DESIGN AND IMPLEMENTATION OF CONTENT DELIVERY PROTOTYPE SYSTEM ARCHITECTURE INTEGRATING ISP

Chao Chen, Fei Gao, Chuan Liang, Chao Li
Process Engineering Research Department,
Petrochina Petrochemical Research Institute, Beijing, China
chenchao010@petrochina.com.cn
gaofei@petrochina.com.cn
liangchuan@petrochina.com.cn
lc459@petrochina.com.cn

ABSTRACT

The availability, reliability have become the key issue of Internet technology because of the growth of Internet traffic and the demand of access speed. Content delivery networks focus on optimizing Internet quality of service. We design and implement a content delivery system for ISP. The system is different from traditional CDN system in two parts: (1) Parsing the network data passively and hierachy deployment; (2) Security monitoring for ISP. The system includes three parts: information collection, cache management and auditing offline. Under the practical evaluations, our deployment machanism, cache management and security monitoring method achieve our initial research objectives.

INDEX TERMS: *overlay network, content delivery networks, deployment mechanism, content management.*

I. INTRODUCTION

With continuous emergence of Internet services with high bandwidth consumption, network structure only depending on high-performance data center cannot provide reliable services for users from around the world, besides, both increasing traffic and asymmetry between capacity and resource of server and that of client contribute to the emergence of content delivery technology. In the late 1990s, network congestion has been the key factor to restrict Internet performance. Therefore, researchers came up with CDN (Content Delivery Network) technology to enhance access experience of users on the edge of Internet [1]. At present, CDN is a kind of important data distribution service on the Internet, which has been widely applied to distribution of multimedia data. As professional and future customizable CDN integrating ISP, it is confronted with several core issues as follows:

A. Implementation of CDN (Content Delivery Network) technology which is easy to deploy

In order to establish ISP-oriented CDN (Content Delivery Network), the first problem to solve is deployment of distribution node, and optimal deployment strategy is not feasible in actual implementation due to complexity of calculations as well as dynamism and heterogeneity of the network. However, hierarchical deployment strategy can be applied. First, ISP's management of topological information can be efficiently utilized to build CDN node overlay network and then, improved RCBF index can be adopted to reduce distributed cache synchronization cost, thus to build low-level interconnection mechanism of integrated CDN.

B. Construction of flexible and easily extensible content management mechanism

As upper-layer core technology on content delivery network, performance of content management directly influences efficiency of the overall distribution system. And content management includes two key technologies, cache replacement and cache push. Due to finiteness of cache space, cache replacement is the key strategy for cache system. This paper puts forward multi-dimension cache replacement strategy by taking advantage of the application layer where CDN is and presents content push method based on degree of resource value by combining physical topological information of ISP. And flexible content management mechanism can be implemented in two aspects including application layer and topological information.

C. Capacity of service quality and security monitoring for ISP intra-domain network

Content delivery network is intended to improve network service performance, and problems it focuses on such as global load monitoring, content routing and content management are all related to performance optimization. However, ISP doesn't only need to enhance user experience and reduce operation cost but also pays attention to the core problem of security. Besides, imperceptibility and group characteristics of network security incidents constantly strengthen, and content delivery system can obtain user's visiting data in the application layer. Therefore, integrated CDN system shall not only include performance optimization as for its design, but have the ability to analyze data security, providing certain security support for ISP.

The paper focuses on the above-mentioned aspects to introduce detailed design of prototype system and includes the following parts: part 1 is about related work; part 2 is about overall design ideas of prototype system; part 3 is about module structure and operating principle in prototype system; part 4 is about design of experiment; and the final part is about summary.

II. RELATED WORK

Exponential growth of Internet traffic and users' demands for network speed have contributed to key problems of Internet technology such as accuracy, availability and reliability of network resource delivery. In the 1990s, due to limit of network access bandwidth, researchers put forward application of proxy cache technology to improve network performance. And proxy technology is intended to deploy proxy mirror server close to content server or some points of presence and to improve quality of access services while reducing load of source end. In proxy service pattern, if the resource that the user requests has already been cached by proxy server, then the user can directly obtain resource from proxy cache pool, directly minimizing resource delivery time. Cache technology has remarkable advantages: it can cut down bandwidth consumption and network congestion by shortening transmission distance of resource in the network, and even when users cannot access the origin server, they can still obtain resources they need through cache server, enhancing resource availability and service dependability. However, because of strong locality of proxy technology, it has a limited effect on promoting efficiency of the entire Internet. Today's proxy cache technology centers on hierarchical cache, i.e. multi-level proxy services are deployed on network edge and backbone as well as core network to improve service performance and to lower bandwidth consumption [2][3]. In the wake of proxy cache technology, researchers put forward the concept of service field [4][5], but the service field is just intended to deploy server cluster and layer 4-7 switch near source server to distribute network requests.

Table 1: Timeline of content delivery networks evolution

	Time	Business object	Brief description
First-generation CDN （Web-based）	1998-2002	Static/dynamic network data;	Distribution mainly for static/dynamic web pages;
Second-generation CDN （Video on demand-based）	2002-2006	Multimedia data;	Distribution mainly for media stream of mass data;
Third-generation CDN （p2p-assisted）	2006-present	Shared data;	Reduce load of server by integrating P2P network;
Fourth-generation CDN （cloud-based）	2009-present	Integrated data;	Solve problems including resource integration and unified management of service by integrating cloud computing platform;

With the continuous advancement of Web technology, proportion of static pages in entire Internet traffic continues to decline, and it is difficult for proxy cache technology to meet content distribution demands in consequence of rapid growth of dynamic interactive data and multimedia shared resource. Besides, a large quantity of network data generated by users cannot rely on cache technology for fast transmission and interaction. What's more, deployment strategy near source end is not a good solution for phenomena similar to flash crowd [6] and Slashdot [7]. In 1998, Leighton et al. of Massachusetts Institute of Technology first proposed a new overlay network structure [8][9], content delivery network, in order to solve problems of time delay when long-distance transmission as well as network congestion on the entire Internet, and then established the world's largest content delivery network service provider Akamai based on this new technology. After more than 10 years of improvement, content delivery network has developed into its fourth generation, and its evolvement process is shown in Table 1.

