



ISSUED DATE | 4 AUGUST 2019

# PROJECT PROPOSAL

Course: SE801 **Software Project Lab**

Submitted by

Tahlil, ROLL 803, year:2015-16

Document Version: **1.0**

**A BLOCKCHAIN  
BASED WEB  
HOSTING  
SCHEME FOR  
CONTENT-  
ADDRESSED  
DISTRIBUTED  
FILE SYSTEM**

Submitted to:  
Coordinator

Ahamedul Kabir,  
Assistant Professor,  
IIT,DU

**Institute of Information Technology  
University of Dhaka  
04-08-2019**



Project proposal on:

# A BLOCKCHAIN BASED WEB HOSTING SCHEME FOR CONTENT-ADDRESSED DISTRIBUTED FILE SYSTEM

## Proposal Submission Letter

4 August, 2019

Dr. B. M. Mainul Hossain

Associate Professor

Institute of Information Technology

University of Dhaka

**Subject:** Submission of project proposal of Software Project Lab 3

Dear Sir,

With due respect, I am hereby pleased to submit the proposal report for Software Project Lab(SPL) 3 entitled "A blockchain based web hosting scheme for content-addressed distributed file system". The report elaborates on the topics of problems with the existing centralized internet ecosystem, the past and ongoing efforts to mitigate the problems and finally the objectives and in-depth analysis of the proposed scheme/framework.

I hope you would appreciate the effort I put into and accept the proposal.

Sincerely yours,

Tahlil (BSSE 0803)

BSSE 8th batch

Institute of Information Technology

University of Dhaka



# Table of Contents

<b>Chapter 1: Introduction</b>	<b>7</b>
<b>Chapter 2: Broad Domains</b>	<b>9</b>
2.1 Internet & WWW	9
2.2 Blockchain	11
2.2 Decentralized file system and web hosting	15
<b>Chapter 3: Problem statements</b>	<b>19</b>
3.1 Problem Statement 1: Problems with centralized internet	19
3.1.1 Servers Can Go Down	19
3.1.2 Servers Can Be Hacked	19
3.1.2 Companies and Government Throttle or Censor Data	19
3.1.4 Companies Monetize Our Data	21
3.2 Problem Statement 2: Problems with decentralized web ecosystem over the internet	21
3.2.1 Mass Adaption	21
3.2.2 Availability	21
3.2.3 Serve big file	21
3.2.4 Users Do not Want the Extra Responsibility	21
<b>Chapter 4: Existing Work</b>	<b>23</b>
<b>Chapter 5: Proposed Scheme</b>	<b>27</b>
5.1 Project Outline	27
5.1.1 Description	27
5.1.2 Data collection	29
5.1.3 Design	29
5.1.4 Implementation	29
5.1.5 Analysis	29
5.1.6 Review	29
5.1.7 Report	29
5.1.8 Advice and policy formulation	29
5.2 Project Objectives	31
5.2.1 Broad Objectives	31
5.2.2 Specific objective	31
5.3 Project METHODOLOGY	33
5.3.1 Data Collection	33
5.3.2 Design	33
5.3.3 Implementation	33
5.3.4 Analysis	33
5.3.5 Review	35
5.3.5 Report	35
<b>References</b>	<b>37</b>

The background of the page is a solid blue color. Overlaid on this is a complex, abstract network diagram. It consists of numerous small, dark blue circular nodes of varying sizes. These nodes are interconnected by a dense web of thin, light blue lines, creating a mesh-like structure. Some nodes are highlighted with a slightly different shade of blue or a thin white outline. The overall effect is a sense of a global or distributed network, possibly representing a blockchain or a web hosting scheme.





## Chapter 1: Introduction

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

## Chapter 1: Introduction

The project that I proposed for the Software Project Lab-3 (SE-801) as a part of BSSE 8th semester course is "A blockchain based web hosting scheme for content-addressed distributed file system". The goal of this project is to design and develop a scheme for a decentralize file system and integrating to a distributed web service over the internet. One of the major tool that will be used to achive this goal is blockchain to ensure security, transparency and availability. Another core concept that will be utilized to exploit the strength of a peer to peer network is to use content-based addressing instead of the traditional location-based addressing. Most of the adopted protocols is inspired by projects like IPFS(Inter Planetary File System) by Protocol Labs, Ethereum blockchain by Microsoft and the decentralized web-like network Zeronet. Some of the main adventages that the project promises to provide is to overcome the problems of a client-server architecture like single point of failure, certain factors of force Majeure(i. e. Political or big tech compnay's censorship), privacy violation and so on. The primary challanges of such a file system adopting peer to peer architcture are several security vulnarebiltiy, data availabilty and storage of big file(s). The subsequent chpaters discuss the broad domains, problem statements, existing work and finally the propsed scheme in that order:

Chapter 2 Broad Domain	Chapter 3 Problem Statement	Chapter 4 Existing work	Chapter 5 Proposed Scheme
			
The broad domains section contains the core ideas, the project revolves around. It discusses: evolution of the Internet and the web, emergence of blockchain, design concepts and challanges for decentralized file system and web hosting.	The problem statements section explicitly states the probems of adopting centralized architecture in the internet ecosystem and hardles to overcome implement a decentralized system.	The existing work section discusses the previous work being done for developing different decentralized framework including some popular ones like Ethereum, IPFS(Inter Planetary File System) and Zeronet.	The existing work section elaborates on different aspects of the proposed project including project outline, project objectives(broad and specific) and project methodology.

