



THE WHITE PAPER OF **SID**

V1.0.1



Before the speech

One, called "decentralized" (Decentralization) thoughts, born from the beginning of the birth of human society. He is eager to get rid of the center and realize direct communication, direct transaction and direct communication between people. He believes that one day we may no longer need centralised institutions.

In the long course of human development, this kind of thought has repeatedly collided, combined with different organizations, environments and carriers, resulting in different social changes. Standing in the current space-time dimension, you can still see the decentralized religious church, as well as the decentralized political system and social organization.

The emergence and large-scale use of information technology have brought technology carriers to the idea of centralization. Tracing the thread of the combination of this idea and technology, a lot of popular and even subversive innovations have been produced: P2P download, CND (p-cnd), distributed computing (cloud computing), social media (we-media), P2P lending, crowdfunding, sharing, blockchain, self-organizing DAO, etc.

With the development of big data industry to today, exciting results have been achieved. Almost all fields are swept by the boom of big data. We can see that with the rapid development of Internet information technology and the gradual maturity of cloud computing and Internet of things technology, the production, collection, storage and processing of massive data are no longer a problem, and the potential of data has been released unprecedentedly. The era of big data is not an instant birth, but an inevitable result when the information society develops to a certain stage.



In the long run, big data will bring about tremendous changes to the whole human society. However, this change can only be achieved through years of technology accumulation and social practice, and its development process also needs the synchronous development and integration of many related technologies.

The value of big data lies in the integration of multi-source data. The current data circulation has seriously restricted the play of the value of big data in the whole society, and the problems of data opening, sharing, circulation and privacy protection have become the biggest bottleneck in the rapid development of big data.

Everyone realizes that "social big data" can only be achieved when different data sources are open and Shared. This goal; Otherwise, data sources that exist independently and are not Shared with each other simply form islands of data. But the obstacles to open data sharing may be greater than people think. The reason is that the current information technology -- database, cloud computing, data center and so on are all based on the idea of centralized service, which will inevitably lead to a high concentration of data and the formation of data monopoly. Therefore, the problem that data monopoly kills data innovation will also exist for a long time. How to find an appropriate balance between data ownership and data sharing will be one of the core issues for the healthy development of the big data ecosystem.

With the development of information technology, big data and blockchain technology, SID is born.

SID is committed to building a comprehensive platform in the field of blockchain for the open, transparent and fair display of information, information and data.



SID is a distributed database system, in which data is automatically backed up at all nodes around the world. With the help of cryptography and game theory, network security is jointly maintained by all nodes around the world. It is difficult for any individual or organization to tamper with system data, so it is a credit system. Through the professional technical team, starting from the security, efficiency, capacity and processing capacity of the public chain, it is possible to expand the blockchain to millions of users and millions of transactions per second, improve the efficiency of intelligent business development, and reshape the blockchain industry ecology.

SID is a blockchain platform for creating and developing communities dedicated to generating, supporting, and promoting new project proposals through a fluid organizational model called the decentralized autonomous organization (DAO). By using an untamable blockchain and smart contract execution, the DAO achieves autonomous and independent organizational governance, management, and operation.

We believe that with the advent of SID, the world of information and data will be transformed.



directory

<i>Before the speech</i>	2
<i>Chapter 1: a brief introduction to blockchain</i>	7
1.1 birth of blockchain.....	7
1.2 development of blockchain.....	8
1.3 characteristics of blockchain.....	10
<i>Chapter 2: project overview</i>	13
2.1 industry status.....	13
2.2SID arises.....	14
2.3SID's design mission.....	16
<i>Chapter three: technical support</i>	18
3.1 information retrieval.....	18
3.1.1 big problems in information retrieval.....	20
3.1.2 search engine.....	23
3.2 data storage.....	26
3.2.1 DAS.....	26
3.2.2 NAS.....	27
3.2.3 SAN.....	28
3.3 fast transmission.....	30
3.3.1. Background of data transmission.....	30
3.3.2. Data transmission of the same network segment.....	31
3.3.3. Data transmission of different network segments.....	32
3.3.4. Several key issues.....	34
<i>Chapter four: platform advantages and characteristics</i>	35
4.1SID digital asset information center advantage.....	35
4.2SID platform features.....	36
<i>Chapter 5: SID economic model</i>	38
5.1 the SID.....	38
5.2 release plan.....	38



5.3 private placement plan.....	39
Chapter 6: team introduction.....	40
Chapter 7: decentralized communities and SID community commissioners.....	41
7.1 decentralized communities.....	41
7.2 community committee.....	41
7.3 rights, obligations and implementation of community committees.....	42
Chapter 8: SID development planning.....	43
8.1 new business model.....	43
8.2 sustainable development.....	43
8.3 commercial application.....	43
8.4 extensive cooperation:.....	43
Chapter 9: SID operation pattern.....	44
9.1 exchanges pay to be listed.....	44
9.2 information priority exposure.....	44
9.4 professional analysis.....	44
9.5 community marketing.....	45
Chapter 10: SID technology application domain.....	46
10.1 decentralized digital asset exchange.....	46
10.2 financial services.....	47
10.3 medical and health care.....	47
10.4 clothing, food, shelter and transportation.....	48
Chapter 11: disclaimer.....	49



Chapter 1: a brief introduction to blockchain

1.1 birth of blockchain

The concept of blockchain was first proposed by Satoshi Nakamoto in 2008, and in the years that followed became the basis for the electronic currency bitcoin.

Core component: a public ledger of all transactions. By leveraging point-to-point networks and distributed timestamp servers, blockchain databases can be managed autonomously. The blockchain invented for bitcoin makes it the first digital currency to solve the problem of duplicate consumption.

Bitcoin's design has become a source of inspiration for other apps.

In 1991, Stuart Haber and W. Scott Stornetta first proposed the encryption protection chain product for the block, followed by Ross J. Anderson and Bruce Schneier & John Kelsey in 1996 and 1998 respectively. Meanwhile, Nick Szabo conducted research in 1998 on the mechanics of electronic currency diversification, which he calls bitkin. In 2000, Stefan Konst published the unified theory of encryption protection chain and proposed a set of implementation plans. The blockchain format was first applied to bitcoin as a solution for making databases secure without the need for an administrative agency to grant credit. In Nakamoto's original paper in October 2008, the words "blockchain" and "chain" were used separately, while they were collectively called "blockchain" when they were widely used, and only became one word: "blockchain" in 2016. In August 2014, bitcoin's blockchain file size reached 20 gigabytes.



1.2 development of blockchain

First generation: bitcoin

Bitcoin is a distributed autonomous system, a perfect combination of applied mathematics and financial economics. If you look closely at the system, you will find that it includes central Banks (issuing money), businesses (paying) and functions as money itself.

