

# Research on Key Technology of Distributed Data Storage in Cloud Computing Environment

Guo-song JIANG<sup>1, a</sup>, Xiao-ling He<sup>2, b</sup>

<sup>1</sup>School of Computer Science, Huanggang Normal University, Huanggang, 438000, P. R. China;

<sup>2</sup>School of Journalism and Communication, Huanggang Normal University, Huanggang, 438000, P. R. China

<sup>a</sup>hustjgs@126.com, <sup>b</sup>hustjgs673521@sina.com

**Keywords:** K-means clustering algorithm; distributed environment; large data set; complexity.

**Abstract.** In view of the large scale and scale of data, cloud computing system faces many problems in scalability and high availability. Aiming at the distributed data storage in the cloud computing environment, a recursive positive N-shaped network model is constructed. Based on the data center network structure, the data management model is given, and the data availability is improved. The purpose of the distribution of the distribution strategy and copy of the strategy. Under the premise of guaranteeing the availability of data, a data migration algorithm based on overlay set is proposed, and the node selection strategy is used to minimize the migration cost. Data migration is carried out to allow more machines to sleep and reduce energy consumption. Through the comparative analysis of the experimental data, the correctness and validity of the proposed network topology, data management model and data migration technology are verified.

## Introduction

Resources in cloud computing include networks, servers, storage, and applications that provide users with universal, convenient, and on-demand network access through these resources that can be quickly used and released without significant management costs. With the rapid development of Internet technology, resulting in the application of data was explosive growth, mass data storage technology has been an unprecedented rapid development. The underlying foundation of cloud computing services is the large-scale data center, tens of thousands of server interconnection, organization, directly affect the cloud computing services, high efficiency and stability. Reasonable network structure is the cloud computing data center high scalability and high utilization of resources an important guarantee. However, with the rapid growth of data storage and transmission requirements, the current network architecture is faced with many problems in terms of scalability, fault tolerance, high performance transmission and cross-regional scheduling [1].

In the cloud computing environment, the underlying data center scale is very large, resulting in high storage costs, where the storage costs not only include a variety of infrastructure costs, including infrastructure in the process of energy consumption [2]. According to research institutions, the utilization of enterprise storage devices is less than 50%, and the utilization rate of equipment is generally low. The solution cost and the growing data storage and computing conflict in cloud computing systems have become a hotspot. Based on the cost, to protect the environment from the point of view, reduce energy consumption has great practical significance and application prospects.

In order to solve the problem of scalability, availability and green energy saving in cloud computing environment, the key technologies to solve distributed data storage include data center network topology design, data management model design and energy saving technology. The data center network architecture is the foundation of cloud computing, which ensures the scalability and resource utilization of cloud computing data centers. First of all, the design of a reasonable network topology, from the high scalability, economy and high efficiency to ensure the quality of cloud computing services. On this basis, the data management technology is studied from the high availability of data. From the perspective of green energy saving, we focus on data migration technology, and finally designed a large data storage scheme with good scalability, fault tolerance and low cost cloud computing environment.

## N-shaped data center network structure

The tree structure is centered on a switch and can be easily extended by adding racks and corresponding switches, but there are also many problems. In order to solve the bottleneck and single point of failure of tree structure aggregation layer, the researchers proposed several improved tree structures, including Fat-Tree and routing methods[3]. The fat tree structure uses this fully connected method to make the upper layer of the tree has a higher bandwidth to solve the problem of bandwidth bottlenecks in the upper layer of the tree structure, but the number of core switch ports limits the scalability of Fat-Tree.

### RNEP network topology

#### Design ideas

In order to eliminate the bandwidth limitation of the switch at the highest level, this paper uses the DCell hybrid structure and B Cube to construct two kinds of data center network structures, and recursively build complex advanced network with low-level network, and ensure the network connection and Classification of routing mechanisms.

The most basic unit of recursive construction is the full connection graph. For its design, with reference to the simple positive n-shape graph, the flexibility of data network construction can be ensured by changing the parameter n to meet the construction needs of different scale data centers. Recursive positive N-shaped network structure RNEP. The network structure refers to the P2P model, the idea of improving the tree structure and the recursive thought. By introducing the variable parameters, recursively constructing and designing the classified route to satisfy the cloud computing data center network Structure for high bandwidth, scalability, fault tolerance, flexibility and green energy and many other needs.