The background of the slide is a complex, abstract network diagram. It features numerous nodes, represented by circles and squares of varying sizes and shades of blue and green. These nodes are interconnected by a dense web of thin, light blue lines, creating a sense of a global or distributed network. The overall color palette is a gradient of blues, from light sky blue at the top to a darker, more saturated blue at the bottom.

## Chapter 2: Broad Domains

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*



## Chapter 2: Broad Domains

Subsections: **Internet, Blockchain, Decentralized file system**

### 2.1 Internet & WWW

Internet with a capital "I" is defined as to global connection of all the interconnected computer network to provide numerous remote services using some predefined communication protocol. But, Internet only refers to the physical infrastructure. The wondrous services internet provide today is only possible for the invention of the web or more formally known as WWW is the huge assembly of information and the protocols to store those information. The WWW was invented by English computer scientist and engineer Sir Timothy John Berners-Lee. The WWW was based on the proposal that was created on his March of 1989 entitled "Information Management: A Proposal" which was originally a concern for the management of general information about accelerators and experiments at CERN.

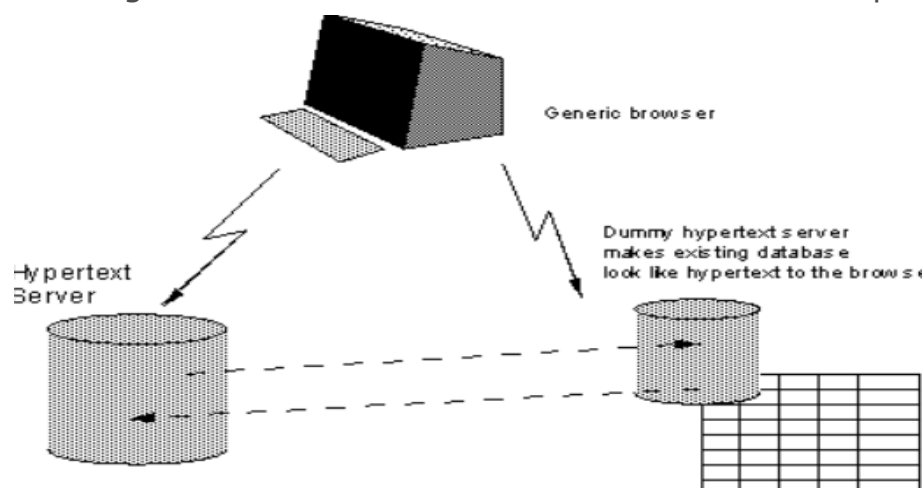


Figure 1: Global Hypertext System proposed by Tim Berners-Lee

source: <https://www.w3.org/History/1989/proposal.html>

He basically tried to solve the problem of lack of flexible in information access. The idea he tried to implement was called "hypertext" Human-readable information linked together in an unconstrained way. The core specific requirements given that time was remote access across networks, heterogeneity, non-Centralisation, access to existing data, private links (basically referring to privacy). These were the main principles while making the first version of the internet that we can call Web 1.0. At this infant state, the web was not widely adopted but slowly growing size and gaining popularity. After a certain threshold a new concern arose that was not mentioned in the original paper and that was security. Through this feature and some technological advancements more and more widely adopted web services started to emerge through the internet. New business model started with beautiful application interface to form and internet began to touch every single part of our

The background of the slide is a solid blue color. Overlaid on this is a complex, abstract network diagram. It consists of numerous small, dark blue circular nodes of varying sizes. These nodes are interconnected by a dense web of thin, light blue lines, creating a mesh-like structure. Some nodes are highlighted with a slightly different shade of blue or are surrounded by a faint, larger circular glow. The overall effect is a sense of a vast, interconnected digital space or a complex data network.

## Chapter 2: Broad Domains

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

life. This version of the web can be referred to as Web 2.0 that we are currently in right now. But there is currently some problem that are being noticed in the Web model 2.0. That are violation of decentralization of information and privacy which were actually addressed in Tim Berners-Lee's original vision of the internet. The company that provides the services now control the data like when we use services from google, facebook, twitter. These companies store our data in a centralized manner. This also in turn raises more security vulnerabilities. So a new vision of A Web 3.0 is being discussed to empower consumers to make them self-sovereign so that they get control over their own data but still be able to enjoy the versatile services. So Web 3.0 ensures self-sovereign money, self-sovereign data, self-sovereign identity, online privacy etc. The technical basis to achieve this Web 3.0 master vision is to introduce robust peer-to-peer framework one of them which is blockchain.

## 2.2 Blockchain

The blockchain is a distributed, shared and effectively immutable ledger that is designed through a chain of blocks in a secure way using cryptographic hash.