In this system, the genius satoshi nakamoto USES asymmetric cryptography to solve the ownership problem of bitcoin, USES UTX0 model to define the concept of "bitcoin", and USES blockchain to solve the problem of distributed transaction verification. The workload proof method (POW) maintains the normal stability and security of the system, and under the action of game theory, all miners maintain a unified blockchain to finally solve the dual payment problem. To this day, bitcoin remains the unassailable digital currency leader.

Litecoin has made few improvements to bitcoin, but it has sparked a wave of algorithmic improvements that have spawned many digital currencies. If bitcoin is seen as a social experiment, many copycat COINS are a bitcoin experiment, and litecoin activates isolation verification, which once again confirms this statement.

Among the numerous tokens, there is one branch of development that is still developing vigorously at present, namely, digital currencies featuring anonymity, among which Dash and Zcash are the representatives.

SID will absorb the advantages of many virtual currency pioneers and apply them to consensus mechanism, algorithm, privacy, etc.



Second generation: chain

When people gradually clarify the relationship between bitcoin and blockchain, their vision is immediately broadened, and they find that bitcoin is only one application of blockchain, and there will be other applications of blockchain. Therefore, the development focus of the whole industry turns from "bitcoin" to "blockchain". As many people now know, blockchain can be thought of as a distributed database whose core feature is the ability to record data and contracts along a timeline that can only be read and written, not modified and deleted.

Bitshares was the first to sound the clarion call for blockchain exploration, and Ethereum really turned the revolution on its head. Ethereum not only inherits many advantages of bitcoin, but also introduces many innovations. It is a smart contract platform as well as a distributed application underlying protocol.

The most fascinating thing about ethereum is its virtual machine (EVM), which is almost "Turing complete". Of course, it also brings huge technical complexity and fault-tolerant cost. It is now widely accepted in the community that the system is far from perfect and there is still a lot to be optimized. The exploration of blockchain and smart contract is still a long way away, which is the origin and opportunity of SID.

Next generation: web

Over the past eight years, the digital currency industry has moved from bitcoin to blockchain. What's next? It is generally believed that after the Internet, the next network is the Internet of things. The Internet of things has been mentioned for many years, and IBM has been working tirelessly on it, but there has never been any substantial progress. Even what is the Internet of things? It wasn't until blockchain came along that the picture was opened.



One reason for this dilemma is that we can't really automate with the technologies of the past, and bitcoin, a proven distributed autonomous system (DAO), is a perfect solution. The DAO provides the genes for automation and thus can serve as a foundation for the Internet of things.

Now the Internet of things and blockchain are always tied together to discuss, looking forward to the ubiquitous Internet of things in the future, which is very exciting. So it is reasonable to believe that the next evolutionary form of blockchain is the network, where numerous blockchains are no longer distinct from each other.

SID is to build a new generation of distributed trust chain network and build a blockchain application platform accessible to all people from all walks of life around the world.

1.3 characteristics of blockchain

Generally speaking, blockchain system consists of data layer, network layer, consensus layer, incentive layer, consensus layer, contract layer and application layer.

Among them, the data layer encapsulates the underlying data block and related data encryption and timestamp and other basic data and algorithms; The network layer includes distributed networking mechanism, data transmission mechanism and data verification mechanism. The consensus layer mainly encapsulates various consensus algorithms of network nodes. The incentive layer integrates economic factors into the blockchain technology system, including the issuance mechanism and distribution mechanism of economic incentives. The contract layer mainly encapsulates various scripts, algorithms and



intelligent contracts, which is the basis of the programmable characteristics of the blockchain. The application layer encapsulates various application scenarios and cases of blockchain.

In this model, the chain block structure based on time stamp, the consensus mechanism of distributed nodes, the economic incentive based on consensus computing force and the flexible and programmable intelligent contract are the most representative innovation points of blockchain technology.

Blockchain has the following characteristics:

decentralized

Due to the use of distributed accounting and storage, there is no centralized hardware or management organization, and the rights and obligations of any node are equal. The data blocks in the system are jointly maintained by the nodes with maintenance functions in the whole system. Thanks to the decentralized nature of blockchain, bitcoin also has decentralized nature.

open

The system is open. Except that the private information of all parties involved in the transaction is encrypted, the data of the blockchain is open to all. Anyone can query the blockchain data and develop relevant applications through the open interface, so the information of the whole system is highly transparent.

autonomy

Blockchain adopts a consensus based specification and protocol (such as an open and transparent algorithm), which enables all nodes in the whole system to freely and securely exchange data in a trusted environment, so that the trust in "people" is changed to the trust in machines, and no human intervention will work.



Information cannot be tampered with

Once the information is verified and added to the blockchain, it will be stored permanently. Unless more than 51% of the nodes in the system can be controlled at the same time, changes to the database on a single node are invalid, so the data stability and reliability of the blockchain are extremely high.

anonymity

As the exchange between nodes follows a fixed algorithm, its data interaction is not required to be trusted (the program rules in the blockchain will determine whether the activity is effective or not), so the counterparty does not need to make the other party to generate trust by means of public identity, which is very helpful for the accumulation of credit.



Chapter 2: project overview

2.1 industry status

The blockchain industry is currently in a stage of rapid development, with entrepreneurs and capital continuously pouring in and the number of enterprises increasing rapidly. We will accelerate the implementation of blockchain applications, promote high-quality development of traditional industries, and accelerate industrial transformation and upgrading. The use of blockchain technology to "reduce costs" and "improve efficiency" for the real economy and promote the standardized development of traditional industries. In addition, blockchain technology is being derived into new business forms and becoming a new driving force for economic development. Blockchain technology is driving a new round of business model change and becoming an important support for building an honest social system. At the same time, local governments have been actively positioning blockchain technology from an industrial perspective, and the policy system and regulatory framework have been gradually improved.

Since 2016, the investment heat in the blockchain sector has increased significantly, with 60 investment events, five times as many as in 2015. 2017 was the peak year for blockchain investment in recent years, with the number of investment events approaching 100. In the first quarter of 2018, the number of investment events in the blockchain sector increased. There were 68. From the current trend, as the speed of the implementation of blockchain technology is also accelerating, the market is becoming more rational, and equity investors are more willing to invest in projects with specific implementation scenarios. It is expected that the investment in the blockchain field will reach a peak this year.

As the value of blockchain is widely recognized, more and more industries



are proposing their own blockchain solutions. From the perspective of application scope, blockchain technology can be applied in almost all industrial scenes, because almost all industrial scenes involve transactions, and all have the need to reduce costs, improve efficiency and optimize the industrial integrity environment, which is exactly the role blockchain technology can play quickly after it is applied.