III. SYSTEM ARCHITECTURE

Integrated content delivery network system is typical application of distributed overlay network, but there are some differences between this system and traditional CDN system:

1. Function of passive parsing and storage. As traditional CDN works with service provider (CSP), all content data is sent to content distribution node from source site. And integrated CDN needs to passively store resources outside the domain requested by users.

2. Node network resource push. Traditional CDN forwards user requests by global load balancing management services to decrease possibility of single-point overload; and instead of being widely distributed on the whole Internet, integrated CDN is deployed in domain network under the jurisdiction of operators, so content passively cached is hierarchically stored in the entire cache node network.

3. Security warning. Traditional CDN only focuses on its own security, and CSP puts source site behind CDN system by building CDN to lower risk of being directly attacked without analyzing the security of the network; and as content delivery system which can be controlled by ISP, the integrated

CDN can carry out offline analysis of in-network events by getting application layer data thus to realize user classification.

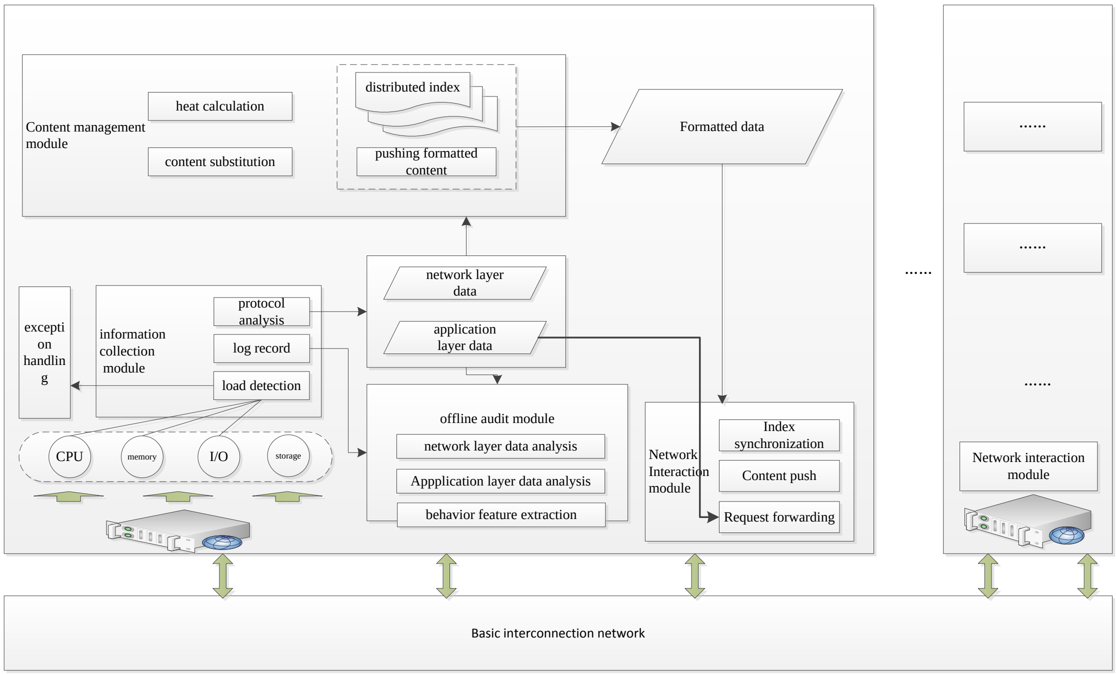

Fig.1 The framework of merged content delivery system

The entire system framework is shown in Fig. 1 and is mainly divided into four modules: (1) information collection module, including protocol analysis of network data, collection of user log as well as monitoring of system parameter; (2) content management module, including integrating and storing extracted data resource, creating index table, computing popularity of resources as well as pushing formatted content; (3) offline audit module, including extraction and analysis of log data, establishment of user trust system based on behavior as well as realization of macrocosmic monitoring and warning for in-network security events; (4) interaction module, i.e. packaged formatted content (index table and hot cached content) push. And the whole data flow is shown in Fig. 2, (1) after user makes a request, and then the request reaches inter-domain cache node layer, (2) the system will store the request and establish a corresponding index, (3) after the server outside the domain returns data, (4) inter-domain cache node first caches the data and computes corresponding popularity, and when degree of value reaches given proportion, (5) the data will be pushed to core cache node, and similar push strategy is also applied to the migration from core cache data to edge cache data, and when changes of cache data exceed the threshold value, (6) index synchronization can be implemented.

