LunarX: Universal Middleware to DApps, Part II

LunarX Founding Team

neo.carmack@lunarion.com
ben.x@lunarion.com
feiben@lunarion.com
https://github.com/LunarX-ONE/White-Paper

Art. No II
ISBN xxx–xx–xxxx–xx–x
Edition 0.6.6

Cover design by Designer: Typesetters.se

Published by neo.carmack@lunarion.com
ben.x@lunarion.com
feiben@lunarion.com
https://github.com/LunarX-ONE/White-Paper
Printed in Beijing, China

Contents

1. Backgrounds

1.1 Problem

With the continuous introduction of internet applications, the internet has been expanding to Asia, Africa and Latin America, and the number of users has also been increasing. But even with the exponential growth of internet in the past decades, there are still 3.5 billion people around the world that do not have access to internet. All these means more data service are needed.

Centralized data platform is today's mainstream data service solution. As the need expands, more and more data centers are put into operation, often at very high costs. At the same time, concerns regarding reliability, security, and privacy have also become more pronounced over the last few years.

To address these shortcomings, the science and technology community has advanced decentralized applications (DApp) that run on peer-to-peer (P2P) networks. However, until now, there is no easy-to-use solution to process a large amount of data in a truly reliable, scalable, flexible and cost-efficient manner.

1.2 Centralized Data Platform

Costs

Data centers are very costly. Since end users will eventually pay for these costs one way or the other, it may worth examining these costs briefly.

Industrial data center is a complex system containing thousands of interdependent devices. Data center operators need to update and optimize the temperature, humidity, air pressure, power usage, fan speed, resources (CPU, Memory, HDD IO) utilization in real time. Therefore, merely the construction and maintenance budget of a Tier-3 data center is usually more than $30 million, evaluated over a period of three years.

	BUILD YOUR OWN
Construction Cost	$24,500,000.00
Floor Space	$300,000.00
Electricity	$5,132,160.00
Electrical Maintenance	$1,298,781.69
HVAC Maintenanance	$145,824.00
Other Systems Maintenance	$75,816.22
24x7x365 Staff	$ -
Redundant Internet Circuits	$201,600.00
	$31,654,182

a tier 3 data center	
Number of Cabinets	100
percentage of Power Consumed	90
total kW of Reduncant Power	1000
Cost Per kWh in Cents	11
Internet Connection in Mbps	100
$24*7*365$	*yes*
Evaluation Period in Years	3

The cost is calculated via https://www.expedient.com/data-center-build-vs-buy-calculator/.

However, not all the data center capacity are utilized efficiently. On one hand, because it is extremely expansive to add marginal capacity, enterprises prone to purchase more capacity than needed, and thus spending money that could have been used to drive business growth. On the other hand, to safeguard data from damage, data centers demand live backup at multiple locations, which makes it even more expensive for enterprises.

Waste

Even in healthy public cloud data centers, at least 30% cloud spend is wasted, due to users' over application for cloud computing resources. This is because users do not have the perfect understanding of network speeds provided by data center in various regions, and therefore often choose expensive data centers and obtain computing resource far exceeding the real requirements.

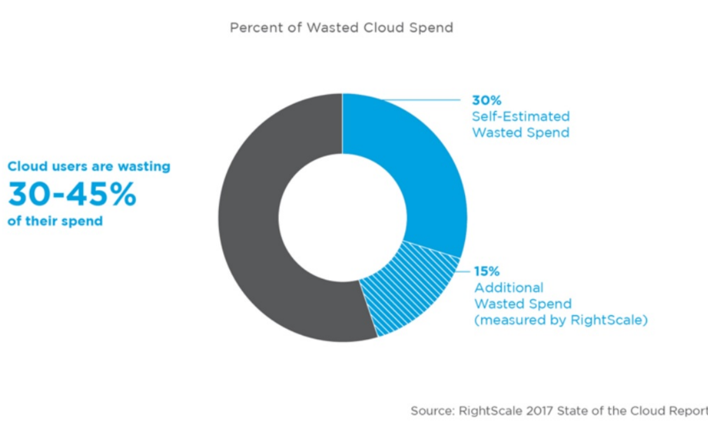


Figure 1.1: Percentage of Waste

Reliability

Even with all the redundant capacity, data center are not fail-proof. A 2016 study by the Ponemon Institute found that unplanned outages for just 63 data centers in one year costed more than 43 million. The average cost of data service outage has reached nearly $10,000 per minute, more than 200% higher than 2010. (Ponemon Institute 2016)