#### Network structure definition and construction methods

The proposed network structure is also recursively constructed. For recursive positive n-shaped network structure, the 0th level network structure all by the server, this layer network only exists within the network transmission, there is no other hierarchical network transmission interaction. In order to take into account the simple and full performance of the recursive structure, here to  $n = 3$  as an example: in the 0-level network structure based on the addition of a switch, and all nodes connected to get Network0'. At this time Network0' as Network1 virtual node, all connected to the three Network0' network switches, get Network1. When the extension is continued, the switch is added to Network1, and the current node is fully connected with the switch. Network1' is used as the recursive unit, and Network2 is obtained by three Network1s.

Similarly, for any n-shaped k-level network expansion, is the first level of the network to add a switch, and k-level network n nodes were connected, and then as a recursive unit, and other  $n - 1$  k Network-wide network connection to form n-shaped  $k + 1$  level network structure. This construction is constantly through the addition of a fixed structure to expand, the entire network structure strictly modular connection, making the network expansion is simple and feasible. At the same time, the internal structure of  $n - 1$  is complete and does not affect the structure of the nth - level structure. After the construction of the nth level, the modification of any low - level internal structure will not affect the superstructure. There is a guarantee of independence between different levels.

#### Routing policy

In the routing strategy, the source nodes  $src(i_0, i_1, i_2, \dots, i_{k-1}, i_k)$ , the destination nodes  $dst(j_0, j_1, j_2, \dots, j_{k-1}, j_k)$ , and the packets sent by src are defined. At the same time, the source packet contains the identification information of the source node and the destination node and the packet ID. Any one of the packets can be uniquely marked by the ID of the source and the source node. When the target node receives the packet, it can also obtain the information of the requesting node according to the source node identification information carried in the packet. The flow of the routing algorithm is shown in Fig1.

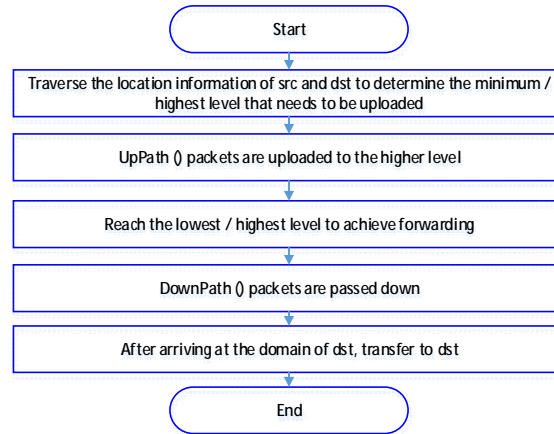


Figure 1 routing strategy flow diagram

The routing strategy algorithm is shown in the following figure. The routing strategy of the k-level RNEP structure can be used to analyze the transmission path length.

```

NetworkRouting (n, k, *srcAdd, *dstAdd)
{
    // Determine the highest level that needs to be uploaded
    for (intCurrentLevel=k; CurrentLevel>=0; CurrentLevel--)
        if (srcAddCurrentLevel!=dstAddCurrentLevel)
            break;
    // transfer packets from src to the highest level switch (SWCurrentLevel)
    UpPath (src, SW0) ;
    for (inttemp=0; temp<CurrentLevel; temp++)
        UpPath (SWtemp, SWtemp+1) ;
    // Through the currentlevel switch at the level of subnet across the subnet into the dst where the subnet,
    // because a high-level switches will be interconnected with several low-level switches, so down the
    // DownPath function in the need to introduce the relative position parametersfor
    (temp=CurrentLevel; temp>0; temp--)
        DownPath (SWtemp-1, dstAddtemp) ;
    // The packet arrives in the base domain where the dst is located
    Path (SW0, dst) ;
}
  
```

### RNEP network structure analysis

As a result of the recursive construction, compared with the traditional tree structure, the proposed network design in the physical structure has a good scalability, can support the network scale exponential growth.

According to the network construction process, we can see that each server needs only to connect with other  $n - 1$  servers and one of the other switches in the domain. Therefore, each server needs to support  $[(n - 1) + 1]$  links. In the actual construction, the structure proposed in this paper is feasible. At the same time, the RNEP network structure can guarantee the lower ratio of the switch and the server, and the number of the switch is  $2n - 1$ , effectively saving energy [5].