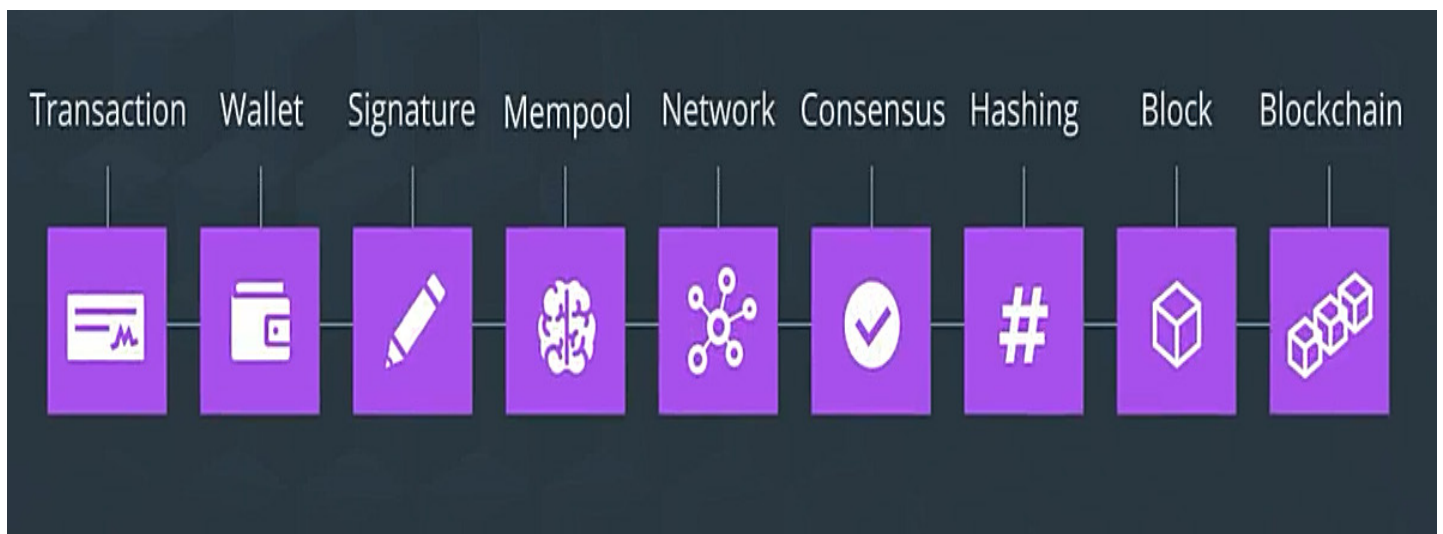


Figure 2: Different components of Blockchain

*source: Udacity Blockchain Developer Nanodegree Program*

As a metaphor we can consider using a blockchain in a banking system. In a traditional banking system people's transaction ledger is maintained in a centralized system. They have complete control over what to do with it and chose whether or not to share with anyone. In a blockchain based system instead of storing the information by a central entity, the ledger is distributed among all users with the same shared copy and whenever a new transaction is made each of the shared copy is updated. This approach has some potential advantages over the previous framework. Firstly,

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, suggesting a distributed system or a network topology. The overall aesthetic is technical and modern.

## Chapter 2: Broad Domains

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

because there is no central storage there is no single point of attack. Then, there is disintermediation meaning for transitioning money to another person we do not have to trust a third party like Visa, Master Card, Bkash or the bank itself for that matter. So we do not have to give our personal info or social security number to others. Also in a centralized bank system things can get complicated when they start handling a lot of transaction which results into delayed transaction time and also additional fee. So blockchain manages to send money faster and cut down the middle-man fee. In this context, we are using a ledger to list a list of transaction because it helps to solve the double spending problem. Double spending problem is when the same money is spent more than once. In actuality the ledger comes in the form of a digital document in blockchain. Blockchain groups, validate and link transactions in a fixed size block. A single block will also contain some meta data like timestamp, nonce (related to hash). When enough transactions are created a new block is made and chained to the previous block. This chain of block is shared among a peer to peer network. But there might be a security concern that someone might tamper with the data within the blockchain. This is where hash function comes in. A block in a blockchain does not only contain the data but a hash of the data (Actually each transaction is stored hashed and those hashes are hashed together from a Merkle Tree). Hash is just a function that takes data of any size and produces a unique fixed sized data. The hash is generated along with an arbitrary number called nonce which makes sure the hash value starts with a certain number of leading zeros. The number of leading zeros corresponding to the difficulty level of the hash. The very first block that is created in a blockchain is called a genesis block. Each block after that has an additional information of storing the hash of the previous block. So this has the advantage of, if someone tries to tamper a block like add or update or delete a transaction the hash of the block changes and the subsequent blocks get invalidated and the change is rejected. But now there should be a mechanism for adding a valid transaction. The rules to add a valid block in a blockchain is called consensus algorithm. Before getting validated the unconfirmed transaction gets stored in a physical storage called mempool. From there they get validated using one of the different consensus algorithms. Some of the consensus algorithms are based on proof of work which was first proposed and implemented in blockchain. It's where there is another entity called miners whose job is to validate transaction to get reward. But two major problems for this approach are huge computation power is needed to validate transaction and there is possibility of miner's monopoly. So a non-miner based consensus algorithm is also developed. Like proof of stake, where it seeks to achieve consensus by voting

The background of the slide is a solid blue color. Overlaid on this is a complex, abstract network diagram. It consists of numerous small, dark blue circular nodes of varying sizes. These nodes are interconnected by a dense web of thin, light blue lines, creating a mesh-like structure. Some nodes are highlighted with a slightly different shade of blue or are surrounded by a faint, larger circular glow. The overall effect is one of a digital or technological network, possibly representing a blockchain or a distributed system.

## Chapter 2: Broad Domains

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*



giving votes to those who have some stakes in the system. Microsoft's etherium uses this mechanism. Another popular one which also does not require any miner is Delagated Byzantyn Fault Tolarence(DBFT) algorithm which is based assigning roles to nodes to help coordinate consensus.

