An Enhanced Load Balancing Methodology with Bio-Metric Feature

Dr.K.Kungumaraj
Associate Professor, School of Applied Science, Sapthagiri NPS University, Bengaluru.
Email : kungumaraj@snpsu.edu.in

ABSTRACT

In load balancing scheme security is another important issue in the aspect of a distributed system. Thus, SSL load balancer is involved in encryption/decryption of data using HTTPS which utilizes Secure Socket Layer (SSL) protocol to provide security on the system. Each request is secured by providing the NTRU file cryptosystem with a Biometric key generation scheme is discussed in this chapter. The next process is assigning the priority to each request by the Intuitionistic fuzzy Inference system to overcome the uncertainty in load balancing.

Keywords: Load balancing, Distribute system, Bio-metric security, Secure Socket Layer.

Date of Submission : January 02, 2025 Date of Acceptance : February 25, 2025

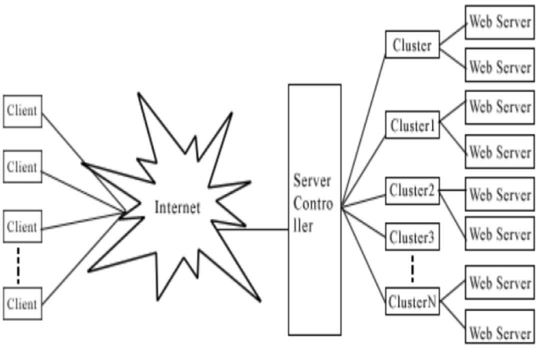
I. INTRODUCTION

Generally, web clusters are termed as a web site that uses more than two server machines which are integrated into a sample place to handle the request of clients. Even though a huge cluster comprises a greater number of web servers with a backend server, it only uses a single name to provide a single interface for the web users. Distribution of request and load balancing is the key terminologies to understand the web server clusters, but prevailing request distribution algorithms are in common static natures, and they don’t have any optimization policies and do not consider the dynamic usage of CPU utilization and memory utilization-based parameters.

Using the different multimedia types of documents, the user request will bring diverse load on web servers. So, the existing stabilized algorithms are not suited for the building of the heterogeneous web server clusters. This research work considers the heterogeneous features of h/w (hardware) and s/w (software) of each node and considers content-based load balancing policies. With growing web traffics at an exponential rate, there is high pressure on web servers based on scalability, performance, and availability of their services.

When the request number rises from a specific site the response time from that website also raises, it’s significant to correctly grip the slow response time delinquent else clients will get unsatisfied and will either trash linking or will not visit that web site again. The most significant anxiety of this plan is to handle inward load dynamically among all occupied treatment of load in a dynamic manner is an actual big contest but the grade of contest increases when working with cluster technology shown in Figure-I.

Figure-I System architecture of heterogeneous web server clusters



II. METHODOLOGIES IN LOAD BALANCING

The primary aim of the load balancer algorithm is to handle both static and dynamic data requests. Server loads increase when more number of client requests generate. The load balancer's responsibility is to predict the change in load and balance multiple resources. Many load balancing methodologies consider the load on virtual servers rather than concentrating on the volume of the prescribed space. Because volume is proportionate with the rate of request arrival. Due to this disregard load is allocated to nodes that have high loads rather than lightly loaded virtual servers. This situation suggests a forecasting scheme to adjust the load for example possible boundaries on the further loads of virtual servers.

The distributed environment is highly available to distribute loads to server clusters with Meta-heuristic schemes. With social behavior, it is possible to estimate the server availability. By this approach balancing with multiple resources is achieved. Resource balancing can be done by implementing various techniques. Each technique involves network construction with sub servers and their allocation is a major part. Some existing approaches perform load balancing but failed in some aspects like failing to redirect requests to provide a proper response despite maintaining efficient navigation of the server. To overcome issues many techniques have come into existence.

Many researchers surveyed on various algorithms on load balancing. One common and foremost technique is the enhanced version of the Domain Name Service. DNS resolves duplication of IP addresses for load distribution. Round Robin DNS is an example of this scheme. The only duplication of IP addresses can be checked but not capable of determining the availability of the server. Some server failures cannot be identified. Failed servers keep on receiving requests resulting in HTTP failures to end-users. This results in several side effects like an increase in loads on the DNS server and network accordingly. Another issue with DNS based schemes is they fail to make decisions to predict load and network topology. The client's requests will be directed to faraway servers rather than reaching to the congested closer server. For achieving better load-balancing in a server cluster many websites use a connection routing mechanism that distributes connections to different server nodes. This maintains a load view on each server and forwards connections to the appropriate server. DNS and connection routing methods can be combined to get the best results. But this is not a solution for finding the nearest server. All connections to the server cluster need to pass through the connection router, if any single point failure happens, lead to failure in handling high loads. Several load balancing approaches are introduced to solve many issues while balancing. Popular load balancing methods are Round Robin with three dispatching policies.