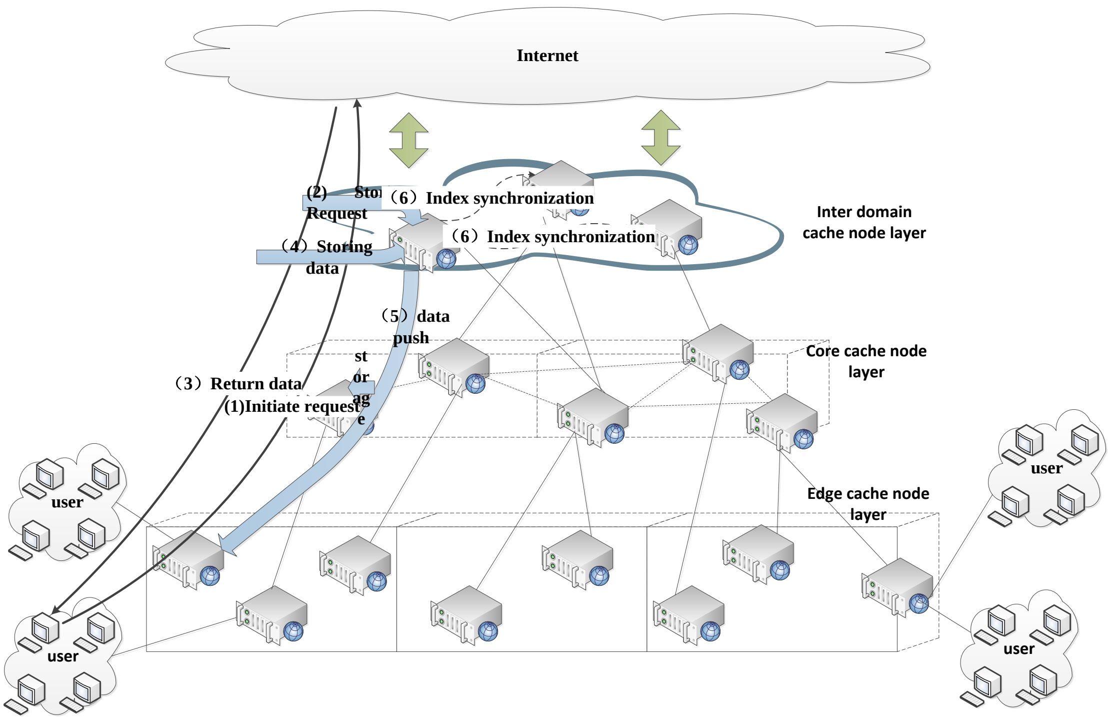

Fig. 2 The data flow of system

IV. MODULAR STRUCTURE AND WORKING PRINCIPLE

A. *Information collection*

Since integrated CDN adopts the method of passive storage to store the data user requests, first protocol analysis shall be carried out for the request data so as to obtain complete data; then load monitoring shall be carried out within the cache node to guarantee stability of the system; finally, log records shall be implemented for information user requests to provide data for offline audit.

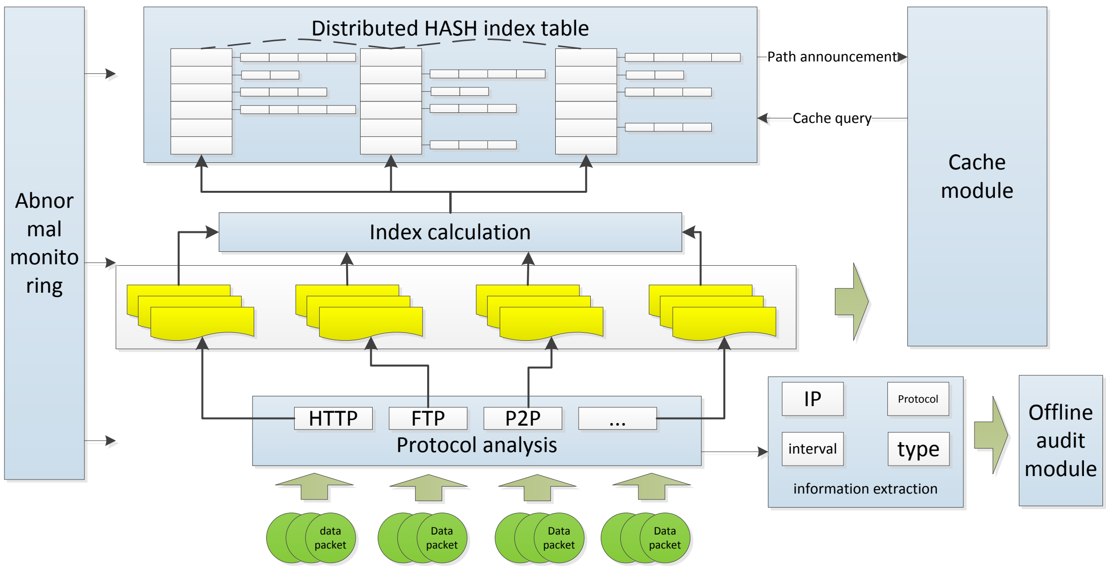

Fig. 3 The module of Data Collection

Details of collection module are shown in Fig. 3. After the data package arrives, classification of the package can be carried out through protocol identification. Then complete data can be generated according to header information of each protocol. Next, these data will be indexed and computed by means of Hash based on URL, and computed values will be stored according to RCBF structure; while in the process of protocol analysis, statistics will be made for user's IP quaternion, adopted protocol type and data file format; finally, the formatted data and log will be transmitted to cache module and audit module respectively; in order to guarantee stability of the system, there is need to build daemon process for monitoring exceptions of key system performance. In addition, new protocols can be freely added into the process of protocol analysis according to new requirements, maintaining scalability of the system.

Information collection module is mainly divided into three sub-modules:

(1) Load monitoring: responsible for monitoring server cluster in each node, including periodical collection of main parameter such as CPU, internal storage, disk, network I/O of the server, then send data feedback to load scheduling controller thus to realize balance among servers within the node.

(2) Protocol analysis: responsible for protocol analysis of data package, with analysis of many protocols included, primarily aiming at analysis of application layer protocol above TCP protocol. With HTTP as an example, GET data package is first received in HTTP, and hash process will be implemented for URL in GET package, then hash value will be inserted in index table.

