Postprint: Practice and Reflection on the Construction of High-Availability Technology Support and Security Systems in the All-Media Context

Authors: He Haiyu

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

Abstract

In the process of continuously advancing the deep convergence and development of media, this study explores and constructs technical methodologies for an all-media communication system, with particular emphasis on conducting in-depth investigations into the high availability aspects of the all-media communication technology architecture. By adopting various Internet-based technologies to build both a high-availability technical support system and a technical security system for all-media communication, a comprehensive high-availability technical system for all-media communication is established.

Full Text

Practice and Reflection on the Construction of High-Availability Technical Support and Security Systems in the All-Media Context

Author: He Haiyu (Dazhong Newspaper Group/Dazhong Daily, Jinan, Shandong 250014)

Abstract: In the process of continuously advancing the deep integration of media development, this paper explores the technical methodologies for constructing an all-media communication system, with a focused investigation into the high-availability aspects of all-media communication technology architectures. By employing various Internet-based technologies, we have constructed a high-availability technical support system and security framework for all-media communication, forming a comprehensive high-availability technical system for all-media dissemination.

Keywords: All-media communication; High availability; Divide and conquer; Technical support system; Technical security system
CLC Number: TN891
Document Code: A
Article ID: 1671-0134(2021)07-018-03
DOI: 10.19483/j.cnki.11-4653/n.2021.07.003

China has become the world' s largest internet user base, characterized by universal internet adoption. With increasingly high volumes of concurrent network access, traditional media production technology systems can no longer support such massive access pressures. In this context, the availability requirements for technical systems within all-media communication architectures have become increasingly critical. Through in-depth practice in constructing all-media communication systems, the question of how to build a highly available technical architecture has gained paramount importance.

Dazhong Newspaper Group has continuously promoted deep media convergence development, constructing an all-media communication system. Driven by technological innovation, the group has gradually achieved technical integration, data connectivity, and shared resources across its newspaper, website, app, and WeChat platform production and distribution systems, eliminating information silos and creating an integrated all-media communication platform. This has facilitated the transformation from traditional media production technology systems to all-media communication technology architectures. Throughout this transformation, we have conducted in-depth exploration into the high-availability aspects of all-media communication technology systems, employing various Internet-based technologies to construct high-availability technical support and security systems, thereby forming a comprehensive high-availability technical system for all-media communication.

1. Practical Principles of High-Availability All-Media Communication Technology Systems

To meet the practical demands of high availability, high concurrency, and easy scalability, our high-availability all-media communication technical support system adopts a "divide and conquer" and "flexible expansion" approach during implementation to satisfy large-scale, diverse application requirements.

1.1 Stability and Reliability

Centered on user needs, the system can rapidly respond to user requests under all circumstances, with substantial concurrent processing capabilities, high throughput, and stable performance, ensuring uninterrupted service and achieving high availability.

1.2 Security and Reliability

The system employs multiple security mechanisms to implement defense-in-depth across physical security, system security, network security, application security, virtualization security, and data security, ensuring platform security.

1.3 Flexible Scalability

Using an elastic scaling framework, the platform' s components support flexible expansion functions. Based on growing business demands, the system can dynamically increase software and hardware application instances, achieving dynamic scaling of application instances in both horizontal and vertical directions.

1.4 Rapid Response

The system can conveniently and quickly respond to new business developments and high-concurrency traffic access, adapting as needed with rapid response capabilities.

2. High-Availability Technical Support Architecture for All-Media Communication

The construction of Dazhong Newspaper Group' s high-availability all-media communication technical support system employs a distributed cluster architecture, logically divided into five layers: the front-end content distribution layer, load balancing layer, Web application layer, middleware service layer, and data management/storage layer. Each layer utilizes virtualization technology, distributed computing, load balancing, application server clustering, message queuing, full-text indexing, distributed caching, distributed file systems, and database master-slave replication to achieve high availability, high concurrency, and easy scalability.

[Figure 1: see original paper] High-Availability Technical Support Architecture for All-Media Communication

The **front-end content distribution layer** primarily enables automatic distribution of local content to network nodes, improving user access experience. CDN (Content Distribution Network) technology can redirect user requests to the most efficient service nodes in real-time based on comprehensive factors such as network traffic, node connection status, load conditions, distance to users, and response time. This enables users to access network content more efficiently, solving problems of slow access caused by limited network bandwidth, large user volumes, and uneven site distribution, thereby improving response speed. By employing CDN technology at the front-end content distribution layer, access pressure is distributed across different network nodes, alleviating the pressure of massive user access on back-end network services. Through "divide and conquer" of access traffic at the front-end layer, high availability is achieved at this stage.

