

# *Research and Implementation of Server Load Balancing Strategy in Service System*

Yunpeng Zhang<sup>a</sup>, Liwei Liu<sup>b,\*</sup>, Xu Jia, Di Xian

*National Satellite Meteorological Centre, Beijing, China*

<sup>a</sup> email: zhang.yun.peng@126.com, <sup>b</sup> email: liulw@cma.gov.cn

*\*corresponding author*

**Keywords:** FTP server; Load balancing; Stress test; Health check; Triangular transmission.

**Abstract:** According to the current situation of the FTP server usage of the Intranet and extranet satellite data archiving systems of the National Satellite Meteorological Centre, this paper carefully analyzes the performance bottleneck of the server. The disaster recovery function of the overall system to cope with unexpected problems was increased and main service object information was counted. From the overall architecture level, the existing FTP service is upgraded and transformed to achieve efficient and secure meteorological satellite data storage and data download FTP server clustering.

---

**Original article, Published date: 2018-05-23**

**DOI: 10.23977/jeis.2018.31011**

**ISSN 2371-9532**

**<https://www.clausiuspress.com/journal/JEIS.html>**

---

## **1. Introduction**

With the continuous increase of satellites operated by the National Satellite Meteorological Centre on-orbit service, the data service pressure on users of meteorological satellite data has followed. Currently, the File Transfer Protocol (FTP) file transfer service for intranet users is provided by the IP address 10.24.16.X, and the FTP service for external users is provided by the demilitarized zone 10.0.65.X. Due to the large amount of data downloaded from weather satellites and the large number of data orders downloaded by the general public, simply relying on a single FTP server can no longer meet the increasing demand for user download services, server processing speed and I/O (input/output input). The output port capability becomes a bottleneck that affects the quality of network services. It is urgently needed for the server to have the ability to provide a large number of concurrent access services. As the service of intranets and extranets of the National Satellite Meteorological Centre, there are certain requirements for performance while maintaining relatively high reliable online time. Under the requirement of 99.99% of the stable reliability of the

operation, the breakdown time is converted to less than 4 minutes per month and cannot exceed one hour per year. It is almost impossible to achieve this standard by the manual operation or cold backup method currently adopted. At the same time, due to the loopholes in the operating system itself, related security issues will also arise, and the urgent need for efficient, stable, and secure FTP data service functions.

## 2. FTP Service Capabilities

When the system is facing a single point of failure and server performance bottleneck, it usually adopts the following two methods to solve: to improve performance by replacing the existing server with a more powerful server, and to build a server cluster by adding servers. However, using both methods alone has some limitations.

The first solution is relatively expensive and does not have good scalability. It requires service interruption during maintenance and upgrades. At the same time, in terms of expenditures, the cost of replacing the server not only includes the cost of the new server, but also the original server can no longer play a role in its fully functional operating state, resulting in a waste of costs. The second solution is an acceptable relatively inexpensive solution. By adding new servers to improve the processing capabilities of the site and working with the original server, the service will not be interrupted during the upgrade. The disadvantage is that each server has a unique IP (Internet Protocol Internet Protocol) address. Users need to remember multiple IP addresses to better access the site. This also causes traffic to not be reasonably averaged among multiple servers.

Therefore, the use of load balancing policies for effective traffic management of FTP servers has become a must. Currently, there are two implementations of load balancing technology, namely software load balancing and hardware load balancing. The difference between the two implementations is shown in Table 1.

Table 1. The difference between load balancing devices

	Software load balancing	Hardware load balancing
Advantage	Simple configuration, Flexible, Low cost.	Independent of the operating system; overall performance is much higher than software; diversification strategies; intelligent traffic management; functions far better than software, some products can be customized; some products have statistical functions; various forms, with independent load balancing The device is integrated in various forms such as a switching device or integrated into a personal computer.
Shortcomings	Software running consumes indefinite resources of the system. The more powerful the module is, the more it consumes. When the connection request is particularly large, the software itself becomes a bottleneck that affects the service quality; the software scalability is not very good. Due to the loopholes in the operating system itself, it often causes security problems.	Higher cost