(3) Behavior statistics: responsible for collection of host behavior within regulated domain through analysis of mirroring of gateway data package and user's request behavior collected by CDN. This mainly includes quaternion statistics of gateway data and statistics of mapping relationship between URL user requests and the user, as well as creation of user's behavior attribute table. As shown in Fig. 4, request behavior attribute table lists statistics of domain names accessed by corresponding source IP, label statistics and the geographical location; and table of behavior statistics lists interactive data volume, interval and duration of source IP and destination IP. And both of the statistical tables are analyzed by audit module.

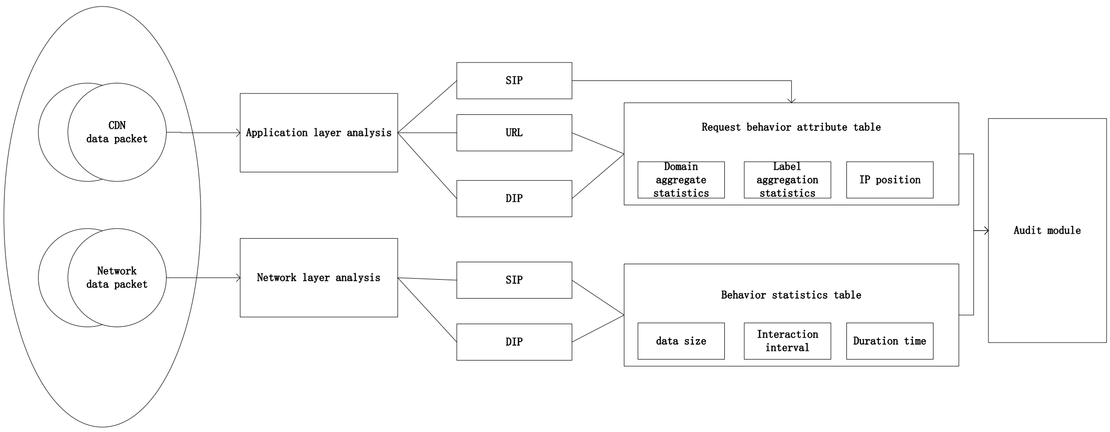

Fig. 4 The procedure of behavior audit

Fig.3 is the module description of inter-domain cache node. With regard to core caching and edge caching, it only needs to parse the users' requests, inquire local index-table and extract information with low complexity.

B. Cache management

After the information collection module transfers the formatted data address to cache module, the cache module needs to set up a mapping table which cache queue applied for mapping cache objects.

As shown in Fig. 5, (1) when the new data arrives, (2) through mapping table of query objects check whether it's already existed or not, and then its heat degree can be calculated according to its density and size, (3) cache queue can be inserted; while the push mechanism will proceed packaging of selected cache data and transfer it to network interaction module via monitoring the changes of cache queue and based on domain valuations and the value of calculating object of users' behavior information.

Cache management module is the core component of the system. It adopts multiprocessing approach to access some shared-classes data to improve its system performance. Besides it also needs to extract domain name information of cache data to set up domain statistical table during the degree of value calculating period.

Cache management module mainly includes two procedures:

Local cache management includes distributed cache service among internal server cluster of each node, of which cache resource localization process is shown in Fig.5. Receiving process of internal server of each node charges the treatment of users' requests, calculates corresponding hash value based on URL and then locates the cache node which stores the data. With regard to the internal cache cluster of single content node, the system employs distributed virtual node storage mechanism to further improve efficiency of cache management. Moreover this system maps single physical storage node into mutiple virtual nodes and makes these virtual nodes into hash ring to achieve cache load balancing.