The **load balancing layer** primarily undertakes the task of balanced traffic distribution from load 均衡 servers. When encountering high-concurrency access volumes, it distributes traffic to different servers according to scheduling algorithms, decomposing high-concurrency access pressure. This logical layer primarily ensures high availability and scalability through the combined deployment of Nginx + Keepalived. Nginx is an open-source, lightweight, high-performance, highly stable, and highly concurrent Layer 7 load 均衡 server software. By configuring listening rules on Nginx, traffic can be distributed to different back-end servers for processing during high-concurrency access periods. Based on business development needs, the system' s service capabilities can be expanded or released by modifying the Nginx load 均衡 module configuration to add or remove servers at any time, flexibly adapting to various application service requirements. To improve Nginx service availability, Nginx is deployed in an active-standby configuration, with Keepalived heartbeat detection technology used to monitor the health status of Nginx servers in real-time. When a server failure occurs in the active-standby pair, Keepalived automatically shields the faulty server and distributes requests to normally operating Nginx servers, eliminating single points of failure and ensuring normal application system operation. Therefore, employing load 均衡 technology at the load balancing layer can further alleviate high-concurrency access pressure, improve application processing performance, increase throughput, enhance network processing capabilities, provide effective automatic fault transfer methods, and achieve service high availability, flexibility, and scalability. Through "divide and conquer" of access traffic at the network request response stage, this stage achieves high availability.

The **Web application layer** primarily responds to browser access requests, providing reliable Web services to users. When front-end browser access requests exceed the processing capacity of Web application servers, cluster technology and dynamic page staticization can improve Web server service capabilities and response speed. Cluster technology generally consists of a loosely coupled collection of computing nodes formed by two or more servers, providing users with a single customer view of network services or applications. These servers can communicate with each other and collaborate to provide powerful computing resources and data services to users. Therefore, in practical application environments, multiple application servers are formed into a cluster, with front-end requests distributed to different servers through load 均衡 technology to handle high-concurrency load pressure generated by simultaneous access from large numbers of users. Additionally, the Web application layer optimizes server resource allocation through separation of dynamic and static resources. When receiving static file requests, the files are returned directly without submitting the request to back-end application servers. For dynamic pages with particularly high access volumes but infrequent updates, they can be staticized—generating a cached static page—to avoid repeated queries on back-end servers. Therefore, applying clustering and dynamic-static separation technologies at the Web application layer improves processing and response capabilities, alleviates pressure

on back-end servers, and enhances system availability. Through "divide and conquer" of Web access traffic at the Web service layer, this layer achieves high availability.

The **middleware service layer** serves as a connecting link between the Web application layer and data storage management layer, accelerating Web application layer response speed, eliminating concurrent access peaks, and alleviating the pressure of front-end high-concurrency access traffic on the data storage layer. The middleware service layer primarily achieves high availability through distributed caching, message queuing, and full-text search technologies. Distributed caching technology provides rapid data reading, dynamic cache node expansion, automatic fault switching, and automatic data 均衡 partitioning, optimizing storage strategies based on content resource access frequency to improve access efficiency. This is an important technology for achieving distributed high availability. Therefore, a distributed caching layer is introduced between the Web application layer and data storage management layer. Using Redis services, frequently accessed hot data is cached. When business requests are initiated, queries are first performed in the caching layer, preventing massive requests from directly hitting the underlying database and thus greatly reducing database queries and pressure. This excellently solves the bottleneck between database servers and Web servers under high traffic conditions, achieving shorter response times and improving data reading speed and carrying capacity. Additionally, a manual data update mechanism is designed in the content publishing backend to meet second-level add/delete/modify requirements. Message queuing is an asynchronous inter-service communication method that can decouple business modules and improve access speed through asynchronous calls. It also queues high-concurrency access requests for processing, smoothing front-end request peaks and alleviating back-end access pressure, allowing the back-end to process at its own speed. Therefore, message queuing is also an important means of ensuring distributed high availability. Full-text search employs Elasticsearch, an open-source, highly extensible distributed full-text search engine that can store, retrieve, and process petabyte-level data in near real-time and supports cluster expansion. It is a popular enterprise-level search engine. When processing large-scale datasets, full-text search systems typically adopt cluster architecture, improving search capabilities through node expansion to meet larger data retrieval demands. To improve system concurrent search capabilities, the number of replicas within nodes can be expanded to meet massive concurrent search demands, achieving fault tolerance and high availability. Therefore, through "divide and conquer" of real-time data flows at the middleware service layer, this layer achieves high availability.