### simulation experiment

Simulation experiments were carried out using NeuroGrid. The experiment was to construct the RNEP network structure and simulate the routing algorithm. Verify the running condition of the routing mechanism without any failure, and analyze the transmission performance through the average path length. Taking the regular hexagon as an example, the experiment produces a network with an initial size of 42 to 1 554 nodes. In the experiment, from a node, to all the other nodes to pass the message, record the sum of the number of hops to obtain the average. For the RNEP structure at  $n = 6$ , the total number of nodes and the average shortest path length are shown in Table 1.

Table 1 related parameters of different levels of RNEP structure while Non node failure and  $n=6$

Current level	K=1	K=2	K=3
The total number of servers	36	216	1296
Total number of switches	6	42	258
Reachable nodes	35	215	1295
Number of unreachable nodes	0	0	0
The sum of the shortest paths	130	1210	9850
Average shortest path	3.714	5.627	7.606

Table 1 shows the average shortest path to any of the remaining nodes when  $n = 6$  in the case of non-node failure. Since the node failure is not considered, the number of unreachable nodes is zero. The average shortest path length is consistent with the algorithm, and the experimental results are better than the theoretical results of the algorithm.

Taking the network structure with  $n = 6$  and  $k = 2$  as an example, the number of nodes is 258. The relationship between average shortest path length and link failure rate is analyzed experimentally. The link here refers to the switch and the server link is broken, rather than the server (node) failure. In the experiment set the link failure rate range of 5% to 30%, 5% increase. The experimental results are shown in Fig2.

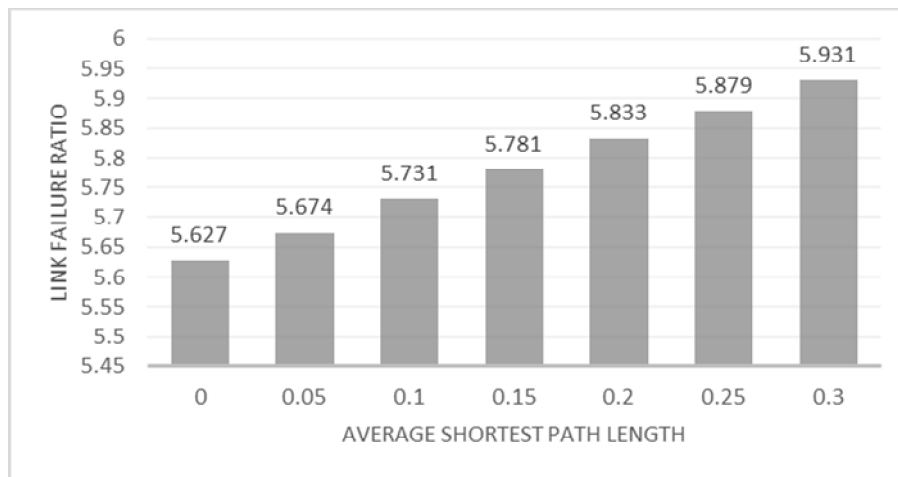


Figure 2 The average shortest path length changes with the link failure rate

It can be seen from Figure 2, when the node failure rate increases, the average increase in hops in about 1 hop, the average path length growth is slow.

The experiment of  $n = 6$ ,  $k = 2$  is carried out to analyze the change of the path failure rate with the increase of the node failure rate and the comparison with the experimental results of the DCell structure. In the experiment, the failure rate of the designed node was changed from 2% to 20% and increased by 2%. The number of failure paths was recorded by flooding to obtain the path failure rate. The statistical results are shown in Figure 3.

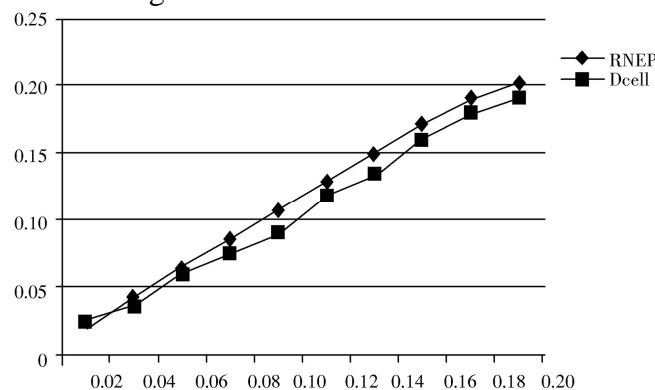


Figure 3 Comparison of structural path failure rate versus node failure rate

## Data management model of RNEP network structure

### Data management model of network structure

Storage Server Clusters enable scalable storage services that dynamically add and remove storage nodes. The primary function is to actually store the data files. Data requestors operate on data through communication with the primary management server and the storage server cluster. The entire system external shield to achieve the bottom, to provide users with fast, transparent, stable and reliable data storage services[6].