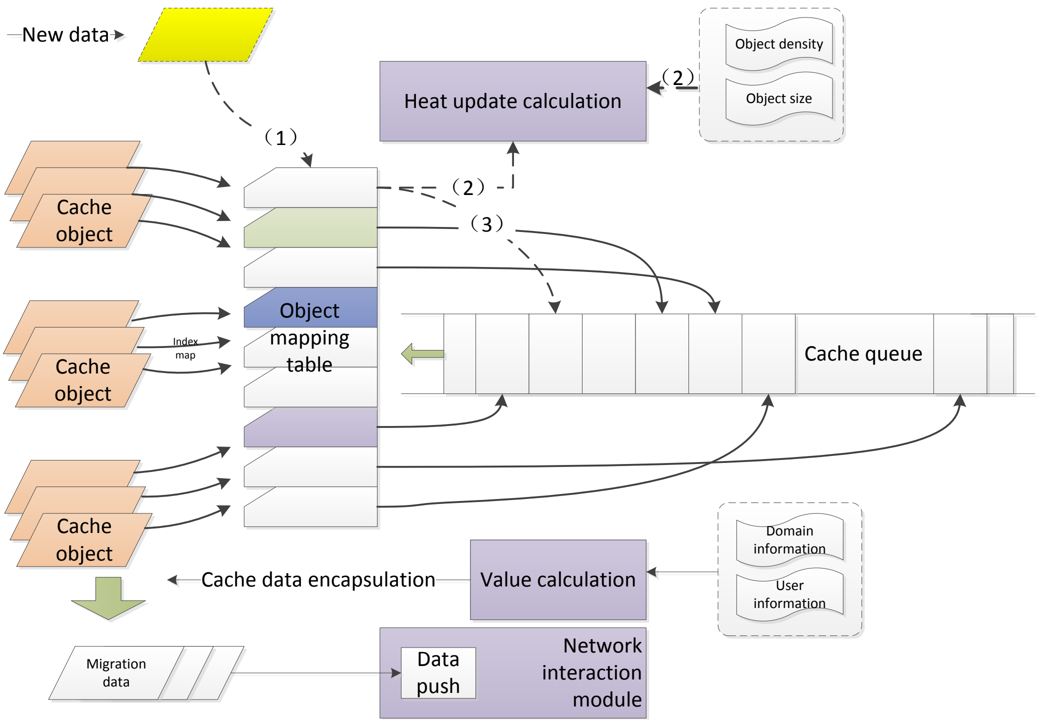

Fig. 5 The procedure of cache management

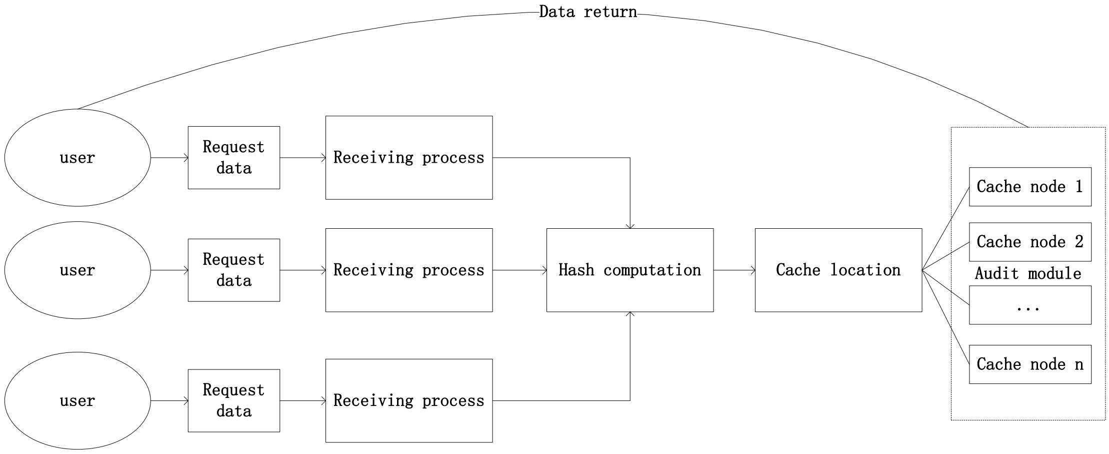

Fig. 6 The procedure of cache localization

(2) As shown in Fig. 6, remote cache management will search whether it exists in the remote node or not via RCBF index when it has a cache miss of users' requests. If it exists, then the request will be transmitted to the remote node. When the remote node receives the request, it will return the corresponding data to the local cache node and then the local cache node will calculate corresponding degree of value and at the same time return the data to users. When the index changes exceed threshold level, the index synchronization will be generated, of which there will be mutual sync for inter-domain nodes without the involvement of intra-domain nodes, for the requested data is only stored in inter-domain nodes for its first arrival while the cache data of intra-domain nodes all derive from inter-domain nodes.

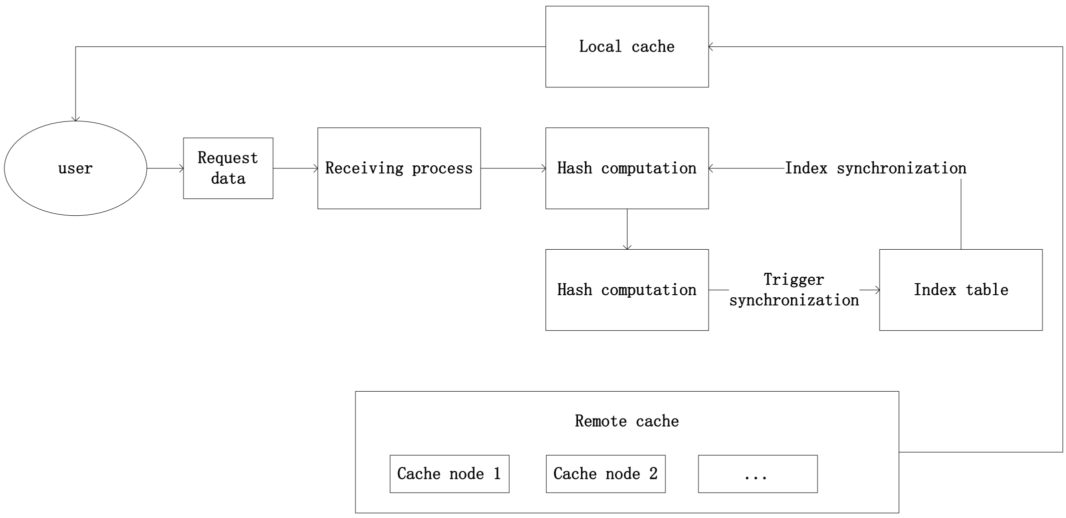

Fig.7 The work procedure of remote cache

C. Offline audit

Different from traditional CDN audit, offline audit is the proprietary module of integrated CDN, and its main job is to analyze and process the accessing behavior of network users under the management of IPS, constructing the exception detection module which based on reliance. The module is independent module which achieves the users' and visiting server's information from the two modules above.

In need of analyzing the interaction behavior of users in the domain, so the data shall be hunted and collected to acquire the statistics of network layer and application layer through information collection module. The whole process could be divided into four steps below, which is shown in Fig.7:

Data processing: network data acquired by mirrored routes has a lot of data partitioning or other data under three-tier protocol. First all these data can be sorted and the statistical information can be acquired based on standard four tuples to set the benchmarks for topology of network behavior.

Topological construction for network interaction: topology is computed through four tuples information. When interactive node information is calculated, nodes will be created only based on IP addresses. The edge weight of topology is calculated by the interactive numbers and interactive duration of nodes.

Behavior clustering: node similarity degree matrix can be calculated through restored connectivity and populated dimension of original interactive topology, and categorized results can be obtained through spectral clustering.

Probability assignment: analysis is made for network packets via traditional malware detection mechanism. If some IP address has malicious behavior, then the credibility of other IP among this IP class will be lowered.

