetworked communications using the familiar and established TCP/IP protocol stack. Typical applications are sensory equipment which, without the need for extra wiring, can collect and relay data to a central host - either on-premise or in the increasingly popular cloud environment. The interconnection of these embedded devices is expected to usher in a new era of automation in nearly all fields. It enables advanced applications such as smart grids, remote surveillance systems, wireless heart monitors and many more.

According to Gartner, there will be nearly 26 billion devices connected to the Internet-of-Things by 2020. ABI Research estimates that more than 30 billion devices will be wirelessly connected to the Internet-of-Things by 2020. Per a recent survey and study done by Pew Research Internet Project, a large majority of the technology experts and engaged Internet users who responded—83 percent—agreed with the notion that the Internet-of-Things and embedded and wearable computing will have widespread and beneficial effects by 2025.

We focus on the following areas by implementing the Streembit system:

Security

Human-to-human communication and Internet-of-Things devices need to have some form of connectivity, resulting in significant security issues that businesses and residential users need to consider. The connected devices require a security protocol and its security policy should be in line with the security policy of the enterprise. These Internet connected devices can easily expose businesses and homes to various security threats. In the meantime there is no common, standardized, robust and proven authorization and access control protocol in existence for Internet-of-Things devices. Of the numerous service providers rolling out devices and services, almost all of them implement their own security protocol, API and infrastructure. Consequently the development of custom security implementations increases the price of the product and the lack of standards increases the security risks. The standard way would be to secure the communication and manage Internet-of-Things nodes from an authorization perspective using the robust, widely adopted and well tested public/private key infrastructure (PPKI). To ensure confidentiality and secure communications, the core part of security should ideally be based on PPKI paired with some kind of PPKI certificate management. Existing systems tend to use custom authorization and access control schemes with domain specific login portals for user name/password based logins instead of using robust PPKI based security. Using the public/private key infrastructure would pave the way for the deployment of a robust and secure authentication and access control scheme. The parties are identified by their public key. The authenticity of messages and the identities of the actors are then verified using PPKI cryptography routines. Using the well-established public/private key infrastructure based security scheme would greatly simplify the authentication, access control and identity management aspects of Internet-of-Things security.

Data Control

From a user perspective, this is one of the more significant barriers to the large-scale adoption of Internet-of-Things. Data control is commonly mistaken for data ownership. In a conventional computing system the issue was who owns the data. In Internet-of-Things the challenge is about deciding who gets access to the data. Enabling access to private data is a serious concern from a privacy standpoint.

Standards

There is almost as much software systems as there are Internet-of-Things devices. The lack of open standards are directly affecting the adoption rate of Internet-of-Things devices as well having a negative impact on the user experience. Well-documented and robust APIs by providers could be a step towards open standards, but so far no such API usable by Internet of Things devices exists. Sensors use custom APIs instead. Domain specific product development such as this dramatically slows down innovation as development resources are allocated to custom software development. This once again increases the price of the product and consequently lowers the adoption rate of these modern devices.

