Key Technologies and Application Analysis of Internet Traffic Optimization Systems: Postprint

Authors: Zhuoqun Li

Date: 2023-10-08T00:00:00+00:00

Abstract

The rapid expansion of Internet users and business scale has propelled the swift growth of data traffic, imposing substantial pressures and challenges upon network construction and operations. Traffic optimization systems have emerged as the enabling vehicle for implementing intelligent conduits for data traffic and distributed technologies. This paper presents the system construction status, encompassing load optimization and traffic optimization systems, conducts a comparative analysis of key technologies for traffic optimization systems, load optimization schemes, and technical implementation solutions, and further optimizes the system's development and deployment.

Full Text

Preamble

Abstract: The rapid development of internet users and business scale has driven explosive growth in data traffic, placing significant pressure and challenges on network construction and operation. Traffic optimization systems have become essential carriers for implementing intelligent channels and distributed technologies for data flow. This paper introduces the construction status of such systems, compares and analyzes key technologies for traffic optimization, traffic optimization schemes, and technical implementation solutions, and proposes optimizations for system development and deployment.

Keywords: Traffic Optimization; Cache; CDN (Content Delivery Network); P2P (Peer-to-Peer); DNS (Domain Name System)

Classification: T33
Document Code: A
Article ID: 1671-0134(2018)01-050-02
DOI: 10.19483/j.cnki.11-4653/n.2018.01.015

Author: Li Zhuoqun

The First Decade of the 21st Century: Evolution of Internet Traffic

The first decade of the 21st century witnessed an accelerating wave of internet and technological revolution, continuously spawning new technological innovations and complex, ever-changing internet applications. These novel applications—including dazzling IPTV online games and IP-based peer-to-peer (P2P) applications—differ fundamentally from traditional internet services such as email, Web, and FTP. Characterized by high bandwidth consumption and complex behavioral patterns, these new applications—particularly those based on collaborative computing technologies, instant messaging, P2P file sharing, and emerging services like streaming media—have indirectly transformed internet application models. Statistical data reveals that P2P applications account for 70% to 80% of total measured internet traffic. According to US statistics, P2P traffic consumed 12% of total US bandwidth in 2006. In 2007, Ellacoya Networks, a provider of IP service control system solutions, analyzed one million North American broadband users and found that HTTP represented approximately 46% of network traffic and 37% of all internet traffic within the measurement scope. Research institution ipoque conducted a study on internet traffic distribution between August and September 2007, discovering that P2P applications occupied 95% of bandwidth. According to 2008 statistics from Chinese operators, P2P traffic accounted for 35%-60% of national total bandwidth during daytime hours and approximately 50%-90% during evening hours. More importantly, the nearly symmetric traffic patterns of P2P applications have also caused congestion across the internet. This situation has significantly impacted and promoted traffic engineering, network planning, and network management for network administrators and service providers.

1. Status of Operator Traffic Optimization System Construction

To improve user access speeds, reduce access latency, enhance quality of experience, lower inter-operator settlement costs, and accelerate the construction of CDN (Content Delivery Network) traffic optimization systems and caching capabilities, network channels and traffic must be optimized. A comparative study of traffic optimization system construction between Shaanxi Broadcasting and China Telecom yields the following conclusions: (1) Shaanxi Broadcasting's CDN network is in the development stage, whereas China Telecom's CDN network has reached maturity, with substantial network coverage scale and abundant resources that have facilitated 3G (third-generation mobile communications) and broadband development while ensuring support for IPTV and promoting other services and businesses; (2) Due to rapid business growth in streaming media and IPTV, China Telecom has also driven continuous CDN net-

work development and expansion. Currently, China possesses relatively few mobile network resources, and without business drivers, CDN network construction has progressed slowly, resulting in lower profitability; (3) Constrained by connectivity limitations, Shaanxi Broadcasting needs to establish a high-speed caching system, particularly network caching. China possesses abundant telecommunications network resources, making it unnecessary to utilize large amounts of external resources. Deploying P2PCache methods and optimizing P2P traffic within the network can help alleviate pressure on backbone network construction.

2. Comparative Analysis of Traffic System Optimization Technologies

2.1 Technical Characteristics of Caching Systems

Service caches include P2PCache and WebCache. Deploying and exporting service caches across multiple networks enables a single Web access point to reach multiple Web services by interconnecting P2P business information, access networks, and content sources. This architecture improves user download and access speeds, reduces settlement costs between networks, and provides users with better perceived service quality. As a local access implementation technology, WebCache is designed for frequently accessed Web content. At the edge of network service caches, it stores Web content to improve user access speeds and reduce access paths. P2P network list content is cached through P2PCache edge cache servers, allowing clients to receive feedback on nearby P2P content sources through these edge servers. Clients can download additional content from multiple locations based on proximity principles, thereby increasing download speeds and conserving network bandwidth. When deploying caching systems, selecting an appropriate redirection mechanism is critical. The redirection mechanism serves as the entry point to the caching system—all user requests must pass through this mechanism to enter the caching system, which then provides local content and services through service caches.