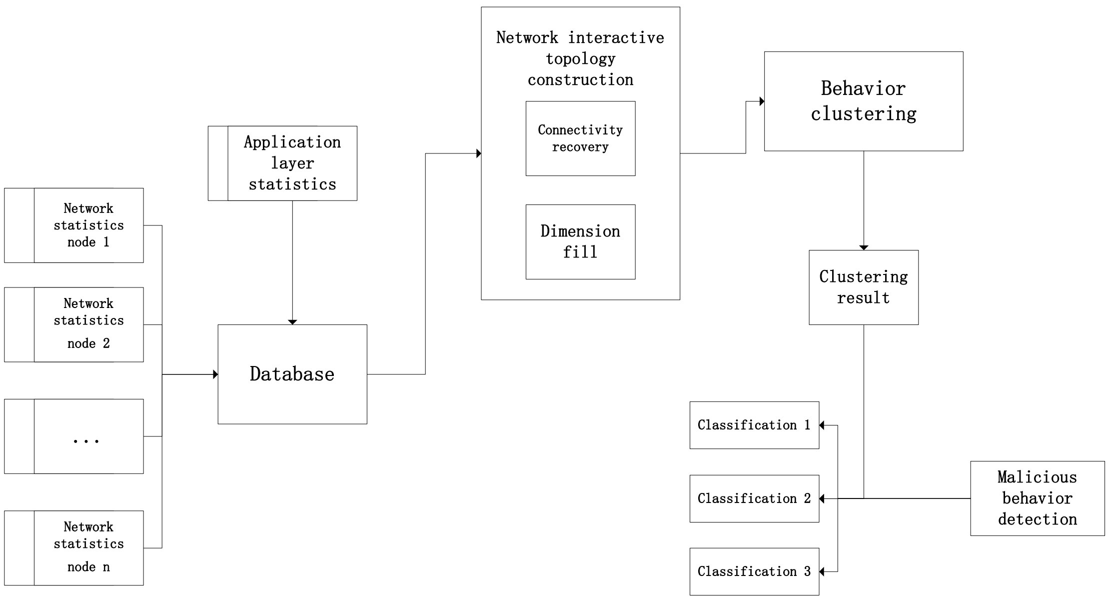

Fig. 8 The procedure of offline audit

D. Node communication

Interaction between each node is carried out by node communication module. Interactive data between nodes mainly includes three types: (1) cache data transmission; (2) index synchronization; (3) audit data transmission. The first two is about the data interaction of cache system and the third is to collect the network layer data gathered by each node to transmit to the offline audit processing node.

In view of efficiency of network interaction, fixed listener port is opened in every inter-domain node to communicate, of which the communication of cache data adopts UDP protocol to send, for the cache data for reliability demands is not high. When some data is lost during transmission, the cache data still can get it in remote node through index; while index transmission adopts TCP protocol, for in consideration of the amount of space occupied by index is not much and the credibility of index is directly relevant to whether it is a cache hit or a cache miss of users' requests and to network traffic and service experience.

Audit data transmission also adopts UDP protocol. It will send the formatted statistics to central audit server and data sent by each node includes behavior statistics and application level statistics. Audit data transmission transfers periodically and each node will summarize data at a regular time.

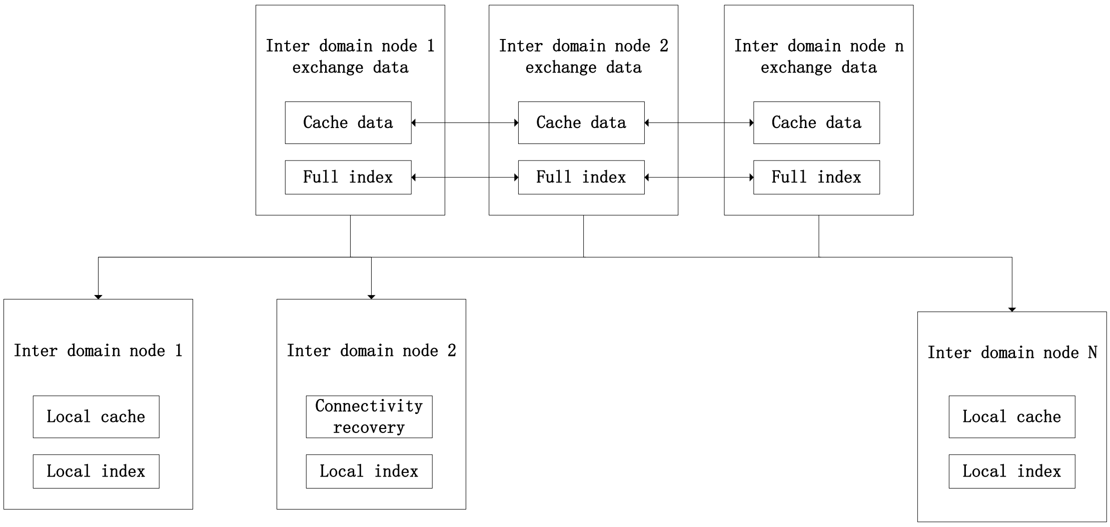

Fig.9 The mechanism of node interaction

The interaction between cache data and index is shown in Fig.8: there is cache data interaction and synchronization of full index for inter-domain nodes, while index and cache data of intra-domain nodes only derives from inter-domain nodes and there is no interactive mechanism for intra-domain nodes. In the actual network, inter-domain nodes are quite few so the cost for interaction between any two nodes is not high. While there are many inter-domain nodes, so if communication is made between them, then the communication cost will be 2n (n is the communication node number), which will cause growth of intra-domain communication flow and will instead reduce bandwidth utilization.

V. SYSTEM EXPERIMENT

Construct a small network by the above realized content delivery system and network topology in the laboratory to verify the system. The gateway of laboratory is similar to inter-domain cache node, for users is like the PC machine in the laboratory which realizes the algorithm proposed by the text in the system.