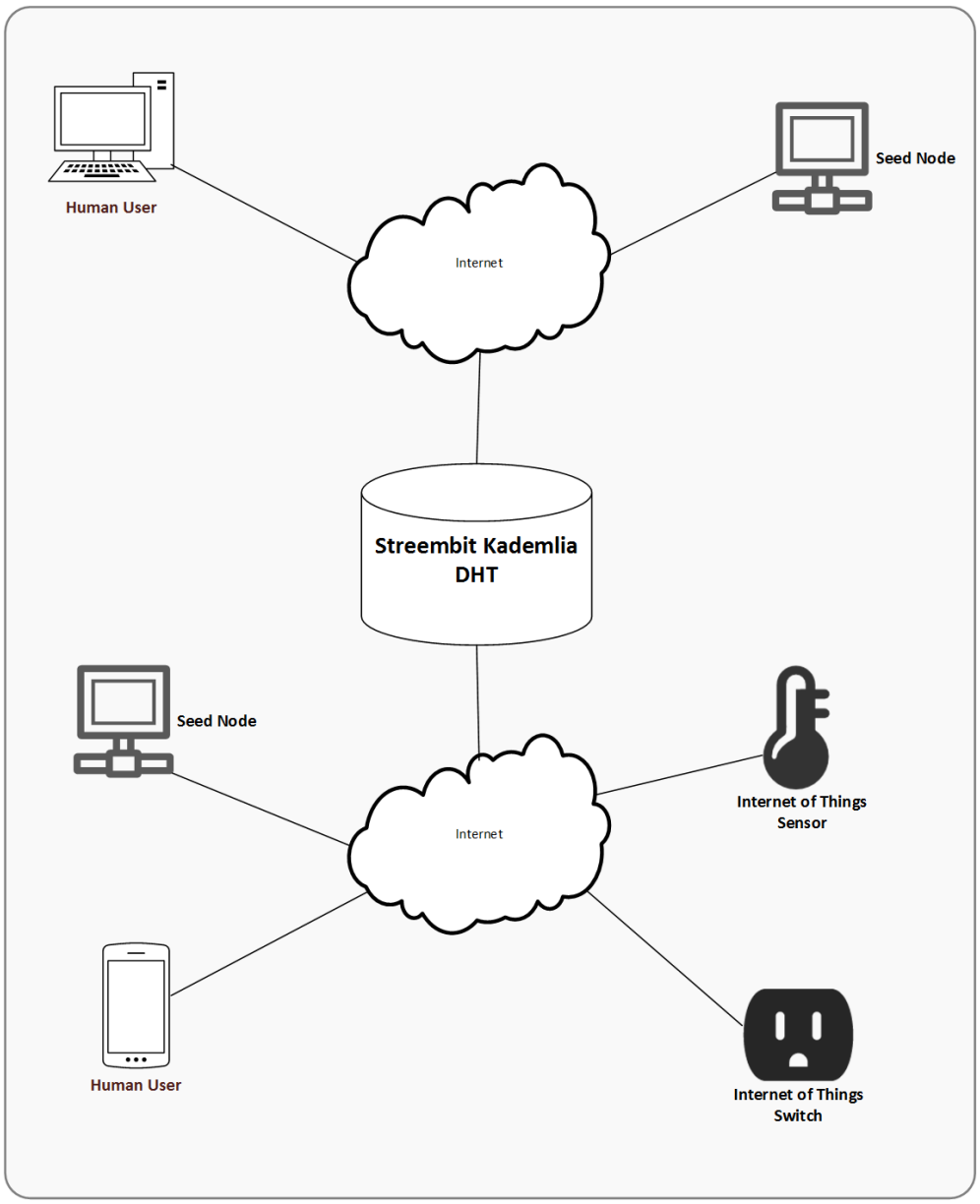
Decentralised, block-chain based technology

Block-chain is a transaction database and ledger shared by all nodes that are normally participating in a digital currency system. A full copy of a currency's block-chain contains every transaction ever executed regarding to the currency. The block-chain can also perform distributed contracts. Such contract management is a method to form agreements between parties via the block-chain.
IDC recently concluded that block-chain technologies could be key tools for confirming data origin and accuracy, tracking updates and establishing true data authority for millions of different data fields. The block-chain is a solid model for establishing an audit trail, in addition to transferring and monitoring distinct entities that represent items of value. As a result, block-chain has the potential to serve as a foundation for improving the authenticity and accuracy of business and government records. A block-chain based system can track activities via a shared record that’s resistant to hacking and unauthorized changes. Once this shared version of the "truth" is established, via a peer-to-peer network, multiple nodes ensure the integrity remains intact even as new records are added.

The Streembit System

The system forms a decentralized peer to peer overlay network to manage connections between human and Internet of Things peers. The participants in the network are the peer nodes. The peer to peer network is scalable and an unlimited number of nodes can participate in the network.

The system performs public private key infrastructure based authentication and access control functions to securely connect peer nodes. The application generates at least one private/public key pair on each peer node including on the Internet-of-Things devices. The actors of the system publish their public keys to the peer to peer network via a Kademlia distributed hash table (DHT). Each peer node knows the public key of the other connected peer nodes. The system identifies peer nodes in the peer to peer network by their public key. To ensure data integrity the nodes sign the messages with their private key. The nodes sign all messages - there are no unsigned messages circulated in the system. (The requirement for signing the messages also help to mitigate the risks of Sybil attacks and DDoS attacks).



Streembit supports two types of network implementations, public and private Streembit networks. The main purpose of both network is the facilitate contact information exchange. The difference between them is the accessibility of the networks.

Public network. Anyone can connect to the public Streembit network. All valid, signed, cryptographically verified messages from any node are registered in the DHT. A node that is connected to the network can receive any messages from any other node. As the main purpose of the network type is to facilitate node discovery, humans and machines can find each other on the network in a decentralized manner, without using a corporate owned centralized system. Nodes publish their availability and network information such as IP address and port information in encrypted form using their contacts' ECDH public key. The DHT acts as a ledger to exchange contact information. The participating nodes of the network are the workers who route the information to any new connecting nodes upon request. The main Streembit network is public.