In terms of transparency blockchain can be of three types: private, public and hybrid. Public blockchains are open and permissionless allowing anyone to participate like bitcoiin blockchain. On the other hand private blockchain has lack of transparency as it can not be seen by t he general public. Although this has the advantage of efficiency in terms of scalability and compliance with regulatory requirements, but has more security vulnarability due to central governance(i. e. Hyperledg-er). Finally, a hybrid blockchain is the combination of public and private blockchain, used by an individual entity that reaps the benefits of both approaches.

## 2.2 Decentralized file system and web hosting

A file system is the entity in a computing ecosystem which controls how data is stored and retrieved. For the purpose of file serving, sharing, editing over the internet a centralized approach had been preferred from the begining meaning everything connects to a central

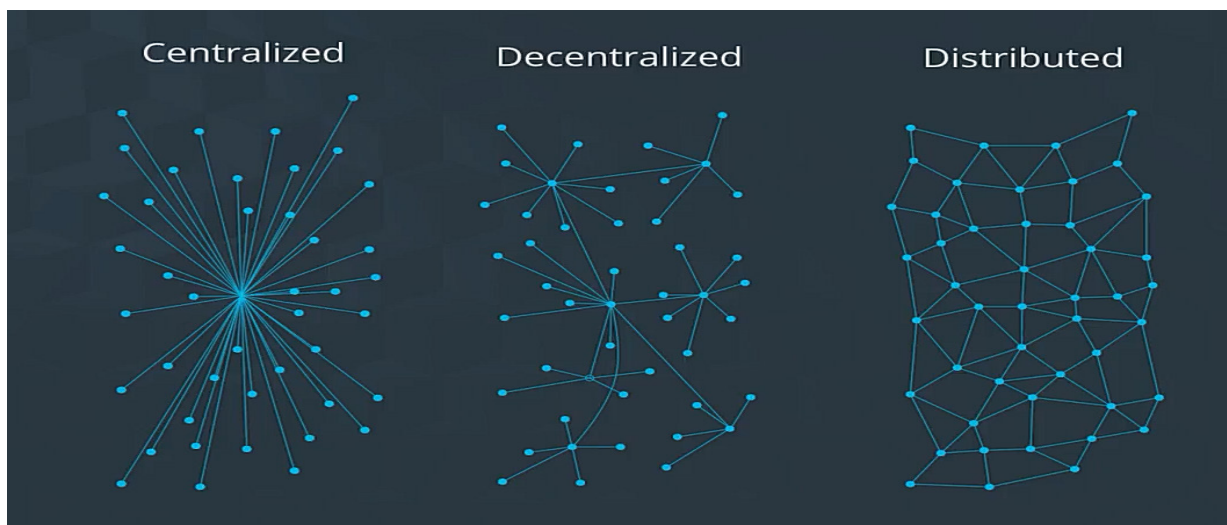


Figure 2: Different types of network

*source: Udacity Blockchain Developer Nanodegree Program*

network owner. The owner could be a person, company or database. But it has a huge downside. That is, the if the central server has to serve a lot of end-user, it fails when the demand for the file is really high even when the server is really powerful. And when the demand is low the server resources are wasted. The same can be said to be true, in case of centralized web hosting. So a more decentralized approach can be considered. A decentralized network is where information(whether file or website) is spread accross multiple nodes. Finally, in a distributed system everyone gets a copy of the information. While blockchain

The background of the slide is a complex, abstract network diagram. It features numerous nodes, represented by circles and squares of varying sizes and shades of blue and green. These nodes are interconnected by a dense web of thin, light blue lines, creating a sense of a global or distributed network. The overall color palette is a gradient of blues, from light sky blue at the top to a darker, more saturated blue at the bottom.

## Chapter 2: Broad Domains

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*



maintain a fully distributed network because the small size of the blockchain. But when it comes to file system and/or web hosting a peer to peer decentralized approach is preferred because of the number and size of files and websites. Through this decentralized approach we can reduce cost and increase throughput. But one potential integrity problem is in a decentralized system a peer node can tamper with a data. This problem can be solved using content based addressing instead of the traditional location based addressing while searching for file/website. When we use a search engine in our browser it retrieves file and sites based on its location as DNS server holds the location of the content we are looking for and try to match it by the stored location path. Content based addressing on the other hand, stores hash of the stored file or web content. That way, we can check the hash of our retrieved file to verify it has not been tampered with. So, security is built-in in this approach.

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, with some lines crossing each other. The overall effect is a sense of a global or distributed network.

## Chapter 3: Problem Statements

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

## Chapter 3: Problem statements

Subsections: **Problems with Centralized Internet , Problems with implementing decentralized web ecosystem over the internet**

### 3.1 Problem Statement 1: Problems with centralized internet

#### 3.1.1 Servers Can Go Down

When a satellite failure in 1998 made 90 percent of all the pagers in the U.S. stop working, one of the largest impacts was on blue collar freelance professions — like plumbers and carpet cleaners — who could no longer be contacted for new gigs. But now that we're even more dependent on the internet, servers that malfunction or are targeted by bad actors could wreak even more havoc across swathes of the world.

#### 3.1.2 Servers Can Be Hacked

Over the past few years hacking is on the mainstream news almost on a regular basis. Centralized server makes way easier for hackers to hack as it gives them a central target. Hacks on truly massive scales have left hundreds of millions (and in a few cases, billions) of internet users hapless to stop their data from being stolen. Worse, most of these companies are in "too big too fail" mode. They will often either hide their data breach from consumers for years — as Yahoo did with its 2013 breach — or they will remain vague about just what type of data was leaked. Even more recently in 2017, the Equifax the Consumer reporting agency company's server was hack and almost half of the entire US population's credit card number was compromised.