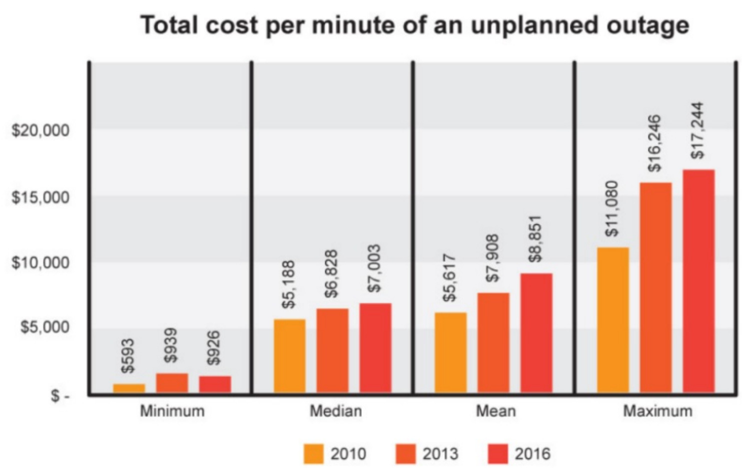


Figure 1.2: cost per minute of an unplanned outage

Source: Ponemon Institute

Security

Database of top enterprises are repeated hacked despite heavy investment to protect data, In 2016, data of 68 million Dropbox users were leaked for sale on darknet markets. In 2017, Equifax, one of the three credit bureaus in the U.S., was attacked and sensitive information for more than 143 million people were stolen.

With the increasing size of users and the increasingly complex network environment, any new function upgrades, or changes to old functions, are accompanied by ever-increasing security testing requirements, and increasingly inevitable risk of attacks.

Privacy

Privacy is one of the often mentioned concerns about centralized platform supported by data centers. The platforms not only have all the knowledge of their users' data, but also share users' identities and characteristics. To use their services, users have no choice but to trust these platforms.

When there is not enough transparency, not all, if any, business are trustworthy. There are plenty of examples where companies monitor and peep at users' data, advertising on users' behaviors and personal information, or even sell users' data directly to marketers.

1.3 Distributed Data Service and Blockchain

Flexibility and scalability

Since the birth of blockchain technology, solutions for distributed computing in an untrusted environment has advanced significantly. Blockchain is a decentralized sequential transaction database. Transaction data are data that describe token generation, transfer and banking. The global consistency of transaction data in blockchain is achieved through proof of work (PoW) or its variants, which prevent double spending, data tampering and other malicious alteration in the network.

However, most data generated in DApp are not about transactions. Nevertheless, they need to be tamper-resistant, traceable, anonymous, autonomous, and scalable. Although we can still put the data into blocks, generating blocks needs to be proved by work and can only be recorded in the blockchain database by global consensus. This is very resource consuming and not an efficient solution for distributed computing.

Alternative includes putting data into FileCoin, Sia, IPFS (InterPlanetary File System) or Storj, but they cannot provide services of traditional centralized data centers, such as structured data processing, vertical scaling, object storage, table storage, CDN, data retrieval, structured query, cache and plug&play with convenient interfaces. Therefore, they are not flexible and scalable for upper-tier applications.

PoW mechanism

By definition, distributed data services require PoW mechanisms to ensure useful storage and computation capacity are provided to complete tasks. This is called Proof of Useful Work, which is used to verify the good faith of users and service providers. For example, Filecoin uses Proof of Replication and Proof of Spacetime to verify useful storage of data are provided to users in a period of time. These PoW mechanisms are used to address Sybil attacks, Generation attacks, and Outsourcing attacks.

Since LunarX provides not only storage services but also many complicated data computing services. And what is more, vulnerabilities related to data computation services but not data storage, such as the Ciphertext-only Attack, are not addressed by current PoW mechanisms. Therefore, a fraud-proof PoW is urgently needed. (See Part I for details)

2. LunarX: Solution for Distributed Data Service

LunarX is a middleware designed to enable distributed data service. It provides a bunch of fundamental protocols that defines infrastructural services which are necessarily supporting the future value network. With its support, DApp can have each node in the network provides services to each other no matter where they are, storing encrypted data, responding to data requests, while earning tokens at the same time. By skipping data centers, LunarX significantly reduces the cost of service.