The main battlefield for the implementation of blockchain technology is the real economy industry field, and the value of blockchain technology will also be reflected in the value increment brought by the scene of the implementation of the industry. At present, the physical flow in the traditional industry has not been widely transformed into information flow. One of the reasons is that many real economy enterprises, especially small, medium and micro enterprises, do not have the means to achieve great actual benefits based on the digital transformation of business, but have to bear a series of risks such as information leakage. Therefore, many enterprises do not actively promote the digital transformation of their businesses, resulting in incomplete, incomplete and unsystematic mapping of physical objects into information flow. Using block chain technology, combined with the Internet of things and industrial application of Internet technology, and based on the new mode of supply chain finance, driven by a large number of trading information, has begun from the offline to chain enterprise management system and equipment LianWangLv began to ascend, digital assets become an important part of enterprise assets, the entity industry business model will also achieve unprecedented depth change.

2.2SID arises

From the perspective of the development of the blockchain industry, SID's ultimate vision is to be able to reshape the value of the blockchain industry, integrate the resources of the blockchain industry by establishing



connections between different blockchain ledgers, and create an ecology of information, information and data sharing from various ecological perspectives.

Create a universal public chain applicable to all walks of life, so that the global blockchain enthusiasts can jointly maintain the blockchain ecology we create.

SID is committed to building a comprehensive platform in the field of blockchain for the open, transparent and fair display of information, information and data.

Liberates the creativity, realizes the value circulation by the technology

SID will realize value and asset circulation, so that the blockchain entrepreneurs with real ideas can complete investment and financing at critical moments. Build a better industry atmosphere and health State, the liberation of creativity, in the current market to solve the public chain is flooding copy intelligent contract but no underlying technical problems, at the same time, to give technical personnel more support, promote ecological improvement.

Build Bridges to the real world

SID is committed to building enterprise-level blockchain products and providing industry solutions. It needs to develop a high-performance and highly scalable blockchain basic service platform, and has the ability to quickly build the upper application business to meet the application scenarios with a large number of users, and can be the same as the real world to help all walks of life.



2.3 SID's design mission

After market research and analysis, we found that there are some problems in the development process of blockchain, and described briefly above, SID can provide a reliable solution for this.

Flexible and easy-to-use blockchain infrastructure

SID provides complete turing-based modular development for developers and users. Developers and users do not need to study cryptography, consensus mechanism, storage and other underlying technical details, using a simple and fast programmable environment directly docking commercial applications, thus reducing the commercial cost of blockchain.

ADAPTS the massive block chain application scene

At the application level, it can be expected that blockchain will serve as the underlying support for institutions and even individuals in various aspects of work and life. SID realizes the cross-ledger transfer of assets through modularity, multi-chain parallelism, intelligent contract and other operational mechanisms, providing an infrastructure for the application of digital assets based on digital currency. We will give full play to the function of blockchain in carrying and delivering value, and give full play to the concept of equality and openness of blockchain.

High-performance drive blockchain commercial landing

Applications from all walks of life have extremely high requirements for performance. SID is committed to solving the performance limitation problem of the existing blockchain by adopting parallel extension protocol and allowing multi-chain parallel operation mechanism through more efficient value transmission protocol to meet the TPS requirements of tens of millions of levels.

A balance between data transparency and privacy



Data privacy and security are extremely important to organizations, but the open and transparent nature of blockchain makes organizations worry. SID ensures the confidentiality and security of business data and resolves the balance between data transparency and confidentiality by properly building the framework and supporting multiple protocols with technology.





Chapter three: technical support

Combined with data storage, fast transmission. Search engine and other technologies to complete data collection and synchronization.

3.1 information retrieval

For Information retrieval is a field concerned with the structure, analysis, organization, storage, searching, and retrieval of Information. From Gerard Salton Information retrieval is a field involving information structure, analysis, organization, storage, search and retrieval.

The main focus of information retrieval is text and text files. The difference between a document and a typical database record:

Documents are mostly text, unstructured. Typical database records are structured with fields and attributes.

Structured data has a clear meaning and format, making it easy to compare and query them. Unstructured data is not so easy, for example:

Now consider a news report about the merger of the two Banks. The story will have some attributes, such as the title and source of the story, but the main content will be the story itself. In database systems, this critical information is typically stored as a single large attribute with no internal structure. Submit a web search citation for this story. Most queries from engines such as Google will take the form of "bank consolidation" or "bank takeover". To do this search, we must design algorithms that can compare the query text to the story text and determine whether the story contains the information being searched. Defining the meaning of a word, a sentence, a paragraph or an entire



news story is much more difficult than defining an account number, so comparing text is not easy. Understanding and modeling how people compare text, and designing computer algorithms to perform such comparisons accurately, are at the heart of information retrieval.

In addition, information retrieval is widely used not only in structured data and text content, but also in multimedia (image, audio and video).

Information retrieval scenario: the user enters the results into the search engine and gets the sorted results.

Examples include the world wide web, enterprise search, vertical search (specific topics), desktop search, and point-to-point search (finding information on a node or computer's network without any central control)

Search and related information retrieval technologies are used in advertising, intelligence analysis, scientific discovery, health care, customer support, real estate, etc. Any application that involves text collections or other unstructured information needs to organize and search for that information.

Searching based on user queries (sometimes called AD hoc search because of the large and undetermined range of possible queries) is not the only text-based task examined in information retrieval. Other tasks include filtering, categorizing, and answering questions. Filtering or tracking involves detecting stories of interest based on an individual's interests and using email or other mechanisms to alert users. Categories use a set of defined tags or categories (for example, Yahoo! Categories listed in Directory) and automatically assign these tags to the document. The q&a is similar to a search, but targeted at more specific questions, such as "what is the height of mount Everest?" The goal of the q&a is to return specific answers found in the text, not a list of documents.



3.1.1 big problems in information retrieval

Several key concerns of information retrieval researchers remain as relevant in the era of commercial web search engines that use billions of web pages, as they did in the 1960s when tests were conducted on collections of documents containing about 1.5 megabytes of text. One problem is correlation. Correlation is the basic concept of information retrieval. In short, the relevant document contains the information that a person is looking for when submitting a query to a search engine. It sounds simple, but one person decides on a particular one. There are many factors that determine whether a document is relevant. These factors must be taken into account when designing algorithms for comparing text and ranking documents. Simply comparing the text of a query to the text of a document and looking for a perfect match, as you do in a database system or in Unix using the grep utility, produces very poor results in terms of correlation. One obvious reason is that language can be used to express the same concepts in many different ways, often using different words. It is called the vocabulary mismatch problem in information retrieval.