The **data storage management layer** is the cornerstone of the entire support architecture, providing structured and unstructured data management and storage functions for the entire system. The underlying data storage employs a combination of database management systems and cache databases. Frequently accessed hot business data with intensive read/write operations is stored in the Redis cache database of the middleware service layer to improve data access

efficiency, alleviate back-end database access pressure, and improve database availability. Structured data is stored using the MySQL database management system. To prevent database service failures or unexpected interruptions, database deployment typically adopts a one-master-one-slave or one-master-multiple-slave architecture. When the master database server fails, switching to slave servers for data access can be completed quickly to minimize downtime and ensure business continuity. Database master-slave synchronization mechanisms ensure data consistency. Another important method for improving data access availability is separating database reads and writes, with the master database handling write operations and slave databases handling read operations, thereby greatly reducing back-end database pressure. When database tables become numerous and data volumes grow large, database and table sharding methods are employed to reduce pressure on the database management system. For massive unstructured data storage, distributed file systems are used for data management and storage. Distributed file systems disperse large amounts of data across different nodes for redundant storage, greatly reducing data loss risk. This redundancy characteristic ensures that partial node failures do not affect overall normal operation, with other nodes recovering damaged data and avoiding single points of failure. Distributed file systems also connect dispersed computers through networks, continuously expanding servers horizontally to increase computing and storage capacity, forming a larger-capacity, more easily scalable storage system. Therefore, through "divide and conquer"of data storage at the data storage management layer, this layer achieves high availability.

3. High-Availability Technical Security System for All-Media Communication

The high-availability technical security system for all-media communication follows the "Information Security Technology—Baseline for Classified Protection of Cybersecurity"issued by the Standardization Administration of China. Through systematic security protection measures and disaster recovery backup systems, a multi-dimensional security protection mechanism is established to form a high-availability technical security system.

[Figure 2: see original paper] High-Availability Technical Security System for All-Media Communication

3.1 Network Domain Protection

Based on the importance of different business systems, various network domains are divided within the platform, including DMZ domains, classified protection domains, and office domains. Businesses with similar security requirements are centrally deployed in the same network domain to achieve network isolation. Simultaneously, defense-in-depth is implemented across multiple dimensions including physical security, system security, network security, application security, virtualization security, and data security to construct a network domain

protection system.

3.2 Boundary Protection

Between network domains, Web application firewalls and intrusion detection systems are deployed according to cybersecurity requirements to provide security isolation. This prevents various illegal intrusion behaviors including SQL injection, XSS cross-site scripting, path traversal, unauthorized access to core files, backdoor isolation protection, command injection, illegal HTTP protocol requests, and Web server vulnerability attacks. The system ensures that data transmission between network domains complies with appropriate security access control policies, safeguarding the security of each network domain.

3.3 Operations Audit

By deploying bastion hosts, direct access from terminal devices to hosts and network equipment is isolated, blocking illegal access and malicious attacks. The system intercepts illegal commands, filters all illegal access behaviors to target devices, and audits and monitors both accidental and intentional internal personnel misoperations for post-incident accountability.

3.4 Database Audit

The database audit system is a security system that performs fine-grained analysis and auditing of user database access behaviors. It provides real-time monitoring, violation response, and historical behavior traceability functions. The system can record database access behaviors in detail, identify unauthorized operations, and provide traceability, offering decision-making support for database security management and performance optimization.

3.5 Disaster Recovery and Backup System

According to Level 3 classified protection standards and actual business requirements, important core system data and business data are backed up both locally and remotely. By establishing an off-site disaster recovery system with remote data backup, automatic service switching, and rapid system recovery functions, uninterrupted service and system stability are ensured.

By employing distributed Internet technologies such as virtualization, load balancing, Web clustering, message queuing, caching, and databases, we have established a distributed, high-availability, high-concurrency, and easily scalable technical support architecture for all-media communication. Through network security measures including network domain protection, boundary protection, log auditing, database auditing, bastion hosts, and disaster recovery backup, we have constructed a multi-dimensional high-availability technical security system. By building both high-availability technical support and security systems within the all-media communication architecture, we have effectively addressed

issues including traffic forwarding, cluster scaling, resource scheduling, security protection, and emergency response under high-concurrency traffic conditions.

References

[1] Tang Dailu. "Qilu Smart Media Cloud" Supporting Deep Media Convergence Development [J]. China Media Technology, 2021(3): 14-17, 39.
[2] Zhang Yanhua. Research on CDN Network Distribution Algorithms [J]. Journal of Zhejiang University of Media and Communications, 2006(3): 39-41.
[3] Geng Xiaoli, Zhang Mang, Yin Yonghong. Research on High-Concurrency, High-Availability Distributed E-Commerce Platform Architecture [J]. Computer Technology and Development, 2021(31): 111-115.
[4] Liu Zengcai, Zhang Fuzheng, Liu Mingzhu, Sheng Zhiguo. Research and Application of Digital Party-Building Platform Based on High-Availability Microservices Architecture [J]. Information Technology and Informatization, 2019(11): 184-188.
[5] Design of Radio Station Converged Media Platform Based on Private Cloud Architecture [J]. Electroacoustic Technology, 2017(4/5): 81-86.

Author Biography: He Haiyu (1977-), male, from Weihui, Henan, Senior Engineer, Dazhong Newspaper Group (Dazhong Daily).

(Editor: Chen Xuguan)

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