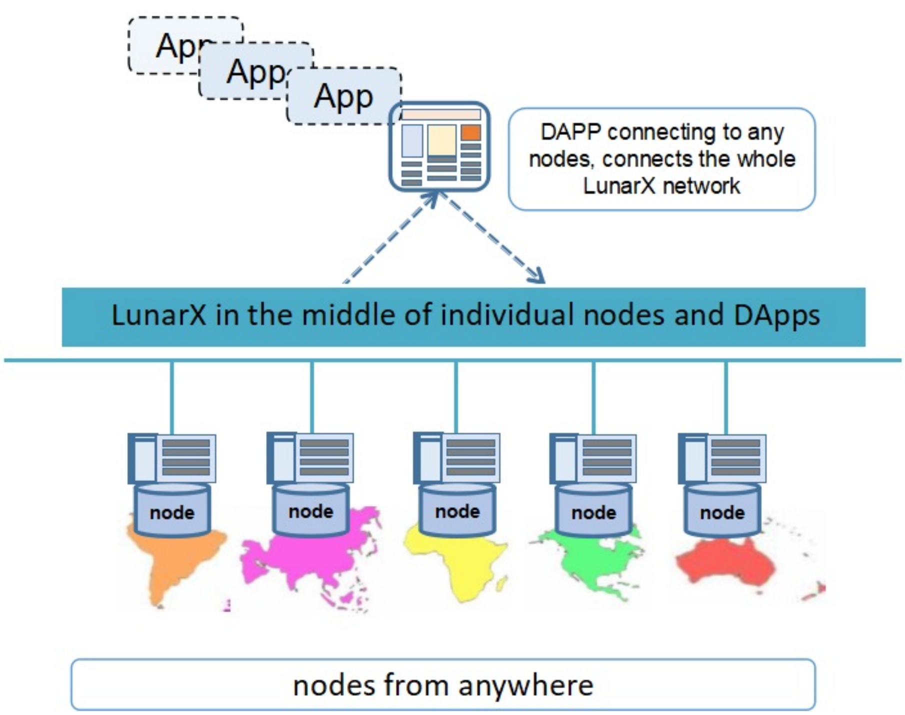


Figure 2.1: Middleware for Dapps

LunarX is a matrix, peer-to-peer network, and bridge.

As a globally sparse and locally dense matrix, LunarX uses this structure to manage the underlying data chain, providing services going beyond storing files like existing solutions. It can store structured data and perform data computation tasks to support information retrieval, structural data analysis, stream media publishing, content delivery, CDN, and more. It also has robust data self-recovery functionality through erasure coding, therefore protecting data from out-of-service nodes or malicious hackers.

As a peer-to-peer network, LunarX enables every node to provide data chain functionality to clients as a part of the whole data service system. Each node can voluntarily contribute to a useful work to earn tokens as return.

As a bridge, LunarX provides a set of standardized fundamental protocols to enable DApps to interact with blockchain and the underlying hardware. LunarX provides abstractions that free DApp developers from knowing the exact implementations of data services in a peer-to-peer network while maintaining transparency.

2.1 Vision

We can imagine the global blockchain network as one giant computer system. In this case, blockchains can be seen as a (ledger) storage and intelligent contracts can be seen as computing capability. Everything depends on a database with chained structure.

However, this system alone will not be sufficient to support the rapid growing needs. Due to the lack of middleware products with good scalability and flexibility, DApps usually develop their own systems to process data. For DApps to strive and utilize the full potential of the system, a uniform, secure, extensible, flexible, and decentralized data processing layer is needed to enable business logic to be implemented quickly.

LunarX fills this gap and offers fundamental building blocks for decentralized, infinitely-scalable, temper proof, anonymous, autonomous and resources efficient data infrastructure for DApps and their users. By doing so, LunarX will greatly reduce the costs of developing and operating DApps, and paves the way for a future where low cost, secure and privacy-protecting distributed data service is the norm.

2.2 Benefits

LunarX's decentralized architecture have the following benefits:

1)Reduce financial cost for data service

By replacing data centers with distributed data service can save resources such as cooling, land occupation and large-scale hardware. In fact, most personal computing tasks are not complicated. The complexity of the data center system lies in the fact that all the users' data are centralized, which requires scheduling, resource allocation, task contention and good hardware to support this computing. From the perspective of a single user, only a small amount of computing capacity is needed to complete relevant tasks.

For example, users may use online notes App only a few times a day. If 100,000 users use it at the same time, say 9 AM when they have their first meetings, the App needs to bear the pressure of data peak and configure good machine clusters to deal with it. However, with distributed data service, this kind of App would not cause data blocking, therefore reducing pressure on the system.

Users eventually pay for data services one way or another. By tapping into idle individual resources to provide services, LunarX helps DApp to avoid data centers and its associated expense. A fraction of the saved expenses would be paid to other individuals in the network for their service.

2)Reduce the cost of developing DApps

Because there is no data management standard, many DApps still manage their data centrally as only a part of transaction data are put into the blockchain. Developers urgently need a universal middleware system that allows them to customize data schema and structures, freely assemble various basic services, vertically scalable, and meet the requirements of storage, query, and retrieval. LunarX provides all these functionalities and it would make the development and deployment of DApp a lot easier and quicker.

3)Ensure data security and privacy

LunarX protects users' data via cryptography schemes, so the data would be safe from unintentional deletion or other more malicious intent. At present, technology communities have already developed many distributed online storage solutions, such as IPFS, Storj, etc. With LunarX, it will be much easier to build a distributed storage system.