It is also important to distinguish topic relevance from user relevance. If the text document is on the same topic, the document is related to the query topic. For example, news reports about tornadoes in Kansas will be linked to queries about "severe weather events." The person asking the question (called the user) may not think the story is relevant. But if she has seen the story before, or if it is a story from years ago, or if it was published in Chinese by a Chinese publishing house, then the story is relevant to the user. User relevance takes these additional features of stories into account.

In order to solve the problem of relevance, retrieval models were proposed and their working conditions were tested. The retrieval model is the formal representation of the process of matching queries and documents. It is the



basis of the ranking algorithm used in search engines to generate a sorted list of documents. A good retrieval model will find documents that may be considered relevant by the person submitting the query. Some retrieval models focus on topic relevance, but search engines deployed in real-world environments must use ranking algorithms that include user relevance.

An interesting feature of retrieval models used in information retrieval is that they usually simulate the statistical characteristics of text rather than the language structure. For example, ranking algorithms often focus more on counting the number of words that appear, rather than whether the word is a noun or an adjective. Higher-level models do contain linguistic features, but they tend to be of secondary importance. The use of word frequency information to represent text began with a pioneer in information retrieval, h. p. 1950s. It was not until the 1990s that textual ideas became popular in other areas of computer science, such as natural language processing.

The other core problem of information retrieval is evaluation. Since the quality of a document ranking depends on how well it matches one's expectations, it is necessary to develop evaluation methods and experimental procedures as early as possible to capture this data and use it to compare ranking algorithms. Cyril Cleverdon developed evaluation methods in the early 1960s, and the two measures he used, precision and recall, remain popular. Precision is a very intuitive measure of the proportion of related documents retrieved (number of related documents retrieved/number of documents retrieved). Recall is the percentage of related documents retrieved (number of related documents retrieved/number of all related documents). When using recall metrics, assume that all relevant documents for a given query are known. This assumption is made in the webThere are obviously problems in the web search environment, but for a small set of document tests, this metric can be useful. The test set for the information retrieval experiment consists



of a set of text documents, a sample of a typical query, and a list of relevant documents for each query (relevance determination). The most famous test set is the set associated with the TREC (text retrieval session) assessment of BBS.

The retrieval model and evaluation of search engines is a very active area, with most of the time currently focused on using large amounts of log data from user interactions, such as click data, which records documents clicked during a search session. Click-through rates and other log data are closely related to relevance and can therefore be used to evaluate searches, but search engine companies still use correlation judgments in addition to log data to ensure the validity of their results.

The third core problem of information retrieval is to emphasize users and their information needs. It should be clear, because search evaluation is user-centered. Search engine users are the ultimate judges of quality. There has been a lot of research on how people interact with search engines, particularly the development of technologies to help people express their information needs. Information requirements are the root cause of the queries people submit to search engines. In contrast to a request to a database system, such as a bank account balance, a text query gives a poor description of what the user actually wants. A word query such as "cat" can be a request for information about where a cat bought or a description of a Broadway music. However, despite the lack of specific cities, word queries are common in web searches. Technical questionnaires, query extensions, and related feedback use interaction and context to re-initiate the query to produce a better ranking list.

This question will appear in this book and will be discussed in considerable detail. We now have enough background to start talking about the main product of information retrieval research – the search engine.



3.1.2 search engine

Search engine is the practical application of information retrieval technology in large-scale text collection. Web search engines are an obvious example, but as mentioned earlier, they can be found in many different applications, such as desktop or enterprise search. Search engines have been around for years. For example, MEDLINE, an online medical document retrieval system, began in the 1970s. The term "search engine" was originally used to refer to dedicated hardware for text search. However, since the mid-1980s, it has gradually been used as the preferred name for a software system that USES "information retrieval systems" to compare queries to documents and produce a sorted list of document results. Of course, search enginesMore than sorting algorithms, we will discuss the general architecture of these systems in the next chapter.

Search engines have many configurations that reflect their design goals. Web search engines such as Google and yahoo must be able to capture or crawl many terabytes of data and then provide sub-second response times for millions of queries submitted every day from around the world. An enterprise search engine (for example, Autonomy) must be able to process various sources of information within a company and use company-specific knowledge as part of search and related tasks, such as data mining. Data mining refers to the automatic discovery of interesting structures in data, including such techniques as clustering. Desktop search engines, such as the Microso Vista™ search feature, must be able to quickly merge new documents, web pages, and emails as users create or view them, and provide an intuitive interface to search for this heterogeneous combination of information. Systems such as Google overlap among these categories and can be used for enterprise and desktop search configuration.



Open source search engines are another important class of systems designed to be somewhat different from commercial search engines. There are many of these systems, and the wikipedia information retrieval page provides links to many of them. Systems of particular interest are Lucene Lemur and the system provided in this book, Galago. Lucene is a popular java-based search engine that has been widely used in commercial applications. Lemur is an open source toolkit that includes a search engine based on an Indri C ++. Lemur is primarily used by information retrieval researchers to compare advanced search techniques. Galago is a java-based search engine based on Lemur and Indri projects. Galago is widely used in the assignments in this book. It is designed to be fast, adaptable and easy to understand, and combines very effective information retrieval techniques.

The "big problems" in search engine design include the problems that information retrieval has identified: effective ranking algorithms, evaluation, and user interaction. However, many of the other key features of search engines result from their deployment in large-scale operational environments. The most important of these features is the performance of the search engine in terms of response time, query throughput, index speed, and so on. Response time is the delay between submitting a query and receiving a list of results, throughput measures the number of queries that can be processed in a given time, and index speed is the rate at which a text document can be converted to a search index. Indexing is a data structure that improves search speed. Index design for search engines is one of the main topics of this book.

Another important performance metric is how quickly new data can be indexed. Search applications typically deal with dynamic, constantly changing information. Coverage measures how much information in the company's information environment is now indexed and stored in search engines,



freshness measures the "age" of information stored.

Search engines can be used with small collections, such as hundreds of emails and documents on a desktop, or very large collections, such as the entire Web. There may be only a few users of a given application, or thousands of users.

Visual classification is very important problem of search engine design. The design for a given application should continue to work as the amount of data and the number of users grows. In section 1.1, we described how search engines are used in many applications and many tasks. To do this, they must be customizable or adaptable. This means that many different aspects of the search engine, such as ranking algorithms, interfaces, or indexing policies, must be able to be adapted and adapted to the requirements of the application.

Practical problems that affect the design of search engines can also occur in specific applications. The best example of this is spam. Spam is generally considered undesirable E-mail, but more generally it may be defined as misleading, inappropriate or irrelevant information in documents designed for some business benefit. There are many types of spam, but one of the types that search engines have to deal with is the spam word put into a document in order to retrieve it in response to a popular query. The practice of "spam indexing" can significantly reduce the quality of search engine rankings and web search engine designers must develop techniques for identifying spam and deleting these documents.