#### 3.1.2 Companies and Government Throttle or Censor Data

The uproar about the loss of net neutrality came down to centralized internet. Internet providers that have enough power over the pipelines used to deliver internet to households have a way to limit the amount of data they delivered so that users would be willing to pay more. According to a survey from ExpressVPN, almost 75 percent of Americans disliked the idea of an internet service provider failing to treat all their data the same. But since those providers are the only gatekeepers to the internet, no market forces are stopping them from sticking a finger on the scale. This problem allow big company to grow more and more and hinders the growth of startups in the same field. One practical example of this unintended incident is the allegation against AT&T as they took bribe from Netflix for better bandwidth than its competitor.

Another problem can censor data in a central system for political and other reasons. Example of this can be Turkey's ban on wikipedia, China's prohibition of Google and

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, suggesting a distributed system or a network topology. The overall aesthetic is technical and modern.

## Chapter 3: Problem Statements

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

Facebook and more recently Microsoft's github banned private repository access from Iranian users in compliance with the US government policy.

### **3.1.4 Companies Monetize Our Data**

All the lengthy forms we never read before clicking "I have read and agree to these terms"? allows massive companies who store our personal data on their servers are not giving us a free account: They're making their money off of selling our data to advertisers who can target us better. In some cases, they are not even doing it legally: A German court ruled in 2018 that Facebook has been illegally collecting data in a breach of consumer law. Because of lack of their privacy concern of end-user facebook lost over 120 billion dollar in share.

## **3.2 Problem Statement 2: Problems with decentralized web ecosystem over the internet**

### **3.2.1 Mass Adaption**

Decentralized internet is the sort of idea that works great once everyone is on board with using it. But until then, the sheer fact that no one is using something tends to function as a Catch-22: No one wants to use it because no one is already using it. As a general rule of thumb, people don't change until the process of changing becomes less painful than sticking with the status quo. And since incumbent centralized networks like Facebook, Twitter and other centralized services already have everyone on them, they have a level of stickiness that's hard to beat.

### **3.2.2 Availability**

This kind of links to the previous problem. If less people are engaged in the service the service will be slow or interrupted. Incentivizing users is a major issue.

### **3.2.3 Serve big file**

The existing decentralized projects mainly struggle with this problem to distribute and maintain files of large size. The protocol designed for distributing these files is complicated and have several challenges.

### **3.2.4 Users Do not Want the Extra**

A system of peer-to-peer protocols would shift the responsibility of maintenance away from the owners of the servers and on to each and every user of a decentralized internet. Which is great for "tech-savvy and politically inclined users" who know how it works and care about their privacy, as MIT pointed out in a report quoted in a Medium article on decentralized networks. But that is hard to explain to everyone.

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, with some lines crossing each other. The overall effect is a sense of a global or distributed network.

## Chapter 4: Existing work

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

## Chapter 4: Existing Work

The emergence of the internet has solved a lot of problems of the world like data availability, data distribution, communication, etc. But today's infrastructure of the internet does not reflect the decentralization, heterogeneity, and privacy of consumers that was originally proposed by the founder of internet Tim Berners Lee, in his 1989 www whitepaper[1]. This creates many challenges such as central point of failure, the overhead of intermediation leading to extra cost both financially and with time, fraud, privacy violation, censorship by political and other entities, etc. One of the approaches that can be taken is to switch from client-server architecture to a peer-to-peer architecture. However, this leads to new technical challenges including some new security vulnerability. A promising technology in this regard could be what was proposed in 2008 called blockchain by the name Satoshi Nakamoto proposed using timestamps of transactions by hashing them into an ongoing chain of hash-based proof-of-work. [2] This idea was implemented to a cryptocurrency in bitcoin but later integrated to other versatile areas. In the field of file storage blockchain technology promises a huge improvement over the traditional centralized file storage system because of no single point failure, disintermediation, low cost, high bandwidth along with the increased security[3]. Additionally, its been proven to better handle the privacy of both the data owner and data consumer[4]. Some blockchain based distributed file storage systems with no centralized authority are Storj[5], Sia[6], IPFS(Inter Planetary File System) & Filecoin [7-8], Swarm[9], etc. Storj and Swarm are deployed on the Ethereum blockchain network [10]. Storj provides an end-to-end encryption approach, and stores a cryptographic hash fingerprint of the file on the Ethereum blockchain while providing a method of verifying file integrity. The consensus protocol used in this system is proof of work so miners are involved in creating new blocks in the blockchain. Although proof of work is run by the successful bitcoin platform, it has its downsides of too much power consumption and minor monopoly problem. Swarm on the other hand based on distributed. swarm token which is a standard ERC20 token (on the Ethereum platform) that allows one to use the Swarm software platform. It uses proof of stake protocol for incentivizing transaction in the blockchain. The SIA project forms storage contracts which are agreements between a storage provider and their client, specifying what file will be stored and at what price. The data are segmented and encrypted separately as ciphertext. The SIA coin is used as the cryptocurrency and user uploads file periodically as a proof of storage to prevent the storage node from deleting the stored file.

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, with some lines crossing each other. The overall effect is a sense of a global or distributed network.

