How to build the Metaverse: A feasibility study of a distributed & decentralized multi user 3D virtual world using IPFS and blockchain technologies.

Feasibility Study

Patrick Tobler

Computer Science and Media, Stuttgart Media University, Nobelstraße 10, 70569 Stuttgart, Germany

Abstract

This paper explores the possibility of a fully decentralized and distributed multi user virtual world. It outlines an architecture that uses IPFS as a distributed storage technology for the content of the virtual world and uses a public blockchain for property governance. It introduces the concept of a Gateway, as an entry point to the distributed & decentralized multi user 3d virtual world and then proves that it is possible to build such a Gateway. It also theoretically explores peer-to-peer trust mechanisms that would allow interactions between users.
At the end the paper outlines a strategy for moving this concept to reality.

Keywords: Distributed, Decentralized, MMO, Blockchain, Consensus, Peer-To-Peer, IPF, Gateway

Term Definitions

Decentralization (n.) - “In a decentralized system, there is no single controlling entity. Instead, control is shared among several independent entities.” [1]

Distribution (n.) - “In a non-distributed (or co-located) system, all the parts of the system are in the same physical location. In a distributed system, parts of the system exist in separate locations.” [1]

Metaverse (n.) - “The Metaverse is a collective virtual shared space, created by the convergence of virtually enhanced physical reality and physically persistent virtual space, including the sum of all virtual worlds, augmented reality, and the internet.” [2]

Blockchain (n.) - “A digital database containing information (such as records of financial transactions) that can be simultaneously used and shared within a large decentralized, publicly accessible network.” [3]

Smart Contract (n.) - “A smart contract is a computer protocol intended to digitally facilitate, verify, or enforce the negotiation or performance of a contract. Smart contracts allow the performance of credible transactions without third parties.” [4]

Gateway (n.) - Client software that connects to a distributed network, provides 3d visual feedback and lets the user interact with the content.

Introduction

Current state of the internet

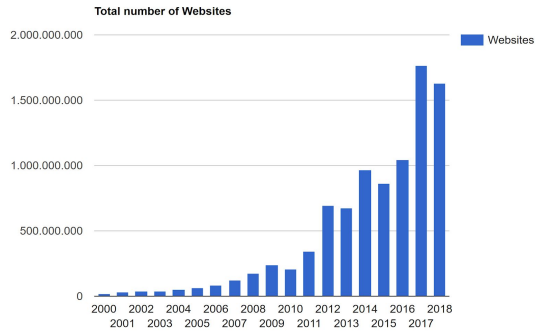
When Tim Berners-Lee first introduced the World Wide Web in 1989 he proposed the idea of a linked data system that interconnects documents using so-called Hyperlinks. The main objective of this data system was to provide an elaborated tool for documentation of large-scale projects at CERN.
To fulfill this objective Berners-Lee set up 9 requirements, he believed to be critical:

- Remote access across networks
- Heterogeneity
- Non-Centralization
- Access to existing data
- Private links
- Bells and Whistles
- Data Analysis
- Live Links
- Non requirements

[5]

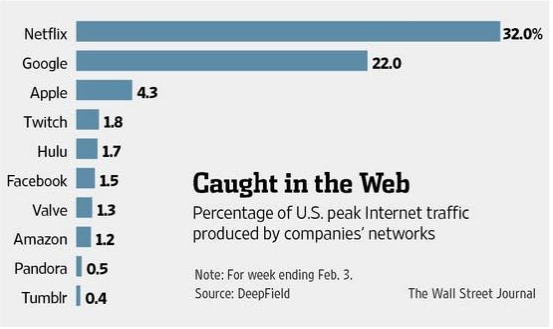
After the introduction of this linked data system the amount of documents stored using the system exploded.

The number of documents linked on the World Wide Web peaked in 2018 with a total amount of over 1.5 billion websites.



Img. 1 -" Total number of websites"

And yet the traffic distribution on the internet paints a very different image. Instead of having a wide spread allocation of traffic, studies show that only a handful of big websites are responsible for most of the traffic on the internet. The common ground for the most traffic-intense websites is the content they provide: Video.