Private network, private hub. Decentralization addresses many infrastructure issues such as high availability and scalability. A decentralized system can achieve that by forming a collaborative network from the participating nodes. The democratic, libertarian concept of decentralized computing assumes that the nodes contribute to the network by routing messages to other nodes. The participating nodes keep alive the decentralized network by interacting with other nodes. A decentralized network provides users with many benefits, but at the same time expects the message routing contribution from the nodes. The issue is, potentially there are millions of nodes in a large decentralized network. On the other hand, a limitless interaction with any node is not allowed id many use cases. For example, it is not optimal nor practical from resources viewpoint if an Internet of Things enabled garage door controller of a family home acts as a full node on the Streembit network. A full node by definition routes any messages to any connected nodes. At the same time, a low power garage door controller device that is dedicated to one task – i.e. open the garage door upon the request of an authorized user – should not perform such message routing. An IoT gateway of a family home or a communication network of a business should exclude actors from the communication that aren't part of their particular use case. Therefore, we have introduced the private network on the DHT concept in Streembit. The nodes of a private Streembit network uses the public Streembit network for a one-time information sharing: they publish their encrypted IP address or multicast DNS name, and then immediately disconnect from the public network to wait for the connection of their private node partners via the private network. Terms of authentication and access controls schemes a private Streembit network is isolated from the public Streembit network. That means only certain peer nodes are allowed to connect to the private Streembit network. A private Streembit network is functioning as a firewall: only preconfigured contacts allowed to connect. The collection of the preconfigured contacts is maintained in a lookup list. The connection of devices and users which are not in the preconfigured list is refused by the private network. The access control is governed by built in authentication functions. Typically, IoT devices within a building, teams, communities or businesses would run a private Streembit hub to establish an even more secure communication mechanism than the public Streembit network does.

Streembit Kademlia DHT

Kademlia is a distributed hash table (DHT) for decentralized peer-to-peer computer networks. It specifies the structure of the network and the exchange of information through node lookups. Kademlia nodes communicate among themselves using UDP or TCP. Streembit primarily uses TCP. A virtual or overlay network is formed by the participant nodes. Each node is identified by a number or node ID. The node ID serves not only as identification, but the Kademlia algorithm uses the node ID to locate values. The node ID provides a direct map to message hashes and that node stores information on where to obtain the file or resource. The node ID on the Streembit network is the composite of the public key and account name. We selected this implementation as using public/private key cryptography allows a simple and at the same time robust authentication and access control (i.e. to identifying the node by verifying the public key based signature on messages). Having the public key validation integrated into the DHT layer allows the filtering of malicious nodes prior the messages reach the application layer.
We extended the generic Kademlia protocol and Streembit adds an extra security layer to the DHT to validate and authenticate messages using the ECC public key of users and devices.

The Streembit distributed has table primary role is to manage user/device discovery on the Streembit network. One of the main problems in IoT device management is how to share the characteristics of a device, like serial number, manufacturer and model with other devices or human users. Sharing device information is required for installation, data retrieval, device management, and device control. Streembit is a network that provides a means to safely install, configure, find and connect massive amounts of IoT devices together, and at the same time minimizing the risk that devices get hijacked. In order to achieve this the Streembit network implements distributed registry that allows simple access to private and public devices without risking their integrity.

Security

The cornerstone of Streembit security is elliptic curve public/private key cryptography infrastructure (PPKI). PPKI allows the implementation of robust security. We use PPKI to identify the entities of the system (based on their public key and PPKI signature), perform authentication, and ensure data integrity (using cryptography signatures).
Using the public/private key cryptography infrastructure paves the way for the deployment of a robust, secure authentication and access control scheme. The parties are identified by their public key. The authenticity of messages and the identities of the actors are then verified using PPKI cryptography routines. Using a PPKI infrastructure based security scheme greatly simplifies the authentication, access control, and identity management aspects of Internet-of-Things security.
The elliptic curve cryptography (ECC) scheme is particularly suitable for IoT devices. The small footprint of ECC allows security modules to be implemented on embedded devices. ECDH is a trusted and proven key agreement protocol. Using the ECDSA digital signature algorithm ensures data integrity.
The system performs public/private key infrastructure based authentication and access control functions to securely connect machines to machines or machines to humans. The Streembit security module of the IoT device must generate at least one public/private key pair for each connected IoT device. Human users generate their ECC public/private key pair when they create their account. The actors of the system publish their public keys to the network, where each entity knows the public key of the other connected human or machine. The system identifies each of the entities by their ECC public key.

The collision resistant SHA-256 hash function is used to create a hash of data which can be signed using the private key to guarantee data integrity, as well as provide information about the originator of the data. To ensure data integrity the messages, control commands, event data, requests, and responses are signed with the ECC private key; the signature can be verified using ECDSA. To ensure integrity of data, each entity must sign all messages with their private key. The messages in the Streembit network are based on the following standards: JSON Web Token (JWT), JSON Web Encryption (JWE), and JSON Web Signature (JWS).

