

# Sundry Online Resource Allocation Technique

**Suraj Indiver**

*M. Tech. Scholar*

*Department of Computer Science & Engineering*

*Sachdeva Institute of Technology, Mathura, India & Dr. A.P.J. AKTU, Lucknow, India*

**Dr. Rajbir Singh**

*Assistant Professor*

*Department of Mathematics*

*Govt. College Tigaon, Ballabgarh, Faridabad, MDU Rohtak, India*

## Abstract

A method for strategic allocation of the resources is proposed here. All the website objects are viewed as resources, and they are allocated depending on the access rights of the person who is requesting, and the availability of the resource. With the rapid growth of the web, the slow response times of the popular websites due to server load and network congestion are viewed seriously because it leads to unsatisfactory customers. Most of the websites have replaced a single server with a cluster of replicated servers, in order to balance the server loads. Load distribution over the replicated servers needs to be controlled in distributed websites. It is a very challenging task to design an allocation strategy for a distributed site without sacrificing all of its benefits. Any solution must address the following issues like load balancing, transparent name resolving, legacy code, standard, scalability, flexibility, geographic distribution, dynamic changes in the server pool and fault transparency.

**Keywords: Online Resources, Online Services, User Categorization, Access Matrix, Circulation Token algorithm**

## I. INTRODUCTION

With rapidly increasing growth of the web, clients trying to access some famous websites are experiencing very slow response times due to overload of the server and network congestion. A method for strategic allocation of the resources is proposed here. All the website objects are viewed as resources, and they are allocated depending on the access rights of the person who is requesting, and the availability of the resource. Allocation of resource is confirmed to all requests by adopting a real time scheduling algorithm (Chandra et al 2003). Each resource is assigned with 'n' number of tokens if it could be allocated to the 'n' number of users simultaneously without any sacrifice in performance. The resource availability is checked by the maximum capacity of that resource. The access rights are identified by the access matrix which determines the rights of the clients about one particular object. The clients are classified and each category of the clients is given different access right. This would be more useful in billing the usage of the valuable resources on a web based applications. The website could be viewed as trusted system where those who have access can only use the resources binding to their access rights. This method ensures that all the user requests are processed with in some deadlines (Czajkowski et al 1998, Gupta 1998).

With the rapid growth of the web, the slow response times of the popular websites due to server load and network congestion are viewed seriously because it leads to unsatisfactory customers. Most of the websites have replaced a single server with a cluster of replicated servers, in order to balance the server loads. To provide solution to the problem of network congestion, few websites have separated the replicated servers geographically throughout the world. This leads to the concept of mirror sites for software archives, typically spread across several continents. Increase in performance of the service availability, congestion, and partial unavailability of the server is achieved in distributed architecture by considering advantage of proximity between clients and servers. Load distribution over the replicated servers needs to be controlled in distributed websites.

It is a very challenging task to design an allocation strategy for a distributed site without sacrificing all of its benefits. Any solution must address the following issues like load balancing, transparent name resolving, legacy code, standard, scalability, flexibility, geographic distribution, dynamic changes in the server pool and fault transparency; Distributed resource algorithms are not well suited for web resource allocation because of its nature of more number of simultaneous access requests (Nancy 1996). The web resources or remote objects can also be accessed simultaneously by more number of clients with a sacrifice in response time and a performance. The Bakery algorithm for mutual exclusion, which uses single writer and multiple readers for the shared variable, is useful. This is working fine for allocating a resource for a single user at any point of time.

## II. ONLINE RESOURCES

In the earlier stages of evolution of the web history, the web resource is all the static addressable documents or files. Now-a-days web resource is everything that could be identified or handled by means of name or address in the whole web or in any particular networked information system. In earlier specification of the web resource, there was no clear idea of separation between the identification and naming as well as addressing and handling. In order to clearly define this separation, one should address the technical, social, linguistic and philosophical issues. The World Wide Web is considered as a network of static addressable objects, which are mostly files and documents that are associated by using uniform resource locators (URLs). The implicit definition of the resource is anything that could be identified. This identification is justified for two unique purposes, naming and

addressing. All the protocols are depending on the addressing of the objects in World Wide Web. The explicit definition of the resource is given in (Request for Comment) RFC 2396 published in August 1998. It defines that a resource is anything that has identity. The examples of the web resource include anything like image, electronic document or service (Eg. Today's weather report, current traffic details of the city and so on).