As shown in Fig.10, cache hit rate in the day peaks at almost 15% while it plummeted below 1% in the night. These hit rates generally are generated by the pop-up ads when it is hanging off-line and downloading. But the hit rate is around 8% during daytime work period which is not a high hit rate, for network traffic in the laboratory mainly is search traffic so there is no cache hit. While during the night both the search and download occupy much of the network traffic and therefore there is a significant rise of the cache hit.

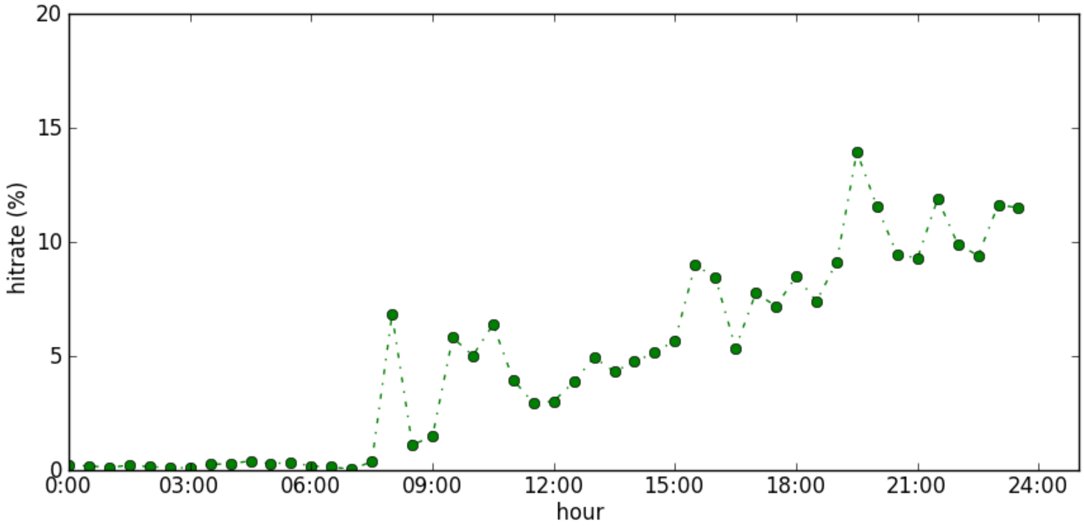

Fig.10 Hit ratio of system

Fig.10 shows that the trend of syn times is similar to that of hit rate. Syn times during weekends are much higher than that in other periods, and syn times in the nights of weekends are also higher than that in the nights from Monday to Friday, which proves that there is a huge change of data in cache system during the weekends, and which brings about the increase of syn times. Although there is a high hit rate in the night, the access frequency of cache misses becomes higher, which prompts the threshold of synchronous counting to be exceeded during short time and thus create the synchronization.

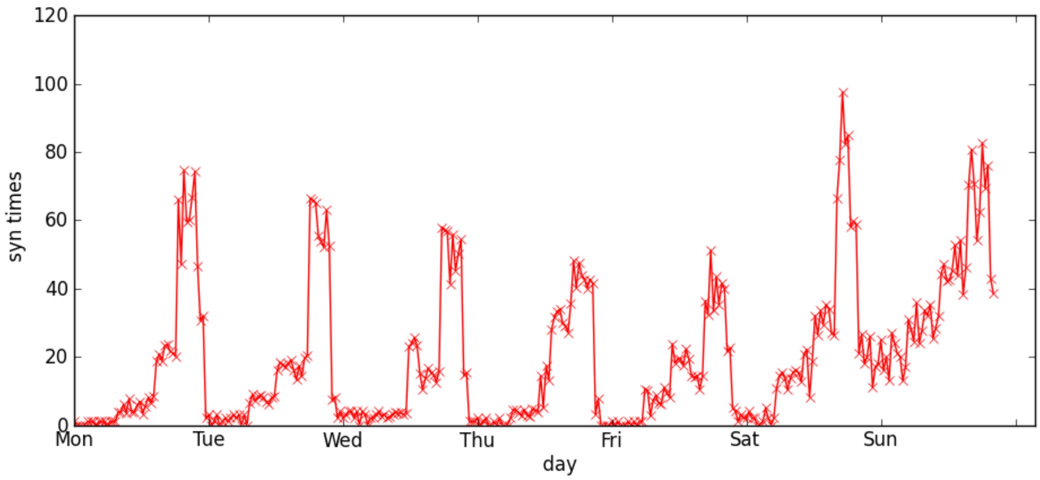

Fig. 11 Synchronization times in a week

The quantity of pushing data is also one key indicator of system performance. When there are many pushing data, it demonstrates that the effectiveness of cache has been enhanced. And these pushing data has shortened the path of users' requests and has decreased intra-domain repetition network flow and thus it will improve the network performance to some degree. As shown in Fig. 12, users' requests present Zipf distribution, so when the quantity of requests increases, the pushing data will increase accordingly.