3.2 data storage

3.2.1 DAS

DAS (Direct Attached Storage) just like our common PC Storage architecture, external Storage devices are directly Attached to the internal bus of the server, and data Storage devices are part of the whole server structure.

DAS storage mode is mainly applicable to the following environments:

1) small networks

Because the network is small, data storage is small, and it is not very complex, the impact on the server will not be great. It is also economical for enterprise users with small networks.

2) geographically dispersed networks

Although the overall network size of the enterprise is large, it is geographically dispersed, and it is very difficult to connect them through SAN or NAS. At this time, the servers of each branch can also adopt DAS storage mode, which can reduce the cost.

3) special application server

On some special application servers, such as Microsoft's clustered servers or raw partitions used by some databases, storage devices are required to connect directly to the application server.

4) improve DAS storage performance

DAS were once considered an inefficient structure for various connections between servers and storage and inconvenient for data protection. Direct-connected storage cannot be Shared, so it is often the case that one server is running out of storage space, while others have a large amount of storage that is idle but not available. If storage cannot be Shared, there is no balance between capacity allocation and usage requirements.



The data protection process under the DAS structure is relatively complex. If network backup is made, each server must be backed up separately, and all data flows must be transmitted through the network. Without a network backup, each server comes with a backup software and tape drive, so the backup process becomes much more complex.

If you want to have high availability DAS storage, you must first be able to reduce the cost of the solution, for example: LSI 12Gb/s SAS, in which there is DAS direct storage, through the DAS can provide good support for large data centers. For large data centers, cloud computing, storage, and big data, all of which put higher demands on DAS storage performance, cloud and enterprise data center data explosionSexual growth can also pushed the market to support more high-speed data access requirements of high performance storage interface, thus LSI 12 gb/s SAS is able to meet the requirements of the performance of growth, it can provide higher IOPS and higher capacity, 12 gb/s SAS improves the performance of the higher writing, and improved the comprehensive performance of RAID.

Compared with direct-attached storage architectures, Shared storage architectures, such as SAN (storage-area network) or NAS (network-attached storage), can better solve the above problems. As a result, we see the process of DAS being phased out faster and faster. As of 2012, however, DAS were still a common pattern for server to storage connections. In fact, far from being phased out, the DAS seem to be making a comeback in recent years.

3.2.2 NAS

NAS (Network Attached Storage) data Storage mode comprehensively improves the previously inefficient DAS Storage mode. It USES a file server which is



independent of the server and developed separately for the network data storage to connect the storage device and form a network. In this way, data storage is no longer attached to the server, but exists in the network as a separate network node and can be Shared by all network users.

Advantages of NAS:

1) true plug and play

NAS are independent storage nodes in the network, independent of the user's operating system platform, truly plug and play.

2) simple storage and deployment

NAS does not rely on a common operating system. Instead, it adopts a simplified operating system designed for user, which is specially designed for data storage. With built-in protocols required for network connection, it makes the management and setting of the whole system relatively simple.

3) storage device location is very flexible

4) easy management and low cost

NAS data storage is based on the existing enterprise Ethernet design, according to the TCP/IP protocol for communication, in the file I/O for data transmission.

Disadvantages of NAS:

1) low storage performance

2) low reliability

3.2.3 SAN

In 1991, IBM introduced ESCON (Enterprise System Connection) technology in S/390 server. It is based on optical fiber media, the maximum transmission rate of 17MB/s server access to memory a way of connection. On this basis,



a more powerful ESCON Director (FC Switch) was introduced, and a set of original SAN system was built.

SAN (Storage Area Network) Storage mode creates the Storage Network. Storage network conforms to the trend of computer server architecture network. The technology behind SAN is the FC Fiber Channel. It is a standard integration established by ANSI for network and channel I/O interfaces. FC technology supports HIPPI, IPI, SCSI, IP, ATM and other advanced protocols. Its biggest feature is to separate the communication protocol of network and equipment from the transmission physical medium, so that multiple protocols can be transmitted on the same physical connection at the same time.

The hardware infrastructure of a SAN is a fiber channel. A SAN built with fiber channel consists of the following three parts:

- 1) storage and backup equipment: including tape, disk and optical disk storage, etc.
- 2) fiber channel network connection components: including host bus adapter card, driver, optical cable, hub, switch, fiber channel and the bridge between SCSI
- 3) application and management software: including backup software, storage resource management software and storage device management software.

Advantages of SAN:

- 1) easy network deployment;
- 2) high-speed storage performance. Because SAN USES fiber channel technology, it has higher storage bandwidth and significantly improved storage performance. The fiber channel of SAN USES full-duplex serial communication principle to transmit data, and the transmission rate is up to 1062.5mb /s.
- 3) good expansion ability. Because of the network structure of the SAN, the expansion capability is stronger. The optical interface provides a connection distance of 10 kilometers, which makes it very easy to achieve physical



separation and storage that is not in the local machine room.

3.3 fast transmission

3.3.1. Background of data transmission

(1) it is now used on the InternetIs the TCP/IP model based on the OSI seven-layer model. TCP/IP model includes five layers, namely physical layer, data link layer, network layer, transmission layer and application layer. The data link layer can be divided into two sub-layers, namely LLC(logical link control layer) and MAC(media access control layer). The cooperation of these layers is the basis of correct data transmission.

(2) ARP protocol (address resolution protocol), whose main function is to convert network layer IP address into data link layer MAC address. There are two ways to map IP addresses to physical addresses: tabular and non-tabular. In Ethernet or on the same LAN, all access to the IP address is translated into a search for the MAC address of the data link layer nic. If host A does not have host B's IP address and corresponding MAC address in the ARP list, it is impossible to reach host B when transferring data.

(3) DNS (domain name server), whose main function is to convert the domain name into the corresponding IP address. In the data transmission of different network segments, host A should first find the network where host B is located according to the IP address and subnet mask of host B, and then find host B according to the MAC address.



3.3.2. Data transmission of the same network segment

Assumption in the same network segment of two hosts A and B want to communication, if you want to send data to B, A must to the IP address of the B and B is obtained and its subnet mask do arithmetic in the network, A will to be the B network number and their comparison, to determine whether B and A in the same network segment, if the same, in the same network segment, if different, is not in the same network segment. If A and B are in the same network segment, but A does not have the MAC address information corresponding to B's IP address, ARP request message is sent in the form of the second layer broadcast, which contains the IP address information of A (source host) and B (target host). All hosts in the same network segment can receive and analyze the ARP message. If the IP address of the target host is found to be different from its own, the message will be discarded. Otherwise, ARP request and response message will be sent to A (source host), whose content includes the MAC address of B (target host).