2.3 Where to begin

Because User Generated Content (UGC) Apps generate most of the internet traffic, we start exploring decentralized architecture with this specific type of App. In order to ensure data security, autonomy and tamper-resistance, this architecture uses a large number of algebraic encryption commutative mapping principles, as analyzed and designed in the part I of this white paper.
The main functional modules we designed in the middleware include:

DHT: Dynamic hash table is used to address a specific resource on the network.,

object storage, file storage: It is very suitable for storage of large objects and needs to support various types, scalability, seamless and flexibility,

table storage: It is mainly used in application scenarios that have no schema and can scale out to store structured data,

naming: Naming rules for a particular computing resource,

archive: It is mainly used for cold data that are not used for a long time, such as massive historical documents and data that will not be used within a few months or many years. But such data still need to be retrievable,

replication: Multiple replications for data redundency, which are mainly implemented through erasure codes in LunarX,

CDN: Data from multiple nodes provides data acceleration services based on physical distance,

data retrieval: Mainly used for analysis of unstructured data, including retrieval of archives,

structured query: Analysis of structured data,

value exchange: used in online selling and buying assets.

These modules are necessary components for upper layer applications. Many of them have already been studied in the part I. In chapter 4, we will look at how these modules serve from the perspective of application development.

3. Business Scenarios

3.1 Miners who serve internet users

The concept of **miners** comes from bitcoin. In lunarX, miners no longer simply provide brutal hash computing as what they are doing in bitcoin, but provide useful data calculation, such as storage, structural analysis, mapping, point query, range query, etc.

in our context, **service providers** are the miners.

Any user can join the LunarX network to be a data service provider, selling his spare computing resource to other peers in the network. The LunarX daemon automatically accepts records from the network and insert them into local table shard. At regular intervals, the service provider earns incentive tokens for his service, as long as automatic proofs show that the data structure is still available and complete, without being altered or damaged.

LunarX employs an blockchain incentive system to mine tokens. During the repeatedly proofing of valid data computing, blocks will be created, and with them a certain amount of tokens as the mining reward will be sent to the provider's wallet address. Also, the blockchain manages exchanges of tokens and services.

For miners,?not only local file systems can be of service, the distributed file systems (DFS) are also good candidates for professional users. There are many industrial level DFS implementations, for example the Hadoop distributed file system(HDFS). Typically, this option is for traditional data center operators.

3.2 Other Roles

The network consists of three major roles: DApp and their users, service provider(miner), tracker. Miner has been introduced above. let's take a look at the other two roles.

DApps and their users:
DApps run on LunarX. Users are those who subscribe the services that DApps published. Both sides are data senders, and the receiver is no longer a private database, but the LunarX open network. For example, a storage business provides a cloud-storage application, its clients may choose to submit files to LunarX, and anyone except the owner has no way to access these files.

Trackers:
Trackers validate transactions between service users and service providers. In addition, each of them drives a bunch of verifiers to audit service providers of the retrievability of the data and availability of the services they deployed. And trackers will be rewarded by tokens for their service as well.

3.3 Business Process

Dapp's supplier A releases the software for users to download and use, the software processes users' data and puts it into LunarX, that is to say, it is put into other individual devices on the internet, but this supplier A has no control over these data. The user pays a certain amount of credit tokens LX. When the service is verified to be valid, these credit tokens LX will be transferred to these individual accounts that provide computing power.

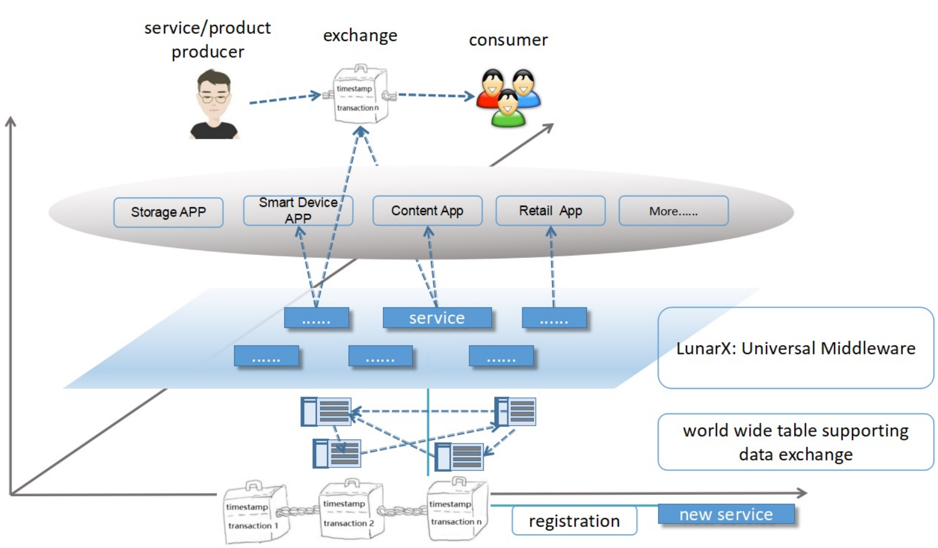


Figure 3.1: Business Process

Miners:

1) Miners accept remote requests, because the data structure is standard, and the data is encrypted, miners do not know the type of service, what the specific data is and who requests. they only know how many computing resources they provide to a request.