The basic premises of the Streembit security are

- Human users and Internet of Things devices use public/private key (PPK) infrastructure and PPK cryptography functions to secure messages
- Each actor of the system must generate a public/private key pair. (Typically keys are generated prior to configuring the device and will be burned into the devices' firmware).
- The device or user publishes the public key to other users of the system.
- The data integrity and authenticity of the messages is guaranteed with PPK signatures.
- Each session between users is secured with strong symmetric cryptography keys.
- All messages between users are secured with 128-bit and 256-bit AES symmetric encryption/decryption.
- The system uses elliptic curve Diffie Hellman (ECDH) key exchange algorithms to facilitate the exchange of session keys.

Protecting the ECC public/private key pair

The ultimate issue of all applications, systems and devices that encrypt data is the protection of the cryptography symmetric key or protection of the private key of the public/private key pair. Similarly to the iPhone and PGP, which both use passcode to protect their keys, Streembit requires users have a passcode to protect their PPKI private key. The Streembit passcode that protects the user's PPK key pair is relatively complex and it requires the following minimum number of iterations with a brute force attack:

$$26^2 \times 10 \times 33 \times 95^4 \times \frac{8!}{4!} \approx 3.05 \text{ x } 10^{16}$$

That means 3,050,000,000,000 (3.05 trillion) times more iterations than the 10^4 iterations is required to break the latest iPhone security. (Of course if the passcode is longer than 8-digit then the number of iterations is even more). Please note, this is related to protecting the PPKI key pair. To protect real time communication, Streembit uses a randomly generated 256-bit AES session key and exchanged between peers with a randomly generated ECDH key

pair. This means that real time communication, which is the main functionality of Streembit, is secure even from the brute force attacks of NSA super computers and large zombie computer clusters of cyber criminals.

Mitigate the risk of Sybil attacks

A known issue with regards to peer-to-peer (P2P) networks and their practical limitation is that they are frequently subject to Sybil attacks. This means that malicious parties can compromise the network by generating and controlling large numbers of shadow identities. A malicious node may present multiple fake identities to a peer-to-peer network in order to appear and function as several distinct nodes.
A Sybil attack is most effective on anonymous and reputation systems where if the malicious nodes outnumber the honest nodes the outcome of the application could be compromised. For example, on a file sharing network a successful horizontal Sybil attack allows the attacker to sniff most of the control messages, hijack the system, and deliver bogus content. Contrarily, Streembit facilitates communication between contacts that you know, such as your family members, teammates, and business partners. You are also aware of the location and identity of your Internet of Things device(s) you control via Streembit, lessening the likelihood of a Sybil attack. Malicious nodes, regardless of their weight and presence on the network, are unable to change the public keys of your contacts (as the public key is based on the cryptographically secured PPK infrastructure).
Additionally, Streembit introduces another layer of security, private networks. Using private networks, the public keys of the contacts are loaded to the seed nodes manually; mitigating further the risk of a Sybil attack as well as a Man-in-the-middle attack (MITM). A Sybil attack, which generally speaking is a serious security issue for P2P networks such as file sharing applications, is much less of a problem for the Streembit P2P system.