To reduce the number of broadcasts, the network device keeps the mapping information of IP and MAC addresses in the cache through the ARP table. In the process of ARP request and response, the corresponding relationship between MAC address and IP address of the other party is saved in the respective ARP table, to be used in subsequent communications. The ARP table USES an aging mechanism to remove mappings between IP and MAC addresses that have not been used in a period of time.

If you pass through the switch, the switch has a database dedicated to the MAC addresses of network CARDS on all ports. According to the working principle of the switch it by analyzing the Ethernet packet header information (including the original MAC address, the destination MAC address, the length of the information, etc.), and achieve the target after B's MAC address, stored in the address table lookup switches (MAC addresses corresponding



port), confirm with the MAC address of network card attached to a port on the which, will then send the packet on the corresponding port, is also corresponding to the target host B.

3.3.3. Data transmission of different network segments

In the data transmission of different network segments, host A does not need to obtain the MAC address of remote host C, but sends IP packet to the default gateway, and the gateway network layer completes the forwarding process. If A (the source host) does not have A cache record of the default gateway MAC address, it will obtain the default gateway MAC address through the ARP protocol, so only the gateway MAC address record can be observed in the ARP table of A, but not the MAC address of the remote host C.

(1) to send A packet to C, if A no C IP address, A first to send A DNS request, router B or DNS domain name resolution server will give response to the IP address of the PC, such A about the IP address of the packet network layer and transport layer information: all the source IP address: A, destination IP address: C.

(2) the following A need to know how to get C, A, will send A ARP address resolution request, send the address resolution request, not in order to obtain the target host C MAC address, but the request is sent on to the router B, then the router B will own MAC address will be sent to the source host A, so A packet of data link layer, all the information, the source MAC address: A MAC address, the destination MAC address: MAC address of router B.

(3) and then the data will reach the switch A, switches, see A packet of MAC address, the purpose of the data link layer is to router B, the packets sent to the router B, router B received packets, first check the packet destination IP address, the network layer according to the routing algorithm to get A



routing table, if in their routing table to C routing, suggests that this is A routing of packets.

(4) the next step is routing. First, router B reorganizes or groups the packets by IP and modifies the header of the data link layer -- changing the source MAC address to its own and the destination MAC address to the MAC address of next-hop router C. The datagram is sent to router c. router C repeats the work of router B until it reaches the network where the destination host C is. Then, LAN communication is used to find the corresponding host C.

(5) the process from router B to router C and then to host C is in a huge network system. In this network, a series of protocols should be encapsulated to complete the correct data transmission.

(6) if the packet is lost during transmission or there is an error in the packet arriving at host C, then the role of the transport layer TCP protocol needs to be played to retransmit the lost packet.

Summary: the process of data transmission is in host A (source host), which is to add header for data layer by layer by using some protocol followed. The purpose of adding header is to find the destination host and how to transfer data safely and correctly to the destination host.

In the target host, it is the process of disconnecting header layer by layer from bottom to top. According to the transmitted header, it decides which protocol to send data to the upper layer, and separates header and data unit correctly to receive correct data.

In a network, the routing protocol does the traveling.



3.3.4. Several key issues

(1) if the MAC addresses of A and C are the same, will it affect the normal communication? The answer is not affected, because the two hosts are located in the LAN is separated by the wan, through the analysis of the contracting process can be seen, there will be no problem. If in the same local area network, then there will be communication chaos. When the data is sent to the switch, the port information will have two identical MAC addresses, and then the data will be sent to both hosts, and the information will be confused. So this is one reason to make MAC addresses unique.

A gateway is a device that sends a packet to a host if it wants to do so. This means that the device must have a routing function or a path to an external network.

In the actual networkTypically, a gateway is served by a router or server.

(3) there are generally at least three ways to identify a computer in a network. The most commonly used are domain address, IP address and MAC address, which correspond to application layer, network layer and physical layer respectively. Network management is generally in the network layer for IP address management, but because a computer's IP address can be set by the user, it is relatively difficult to manage, MAC address can not be changed, so the IP address and MAC address together management has become a common way of management.



Chapter four: platform advantages and characteristics

4.1SID digital asset information center advantage

High-performance support

SID's trading system USES memory matching technology and is certified by the authoritative evaluation center. The order processing speed is as high as 1.4 million orders per second, which is hundreds or even thousands of times faster than most systems in the industry and can ensure no delay or lag in order processing. The system has rich and successful application experience in data synchronization of other similar exchanges, and can support smooth and stable operation of exchanges with more than 20 million online users at the same time.

The safe and stable

SID USES an advanced multi-tier, multi-cluster architecture. The design of multi-layer architecture greatly improves the performance, security, stability and scalability of the system. Functional deployment and version update do not need to be shut down to ensure the operation experience of end users to the greatest extent.

Supporting financial products

SID platform initially supports multiple exchange data synchronization, spot trading mode, online leverage trading and futures trading. When platform liquidity reaches a certain level, we do it through the introduction of anonymous transaction functions like ShapeShift, thereby establishing a decentralized information platform and ensuring its liquidity through our own platform.

And later stage will go up gradually other high grade currency. If you have



an agent Bitcoin hopes to be traded on our platform, and there will be more opportunities to participate in our community co-governance.

Liquidity support

SID has rich resources and many partners in the industry. Before the system goes online, we have reached a consensus with a number of foreign mines, investment funds and large trading houses. After the SID platform goes online, their transaction information will be synchronized to our platform. Provide sufficient liquidity for the platform and bring users a good experience.

Multilingual support

The first online version of SID has been bilingual in Chinese and English, and will gradually support other languages including Japan and South Korea in the later stage, so as to clear language barriers for building a world-class blockchain asset trading platform.

Full platform client support

SID provides support for full-platform clients, including the following:

WEB DuanLiuLanQi

Android KeHuDuan

IOS KeHuDuan

PC KeHuDuan

4.2SID platform features

Open and transparent

SID will be open and transparent in real time on a global scale. Through the continuous development and research of blockchain technology, we have made sufficient preparations for creating an open, transparent and co-governance digital asset information center, and made this possibility a reality and



open to the public.

The highest level of trading system

SID transaction synchronization system is the leader in the global block chain data synchronization, because it can achieve the highest level of stability and fast, ensure the efficient. SID provides security-level advanced algorithms that support professional trading orders such as GTT and GTC, and provides specialized support for traders through the integrity JVM, while lightweight client and network mitigation design ensures SID can reach 2.4 million transactions per second.

Robust security

At the beginning of SID development, most attention is paid to the security of digital assets, including transaction processing and information processing.

SID design by the IBC information processing system, and based ON the multiple signature, layered architecture, such as safety design, the use of hot and cold purse, ensure the SID is enough to cope with long-range attack, corrosion attack ON the market at present, most of the attacks, and adoption of zero knowledge proof and PKI – ON – BLOCKCHAIN services provide key is also an important protective line, and the external audit ON a regular basis and public information. SID digital asset exchange white paper



Chapter 5: SID economic model

5.1 the SID

SID promotes tokens as sids, with a constant total issuance of 100 billion and a promise never to issue more.