Aiming at the problem of performance, scalability and reliability existing in the current system, a hardware load balancing device is referred to as a solution. Therefore, using a multi-server architecture, the use of hardware load balancing technology to manage FTP server traffic distribution is the most reasonable solution. Because of the special nature of the FTP service,

adopting four layers of traffic to share multiple algorithms can reduce server load. Multiple servers and load-balanced hardware management can perfectly solve the server's multiple requests to the same address initiated by the NAT (Network Address Translation) class and are not limited by the system itself. The hardware load balancing device has become a necessary means to guarantee the Service Level Agreement (SLA). First, the system's vulnerabilities will not affect the hardware itself; secondly, in the new server deployment, the load balance of dual HA (High Available cluster) hot standby and user connection information backup mechanism, in a single load Equalization itself can still guarantee the reliable operation of the entire system due to hardware and software failures.

### **3. Load Balancing Technology**

#### **3.1. Load Balancing**

The load balancing technology is to virtualize the virtual server IP (VIP virtual Internet Protocol virtual address) to virtualize the application resources of multiple real servers at the back end into a high-performance application server. Through the load balancing algorithm, a large number of application requests from the client are requested. The server assigned to the back end processes it. The load balancing device continuously checks the application state on the server and automatically isolates the invalid application server, thereby realizing a cheap, effective, simple, highly scalable, and highly reliable application solution.

#### **3.2. FTP Load Balancing Health Check Mechanism**

The load balancing system uses a comprehensive health check mechanism, including Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP), and hypertext transport protocol (HTTP). Make accurate judgments and ensure that the services provided are correct. The ICMP health check mechanism is preferred for FTP load balancing.

#### **3.3. Load Balancing Algorithm**

Currently, the basic network load balancing algorithms include the weighted polling algorithm (polling algorithm with a weight of 1:1), the shortest response time algorithm, and the weighted minimum connection number algorithm (the minimum connection number algorithm when the weight value is 1:1). And other algorithms, managers can adjust the weight between servers according to actual needs.

#### **3.4. FTP Load Balancing Mode**

The load balancing device pair and the packet processing mode are divided into two basic operation modes: a simple four-layer processing based on the port number for simple application request forwarding, and a content-based request forwarding and processing seven-layer processing. FTP load balancing uses four layers of processing.

### **4. Hardware Load Balancing Test Results Analysis**

#### **4.1. Test Equipment**

Test equipment is shown in Table 2.

Table 2. The difference between load balancing devices

Equipment to be tested	Load balancing 1	Load balancing 2	other devices
Product name	AppDirector	AppDirector	Pressure tester
Model	OnDemand Switch 1	OnDemand Switch 1	

#### 4.2. Network Environment and Tools

The network environment used during the test is an internal non-exclusive network with a bandwidth of 1000 Mbps. System test tools for the Windows system FTP client and FTP download software.

#### 4.3. Network Environment and Tools

Through this test, it can be found that the internal download service of the center needs to ensure that the normal user access and improve the transaction processing capability of the background server are the basic requirements for building high-quality services. In the actual business system construction, large files (>5MB) use FTP to provide external download services. Through the use of load balancing triangle transmission, data transmission delay can be effectively reduced, and user download speed can be improved; for files smaller than 5MB, it can be used. Web publishing, combined with load balancing, reverse proxy caching, and connection multiplexing, can greatly improve the background server's transaction processing capabilities and user request response speed, and ensure the user's business access.

Health check mode: The session maintenance mode of the load balancing device is mainly the source address session maintenance; the deployment mode is the triangle transmission mode; the server smoothly enters and exits; the service access is maximized when the service is flexibly expanded; the load balancing cluster switch ensures security and stability. The main health check method is the ICMP health check method. Check whether the network is connected. Cut off the network (unplug the network cable) to check whether the load balancing device can correctly reflect the server status and cut services to other servers.

Comparing the difference of packet loss between the load balancing devices in the stress test environment; comparing the throughput performance of the load balancing devices; testing the performance of different load balancing algorithms under different pressure environments; using the polling algorithm, due to various Internet The link throughput may be different, so different links can be assigned different weights. Based on a 1:1 ratio, data traffic is polled to each link.