The service management node manages the block server within the cluster, monitors the operation of the data block server, and provides the location of the data block server for data creation and data access. The metadata management module is responsible for managing and maintaining the metadata of the file system. The system monitoring module is used to view the running status of the current system. The load balancing module is responsible for assigning the user's request.

### copy management method

Depending on the location of the replicas, the availability and consistency of the replicas are not changed during the access process. Replicas on multiple storage nodes in the same domain can maximize the use of network bandwidth within the domain, prevent machine failures, and disk damage Impact, but also reduce the inter-domain write operation of the network communication [2].

When a new data is uploaded to the data management model, the client presents a copy request to the service management node, and the service management node places the copy with the following policy:

(1) the default number of copies is 3, the service management node according to the storage node factors, the network within the network conditions and the client's location information for the first copy of the node location selection, storage node factors include: disk utilization is low; Limit the number of "most recent" created by each storage node.

(2) The second copy selects the optimal storage node in the same domain according to the storage node factor.

(3) The third copy is selected according to the location selection strategy of the first copy in different domains. This strategy reduces the data transmission between domains and improves the efficiency of the write operation.

Two of 3 copies are in different domains to ensure the reliability of a copy. The copy selection algorithm flow chart, shown in Figure 4.

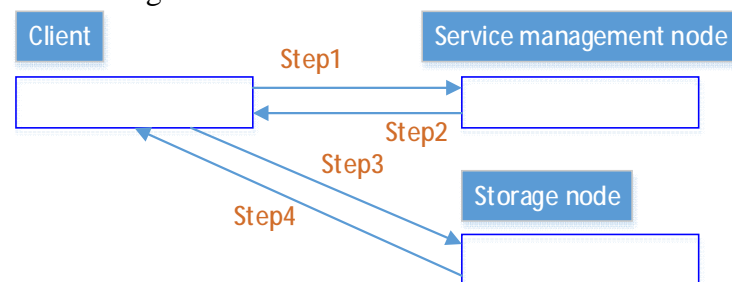


Figure 4 copy selection algorithm flow chart

### simulation implementation and analysis

Refer to the OptorSim simulator to extend the capabilities of the cloud simulator CloudSim and perform a simulation experiment on the extended simulator. A copy of the placement and replica selection strategy based on the RNEP data management model, a comparison strategy for HDFS replica placement strategy and a copy selection strategy; HDFS replica placement strategy: if the client has a data node, then the random will Copies are placed locally at the local node, otherwise a suitable data node is randomly selected throughout the cluster as the local rack node at this time, the first copy is placed on the local rack; the second copy is placed on the local rack Another node; the third copy on the different rack nodes. The HDFS replica selection strategy for the client always chooses to read the required data blocks from the nearest available data node.

Each Datacenter contains 10 Hosts with a maximum power of 250 W, a processor capacity of 2 000 MIPS, 2 GB per host memory, 100 Mb / s host bandwidth, 100 data blocks per host, and each data

The block has three copies. The first data center simulates a data management model based on the RNEP network architecture, and the second data center simulates the HDFS strategy.

(1) Comparison of the copy read time: copy of the time is read from the client to read the request to read the copy of the time spent reading. In the two data centers were randomly selected a client using their respective strategies for 100 random read operation of the average time to compare, as shown in Figure 5.

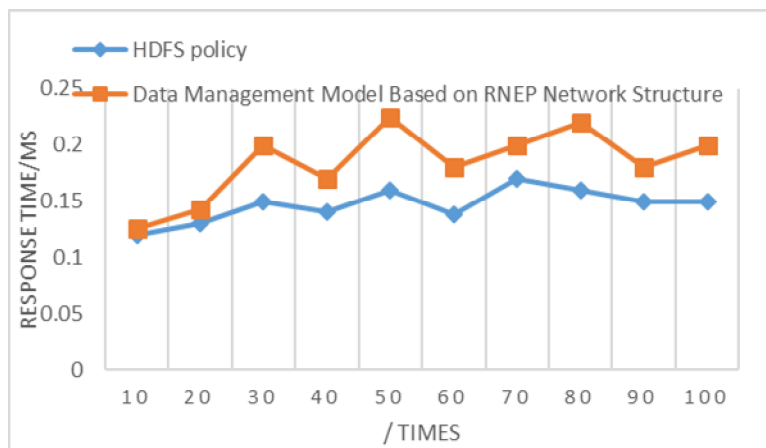


Figure 5 Average read time for both strategies

As shown in Figure 5, the average time for replica reads based on the RNEP data management model is significantly less than the average read time of the HDFS policy.