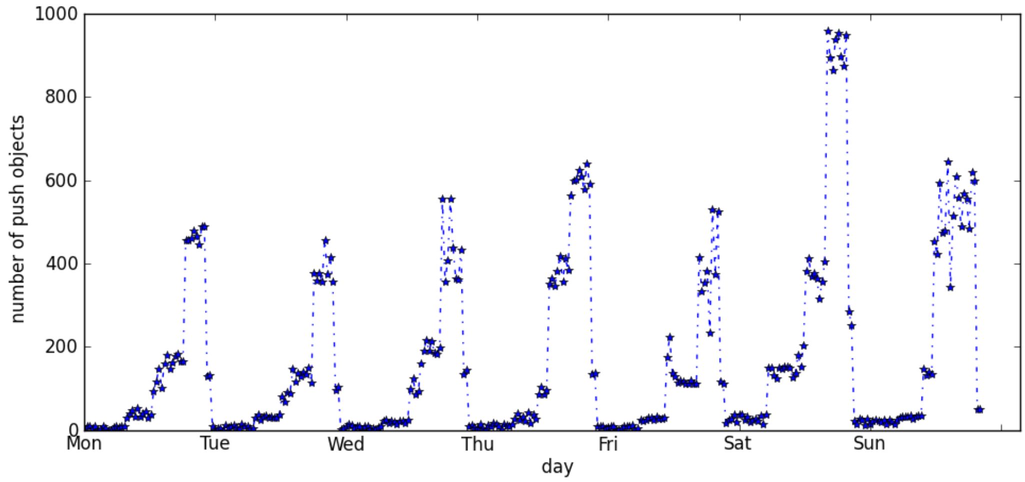

Fig.12 The number of object pushed in a week

One key function of the system is to make early warning for intra-domain security. Since there is no actual group cyber attacks in the laboratory, DDoS attack is simulated in some PC machines and which is compared with FTP downloading. As shown in Fig.13, nodes aggregation degree in (a) is much higher than that in (b). Although FTP also is many to one users' behavior, its aggregation degree is not high, for FTP users have other interactive behaviors such as instant messaging, SSH access, network file sharing and so on and all these have weakened many to one behavior mapping. While DDoS attack presents instinct group characteristics.

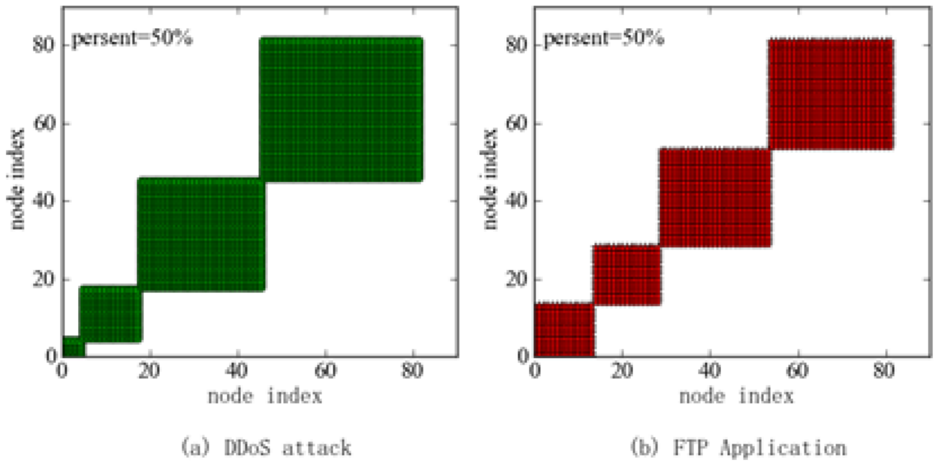

Fig. 13 Comparison of two events

VI. SUMMARY

This paper has designed and implemented the CDN prototype system integrating ISP. It presents the whole system architecture and the whole network data process of the system, and it introduces implementation details of each module, including users' data parsing and construction of index mechanism in the information collection module, and objects mapping and cache queue management procedures in the cache management module. The core component of the system is the cache management part. The experiment result shows that hit rate in the actual network testing can reach up to 15%, which is comparatively ideal but with less synchronization times and low cost of consumption. And the content delivery system integrating ISP has been implemented through the method and system design proposed in this paper.

REFERENCES

[1] Pallis G and Vakali A. Insight and Perspectives for Content Delivery Networks [J]. Communications of the ACM, 2006, 49(1): 101-106.
[2] Gadde S, Rabinovich M, Chase J. Reduce, reuse, recycle: an approach to building large Internet caches[C]. Proceedings of 6th Workshop on Hot Topics in Operating Systems. Cape Cod, Massachusetts, USA. 1997. 93–98.
[3] Gao L, Dahlin M, Nayate A, et al. Improving availability and performance with application-specific data replication[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(1):106-120.
[4] Davison B D. Web caching and content delivery resources[EB/OL]. [2013-10-24]. http://www.Web-caching.com
[5] Day M, Cisco, Cain B, et al. RFC3466 (A Model for Content Internetworking (CDI))[S/OL]. Internet Engineering Task Force RFC, 2003. [2013-10-24]. http://tools.ietf.org/html/rfc3466
[6] Jung Y, Krishnamurthy B, and Rabinovich M. Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites[C]. Proceedings of the 11th International World Wide Web Conference. Hawaii, USA, 2002. 293–304.
[7] Lampe C, Resnick P. Slash(dot) and Burn: Distributed Moderation in a Large Online Conversation Space[C]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Vienna, Austria, 2004. 543-550.
[8] Krishnamurthy B, Wills C and Zhang Y. On the Use and Performance of Content Distribution Networks[C]. Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement. San Francisco, USA, 2001.169–182.
[9] Lazar I and Terill W. Exploring Content Delivery Network[J]. IEEE IT Professional, 2001, 3(4):47-49.

AUTHORS

Chao Chen was graduated from Harbin Institute of Technology, he has got master degree in software engineering. Now he works as a network engineer in Petrochina Petrochemical Research Institute, proxy cluster management is part of his work.

Fei Gao was graduated from Fushun Petroleum Institute, he has got master degree in petroleum engineering. Now he is a director of Process Engineering Research Department, the department has an important responsibility for information construction.

Chuan Liang was graduated from Chinese Academy of Agriculture Sciences.He is an engineer works in Petrochina Petrochemical Research Institute now.He works on normal operations of network environment and unified constructional systems.

Chao Li was graduated from Heilongjiang University, he has got bachelor degree in petroleum engineering. Now he is a vice director of Process Engineering Research Department, the department has an important responsibility for information construction.