#### 4.4. Algorithm Test

In this test, the RR method was used to provide load balancing services to the FTP server. However, when using IE to access FTP, there was no round-robin effect. By using the FileZilla browser and the FTP command provided by windows, it was very good. Show round-robin effect; after capturing the packet test, this phenomenon is associated with the browser accessing FTP by default using the anonymous user for login authentication, as shown in Figure 1.

1	16:30:55.668565	192.168.110.253	192.168.110.248	FTP	Response: 421 No Transfer Timeout (300 seconds): closing control connection.
2	16:31:00.262753	192.168.110.253	192.168.110.248	FTP	Response: 220 10.24.16.3 FTP server ready
3	16:31:00.262820	192.168.110.248	192.168.110.253	FTP	Request: USER anonymous
4	16:31:00.285740	192.168.110.253	192.168.110.248	FTP	Response: 331 Password required for anonymous.
5	16:31:00.285821	192.168.110.248	192.168.110.253	FTP	Request: PASS IEuser@
6	16:31:00.327834	192.168.110.253	192.168.110.248	FTP	Response: 530 Login incorrect.
7	16:31:00.455136	192.168.110.253	192.168.110.248	FTP	Response: 220 10.24.16.4 FTP server ready
8	16:31:00.455211	192.168.110.248	192.168.110.253	FTP	Request: USER anonymous
9	16:31:00.521712	192.168.110.253	192.168.110.248	FTP	Response: 331 Password required for anonymous.
10	16:31:00.521788	192.168.110.248	192.168.110.253	FTP	Request: PASS IEuser@
11	16:31:00.733822	192.168.110.253	192.168.110.248	FTP	Response: 530 Login incorrect.
12	16:31:00.759924	192.168.110.253	192.168.110.248	FTP	Response: 220 10.24.16.3 FTP server ready
13	16:31:00.759979	192.168.110.248	192.168.110.253	FTP	Request: USER lbtest
14	16:31:00.769554	192.168.110.253	192.168.110.248	FTP	Response: 331 Password required for lbtest.
15	16:31:00.769611	192.168.110.248	192.168.110.253	FTP	Request: PASS lbtest
16	16:31:01.155911	192.168.110.253	192.168.110.248	FTP	Response: 230 User lbtest logged in.
17	16:31:01.156155	192.168.110.248	192.168.110.253	FTP	Request: opts utf8 on
18	16:31:01.156900	192.168.110.253	192.168.110.248	FTP	Response: 501 OPTS: utf8 not understood
19	16:31:01.156950	192.168.110.248	192.168.110.253	FTP	Request: syst
20	16:31:01.157524	192.168.110.253	192.168.110.248	FTP	Response: 215 UNIX Type: L8
21	16:31:01.157599	192.168.110.248	192.168.110.253	FTP	Request: site help
22	16:31:01.158150	192.168.110.253	192.168.110.248	FTP	Response: 214-The following SITE commands are recognized (* =>'s unimplemented).
23	16:31:01.275953	192.168.110.253	192.168.110.248	FTP	Response: HELP
24	16:31:01.276061	192.168.110.248	192.168.110.253	FTP	Request: PWD
25	16:31:01.281340	192.168.110.253	192.168.110.248	FTP	Response: 257 "/ARSFTPREAL/LBTEST" is current directory.
26	16:31:01.325302	192.168.110.253	192.168.110.248	FTP	Response: 220 10.24.16.4 FTP server ready
27	16:31:01.325387	192.168.110.248	192.168.110.253	FTP	Request: USER lbtest
28	16:31:01.391390	192.168.110.253	192.168.110.248	FTP	Response: 331 Password required for lbtest.
29	16:31:01.391453	192.168.110.248	192.168.110.253	FTP	Request: PASS lbtest
30	16:31:02.369489	192.168.110.253	192.168.110.248	FTP	Response: 230 User lbtest logged in.

Figure 1. Packet capture test

By default, IE uses anonymous for secondary login authentication. If the login fails, lbtest user is used for two login authentication to log on to the FTP server; that is, if you use IE to access (or refresh) the FTP server once, it is established with the load balancer. Four TCP connections; while the load balancing configuration for the FTP server counts and loads are based on the number of FTP controlled 21-port TCP connections; you can see through packet capture that load balancing actually performs a normal load on IE access.