SID is the public chain. Meanwhile, SID, as the cryptographic rights and interests certificate that can be circulated, is an important part of the general economy era and an important guarantee for the representatives of SID rights and interests and the transparent co-governance of the community. SID holders will enjoy the following rights and interests:

Description of rights and interests:

Participating sids involve SID holders in important transaction decisions through the virtual machine.

Members of the SID community for election and co-governance will be replaced on a regular basis, SID holders will participate in the election of committee members, and the transparency and fair governance of the exchange center will be monitored. Income from a proportional SID is allocated to the SID holder on a fixed proportional basis.

5.2 release plan

Token name: SID

Total circulation: 1 billion

Distribution plan:

5% for cornerstone investment.

10% of the funds are used for private placement, and the funds after private placement are used for the development, operation and marketing of SID.



15% is used to reward the human, resource and technical support of the founding team, early investors and development team in the development process (lock the warehouse for 3 years, release 30% annually after 3 years).

30% of the SID reserve funds will be used by decision-making committees to stimulate ecological construction.

40% of the SID will be distributed to the community gradually, and this part of SID will be used to publish audit reports to the community every year.

5.3 private placement plan

Total collection: 600000USDT

Raising limit: (China proper USDT) (Twenty Founding Nodes)
1000USDT since 200000

Issue price: 1SID=1Cent, 1000USDT as a copy, 1USDT=100SID

To send usdt, you need to add your SI address in the comment bar (the address at the beginning of the ox in the ETF wallet)

Receive usdt address: 35 hb3conjkze5dy71mtsqs2nm4a4zuacms

If private placement is reached, usdt will be refunded

Duration of private placement: 30 day



Chapter 6: team introduction

SID center, founder and executive vice President of Striim

Bachelor's degree in computer engineering and master's degree in management. He is a successful entrepreneur with 19 years of experience in various fields. Several multi-million dollar complex projects were successfully promoted with excellent results.

He is enthusiastic about blockchain and cryptocurrency digital technology. Khal's expertise ranges from, but is not limited to, blockchain business integration, strategic planning, project management, architecture solutions, process improvement, product management, corporate communications, and cross-functional leadership.

Mark Cunningham

He has been involved in business intelligence since 1992, when his family business began developing Crystal Reports, an early and highly successful Windows reporting tool.

Gerge Congle

Internationally renowned blockchain expert, engaged in blockchain for more than 10 years, has participated in a number of blockchain projects, in-depth research on the underlying architecture of blockchain, proficient in writing complex intelligent contract system.

Viayator Saiffe

As a well-known software engineer in SiliconValley and an expert in open source code compilation, he used to be the chief technology expert of Google.



Chapter 7: decentralized communities and SID community commissioners

7.1 decentralized communities

SID advocates the establishment of a decentralized community of self-governance and self-development, and the authority and profit of the platform are fed back to SID community users through the license system.

SID is committed to creating an autonomous, transparent, and universally profitable digital asset trading platform that will use the mining model, the hottest trade in the market, and reward users with more than half of SD's mining fees. This system will promote the cohesion of users to the community and promote the self-development of the community.

As a decentralized community, this community is not affiliated to an organization, and the community operation will vote on a proposal to determine the direction of SID community committee in the best interest of the public.

7.2 community committee

SID community committee is a distributed management organization elected by all SID holders, which solves the problem of voting apathy (the majority of voters do not vote or vote at will) and improves the decision-making efficiency of the whole chain.

The transparent nature of blockmoney will ensure that SID community commissioners always act in accordance with the interests of each community member. In addition, the SID community committee will be overseen by a community monitoring team to ensure that every proposal and vote is fair and



transparent.

7.3 rights, obligations and implementation of community committees

The SID community committee can initiate proposals to run the SID public chain. The proposal was 50%

The above voting support can adjust or upgrade the running parameters of the basic chain.



Chapter 8: SID development planning

8.1 new business model

SID is committed to creating a globally influential platform blockchain through the innovation of technology, product function and operation, and finally integrating the blockchain into the social ecosystem, improving the collaboration efficiency and reducing the operation cost to create a new business model for global economic growth.

8.2 sustainable development

SID establishes an open governance structure for scientific management of intelligent algorithms. Introduction of supervisory and supervisory functions to strengthen the governance of open networks with global node participants.

8.3 commercial application

SID analyzes and screens potential application industries, selects appropriate industries to promote the application of SID public chain technology, enables enterprises to develop and apply SID public chain, and promotes SID publicThe sustainable development of the chain.

8.4 extensive cooperation:

SID provides a completely parallel interface to various resources of enterprises, business, technology and government to maximize resource sharing, make the most efficient use of resources and realize the coordinated development of society.



Chapter 9: SID operation pattern

9.1 exchanges pay to be listed

Use TOKEN credits from SID purchase center to synchronize data.

Many emerging exchange hopes to have more opportunity to present in front of the circle money, SID information center will have docking of the exchange of real-time trading situation analysis and ranking ranking, let circle money can more intuitive analysis of the dynamic through the exchange and the implementation of the various currencies exchange active situation, to provide digital currency authority player data analysis.

9.2 information priority exposure

Use TOKEN credits of SID purchase center, consumption credits and bidding exposure.

Through storage and transmission technology, the most comprehensive market news, information, comprehensive display, news, information synchronization.

9.3 search ranking

TOKEN credits, consumption credits and competitive ranking analysis of SID purchase center are used.

Digital asset data ranking collects comprehensive data of thousands of digital currencies, showing the market value and rise of the same currency Drop, hot search the ranking of different data.