- Round Robin
- Weighted Round-robin with response time as its weight
- Minimum connections with limits
- Low Bandwidth Technique
- Custom Load Method
- Minimum Response time Method

Round - Robin is a simple algorithm that distributes client requests over a group of servers. Requests are allocated to the server on a turn basis. The algorithm informs the load-balancer to go to the first position for issuing requests to servers.
Weighted Round-Robin is an upgraded algorithm for the round-robin method where response time for each server is determined constantly to know which server takes the next request.
Minimum Connections with Limits maintains a record of all the available servers with how many connections that each server maintains. The server with the lowest number of connections gets the next request. A round-robin is an effective approach for distributing workload amongst servers with equal processing capacities. When servers are vary in their processing capacity, use response time or count of active connections as selection criteria.
Low Bandwidth Technique selects backend servers depending on the consumption of the bandwidth of the server. A server with the lowest bandwidth is selected for service. Similarly, the least packet method also works. The virtual server uses bandwidth value and the total value of the bytes. In this method, the server that transmits fewer packets is selected by the load balancer.
Custom Load Method chooses backend servers based on loads of requests. To calculate server load usage of CPU, servers' response time, memory is considered to estimate the server load. Many software firms and enterprises use this algorithm frequently and efficiently to make use of resource utilization. Custom load balancing is applicable only when the network traffic is known and when the traffic is stable. When traffic is uneven and frequent changes are not suitable to use custom load balancing.
The minimum Response time Method allocates requests to the backend server which has the least number of ready connections and minimum response time. This algorithm ensures faster response time for end-users.

Other than these there are several other load balancing algorithms. There are few limitations in the existing approaches as

- Users waiting time is high to retrieve data
- Latency issues
- Low Throughput
- Tedious process

A Load balancing method should be developed in a way to overcome all the issues. The fundamental features that a load balancer should maintain is

- Adaptability
- Availability
- Reliability

Adaptability or scalability is the system’s ability to handle all the loads to get acceptable performance. The other two features' availability and reliability are closely relevant to each other.

Availability is the system's presence and its service available to clients. When multiple servers are in use, if any server goes down the client’s request must be processed by other servers. This is measured as availability. Though failure occurs it can process the work without interruption is reliability. But in some conditions, if database software fails on one server, that server might be reliable, but the combination of software and server may not be available. In this condition, a single machine cannot meet the necessary adaptability, availability, and reliability. A new system of research can be against these issues.

III. PROBLEM STATEMENT

The SSL load balancer behaves like a server-side SSL reach point for interconnection with clients, it performs both cipher and deciphering of requests and responses. The process involved in the security scheme varies relying on the load balancer and the server.

- If the server and load balance are in the same firewall then the SSL load balancer will involve in the decryption of request and pass it to the server as plain information. At the equivalent time, the response sent by the server is encrypted and forwarded to the client.
- In case if the network among load balancer and server is not secured then the SSL load balancer is structured such that after the request from the user is received by SSL, it decrypts the needed information and re-encrypt the request before passing it to the main server. The process is reversed for the process of responding from the server to the client.

IV. NEED FOR BIOMETRICS

Biometric methods are one of the effective ways of identifying an individual. It is based on the physiological behavioral feature. So, it is applied for authentication and identification methods. It provides distinctive information and will be unique for each individual. This type of authentication is needed in sharing sensitive information like e-banking transactions. [60] With the use of biometric technology high individual identification accuracy is achieved. There are very few chances of getting damaged or cannot be changed suddenly. In the proposed methodology ear feature is chosen for key extraction.

The usage of biometrics has many advantages other than security which include:

- Difficult to forge or crack unlike passwords
- Convenient and easy o use
- Non-conveyable
- A small change in users lifestyle
- Minimal storage templates

V. RESEARCH METHODOLOGY

NTRU is the first asymmetric key cryptographic system that is not dependent on the discrete logarithmic or factorization problem. It is also denoted as R=Z[χ]/(χ ^M-1). Alternatively, NTRU is a lattice-based which is a substitute to ECC and RCA and it solves the problem using the shortest vector in a lattice.

(a) Key Generation

- Arbitrarily select a polynomial $e \in LT_e$ such that e is invertible in modulo r and modulo s.
- Calculate $e_r \equiv e^{-1}$ modulo r & $e_s \equiv e^{-1}$ modulo s
- Arbitrarily select a polynomial $h \in LT_h$.
- Calculate $k \equiv h * e_s$ (mod s).
- Distribute the public key (M, k) and the set of parameters r, s, LT_e , LT_h, LT_t, and LT_v.
- Preserve the private key (e, e_r).