Test mode: Set up 150 concurrent users (Note: The current number of mysql maximum users is set to 150, the maximum number of users to test the number of concurrent users), the duration of 2 minutes, the test file is: 8M, 200M, 500M HDF The format of weather satellite data and the minimum number of user algorithms are shown in Table 3. The server throughput and performance results are shown in Table 4.

Table 3. The difference between load balancing devices

File size	Transaction average response time	Transactions/sec	Load balancing CPU average	Server CPU average	Average server throughput	Running time	Failure
8M	80s	1.3	2%	6.7%	12M/s	02:54	0
200M	593s	0.2	3%	7.0%	12M/s	11:37	0
500M	1407s	0.1	3%	7.1%	12M/s	24:04	0

Table 4. Server throughput and performance results

File size	Throughput (average)		Server CPU (average)	
	8M	200M	8M	200M
10.24.16.27/28	11.6M/s	11.4M/s	8.1%	9.0%
Load balancing	12M/s	12M/s	6.7%	7.0%

According to the analysis and test data, after using the load balancer, the server throughput of the same size test file is slightly higher than that of using separate server throughput, but the throughput size does not change much; the CPU pressure is lower than that of using a separate server by approximately 15% to 23%. . According to the test, it can be concluded that the server throughput is basically independent of the test file size. In addition, due to the use of triangular transmission mode, the load balancer does not have CPU processor and network bandwidth pressure and will not become a bottleneck.

## 5. Conclusions

The existing single server environment was replaced by four new FTP server clusters and two load balancers. The weather satellite data product of the re-established National Satellite Meteorological Centre service network and the FTP service environment for downloading external data services were adopted. The above hardware deployment and corresponding setup tests we can draw the following conclusions.

Hardware load balancing is more in line with current server performance requirements. Different deployment methods have different roles, but the triangular transmission mode is more in line with current business requirements for high-traffic and high-bandwidth configurations. The appearance of the health check method not only ensures that the service will never be downed, but also can clearly locate the problem through the health check when there is a problem with the service. Convenient supervision. According to testing, the throughput of stand-alone server compression and use of load balancing equipment is not much improved, so when the throughput is much smaller than the network bandwidth, the load balancing will not greatly improve the throughput. According to the test, the CPU pressure of a stand-alone server is significantly higher than that of a load balancing device.

In summary, the load balancing device improves the performance of the original FTP service by adding new devices and introducing load balancing. It effectively solves the problem of insufficient single-server processing performance, insufficient scalability, and low reliability. Network congestion problem, achieving geographical independence; providing users with better access quality; improving server response speed; improving the effective utilization rate of servers and other resources; avoiding single points of failure in key parts of the network; improving server bandwidth and system Overall throughput; At the same time load balancing through a variety of means to correctly determine the health of the server, identify different sessions, support session retention function; and have statistical reporting capabilities, to understand and analyze the management of the site operations, access, user distribution, etc. stand by. The load balancing device is particularly suitable for places with poor network environment. Due to the limitation of the network environment, load balancing will play a significant role in throughput, failures, and other aspects.

## References

- [1] Qian J M, Sun A L, Xu Zhe, et al. Fengyun Series Meteorological Satellite Data Archiving and Service System[J]. *Journal of Applied Meteorological Science*, 2012, 23(3):369-376.
- [2] Akbari H, Berenbrink P, Sauerwald T. A simple approach for adapting continuous load balancing processes to discrete settings[J]. *Distributed Computing*, 2016, 29(2):143-161.
- [3] Wang S, Zhang J, Huang T, et al. Flow distribution-aware load balancing for the datacenter[J]. *Computer Communications*, 2017, 106:136-146.
- [4] Pham C, Tran N H, Do C T, et al. Joint Consolidation and Service-Aware Load Balancing for Datacenters[J]. *IEEE Communications Letters*, 2016, 20(2):292-295.
- [5] Luo J, Rao L, Liu X. Spatio-Temporal Load Balancing for Energy Cost Optimization in Distributed Internet Data Centers[J]. *IEEE Transactions on Cloud Computing*, 2017, 3(3):387-397.