2.2 CDN System Technical Characteristics

Internet-based CDN consists of geographically distributed CDN nodes. By utilizing content delivery centers and Global Server Load Balancing (GSLB) for coordinated scheduling, the system achieves the goal of caching contracted content sources, enabling users to quickly locate recently accessed networks. Content distribution technology and content routing technology constitute crucial CDN technologies. CDN content paths are distributed to optimal CDN nodes through load balancing technology. Various methods can be employed to select optimal nodes, such as considering node load, service availability, or proximity. Content routing can utilize either redirect routing or DNS routing, with CDN service caches receiving user requests for resource access.

3. Application Scenarios and Implementation

3.1 Caching System Application Scenarios and Technical Implementation

Caching systems enable passive caching and accelerated content delivery, making them suitable for large HTTP objects (multiple formats of streaming media downloads), small HTTP objects (images, web pages, and other files), Baidu Cache, and P2P content (Thunderbolt, Vogue, QQLive, video). In different application scenarios, the following technical implementations can be employed for caching system redirection mechanisms based on deployment requirements and actual network conditions: (1) For Web request domain name access, intelligent DNS can be used to develop scheduling methods. DNS can be forwarded to the caching system using DNS forwarding methods, or redirected DNS can be utilized to obtain DNS messages and perform DNS redirection on the caching system; (2) For video links and URL requests for large HTTP objects, HTTP requests can be redirected to the caching system through DPI (Deep Packet Inspection) via intelligent DNS; (3) For Peerlist-based P2P application requests, the redirection DPI method is used to monitor and observe the cached P2P resource locations on service caches, redirecting query requests to the caching system. Through this technology, users can obtain domain user information, cached resources, and cache nodes.

3.2 CDN System Implementation Methods and Application Scenarios

CDN systems are particularly suitable for caching and downloading web pages and HTTP streaming media. Based on different application scenarios for content routing technologies, the following implementation methods can be employed for CDN resource scheduling: (1) As a content acceleration system, CDN can authorize CP (Content Provider) domain names for resolution (or utilize LocalDNS to obtain and analyze specific domain names), with domain name analysis and completion performed through CDN's GSLB; (2) For DNS routing, domain names are resolved through the IP address of service caches. For redirect routing, domain names are resolved through the IP address of GSLB, with endpoints requesting and accessing GSLB content. GSLB selects caches and returns the service cache's IP address using 302 redirection; (3) When users click on desired content, endpoints request content from service caches. If the content is not available, service caches retrieve it from the source and return the requested content to endpoints.

4. Deployment Recommendations and Development Approaches for Traffic Optimization Systems

To promote the overall development of network traffic optimization systems, content sources should first be introduced to improve CDN network content caching technology and promote IDC (Internet Data Center) construction. By caching as much content as possible, networks can retain user traffic. Addition-

ally, CDN intelligent systems should be gradually installed to expand content distribution network capacity, promote internet business development, improve user quality of experience, and conserve internet network bandwidth. During different stages of IDC construction, operators should deploy traffic optimization systems following these principles: In the early development stages of IDC construction with limited content sources, CDN nodes should be centrally built at provincial centers to introduce content sources through ICP (Internet Content Provider) agreements, supplementing IDC content sources. Simultaneously, centralized provincial center construction should supplement IDC and CDN resources, gradually transitioning from leased CDN and cache access to self-built CDN and caching infrastructure.

References

[1] Sun Ding. Survey Shows HTTP Network Traffic Exceeds P2P for First Time in Four Years [EB/OL]. http://www.techweb.com.cn/news/2007-06-20/211658.shtml. 2007/2017.

[2] Zhang Qiang. Expert Says P2P Traffic Accounts for Over Half of Domestic Total Bandwidth [EB/OL]. http://blog.163.com/login.do?err=403. 2007/2017.

[3] Driver Home. Latest 2008 P2P Traffic Monitoring and Management Solutions [EB/OL]. http://tech.hqew.com/fangan_{461316}. 2008/2017.

[4] Cai Xin. Mobile Internet Traffic Approaching Fixed Internet Levels [EB/OL]. http://www.doc88.com/p-304514267783.html. 2010/2017.

(Author's Affiliation: Xi'an University of Posts and Telecommunications)

Note: Figure translations are in progress. See original paper for figures.