(b) Encryption

- Signify the message as a polynomial $v \in LT_v$.
- Arbitrarily select a polynomial $t \in LT_t$.
- Encrypt v with the public key (M, k) using the rule $\hat{E} \equiv r * t * k + v$ (mod s).
  - // “v” is a message

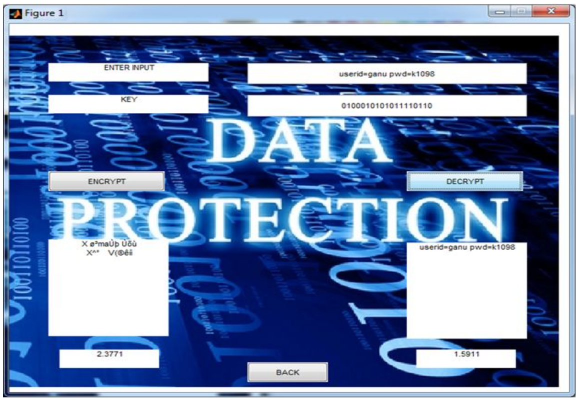
(c)Decryption

- The receiver calculates $b \equiv e * g$ (mod s).
- Through a centering process, change b to a polynomial with coefficients in the interval $[-\frac{s}{2}, \frac{s}{2}]$.
- Calculate $v \equiv e_r * b$ (mod r).

This research is to generate a private key biometric system is used to derive the private key involved in encryption and decryption. The ear hole and the outer lobe edge, interior curves of the helix, antihelix are used to determine the distance by using Euclidean distance formulation. Segmentation and filtering of the ear image are done using adaptive thresholding and masking techniques respectively.

V. RESULTS AND DISCUSSION

A ntru crypto screen for encrypting username and password and its decryption is depicted in Figure – II.



ENTER INPUT: userid=ganu pwd=k1098

KEY: 0100010101011110110

X ø³maÙþ Ùõù
X^* V(®êï

userid=ganu pwd=k1098

2.3771

1.5911

Figure – II: Ntru key system for sensitive data

This research work transforms the ear features into binary format in the representation of strings of RNA. It is performed by using the RNA coding approach. The coding of RNA string is denoted in Table-I.

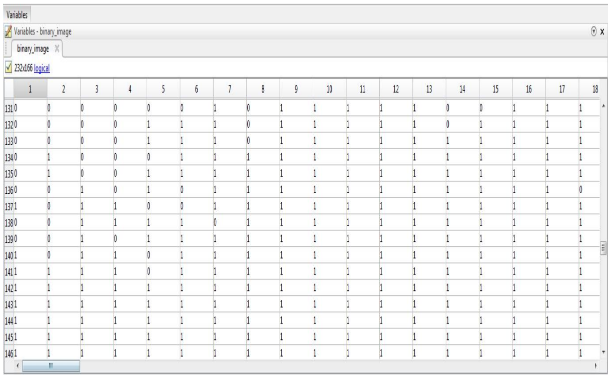
The concept of the RNA conversion is if the ear binary string is ‘00’ then it is converted to AA or if it is ‘11’ then it is transformed to ‘GG’, for the binary string ‘01’ then it is converted to ‘U’ and the binary-valued string ‘10’ is converted to ‘C’. This process continues to convert all the binary information of ear features extracted into an RNA sequence. In case if the length of the ear binary data is not even, the last two rows help us to convert correctly. By using this internal conversion it is highly impossible to identify the key.

Table - I: Binary data – RNA Coding Technique

Binary Data	RNA Data
00	**AA**
01	**U**
10	**C**
11	**GG**
0	**A**
1	**G**

The binary format of the image is as in the below Figure – III.

Figure – III: Binary values of ear image



232x166 logical

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
131	0	0	0	0	0	0	1	0	1	1	1	1	1	0	0	1	1	1
132	0	0	0	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
133	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1
134	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
135	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
136	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0
137	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
138	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
139	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
140	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
141	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
142	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
143	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
144	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
145	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
146	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

The sample binary string representation is supposed to be in this format :

0010100100100100100100100100100100100010110101
0010100100100101001001010111011010000001001010
1001101010101001001001001010010111111100101110
1011001001001001010010010100010010000001011110
100100100110110.

The value of ear binary string is converted to code as specified in the proposed RNA coding technology, as follows:

AACCUAACUAACTAACUAACUAACGGUUACCU
AACUUACCGGCGGUAAAAUAACCAAGGUUUU
CUAACUAACCUUGGGGCUUGGUCUAACUAAC
CUAACCAAUAAAAUUGGGGUAACUAUCGGA

This work chooses an RNA secret key with a key length of 512 bit randomly from the resultant RNA code based ear feature extraction.