## Chapter 4: Existing work

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*



The IPFS is more of a generic peer to peer file system with the potential to replace the HTTP protocol itself with secure file distribution with the usage of hashing each uploaded file. For incentivizing the users in getting involved in contributing to the IPFS network it also developed a native crypto coin called filecoin. Although by itself IPFS does not provide powerful cryptographic privacy protection like another peer to peer file system. Normally, in a traditional blockchain application, use asymmetric encryption with both public and private key. Usually, the private key is generated using cryptographically secured random number generator algorithm and importantly required no to be shared with anyone. From the private key, using elliptic curve multiplication public key is generated[11]. The problem is the secrecy of the private key over time. Another problem is, with a huge quantity of encrypted files, to find the required file at a reasonable time for the data consumer. Both of this issues can be tackled using Cipher-Policy Attribute-Based Encryption(CP-ABE) scheme with the secret key is related to attribute set, the ciphertext is related to specific access policy providing fine-grained access control for data consumer and also with multi-keyword search and supporting attribute revocation.[12]. The CP-ABE scheme was first proposed by Bethencourt et al. [13], but his proposed framework was not fully cryptographically secured. The system was later evolved to pass the complexity assumption of Decisional Diffie-Hellman Assumption by Li et al.[14]. Traditionally, this scheme for CP-ABE with multi keyword search is adopted by a centralized cloud storage system. Later, S. Wang et al were the first to develop such a scheme for decentralized storage systems.[15]. In addition to distributed file storage, peer to peer network is also being used in the distributed website and web application hosting.[16]. A practical example of this distributed hosting is Zeronet.[17].

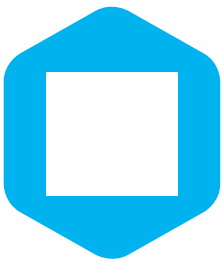
The background of the page is a solid blue color. Overlaid on this is a complex, abstract network diagram. It consists of numerous small, dark blue circular nodes of varying sizes. These nodes are interconnected by a dense web of thin, light blue lines, creating a mesh-like structure. Some nodes are highlighted with a slightly different shade of blue or are surrounded by a faint, larger circular glow. The overall effect is a sense of a global or distributed network, typical of blockchain or web hosting architectures.

## Chapter 5: Ptoposed Schema

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

## Chapter 5: Proposed Scheme



### Subsections

● Project Outline	subsection 5.1
● Project Objectives	subsection 5.2
● Project Methodology	subsection 5.3



### 5.1 Project Outline

#### 5.1.1 Description

The framework proposed here is to develop a peer to peer file distribution system with a web interface also hosted in a decentralized manner. There are two blockchains to be maintained one is a hybrid blockchain for the structure of the distributed file system and another public blockchain for the deployment of the smart contract. One smart contract is one block on the chain will represent one service. A single smart contract will correspond to a separate blockchain for a distributed file system. A single entity can create their blockchain for providing service under a specified name. The smart contract can be created and deployed using a web interface instead of a programming language and command line interface like in the Ethereum Ecosystem. In the system, the service provider can set up the ABE scheme for data consumer, application type, etc in the smart contract. There are two types of node in the network data/service provider node and data/service consumer node that would be called a provider node and consumer node respectively. The provider node is able to create services and both the provider and consumer node plays a role in serving according to the primarily defined smart contract. Beyond this, there is no discrimination among peers in terms of network bandwidth, anonymity and/or ownership as equity is built-in to the smart contract. So, trust is embedded in the network. Getting registered to service is as easy as browsing to the distributedly hosted web application although privacy features like hiding IP address is maintained through the smart contract. Accessing the file system require registration which is secured adopting bilinear pairing using elliptic curve digital signature algorithm(ECDSA). The process for registration for provider node and consumer node is defined at the smart contract level. The possible services that can be provided are public file sharing, end-to-end file sharing, file searching, upload, and download file, etc.

The background of the page is a solid blue color. Overlaid on this is a complex, abstract network diagram. It consists of numerous small, dark blue circular nodes of varying sizes. These nodes are interconnected by a dense web of thin, light blue lines, creating a mesh-like structure. Some nodes are highlighted with a slightly different shade of blue or are surrounded by a faint, larger circular glow. The overall effect is a sense of a global or distributed network, possibly representing a blockchain or a web hosting infrastructure.

## Chapter 5: Ptoposed Schema

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

### **5.1.2 Data collection**

Data related to the IPFS scheme and standards regarding ABE technology, blockchain underlying cryptography, consensus protocol. These are easily available on the internet as open source projects and research papers.

### **5.1.3 Design**

The design decision is needed to be made regarding block metadata and consensus protocol in both the blockchain. For the smart contract blockchain, a mix of Delegated Byzantine Fault Tolerance(DBFT) and proof of concept would be used as the consensus algorithm. The distribution file system should follow the scheme of IPFS with Proof of Spacetime and Proof of Replication consensus protocol although unlike IPFS would be designed on top of the HTTP protocol. The other design works involved are in the area of smart contract web framework, registration scheme using ABE and bilinear pairing, the user interface for offered services.

### **5.1.4 Implementation**

According to design, the distributed file system, smart contract blockchain, web interface for system setup and UI for the offered system is to be implemented.

### **5.1.5 Analysis**

The implemented system is to be analyzed regarding the feasibility of deployment in real life scenario, performance, data availability, and integrity.

### **5.1.6 Review**

After proper analysis, the system is to be reviewed to find out any design or implementation change is needed and impose those changes accordingly

### **5.1.7 Report**

The report is to be prepared with the literature review, steps taking in building the system in-detail, result obtained and future scope of the research.