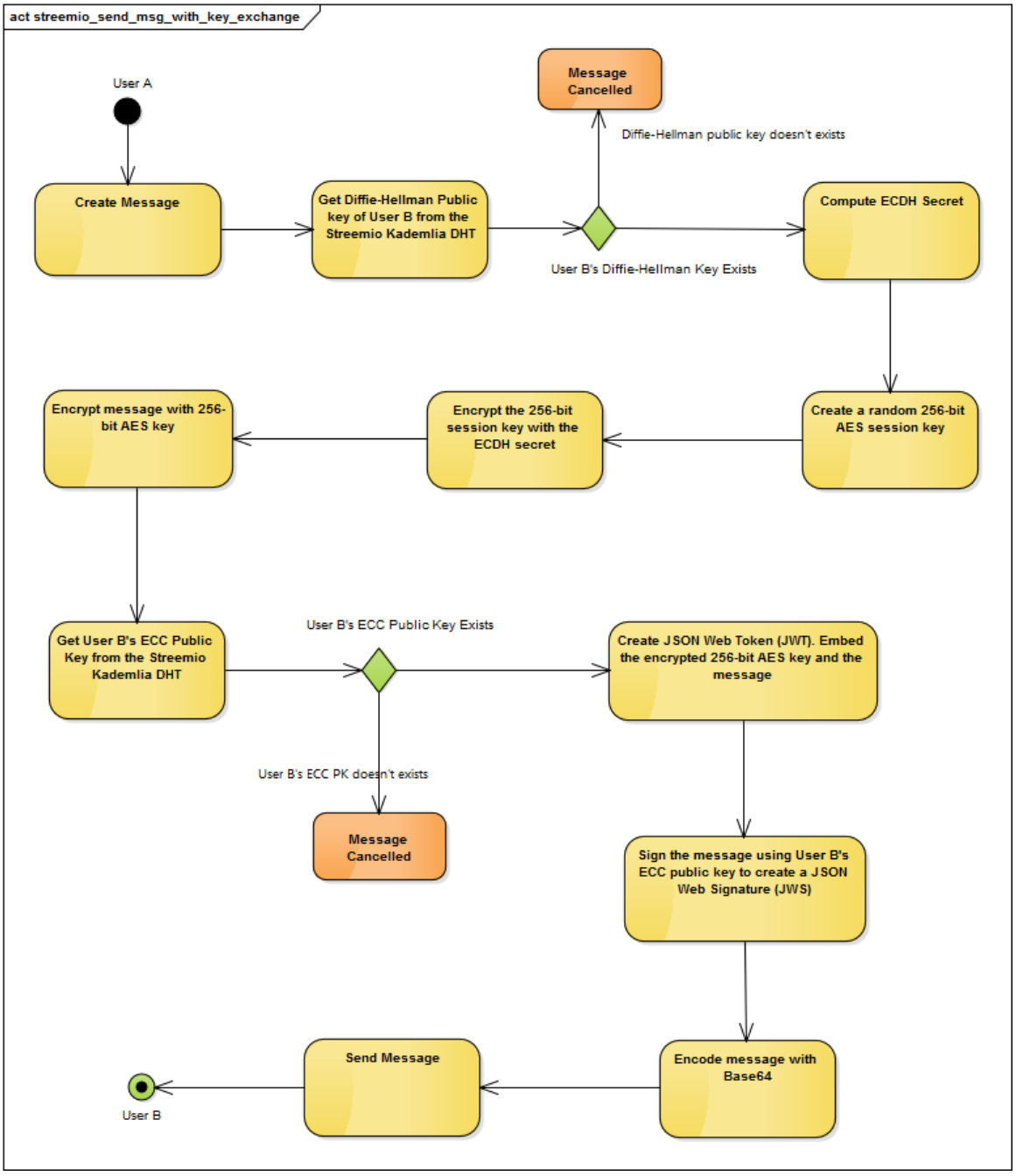
Support for GCM

Streembit supports and implement GCM. Galois/Counter Mode (GCM) is a block cipher mode of operation that uses universal hashing over a binary Galois field to provide authenticated encryption. It can be implemented in hardware to achieve high speeds with low cost and low latency. Software implementations can achieve excellent performance by using table-driven field operations. It uses mechanisms that are supported by a well-understood theoretical foundation, and its security follows from a single reasonable assumption about the security of the block cipher. There is a compelling need for a mode of operation that can efficiently provide authenticated encryption at speeds of 10 gigabits per second and above in hardware, perform well in software, and is free of intellectual property restrictions. The mode must admit pipelined and paralellized implementations and have minimal computational latency in order to be useful at high data rates. Counter mode has emerged as the best method for high-speed encryption, because it meets those requirements. However, there is no suitable standard message authentication algorithm. This fact leaves us in the situation in which we can encrypt at high speed, but we cannot provide message authentication that can keep up with our cipher. This lack is especially conspicuous since counter mode provides no protection against bit-flipping attacks. GCM fills this need, while no other proposed mode meets the same criteria.

How secure the cryptography of Streembit?

Both 128-bit and 256-bit AES symmetric encryption, the cryptography schemes of the messages in the Streembit network, are safe from any brute force attacks. Bruce Schneier points out: "*These numbers have nothing to do with the technology of the devices; they are the maximums that thermodynamics will allow. And they strongly imply that brute-force attacks against 256-bit keys will be infeasible until computers are built from something other than matter and occupy something other than space.*"

The following UML diagram describes the key exchange using ECDH and securing the message with 256-bit AES symmetric key between "User A" and "User B"