VI CONCLUSION

This Research paper explains the main module involved in the process of security-based load balancing. The predominant process of security services is explained with various techniques combined to develop a strong security model, during the transmission of data and load balancing efficiently. The Ear feature extraction is the biometric technique played a key role in generating the private key or secret key with the representation of RNA coding to make the brute force attacker a tough challenge. The NTRU is a primary cryptographic scheme combined with the Shamir secret sharing policy for providing more confidentiality, availability, and accountability in a distributed environment. After the security process then proposed work moves to the next process load balancing of requests.

REFERENCES

[1]. Alkhudhayr .F, S. Alfarraj, B. Aljameeli and S. Elkhdiri, 2019, "Information Security:A Review of Information Security Issues and Techniques," International Conference on Computer Applications & Information Security (ICCAIS), pp. 1-6.

[2]. Amit Gajbhiye, Dr. Shailendra Singh, 2017, “Global Server Load Balancing with Networked Load Balancers for Geographically Distributed Cloud Data-Centres”, International Journal of Computer Science and Network, pp. 682-688.

[3]. Bora A. and Bezboruah, T., 2020. “Some Aspects of Reliability Estimation of Loosely Coupled Web Services in Clustered Load Balancing Web Server”. In Critical Approaches to Information Retrieval Research, IGI Global, pp. 198-209.

[4]. Challa N. and Pradhan, J., (2007). “Performance analysis of public key cryptographic systems rsa and ntru”. International Journal of Computer Science and Network Security, 7(8), pp.87-96.

[5]. Dave .A, B. Patel, G. Bhatt and Y. Vora,2017, “Load balancing in cloud computing using particle swarm optimization on Xen Server,” International Conference on Engineering (NUiCONE), Ahmedabad, 2017, pp. 1-6.

[6]. Grudenic I. and N. Bogunovic, 2011, “Computer cluster scheduling algorithm based on time bounded dynamic programming,” Proceedings of the 34th International Convention MIPRO, pp. 722-726.

[7]. HaiTao Dong, Lei Song, Jin-Lin Wang, and Jun Yang, 2016, “SSLSARD: A Request Distribution Technique for Distributed SSL Reverse Proxies”, Journal of Communications Vol. 11, pp. 374-382.

[8]. Mainak Adhikari and Tarachand Amgoth, 2018, “Heuristic-based load-balancing algorithm for IaaS cloud", Future Generation Computer Systems, volume 81, pp. 156 – 165.

[9]. Neetesh Kumar, Deo Prakash Vidyarthi, 2019, “A Hybrid Heuristic for Load-Balanced Scheduling of Heterogeneous Workload on Heterogeneous Systems”, *The Computer Journal*, Volume 62, Issue 2, February 2019, pp. 276–291.

[10]. Rahul Godha , Sneh Prateek, 2014, “Load Balancing in a network”, International Journal of Scientific and Research Publications, Volume 4, Issue 10, pp. 1-3.

[11]. Shaoyi Song, Tingjie Lv, and Xia Chen, 2014, “Load Balancing for Future Internet: An Approach Based on Game Theory”, Hindawi Publishing Corporation Journal of Applied Mathematics, Article ID 959782, pp. 1-11.

[12]. Shweta R., Niharika G., 2015, “A Clustered Approach for Load Balancing in Distributed Systems”, Global Technical Campus, India, International Journal of Mobile Computing &Application (SSRG-IJMCA), Vol. 2, Issue 1, pp. 1-6.

[13]. Suresh .V.M, Karthikeswaran .D, Sudha V.M, Murali Chandraseker .D, 2012, “Web server load balancing using SSL back-end forwarding Method”, IEEE-international conference on advances in engineering, science and management pp: 822 -827.

[14]. Tian .W, Y. Zhao, Y. Zhong, M. Xu and C. Jing, 2011, “A dynamic and integrated load-balancing scheduling algorithm for Cloud datacenters”,IEEE International Conference on Cloud Computing and Intelligence Systems, Beijing, pp. 311-315.

[15]. Yang, H.A., Sun, Q.F., Saygin, C. and Sun, S.D., 2012. Job shop scheduling based on earliness and tardiness penalties with due dates and deadlines: an enhanced genetic algorithm. The International Journal of Advanced Manufacturing Technology, 61(5-8), pp.657-666.

[16]. Zhu X.M. and Lu P.Z., 2009. Multi-dimensional scheduling for real-time tasks on heterogeneous clusters. Journal of Computer Science and Technology, 24(3), pp.434-446.