Img. 2 - "Percentage of U.S. peak Internet traffic produced by companies networks"

Another easily observable effect the internet had was the democratization of content creation. Instead of a minority of people creating content for the majority of consumers the internet shifted this heavily by introducing the ability to create all kinds of content more easily, as well as the ability to share that content with a potentially enormous target group.

Combined with the shortened attention span that the overload of information on the internet introduces [6] this makes for the new phenomenon of **micro-content-creation**. Suggesting that a large

group of creators produces an even larger amount of very tiny pieces of content.

To do so the creators flock to the most popular platforms for specific content types. From Youtube, Tik Tok and Snapchat for video content, to Twitter, Facebook, Instagram and Tumblr for text and image content.

Current state and future of 3D media consumption

Computer software specifically designed for creating 3D graphics began appearing in the 1970s. Since then the development of 3D graphics has been on a steady upwards curve. Video Games have been dominating the utilization of 3D graphics for years and are now a major medium that drives the development of graphics hardware.

With the development of stronger and more capable hardware previously unthinkable possibilities for 3D media consumption open up. Especially Augmented Reality and Virtual Reality devices have shown promising prospects in that regard.

“VR and AR have the potential to add $1.5trillion to the global economy by 2030.” [7]

Future of the Internet

Now that the question of what the internet looks like today is answered, it’s finally time to turn to the much more interesting question of what the internet will look like in the future.

There are two main observations that lead to a possible outlook on the future of the internet:

1. It’s clear that the behaviour of average people on the internet is rapidly changing from being mostly consumption focused to the creating and sharing their own content on the internet.
2. Due to steady advances in Virtual Reality technology, Augmented Reality technology and graphics hardware the world is transitioning from a 2D approach to using technological devices to a 3D one.

These observations allow us to conclude that the future of the internet will focus on currently emerging technologies like AR & VR, as well as having a very content generating approach. Quite similar to the popular term **Metaverse**.

Risks

This Metaverse-like outlook on the future of the internet brings unthought-of prospects on the future of the life of everyone but also bears enormous risks.

While it’s close to impossible to predict how exactly this future internet might look like in detail, it is certain that it will reach every corner of our lives and will have a deep impact on the world.

Even though the current state of the internet is clearly displaying that human behaviour on the internet is developing into a very creation focused direction, it is also showing that the internet is developing into a very centralized and narrowly focused direction.

The combination of an internet that is deeper intertwined with the lives of the people using it, as a result of a creation focused behaviour & the rapid development of 3D media technologies, with the internet being more and more centralized will lead to a dramatic shift how the world creates and distributes information and might possibly lead to a very authoritarian internet controlled by few.

Solution

So what is the solution to all of this? How can we create an internet that is ready for the creation and sharing of ever-increasing complex 3D media while not having to rely on enormous platforms that have complete oversight over what is being created and distributed?

Ideally this '3D internet' should fulfill the following criteria:

- Strong decentralization
- Accessible to everyone
- Sharing media is possible without relying on platforms
- Interaction between content & user and content & content is possible
- Uncensorable
- Not regulateable

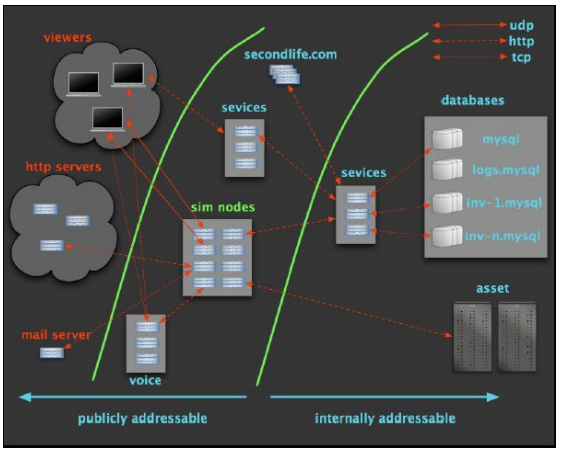
This paper is the documentation of a feasibility analysis with the goal to test if it's feasible to create a system that fulfils the above criteria using state-of-the-art Peer-2-Peer and other decentralization technologies.

Related Projects

In order to better understand how such a system might look like it's helpful to analyse projects with a similar context.

Second Life

Being the most prominent virtual world project to date, Second Life has had success by being a platform completely reliant on user generated content. Content can be created using different tools and imported into the virtual world. There it can be sold using the games currency.