2) When the auditor verifies that the service is valid, or after verifying that it is valid for a period of time, the miner will receive the service confirmation message (sent by the smart contract) and then get the payment.

DApp's suppliers:

1) Create a table in LunarX, define the data structure(schema) of its business.

2) Publish the software to its user.

3) Users submit data, use data, or analyze data via the software.

4) The supplier draws a commission from the user's computing resource usage fee paid to the miners. If the user agrees, this ratio will be determined by the supplier himself.

5) If several suppliers provide similar services, users will choose the best products to use, just like the traditional software market.

DApp's users:

1) The user downloads the software, which has the functions or services he needs.

2) When the auditor verifies that the service is valid, or after verifying that it is valid for a period of time, the user will receive the service confirmation message (sent by the smart contract) and then pay for the remote resources consumed during the use of software.

3) Users can pay with LX or mainstream tokens, depending on the type accepted by the miner and the DApp, who provide the tool and resource. LunarX will read the exchange rates from major exchanges to help users exchange accurately.

4. Technical Architecture

The discussion in this chapter is based on fundamental framework introduced in the part I of this white paper. The purpose here is to explain how lunarX serves applications technically.

4.1 Two-layer protocol stack

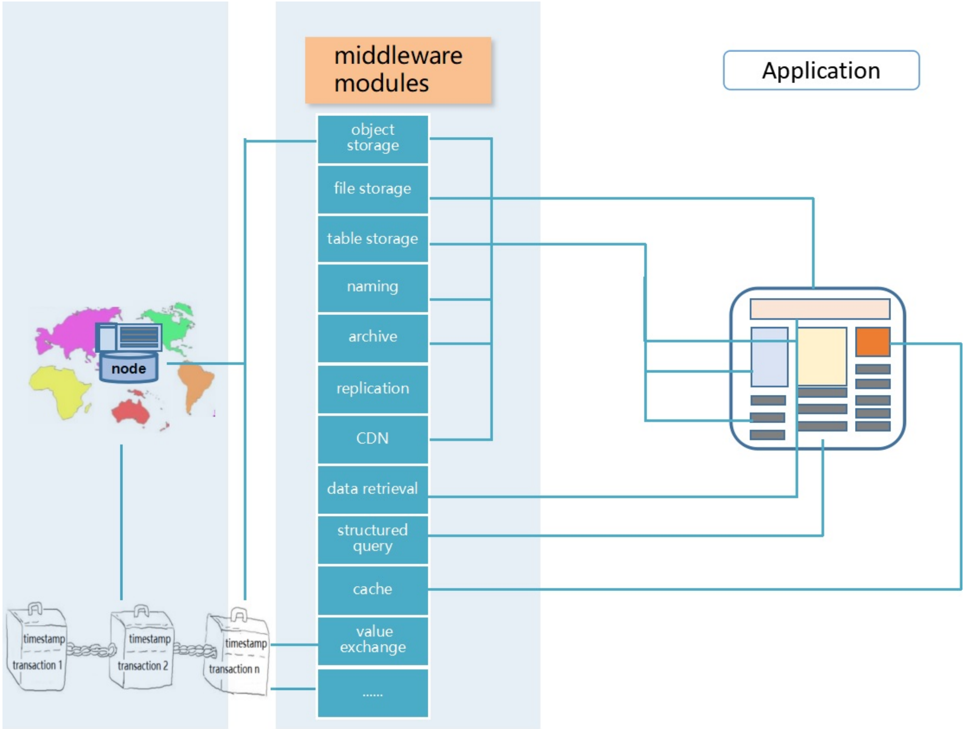


Figure 4.1: modules in the two layer architecture

The main functional modules we designed in the middleware include: DHT, object storage, file storage, table storage, naming, archive, replication, cdn, data retrieval, structured query, cache and value exchange (used in online selling and buying assets). Most of them have already been discussed in detail in part I, we briefly introduce some common and important modules here again.

4.2 DHT

Every peer has a pair of public key and private key. When this peer is on line, it has a long string (the hash of the public key) representing the unique identity of this peer. Other peers encrypt the message with the public key, the receiver decrypt it and know the content/instruction/value that the message bears. Also, public key and its hash are used to validate the authenticity of a peer. If the hash of the public key of peer is not equal to the id of the peer, the connection will be severed.