(2) Comparison of load balancing: In the two data centers randomly selected a data block, of which seven clients from the selected data block in a copy of the relatively close, the other three clients located in the far node, and then this 10 clients randomly send read requests to this block, using their respective policies: selecting the HDFS policy closest to the client distance and the replica selection strategy based on the RNEP network structure data management model. The load is expressed as the number of client-side read requests received by the storage node where the replica is located, and the total amount of load in the 3-minute storage node is recorded every 3 minutes.

The three replicas of the HDFS policy are not balanced, and the client chooses the nearest replica so that the client's access is gathered in a copy, resulting in a high load on the storage node where the replica is located and the storage load on the other two copies, resulting in cluster load is not balanced. Based on the RNEP data management model, the three replica load gaps are not large, and the effect of load balancing is achieved. Experiments show that the data management model based on RNEP network structure can improve the access performance of data, make reasonable use of network bandwidth and make the system load balance.

### Data migration method based on overlay set

In the cloud storage environment, data migration is a complex task, in order to better energy saving, to meet the premise of data availability, sleep more machines, save more energy[7]. In this paper, we propose a data migration algorithm based on overlay set. Before data migration, we need the following information: source server, data object, target server, migration cost. The flow chart of the data migration algorithm based on the overlay set is shown in Fig6.



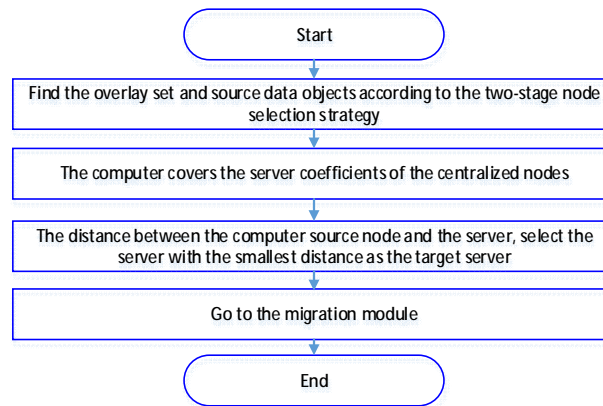


Figure 6 Flow chart of data migration algorithm based on overlay set

In the extended function of the Cloud-Sim simulator on the simulation experiment to achieve based on the overlay set of data migration algorithm[8], Cloud-Sim energy consumption model based on CPU utilization to calculate energy consumption, by modifying the Power Energy Model used energy model calculation Energy consumption, and set the sleep power consumption of the host is 0.

At every 1000 s, the energy consumption of the statistical data center, the energy consumption of various strategies compared to the data shown in Figure 7. Based on the data migration strategy of the overlay set, although the energy consumed in the pre-data migration process increases the total energy consumption, the dormant nodes after the data migration allow the data center to consume less energy than the data migration strategy Energy is less, and with the increase of time this advantage is more and more obvious.

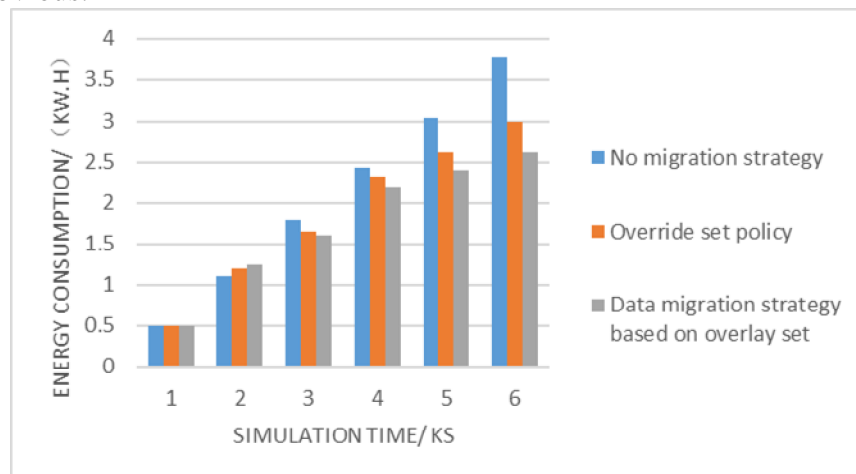


Figure 7 Comparison of the energy consumption of each strategy

The number of nodes in the data center is 20, 100, 300 and 500, respectively, as shown in Figure 8, given the comparison of the three different methods in terms of power consumption, where the cluster scale represents the physical Number of nodes.