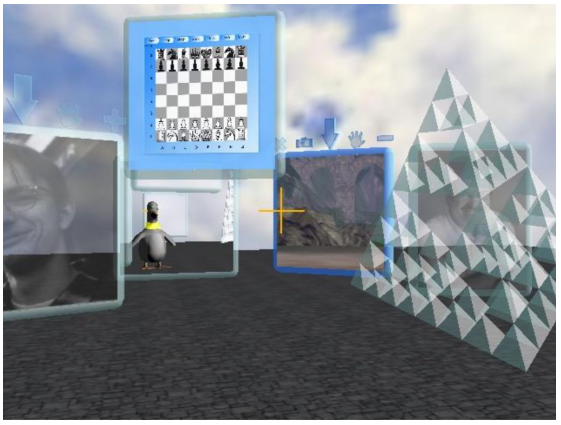
Img. 3 - "Second Life's architecture"

Second Life created a long-living virtual world with up to a million active users and a sustainable marketplace for virtual goods.

The development of the project is mostly built on open source technologies and almost all of the source code is freely available. It's structured using a traditional client-server-architecture. Servers can be hosted by anyone and clients can be written by everyone using the defined standards to access the servers. [8]

Croquet Project

The Croquet Project was a software project that intended to allow for the creation of multi user collaborative Peer-To-Peer applications:



Img. 4 - "Croquet Multi User Environment"

“Croquet supports communication, collaboration, resource sharing, and synchronous computation among multiple users. Applications created with the Croquet software development kit (SDK) can be used to support highly scalable collaborative data visualization, virtual learning and problem solving environments, 3D wikis, online gaming environments (massively multiplayer online role-playing games), and privately maintained or interconnected multi user virtual environments.” [9]

Croquet introduced a new synchronization architecture, called TeaTime. This architecture allows for massive amounts of users interacting with each other in real time. The key elements of this architecture include:

- A coordinated universal timebase embedded in communication protocol
- Replicated, versioned objects that unify replicated computation and distribution of results
- Replication strategies that separate the mechanisms of replication from the behavioral semantics of objects
- Deadline-based scheduling extended with failure and nesting
- Coordinated, distributed two-phase commit that is used to control the progress of computations at multiple sites, to provide resilience, deterministic results, and adaptation to available resources
- Uses distributed sets

[10]

Based on this approach have been a number of further works. One of them is Open Cobalt, a free and open-source software platform for constructing, accessing, and sharing virtual worlds both on local area networks or across the Internet, with no need for centralized servers. [11]

Recently most of the efforts of advancing the Croquet Project have been getting less, except for one company, the Croquet Corporation, that is trying to advance the Croquet Project by launching a live collaboration platform based on the Croquet Project. [12]

Solipsis

Solipsis is a proposed decentralized architecture for virtual environments. It focuses on the distribution of computing work, using a peer-to-peer approach.

Solipsis does not outline a strategy for creating a failure tolerant network that is resistant to malicious attacks. [13]

Concept

The following core problems need to be solved in order to create a fully distributed multi user virtual world that satisfies the previously defined criteria.

Storage

In order to create this decentralized and distributed 3d world there needs to be a large scale network beneath the surface to which users connect and download files from or upload their own content to. Since this network can not rely on a classic Server-Client architecture, to prevent the network from centralization, it has to be based on a Peer-To-Peer approach.

Ideally every connected user expands the network by functioning not only as a client but also as a host and thus expanding the distribution of the content on the network.

Proposed solution: The underlying storage system of the distributed & decentralized 3d world relies on a

distributed file network that is easily connectable and accessible.. Software that connects to the network expands it by serving as a client and host at the same time.

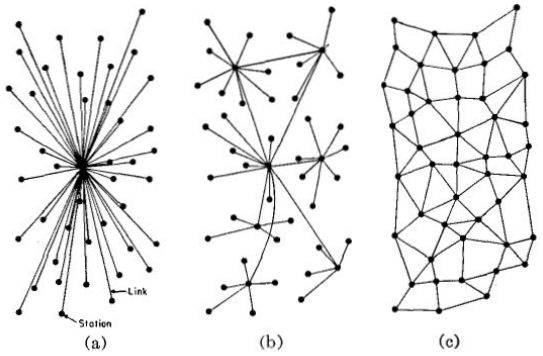


Fig. 1—(a) Centralized. (b) Decentralized. (c) Distributed networks.