Peers are organized in a DHT network, e.g Kademlia DHT and its improvement S/Kademlia DHT, which stores nodes and resource locations throughout the network. Kademlia DHT uses XOR(exclusive or) operator to calculate the logical distance between any two peers. All nodes in a massive network according to the XOR metric construct a prefix binary tree, hence has the logarithmic complexity in searching through. Suppose we have 30 million nodes, in the worst case, it only costs 25 hops to reach the searched node.

A much more encouraging advantage of Kademlia DHT is the ability against the DOS(denial of service) attack. Even if some of nodes is flooded, routing procedure will seek around the jammed area, therefor it will have limited effect on network availability.

4.3 Data Storage and Query

This module includes file storage, object storage, table storage, archive and replication. These data are stored as data rows in the underlying storage. Each row has many entries. For example, a P-2-P storage DApp has each data row which records user's files and metadata including author, date, size, price, etc.

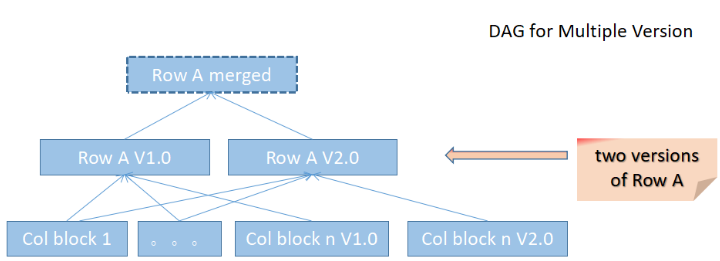


Figure 4.2: Locally DAG for Managing Data Rows

Multiple versions of a data row are organized in a DAG. And for big rows, such as 5GB size, they will be stored in a merkle tree for the retrievability.

Querying on encrypted data is another essential component. As long as the properties of encryption commutative mapping are satisfied, this function can be safely performed on remote anonymous nodes.

The diagrams here can be found in part I.

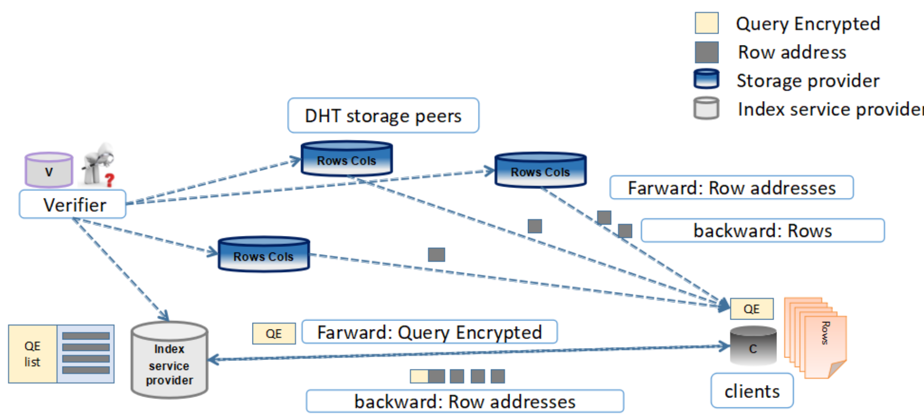


Figure 4.3: P-2-P Service Engine

4.4 Naming

Nodes are identified by an ID, which is the hash of the peer's public key, unique but not human readable. During first hand shake, peers exchange their public key, and compute the hashes to check if the node id equals to its hash of the public key. If not, the connection will be severed immediately.

We will locate a resource by its readable name, not its key, like what current internet domain is:

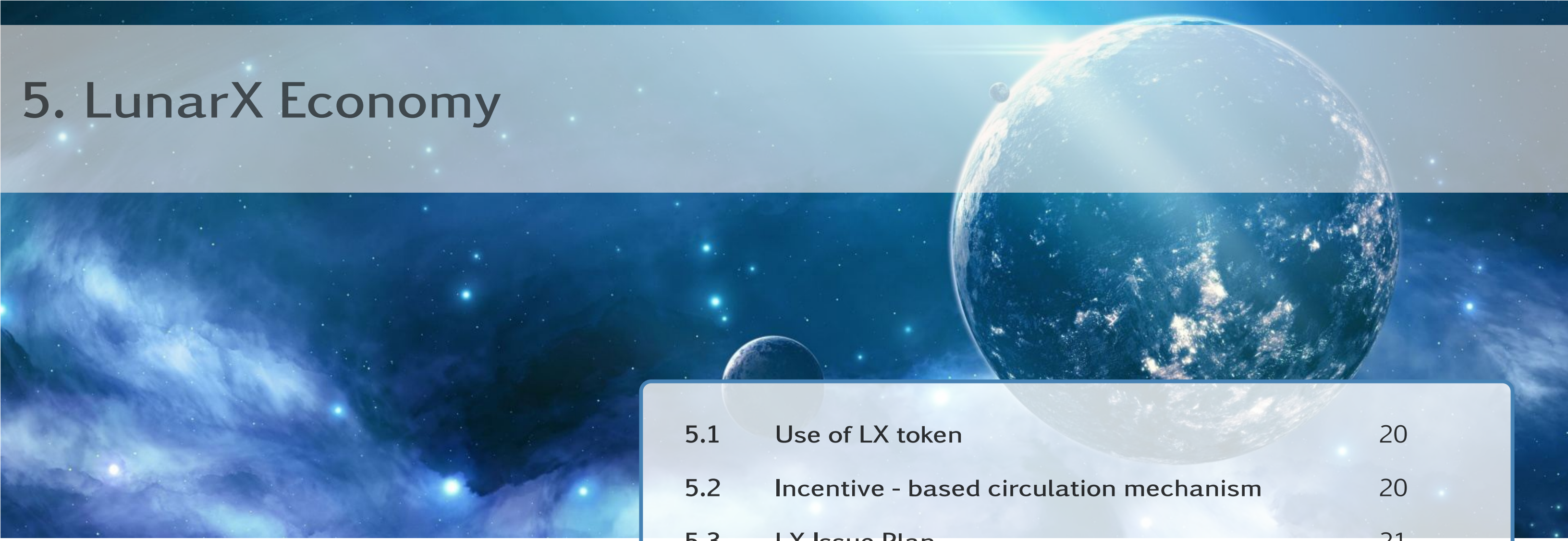
$LunarX/content_name/SXcdi86wsDkKyYUB......$

Internally it is a hashmap from name string to an ID, user can query an ID by the name. Using linear hash because it is space saving, and grows almost linearly along with the increase of elements. Here are two basic operation of a hashmap:

Put(name, ID)
ID = get(name)

For the concern of performance in put and query of massive sites, we use binary linear hash storage to manage names. If a node has a higher configuration, a k-v database will be a much better choice.

5. LunarX Economy



5.1 Use of LX token

The LX token will be devoted to research and development, team expansion, community operations and marketing in the agreed ratio. In order to make the leading creation of science and technology, our team will gradually release reserved LX tokens to invest in R&D, reward the blockchain developers who have made outstanding contributions, and incubate the potential scientific and technological direction.

5.2 Incentive - based circulation mechanism

Who will make profits? The following participants will earn tokens, not only LX, but also mainstream cryptographic currencies:

1) The miners;

2) The auditors who audit whether there are forgery nodes;

3) DApps' suppliers;

4) Users who create valuable services/products/contents/tools to the internet via DApps;

In the early version of LunarX, we only support the mainstream BTC, ETH, and our partners' tokens. We will have periodic bulletin boards telling users how many tokens a registered service can exchange and the exchange rate between tokens. The pricing of services is determined by the DApp/service suppliers.

who will pay?

1) LunarX system will generate tokens for miners after proofing the validity of data services;

2) DApps' users, who seek useful services for their own needs;

5.3 LX Issue Plan

LX issued a total of 100 million, each ETH convertible 3790 LX, no crowdfund. The token will be logged into the mainstream exchange in the short time.

[**Token total**] 100 million.
[**Crowdfunding**] NO ICO.

Token Distribution

25%=25 million of core team,
55%=55 million of community and miners, for rewards of contribution,
20%=20 million of early investors.

6. Road Map

Readers should understand that with the development of technology and market, the planning in this white paper will also make corresponding adjustments. The only thing that will never change is our vision. We will keep working hard to create a better internet data service ecosystem.

The engine part of LunarX started in the fourth quarter of 2016, and we did not start writing white papers (Q2, 2018) until all the basic modules were developed.

6.1 Q4, 2016 – Q2, 2019

TABLE 6.1 LunarX engine development (before 2018)

2016 Q4	Start planning for the IO engine part of LunarX.
	IO modules(memory and disk).
2017 Q1	Virtual File system for block storage.
	Index part for various data structure, including linear hash table, B+ tree, inverted table,etc.
2017 Q3	Stability test of basic modules.
	Start design active garbage collector.
2017 Q4	Data chain engine interface.
	compressed storage, performance improvement.
	Finish active garbage collector. Stability test.
	Apply the engine in time serial analysis for testing the performance and scalability in real scenario.

TABLE 6.2 White Papers

2018 Q2	Part I: Technique white paper, architecture, fundamental module planning.
	Part II: Business, marketing and token economy white paper.
	Website and github.