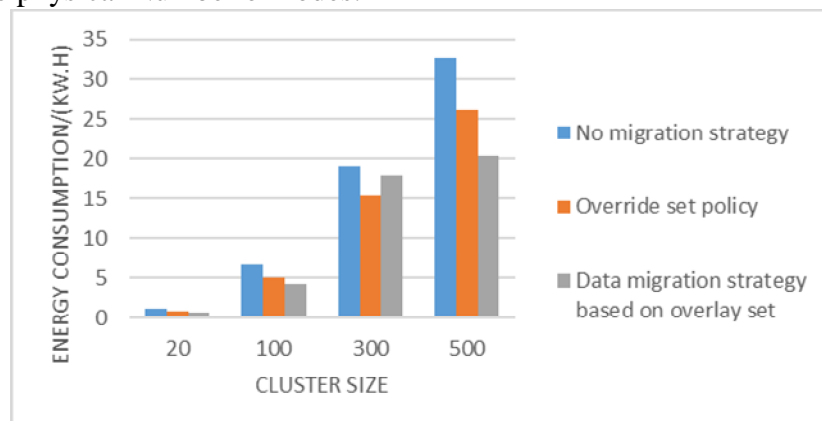


Figure 8 Comparison of energy consumption at different sizes

The data shows that the larger the scale of the cluster, the more energy savings from the data migration method based on the overlay set. This is due to the data migration method based on the overlay set. In the larger cluster, the number of dormant machines is reduced, which reduces the energy consumption of the data center.

## Conclusion

Cloud storage through cluster applications, grid technology or distributed file system and other functions, the network will be a large number of different devices together to work together to provide external services. In this paper, we study the network structure, data management model and data migration algorithm in cloud storage, and propose a recursive positive N-ray network topology and its routing mechanism design. At the same time, we design a data migration method based on overlay set. So that the cloud computing system is easy to expand, improve access performance, rational use of network bandwidth, reduce system energy consumption, both an important academic value, but also has broad application prospects.

## ACKNOWLEDGMENT

This work is supported by Research project of Hubei Provincial Education Department (No.D20152903,15Y159), Foundation of Huanggang normal university(No. xfg2015A11, 2014015103, 201616303), Thanks to the reviewers for the valuable comments helping to improve the quality of the manuscript.

## References

- [1] (1)GREENBERG A, HAMILTON J R, JAIN N, et al. VL2: a scalable and flexible data center network [J]. Communications of the ACM, 2011,54 (3): 95-104.
- [2] (2)CHOU J, KIM J, ROTEM D. Energy-aware scheduling in disk storage systems [C] // Proceedings of 2011 31st International Conference on Distributed Computing Systems. [SI]: IEEE, 2011: 423-433.
- [3] (3)ABU-LIBDEH H, COSTA P, ROWSTRON A, et al. Symbiotic routing in future data centers [J]. ACM SIGCOMM Computer Communication Review, 2011,40 (4): 51-62.
- [4] (4)AVERSA L, BESTAVROS A. Load balancing a cluster of Web servers using distributed packet rewriting [C] // Proceedings of IEEE International Conference on Performance, Computing and Communications. Phoenix: IEEE, 2000: 24-29.
- [5] (5)Zuo Xiaojing, Wang Rong. A Distributed Storage Strategy for Massive Data Based on Cloud Computing [J]. Henan Science, 2015,33 (12): 2135-2140.
- [6] (6)Cheng Fu-chao, Miao Fang, Chen Ken. Adaptive Distributed File System Metadata Management Model [J]. Computer Engineering and Design, 2014,35 (3): 867-874.
- [7] (7)WU Dan, WANG Gai-yun, LI Xiao-long. Weature data clustering algorithm based on minimum covering set [J]. Computer Engineering, 2012,38 (2): 97-99.
- [8] Liu Yunfei. Hadoop-based data migration and storage research [D]. Beijing: Beijing University of Posts and Telecommunications, 2015.