### **5.1.8 Advice and policy formulation**

The introduction of Blockchain technology creates new challenges for both representatives at small-to-large organizations and the decision makers at the regulatory agencies. This is mainly because blockchain technology itself is as revolutionary as entering the next stage of the internet. But the open and decentralized nature of blockchains and the absence of unipolar governance means that regulatory issues need deeper analysis than previous technologies. [18] Core blockchain fundamentals like smart contracts, consensus protocols are needed to be aligned with the local and global legislation. The emergence of these newly found concepts will inevitably outdated the existing

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, suggesting a distributed system or a complex network structure. The overall aesthetic is technical and modern.

## Chapter 5: Ptoposed Schema

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

laws and regulation of the global and local governing agencies in many cases. This, in turn, means either enter new policy is to be thought and deployed or provide an adequate understanding of the newly adopted technology for the people involved (i. e. through training) without alteration of the existing laws. In deploying the proposed framework these points are needed to be noted. In addition, the framework avoided using cryptocurrency for incentivizing transactions in the blockchain for adopting the system in government agencies where cryptocurrency is banned and the person who adopts this system is discouraged to integrate digital currency into the system.



## 5.2 Project Objectives

In this subsection, it will discuss the key goals of the proposed research work, both the broader objectives and the specific ones.

### 5.2.1 Broad Objectives

The principal objective of this research is to develop a more non-centralized version of the internet that was proposed in the first place[1] by leveraging the potential of blockchain with the primary focus on a distributed file system that largely outperforms the traditional centralized system of file distribution.

### 5.2.2 Specific objective

- ☐ Develop a decentralized web-hosting scheme.
- ☐ Build a distributed file system on the internet following the principles of IPFS[7].
- ☐ Use blockchain to ensure secure file transfer, secure web-surfing, secure user identity.
- ☐ Content-based addressing instead of the tradition location-based addressing when it comes to surfing the internet.
- ☐ Ensure proper anonymity of the web-surfer.
- ☐ Cut cost in serving more people both in terms of finance and time by instead of buying more powerful servers, leveraging the power of a peer to peer network.
- ☐ Remove intermediation, thereby cutting cost and saving time.
- ☐ Prevent fraud on all levels.
- ☐ Apply smart contract to prevent any kind of censorship.
- ☐ Utilize non-centralization to avoid force Majeure factors i. e. natural disaster.

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent, while others are smaller. The connections form a dense, interconnected web, suggesting a distributed or decentralized system. The overall aesthetic is technical and modern.

## Chapter 5: Ptoposed Schema

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*



- ☐ Make a peer to peer system that could scale from small startups to big organizations.
- ☐ Stay aligned with the possible use cases of deploying in a government agency specifically for the people's republic of Bangladesh.



## 5.3 Project METHODOLOGY

In this section, it would discuss the scientific methods that will be used in every process of the research.

### 5.3.1 Data Collection

All the data that would be collected for conducting the research are from the secondary data sources. The knowledge of the open source project that is acquired is mostly from their official whitepaper or presentation derived from the respective official website. The research papers that were consulted were either from the "IEEE Access" the official digital library of IEEE only be accessed by an IEEE member or from the respective university website.

### 5.3.2 Design

The prototype for the proposed framework will be designed using proto.io[19] a fully interactive web platform for designing high-fidelity prototype. The prototype design will be based on the knowledge gathered from open-source projects i. e. IPFS[7] and research papers. Also, the scalability of the system has to be kept in mind using techniques like a smaller chain, big blocks, compressing data on the blockchain, layering in blockchain(i. e. plasma in Ethereum)[22], sharding, improving consensus protocol.

### 5.3.3 Implementation

For implementing the backend and frontend of the proposed system this research will use Google's dart[20] and dart framework flutter[21] along with the language of the web HTML, CSS, and javascript. The Integrated Development Environment(IDE) used will be Visual Studio Code.

### 5.3.4 Analysis

For analyzing the built framework a checklist is to be made that covers the minimum requirement to deploy it in an actual network. the checklist should highlight the performance, security and privacy factors. The methods adopted to test each point of the checklist should be well defined. A draft checklist is presented below:

- ☐ The system works as it was designed to work:
  - Decentralized web hosting
  - Distributed file system.

The background of the slide is a solid blue color with a complex, abstract network diagram overlaid. The diagram consists of numerous small, dark blue circular nodes connected by thin, light blue lines. Some nodes are larger and more prominent than others, and the connections form a dense, interconnected web that suggests a distributed system or a blockchain network. The overall aesthetic is technical and modern.

## Chapter 5: Ptoposed Schema

SPL3 Project proposal

*A blockchain based web hosting scheme for  
content-addressed distributed file system*

- ☐ Smart contract(s) works as it was intended
- ☐ File access time meet a certain benchmark
- ☐ Privacy criteria(i. e. IP address hidden) met

### **5.3.5 Review**

According to the result of the analysis, any alteration of the design choice and/or implementation would be considered to meet certain criteria that were not met. Again, the system would be tested against the created checklist. This would be an iterative process.

### **5.3.5 Report**

Finally, the report is prepared following the IEEE standard using IEEE latex template.

The background of the slide is a complex, abstract network diagram. It features numerous nodes of varying sizes and colors (dark blue, light blue, green, and yellow) interconnected by a dense web of thin, light blue lines. Some nodes are highlighted with larger, semi-transparent circles. The overall aesthetic is technical and digital, suggesting a network or blockchain theme.