Img. 5 - "Centralized vs. Decentralized vs. Distributed"

Access

A crucial part of creating a distributed & decentralized 3d world is having a common set of rules that specify how to access the network and are implemented in every software accessing the world. Establishing these rules in form of specifications enables a multitude of independent client tools that are able to access the world in different ways.

The most obvious way of accessing the distributed & decentralized 3d world would be through a browser-like software that connects to the network and provides a graphical feedback.

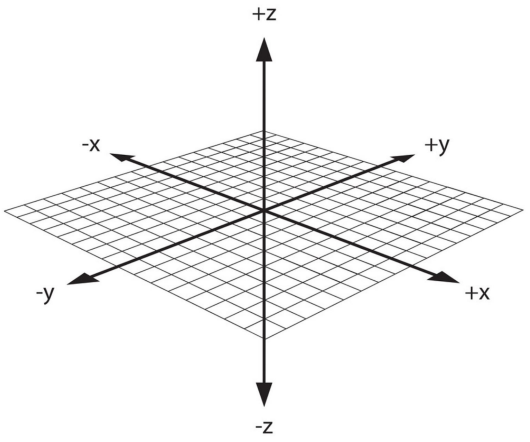
To provide a more fitting name this paper refers to the term **Gateway** when referring to the above explained software concept.

Proposed solution: Public specifications for how to exactly access the network.

Location

The internet as it is has a very clean organizational structure. Because it was designed to serve 2D data it is easily structurable into completely separate documents that are only connected with each other via hyperlinks.

Unfortunately this 2D concept is not a natural fit for 3D media. To convert it there would have to be a fourth dimension that connects separate pieces (or documents) of 3D media. A more natural fit for 3D media is to use the **world space** concept and place everything in the same enormous (or even infinite) 3D space. Similar to the real world.



Img. 6 - "3D coordinates"

Proposed solution: Every 3D object is given a X-, Y- and Z-Coordinate of this shared world space to clearly identify its location.

Property

To avoid having multiple 3D objects at the same position there needs to be a method to handle properties. Similarly to the way that URLs can't be owned twice.

To do so the world space is subdivided into smaller **chunks**. Each chunk can be allocated to an owner. The owner is the only one that has the power to decide which content a chunk should hold.

To view, change or create property rights there needs to be some kind of trusted system that every user can openly access.

Traditionally this would be solved with a database or some other central instance. In order to create a distributed system there has to be a trusted instance that is not reliant on one single actor.

Proposed solution: The property rights of specific parts of the world space, called chunks, are stored on a distributed and decentralized ledger in the form of a public blockchain.

Retrieval

Once the property rights of a specific chunk in the world space have been retrieved from the public blockchain, the client should be able to identify the internet addresses of all objects that have been placed inside the chunk, as well as the exact location of the objects inside the chunk.

A possible way of realizing this would be to store the location and addresses of all objects inside the public blockchain. Due to the nature of the blockchain concept this is not a scalable approach since blockchains have a limited amount of data that can be stored on them. Because of this restriction it makes more sense to store all object information about location and addresses in a markup file and only store the address of the markup file in the blockchain.

Once the client retrieves the address of the markup file over the blockchain it accesses this file and retrieves addresses and locations of all objects in the chunk. The client then downloads all objects via provided addresses and places them in the chunk.

A significant restriction of this approach is that in order to change the content of the markup file, a new markup file has to be uploaded to the distributed file network and referenced in a public blockchain transaction.

Proposed solution: Every chunk's objects information (addresses and locations) are stored in a markup file. The address of the markup file is stored in the public blockchain.

User-World Interactions

Interactions between the user and the environment play a significant role when creating a compelling world. Because of this, all 3d objects have a need for being able to receive different physical properties.
Some objects could have physical boundaries through which the user cannot move, other objects might be bouncy or triggerable (without actually having physical boundaries).

These properties, as well as just cosmetic properties like the colour of (parts of) the object should be changeable due to certain conditions surrounding them.