Internet-of-Things Device Handling

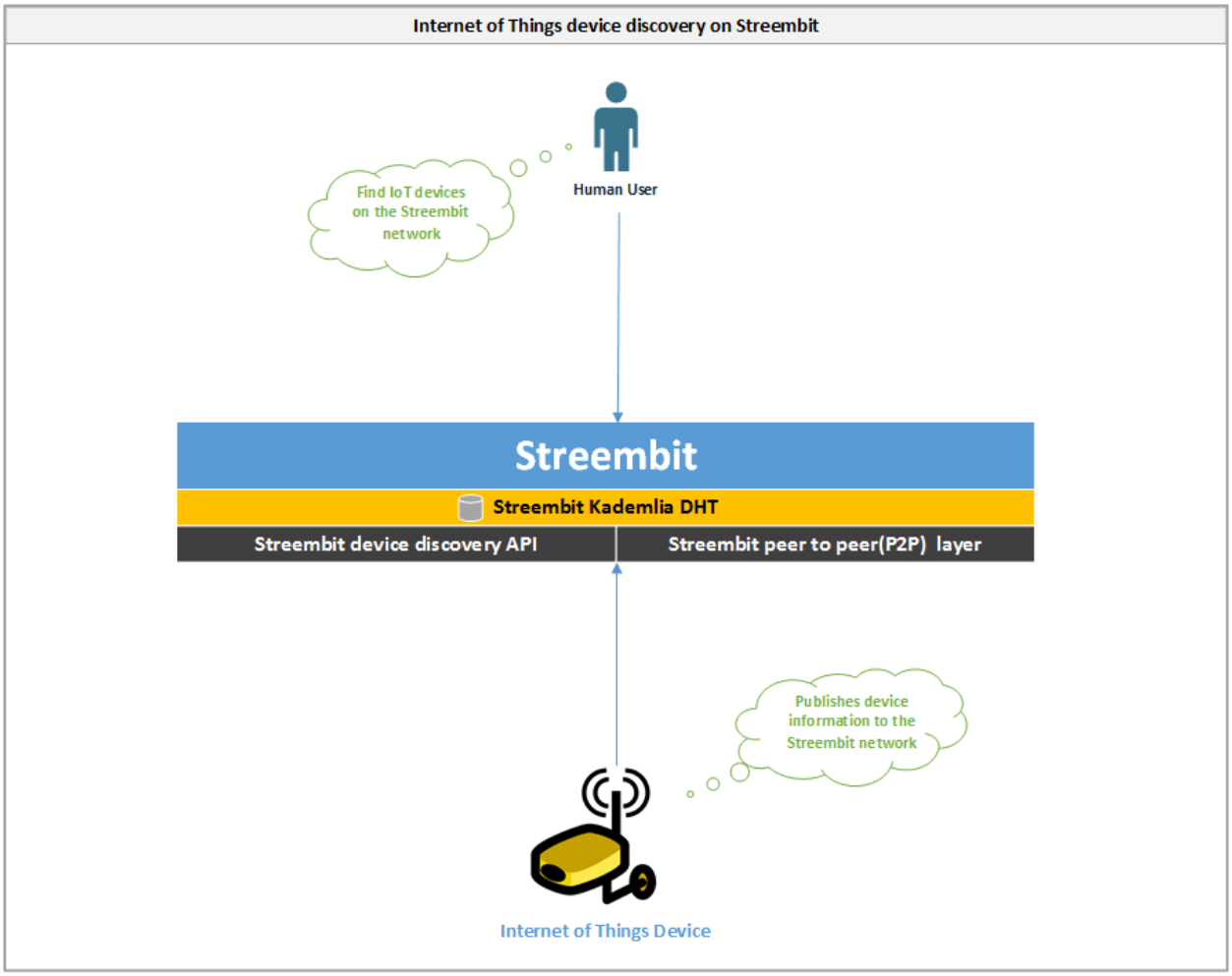
It is troubling that robust security was not taken into account with the majority of Internet of Things implementations. The data of IoT devices are being harvested in an automated fashion but who has access to the data? What functions can a human or another machine execute on the IoT device? Is my office door actually being opened by a former employee who is not supposed to enter the premise anymore? Is the data being forwarded by the IoT device compromised at all during its way to the end-user, never mind whether or not it was sent by the actual device and not in fact by an intruder?

Streembit manages device discovery as well as authentication, access control and provisioning of devices without using a centralized authority server. The system facilitates device discovery and device control in a peer to peer manner. For many IoT use cases a decentralized, P2P network topology is the most secure, robust, scalable and reliable method of operation. Instead of using expensive corporate owned third party cloud systems, Streembit manages Internet of Things devices without the need of proprietary centralized cloud platforms.
Centralized, corporate owned cloud is certainly an easier way to build out IoT platforms. However, the owners and authorities in these topologies all have an influence upon the network and can be exploited. They can ban devices, spy on devices and compromise data integrity of devices. In fact, government can order them to do any or all of these. In the near future, the doors of your garage and home, air conditioning units and your home security system will be fully internet connected. You will be able to control your home automation system from your mobile phone. It is essential that only you can control your IoT devices. Streembit excludes third party service and cloud providers from the ecosystem to give full control to the end users over the devices.

The Streembit IoT implementation is based on open standards. The Streembit developers contribute to the standardization process of W3C Web of Things Interest Group and mirror all W3C open IoT standards in the Streembit source code.