TABLE 6.3 LunarX Product Road Map

2018 Q2	architecture, fundamental module development.
2018 Q3	Hardware design, interface release.
	Software architecture improvement. Test fundamental modules.
	Internal data chain testnet online. Start to build and grow LunarX community.
	Middleware protocol release.
2018 Q4	Storage application, which is the first DApp release.
	Name Service.
	Content site building Service, CDN.
	live streaming service.
	complete hardware development.
	by the end of Q4, LunarX main data chain network online. Join test with partners.
	Miner release.
2019 Q1	More DApps plugin. Begin to connect main stream blockchains to LunarX data chain network, including: BTC, Ethereum, Zilliqa, EOS, and partners' public chains.
	Begin to test LunarX on Skywire.
2019 Q2	Release standard data chain interface.
	Personal AI assistant trained by individual secret data.
	plugin more DApps.

7. Team

7.1 team background

Ben Fei, Funder of LunarX. Years of asset management, product management experience.

Exp:

Funding internet company, as CEO, won "Outstanding Enterprise Award" on Demo China Summit 2016, Beijing.

Former IBMer at IBM CRL, PBC(Personal Business Commitment) top 5% once, top 30% twice.

Edu: B.S/Ph.D, Fundamental Math, Tongji University, Shanghai, China.
LinkedIn:https://www.linkedin.com/in/ben-fei-813867113/

Jinlei Ben Xiong, Funder of LunarX. Successful entrepreneur.

Exp:

Senior Vice President of Listed Company, Product/MediaStrategy(Fractalist, code: 870104),
Co-publisher, MIT Technology Review Chinese Edition,
Co-publisher, Science Times Magazine from NYTimes,
Founding Telecom Servicing Company, acquired by a listed company(code:837272)
Sales Manager, Hewlett Packard,
Sales/Marketing, SONY VAIO,

Edu: MA, University of Sheffield, England.
LinkedIn:https://www.linkedin.com/in/ben-xiong-802a496/

Yan Wang, co-Funder

Exp: worked in China Academy of Social Sciences.
Join in the early blockchain field as angel investor. Used to be advior of several overseas blockchain projects. Founder of SPO project.

Edu: M.S at China Academy of Social Sciences(the top research institute in China).

Ryo Umezawa, co-Funder

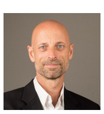
Exp:
Ryo has been with J-Seed Ventures since 2005, participating in several partner companies where he frequently assisted in starting up the businesses, marketing and business development. He has experience in the mobile, online marketing, advertisement and platform businesses.

Principle, East Ventures. East Ventures is a seed-early stage venture capital firm with offices in Singapore, Indonesia and Tokyo. It was founded in 2010, and have invested in more than 200 companies across Asia and US, which include internet startups on commerce, social, game, SAAS, and mobile service.

Director, J-Seed Ventures, Tokyo-based venture incubator. The company also manages Venture Generation, a startup community in Tokyo, and assists foreign clients to establish operations in Japan.

President & CEO, Hailo. Hailo is the evolution of the hail - a free smartphone app which puts people just two taps away from a licensed vehicle, and lets drivers get more passengers when they want them.

Edu: BA, International Business and Economics, Sophia University .
LinkedIn:https://www.linkedin.com/in/ryoumezawa/



Simon Hobbs, CTO

Exp: VP Engineer of Skycoin.
28 years of high-level technical experience in delivering complex, high-volume IT projects on time and on budget, He began his career as an engineer on missile guidance software, and later moved into financial services, l**eading IT projects for Barclays Home Finance Mortgage, Swiss Bank and Cedel Bank**.

Edu: B.S, Computer Science, University of Hertfordshire, England

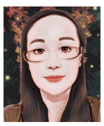
Young Jun, Core Member

Exp: Amazon (the leading cloud computing unicorn).
He has rich experience in cloud computing, big data, e-commerce platforms, is excellent at internet architecture and leads the technical progress.

Edu: M.S at Tsinghua University, Computer science, Beijing, China.

Cheng Zhang,

Exp: Over 18 years of work experience in the IT industry. From 2001 to 2007 as a senior software engineer in the software outsourcing industry in Japan. Since 2007, as the **Co-Founder** & **CTO**, he founded **the largest mobile game company** in Xi'an(Capital of Shanxi Province). Since then, he has Co-Founded a number of Internet companies, focusing on the development of apps and games for mobile platform. He has great experience in team management and project research and development.



Yenny.Z

Exp: More than 10 years as editor-in-chief of traditional media and new media, management experience and planning experience of large-scale activities, and has been involved in blockchain industry for 3 years.
Worked in many top media such as <<The Mirro>> and <<Beijing news>>, participated in the planning and reporting of major national events, and won more than 10 awards in the field of news. Have rich practical experience in new media operation, brand public relations, market expansion and community operation.

Edu: B.A, Renmin University of China, Beijing, China.

Vincent, Leader of new media operation

Exp: project manager at the head office of a bank. Promoted the progress of several intra-bank projects. He is a geek of block chain, used to hold media and community operation positions in two top 70 block chain companies, and have a deep understanding of block chain industry.

Edu: M.S, machine learning, Chinese academy of sciences.