Proposed Solution: To provide hosts with the option to add interactability between the user and the environment to their chunks, there needs to be one (or more) common scripting language(s) specified. This/These scripting language(s) are interpreted by the Gateways, comparable to Javascript in the Web. The scripting language(s) require access to the objects of their chunk and the player, aside from their normal functionalities.

User-User Interactions

While interactions between the user and the environment are relatively easy to realize, the case becomes a lot more complicated once there are multiple users that should be able to interact with each other and possibly the environment as well.

In this case there needs to be some sort of trust-mechanism that ensures that no user is

exploiting the others by using injected code or any other malicious attack.

Solution: A flexible amount of users in the same chunk connect and form a peer-to-peer network where they use trust mechanisms to provide security in the correctness of the decision making of each user.

Copyrights & Economy

The underlying public blockchain makes it very easy to follow the trace of content back to its original creator. But having a distributed file storage system also makes it very easy to illegitimately use and distribute the content.

So in order to prevent content from being stolen the only reasonable solution would be a Gateway layer solution based on the implementation of the Gateway.

A Gateway could easily implement an algorithm that checks the legitimacy of content before displaying it. This system would work together with a predefined smart contract system that would allow creators to transfer rights for usage of the content to other users.

The implementation of this smart contract validation system would heavily depend on the Gateway developers. It is entirely possible to create Gateways that completely ignore this solution.

There is a societal dilemma hidden in these lines. Society will have to find an answer to the question if this distributed & decentralized multi user 3d world should have this Gateway layer type of censorship or if the current copyright rules might not be adequate for the future.

Another big problem with abandoning any type of validation of legitimacy of content is that it'd make it nearly impossible for creators to sell content, meaning that there won't be an active economy inside the network.

Proposed solution: There could be a Gateway layer standard of validating copyright rules by tracing a 3d object back to its first instance in the network. Expanding on this, the rights for usage of the 3d object could be changed (transferred, sold, rented, …) using smart contracts.

This approach opens up a lot of societal questions that can not be answered at this point, thus making it impossible to conclude to a final solution of this problem.

Technology

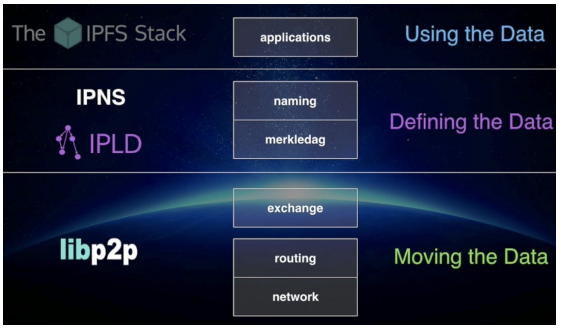
To realize a distributed & decentralized multi user 3d virtual world there needs to be a strong technological foundation. And yet there are multiple technologies that could work with the underlying architecture of the system. The strength of the proposed architecture lies in its openness, meaning that it's possible to realize it with a wide variety of technologies.

Following are some promising technologies that could be used for an early prototype of the proposed architecture.

The Interplanetary File System (IPFS)

The Interplanetary File System, or IPFS is a peer-to-peer networking protocol used for data sharing on the distributed web.

The underlying technology behind IPFS is a combination of sophisticated technologies, most importantly Distributed Hash Tables, Block Exchanges, Version Control Systems and Self-Certified File Systems.



Img. 7 - "IPFS Stack"

Content Based Addressing

In contrast to other networking protocols, like the Hypertext Transfer Protocol (HTTP) it uses a different way to address content on the network. Instead of using a location based addressing system, e.g. Uniform Resource Locators (URLs), IPFS uses a content based addressing system. This means that instead of addressing the location of where a file is stored, in IPFS you simply address the file itself.

IPFS uses cryptographic hashing in order to generate the unique content based addresses of the files. This method helps build trust between the members of the networks. If a file was being tempered with in any way, the resulting cryptographic hash would be completely different from the expected hash. Modifications can be verified simply by looking at the hash of the file.

For this purpose IPFS uses the concept of Content Identifiers (CIDs) to access the files on the network. A CID consists of the cryptographic hash and a codec, which holds information about how to interpret that data.

Using this content based addressing system to retrieve files from IPFS means that instead of asking a single server (peer) directly for the files you ask the network as a whole for the specific files, using the CID. The network detects peers that are currently available and hosts the specified files. It then connects us to one of these peers to transfer the files.