Internet of Things device discovery on Streembit

Streembit allows IoT application providers and context producers to register their IoT Objects on the decentralised Streembit network, and in turn allow context consumers to discover them in a secure and peer to peer manner.

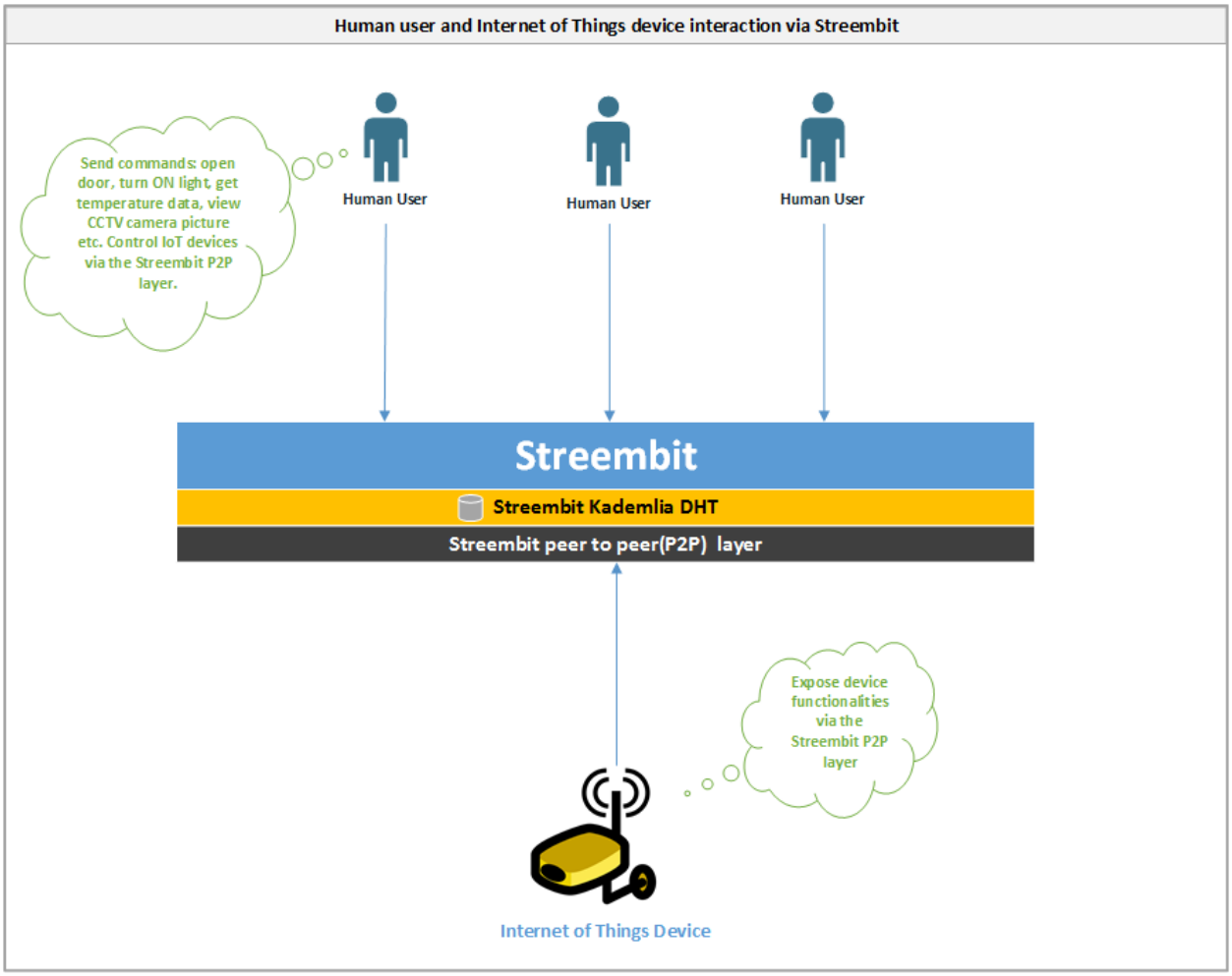


Example data structure for device information:

```
{
    "account": "058dd8eb72d1936df31eebdd9afb49e1415787e3",
    "public_key": "03e9a353b86cc482af497a8afb56209ee35eaa80bee228c2af8ce48747b90c50e7"
}
```