9.4 professional analysis

Use TOKEN credits of SID purchase center to obtain more professional analysis

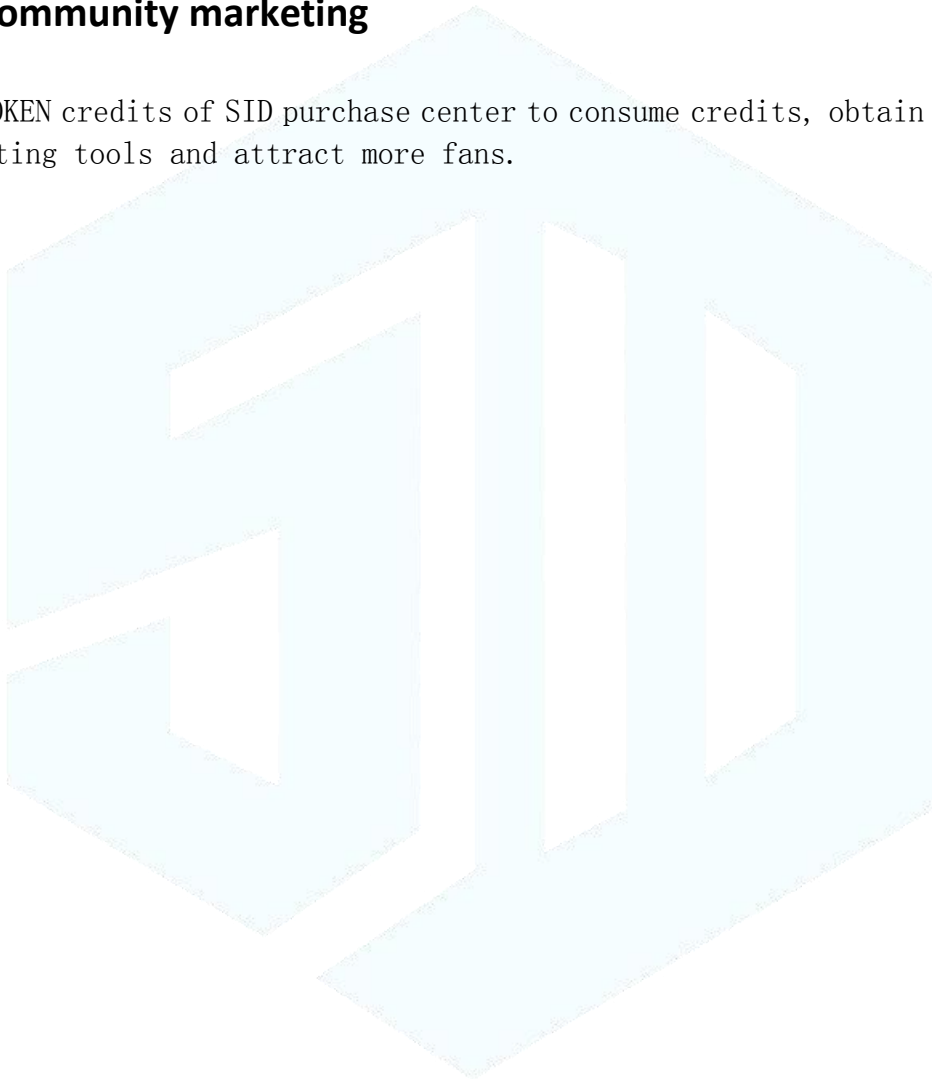


and evaluation data.

Collect data from hundreds of exchanges to show the performance of the same currency in different exchanges.

9.5 community marketing

Use TOKEN credits of SID purchase center to consume credits, obtain community marketing tools and attract more fans.





Chapter 10: SID technology application domain

SID blockchain's high performance, rich supporting functions on the chain and massive data on the chain will lay a solid foundation for its large-scale commercial application. SID blockchain builds a perfect set of infrastructure for the data economy. Many commercial applications can be based on SID blockchain, to provide users with quality products and services.

The blockchain technology solutions provided by SID can solve the problems of data ownership, control, income ownership and data leakage in various fields. Data authenticity problem and data link on the incentive problem and so on. In the near future, the big data economy ecology will be built on the basis of SID blockchain, covering but not limited to personal data rights management, financial services, mobile social, entertainment and games, medical and health care, clothing, food, housing and transportation, etc.

When these applications are created, they will actively join the full SID ecosystem.

10.1 decentralized digital asset exchange

Since the introduction of cryptocurrency, its huge potential has attracted the interest of millions of exchanges around the world, and traditional centralized trading institutions cannot adapt to the rapid development of blockchain technology. The open and transparent nature of blockchain-based exchanges has created an urgent need for a fully decentralized exchange in the cryptocurrency market.

This decentralized cryptocurrency exchange needs to provide a transparent and verifiable trading platform for users. For small and medium enterprises



to provide a completely different solution from the previous centralized exchange. Both parties of digital asset transaction can directly conduct point-to-point transaction and exchange on the basis of the advantages of protecting privacy and curbing fraud. Combined with SID blockchain's whole-network consensus management system, decentralized transactions and information acceleration equivalent points on the chain, it effectively solves network congestion and transaction credit issues for enterprises in the Internet finance field, such as loans, consumer finance and Banks.

10.2 financial services

Financial sector is the earliest field where blockchain technology falls into place. Currently, blockchain technology is also serving more and more extensive financial fields, including banking, insurance, equity transaction, etc. SID blockchain breaks through network congestion and transaction credit granting in previous financial transactions by launching decentralized digital asset exchange. SID decentralized digital exchange provides real and effective risk control services based on the fact that data on SID public chain cannot be tampered with and is more transparent.

10.3 medical and health care

Health is the most important wealth for everyone, and the medical system is an important industry that goes deep into people's livelihood. Its huge scale and influence enable blockchain technology to grow. As an untamable ledger, SID blockchain allows patients and hospitals to trace the origin, medicinal ingredients, manufacturers and wholesalers of the drugs they use. This brings effective solutions to the problems of verifying the authenticity of drugs, eliminating fake drugs, and cracking down on the circulation of illicit drugs.



If this series of blockchain system is implemented correctly, it is a solution that will benefit the country and the people and make great achievements.

10.4 clothing, food, shelter and transportation

In addition to these areas, the SID ecosystem will roll out tens of thousands of applications in the future to serve global consumers. SID's high performance, abundant intelligent contract on the chain and millions of data on the chain can help the gradual implementation of blockchain applications in various fields, including retail, catering, food production, real estate, vehicle travel and other industries.



Chapter 11: disclaimer

Trading and investing in blockchain cryptocurrencies (also known as digital assets) involves significant risk of loss and is not appropriate for every investor. Cryptocurrency valuations can fluctuate wildly depending on market changes; As a result, users may suffer losses in excess of their original investment.

Users will be responsible for the risks they suffer and the financial resources they use. Users should not participate in trading without fully understanding the nature of the market and the potential losses. If users are not fully aware of these risks, they must seek professional advice from their financial advisers.

This document is for reference only and describes the functions on SID blockchain. This is not to be construed as giving legal advice or any assurance as to the trading decisions made by users.

The SID team has disclosed the risks to the user. It is the responsibility of the user to identify and decide which actions to take in accordance with the legal provisions of the relevant jurisdiction.

The SID expressly disclaims liability for any damages, including, but not limited to, the user's own decisions

The following regulations:

- (1) Economic losses caused by users performing transactions on the platform
- (2) Any errors, omissions or misinformation arising out of the personal understanding of the user
- (3) Any loss caused by the trading of various types of blockchain assets by individuals and any behavior resulting therefrom



THE WHITE PAPER OF **SID**