Merkle-DAGs

Merkle-DAGs (Directed Acyclic Graph) are basically a collection of linked nodes. Each of the node points to other nodes via their content addresses (hashes).

This data structure is used in IPFS so connect all nodes and navigate the network using the content addresses.
[14]

Conclusion

IPFS can be used as an ideal storage technology for a large scale distributed & decentralized 3d virtual world.

Unity

To render the distributed & decentralized 3d world with all its objects and users there needs to be a client that is able to utilize a render engine built for this purpose.
Luckily a lot of modern game engines provide that functionality.
Every game engine has slight variations in their render techniques, meaning that there is no perfect engine. Gateways could differ in what kind of

rendering technique they use. Some of them could be built on top of commercial game engines, others could implement their own rendering engine.

Unity is a good choice as development platform for a Gateway because it has a lot of cross-platform functionalities. Meaning that the same code base used for one platform can be used for many other platforms as well. This is particularly interesting when it comes to porting the Gateway to VR or AR devices. [15]

Questions

In order to build a prototype Gateway and find possible mistakes in the architecture the following questions need to be answered:

Can you run a node via Unity and if not how does this influence the architecture?
Answered: Yes. Not only are there existing Open Source implementations available [16] but its also possible to write an IPFS impementations in C# (Unity) from scatch.

Is it possible to download files from IPFS using Unity and dynamically load them into the scene?
Answered: Yes. By testing it with multiple actors this could be confirmed.

What are the minimal requirements to run a node and connect to the 3d world using a Gateway?
The requirements for a Gateway basically split up into two parts.

- **The IPFS Node:** The Gateway has to be able to set up a stable node.The IPFS node will be downloading, uploading and deleting content dynamically. For this it requires a good internet connection and a good amount of free disk space.

  A minimal requirements recommendation from Cloudflare is the following:

  - 2 gigabytes of RAM
  - 10 gigabytes of disk space
  - 1 terabyte of bandwidth per month

  [17]

- **Client responsible for rendering the virtual world:** This heavily varies on external factors like the screen size, platform, graphical fidelity, as well as amount of 3d objects being rendered at the same time. So it's really hard to make a conclusive statement. There needs to be more testing in this direction.

Is there a way to find all hosts that are currently hosting a specific chunk?
In order to create a peer-to-peer connection that allows users to interact with each others by using some sort of decentralized computing & trust mechanism, there needs to be a way to connect all of these users. For that the first step is finding all of the users currently located in the same chunk.

This problem can either be fixed by using IPFS infrastructure or by making changes to the overall architecture of the network.

Which consensus mechanisms can be used to allow for peer-to-peer multi user interactions?
Thats a very big question that a lot of research has been done for (e.g. Raft Consensus Algorithm [18]). To find the most fitting answer to this question in the context of a distributed & decentralized multi user 3d virtual world is one of the key elements for bringing this concept to reality.

Conclusion

This paper underlined the need for a distributed & decentralized multi user 3d virtual world. It then proposed an architecture that could realize the specifications given for such a distributed & decentralized virtual world and continued on with an overview of possible technologies used to implement this architecture.

In the end it's possible to conclude that such a distributed & decentralized multi user 3d virtual world can absolutely be realized using IPFS and blockchain technolgies.

But there are also many more open questions, that have to be answered before truly creating such a system. Especially surrounding the topic of distributed computing and trust in peer-to-peer connections, there has been a lot of research in the past but that also means that it's quite difficult to find the perfectly fitting solution for the system.

Another topic that is yet to be solved has to do with the scalability of the system. The scalability of the whole system is deeply connected to the scalability of the IPFS network. Because of this there needs to be more research surrounding the core structures of the IPFS network in general. While it in theory is limitless, there are some restrictions that could (at the moment) hold the advancement of the distributed & decentralized 3d world back.

Outlook

Based on the learnings of this paper there are some specific steps that can be taken to advance the concept of this paper and truly create the Metaverse.