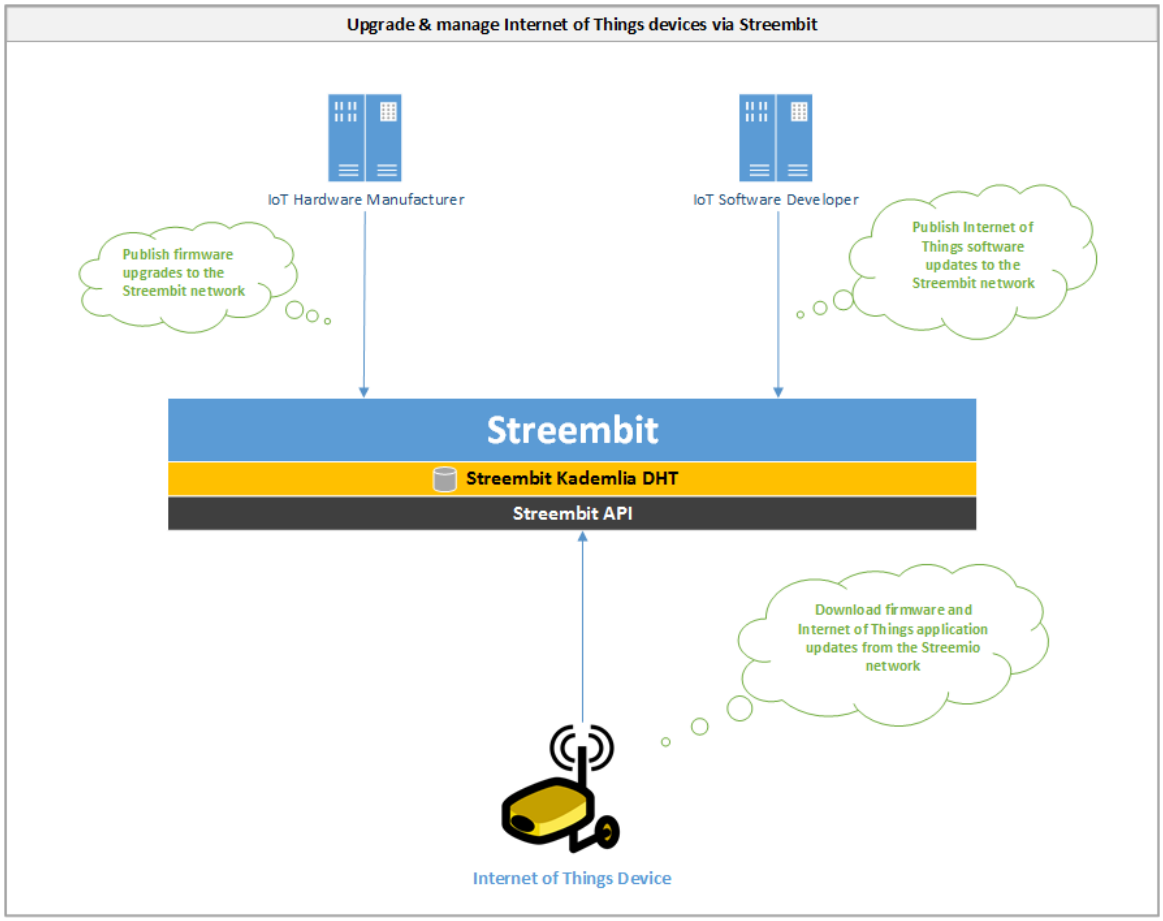
Control Internet of Things devices via Streembit

- The Internet of Things devices and human users communicate with each other directly in a peer to peer manner
- The data is end to end encrypted between the human user and IoT device. The encryption key is shared only between the user and device – never with any third parties.
- The device exposes functionalities via the Streembit network and user interface using W3C WoT standards
- The user interacts with the device via the Streembit P2P layer and UI. For example, opening a door, turning ON a light, controlling motor speed, getting temperature data, viewing CCTV pictures etc. all peer to peer, without a centralised solution and via the Streembit user interface.



Upgrade and manage Internet of Things devices via Streembit

- Hardware and software providers upgrade Internet of Things devices on the always up and running Streembit network.
- Internet of Things device manufacturers and software designers publishes firmware and software updates via the Streembit network.
- Internet of Things devices run the Streembit system and ensure the origin and data integrity of the updates by verifying the public key of the publisher.



References

https://github.com/w3c/web-of-things-framework
https://github.com/w3c/web-of-things-framework/blob/master/security.md
http://www.information-management.com/news/big-data-analytics/blockchain-could-be-key-tool-for-tracking-data-origin-and-authority-10028871-1.html
http://www.mindspring.com/~dmcgrew/gcm-nist-6.pdf
http://xlattice.sourceforge.net/components/protocol/kademlia/specs.html
https://pdos.csail.mit.edu/~petar/papers/maymounkov-kademlia-lncs.pdf
https://en.bitcoin.it/wiki/Blocks
https://www.igvita.com/2014/05/05/minimum-viable-block-chain