## References

SPL3 Project proposal  
*A blockchain based web hosting scheme for  
content-addressed distributed file system*

## References

- [1] Tim Berners-Lee, "Information Management: A Proposal", in CERN March 1989 [Online] Available: <https://www.w3.org/History/1989/proposal.html>
- [2] Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", on October 31, 2008, [Online] Available: <https://nakamotoinstitute.org/bitcoin/>
- [3] Hoang Giang Do, Wee Keong Ng, "Blockchain-Based System for Secure Data Storage with Private Keyword Search", in 2017 IEEE World Congress on Services (SERVICES) [Online] Available: <https://ieeexplore.ieee.org/document/8036727>
- [4] K. G. Zyskind, O. Nathan, and A. S. Pentland, "Decentralizing privacy: Using blockchain to protect personal data" in Proc. Secure. Privacy Workshops (SPW), May 2015, [Online] Available: <https://enigma.co/ZNP15.pdf>
- [5] S. Wilkinson, T. Boshevski, J. Brandoff, and V. Buterin, "Storj a peer-to-peer cloud storage network," [Online]. Available: <https://storj.io/storj.pdf>
- [6] David Vorick, Luke Champine, "Sia: Simple Decentralized Storage" in Nebulous Inc., November 29, 2014 [Online] Available: <https://sia.tech/sia.pdf>
- [7] Juan Benet, "IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3)" [Online] Available: <https://ipfs.io/ipfs/QmV9tSDx9UiPeWExXEeH6aoDvmihvx6jD5eLb4jbTaKGps>
- [8] "Filecoin: A Decentralized Storage Network" in Protocol Labs, July 19, 2017 [Online] Available: <https://filecoin.io/filecoin.pdf>
- [9] Philipp Pieper, Timo Lehes, "SWARM FUND White Paper The Blockchain for Private Equity" in March 2018 [Online] Available: <https://docs.swarm.fund/swarm-whitepaper-eng.pdf>
- [10] G. Wood, "Ethereum: A secure decentralized generalized transaction Ledger", [Online] Available: <https://gavwood.com/paper.pdf>
- [11] Bos, Joppe & Alex Halderman, J & Heninger, Nadia & Moore, Jonathan & Naehrig, Michael & Wustrow, Eric. (2014). Elliptic Curve Cryptography in Practice. 8437. 10.1007/978-3-662-45472-5\_11. [Online] Available: [https://www.researchgate.net/publication/274651795\\_Elliptic\\_Curve\\_Cryptography\\_in\\_Practice](https://www.researchgate.net/publication/274651795_Elliptic_Curve_Cryptography_in_Practice)

The background of the slide is a complex, abstract network diagram. It features numerous nodes, represented by circles and squares of varying sizes and shades of blue and green. These nodes are interconnected by a dense web of thin, light blue lines, creating a sense of a global or distributed network. The overall color palette is a gradient of blues, from light sky blue at the top to a darker, more saturated blue at the bottom.

## References

SPL3 Project proposal  
*A blockchain based web hosting scheme for  
content-addressed distributed file system*

- [12] Wang S, Yao L, Zhang Y (2018) Attribute-based encryption scheme with multi-keyword search and supporting attribute revocation in cloud storage. PLoS ONE 13(10): e0205675. <https://doi.org/10.1371/journal.pone.0205675> [Online] Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0205675>
- [13] Bethencourt J, Sahai A, Waters B. "Ciphertext-Policy Attribute-Based Encryption" in 2007 [Online] Available: <https://ieeexplore.ieee.org/document/4223236>
- [14] Jiguo Li, Haiping Wang, Yichen Zhang and Jian Shen, "Ciphertext-Policy Attribute-Based Encryption with Hidden Access Policy and Testing," KSII Transactions on Internet and Information Systems, vol. 10, no. 7, pp. 3339-3352, 2016. DOI: 10.3837/tiis.2016.07.026[Online]Available: <http://www.itiis.org/digital-library/manuscript/1408>
- [15] Shangping Wang, Yinglong Zhang , Yaling Zhang, "A Blockchain-Based Framework for Data Sharing With Fine-Grained Access Control in Decentralized Storage Systems" [Online] Available: <https://ieeexplore.ieee.org/document/8400511>
- [16] Svebor Prstačić ; Mario Žagar, "A model for web application and web service peer-to-peer hosting network architecture" in 2013 [Online] Available:
- [17] "ZeroNet: Decentralized web platform using Bitcoin cryptography and BitTorrent network"Presentation,[Online]Available: [https://zeronet.io/files/ZeroNet\\_Presentation.pdf](https://zeronet.io/files/ZeroNet_Presentation.pdf)
- [18] Jason Potts, "Public Policy for Cryptocurrency and Blockchain Technology", [Online] Available: [www.ippapublicpolicy.org/panel/pdfPanel.php?panel=225&conference=7](http://www.ippapublicpolicy.org/panel/pdfPanel.php?panel=225&conference=7)
- [19]"Proto.io-Prototypethatfeelreal",<https://proto.io/>,lastaccessed:10:55AM6/9/2019
- [20]"Dartprogramminglanguage|Dart",<https://dart.dev>,lastaccessed:12:26PM6/9/2019
- [21] "Flutter - Beautiful native apps in record time",  
<https://flutter.dev>, last accessed: 12:26 PM 6/9/2019
- [22] "Understanding Plasma | Ethereum Scaling | State Channels vs. Plasma",<https://education.district0x.io/general-topics/understanding-ethereum/understanding-plasma/> last accessed: 2:38 PM 6/9/2019