- Develop and publish the formal specifications for the Markup-Language used to store 3d objects location.
- Develop and publish the formal specifications for the smart contracts that are responsible for tracking the ownership of chunks.
- Develop and publish the formal specifications for a scripting language that is universally applicable for user-environment interactions.
- Build a Gateway prototype that implements these specifications.
- Decouple the research for a Gateway layer solution to distributed computing and trust mechanisms in peer-to-peer connections. Then develop and publish specific formal specifications for this.

References

1. "Whats the difference between Decentralized and Distributed?", Julia Poenitzsch, 3.8.2018, *Source:* medium.com/nakamo-to/whats-the-difference-between-decentralized-and-distributed-1b8de5e7f5a4
2. Smart, J.M., Cascio, J. and Paffendorf, 2007
3. "Blockchain.", *Merriam-Webster.com Dictionary*, Merriam-Webster, Accessed 15.03.2020, *Source:* www.merriam-webster.com/dictionary/blockchain
4. "Smart Contracts: The Blockchain Technology that will replace Lawyers", Ameer Rosic, 2017, *Source:* blockgeeks.com/guides/smart-contracts
5. "Information Management: A Proposal", Tim Berners-Lee, CERN, 03.1998, *Source:* www.w3.org/History/1989/proposal.html

6. “Accelerating dynamics of collective attention”, Philipp Lorenz-Spreen, Bjarke Mørch Mønsted, Philipp Hövel & Sune Lehmann, 15.04.2019, *Source:* nature.com/articles/s41467-019-09311-w
7. “Seeing Is believing”, PwC, 2019, *Source:* pwc.com/seeingisbelieving
8. “Second Life’s strange second life”, Chris Stokel-Walker, 24.09.2013, *Source:* theverge.com/2013/9/24/4698382/second-lifes-strange-second-life
9. “Open Cobalt”, Accessed 15.03.2020, *Source:* organicdesign.nz/OpenCobalt
10. “Croquet - A Collaboration System Architecture”, Alan Kay et al., 03.2003, *Source:* researchgate.net/publication/4029258_Croquet_-_A_collaboration_system_architecture
11. “Open Cobalt - Wikipedia”, Accessed 15.03.20202, *Source:* en.wikipedia.org/wiki/Open_Cobalt
12. “Croquet.io”, Accessed 15.03.2020, *Source:* croquet.io
13. “Solipsis: A Decentralized Architecture for Virtual Environments”, Davide Frey, Jérôme Royan, Romain Piegay, Anne-Marie Kermarrec,Emmanuelle Anceaume, Fabrice Le Fessant, 5.11.2008, *Source:* hal.inria.fr/inria-00337057/document
14. “IPFS- Content Addressed, Versioned, P2P File System”, Juan Bennet, Accessed 15.03.2020, *Source:* ipfs.io/ipfs/QmR7GSQM93Cx5eAg6a6yRzNde1FQv7uL6X1o4k7zrJa3LX/ipfs.draft3.pdf
15. “Unity - Core Platform”, Accessed 15.03.2020, unity.com/products/core-platform
16. “IPFS for Unity”, Accessed 15.03.2020, *Source:* github.com/samueljlopes/IPFS-for-Unity
17. “IPFS-Gateway - Setting up a server”, Cloudflare, Accessed 15.03.2020, *Source:* developers.cloudflare.com/distributed-web/ipfs-gateway/setting-up-a-server
18. “The Raft Consensus Algorithm”, Accessed 15.03.2020, Source: raft.github.io

Images

1. “Total number of websites”, Accessed 15.03.2020, Source: internetlivestats.com/total-number-of-websites/
2. *“Percentage of U.S. peak Internet traffic produced by companies networks”, Accessed 15.03.2020, Source:* de.statista.com/infografik/15674/zusammensetzung-des-weltweiten-downstream-internet-traffics
3. “Second Life’s architecture”, Accessed 15.03.2020, Source: infoq.com/news/2008/12/Second-Life-Ian-Wilkes/
4. “Croquet Multi User Environment”, Accessed 15.03.2020, Source: worrydream.com/refs/Smith%20-%20Croquet%20-%20A%20Collaboration%20System%20Architecture.pdf,
5. “Centralized vs. Decentralized vs. Distributed”, Accessed 15.03.2020, Source: medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274
6. “3D Coordinates”, Accessed 15.03.2020
7. “IFPS Stack”, Accessed 15.03.2020, Source: decentralized.blog/img/ipfs-stack.jpg