The definition of the web resource is considered as stated above. Everything which includes electronic documents, web services, any addressable files etc are considered as a resource available in World Wide Web (WWW). All web resources are identified by using the Uniform Resource Identifiers (URIs), which is assigning a unique name or address.

### III. ONLINE SERVICES

Service components that are developed have to be integrated into composite services easily and flexibly, irrespective of its development language, platform, location, process and speed of execution. Because of the unpredictable nature and dynamics of the computing system, offering quality services to meet and satisfy the customer's requirements is a challenging task (Menasce 2002, Kreger 2003). Exchange of data between different computer applications using internet protocols is enabled by web services (Gottschalk et al 2002). A formal contract between a customer and a service provider must be established by defining all the aspects of the services which includes security, performance and availability. Such a contract is normally termed as service level agreement (SLA). The types of the security services are identity security and behavior security (Kaiqi Xion and Harry Perros 2006, Miller 1996).

- Identity security deals with authorization and authentication between a service provider and a client, data integrity and confidentiality.
- Behavior security deals with the trust among multiple resource sites, trust on the computing results provided by the sites and trust between the resource sites and its customers.

Performance component of the service agreement deals with the following two aspects

- 1) Response time is defined as the time for a service request to be satisfied.
- 2) Throughput is defined as rate of service that a service provider could offer.

### IV. NEED OF THE PROPOSED TECHNIQUE

Here a model is proposed for providing web services which could meet the customer satisfaction in terms of response time and throughput. The model finds its application in the following:

- 1) Billing the customer based on his usage on the web resources.
- 2) The website administrators could update and maintain the sites based on the popularity and demand of the resources.
- 3) Allocation of the resources to the authorized person only.
- 4) The authorized person can access the resource within his access rights or in other words the access rights for the resources could be restricted for different class of customers.

### V. PROBLEM STATEMENT

Every website object is viewed as a resource. The objective is to grant access permission to the clients for a specific resource requested if the clients have the access rights, to ensure the authorization and allocation of resources to those clients should not degrade the overall performance of the web server. The proposed model also provides a means of redefining the access matrix based on the demand of the remote object. Suppose the web application, website, or web service is hosted in a cloud environment, more processing and compute power is offered from the cloud provider. Huge number of websites of different size and memory would share the processing and computing power offered by the cloud. When a website becomes more popular and famous suddenly, the cloud could direct more number of individual computers to serve pages for the website, automatically and more money should be drawn for the additional usage. Since the usage of the cloud becomes less as a consequence of the web site lose its popularity, the amount of money will be less. Cloud computing becomes popular for its pay and use pricing model. Here a model is introduced in which the popularity of the web resources is computed based on the probability of the success of the request for the particular resource.

### VI. USER CATEGORIZATION

The visitors or customers of any website are distinguishable and have different taste and nature. The visitors' needs are different for individuals and the success is behind how the websites are dynamically providing service to every individual. Hence the user categorization is very much important and can be simply done by making the user to fill his personal needs and interest during the registration process.

This could be elaborated with an example of the matrimonial sites. The visitors are categorized as

- 1) Male seeking for a Female (Priority User)
- 2) Female seeking for a Male (Priority User)
- 3) Male seeking for a Female (Free User with Limitations)

4) Female seeking for a Male (Free User with Limitation)

For simulation purpose the user categorization is done by using multinomial distribution.

**A. Multinomial Distribution**

It could be stated that the multinomial distribution is a general form of the binomial distribution in probability theory. It is a categorical distribution. The possible outcomes are fixed and finite say k. Each trial results in exactly one of the possible outcomes with probabilities p1, ...,pk such that pi ≥ 0 for i = 1, ..., k and Σ p • = 1, and there are ‘n’ independent trials. (Papoulis 1984)

Let the random variables Xi point to the number of times outcome ‘i’ was observed over ‘n’ trials. The vector X = (X1, X2, ..., Xk) is following the multinomial distribution, with parameters n and p, where p = (p1, ..., pk). Probability mass function of multinomial distribution is given by Equation 5.1.

$$f(x_1, \dots, x_k; n, p_1, \dots, p_k) = P_r (X_1 \text{ and } \dots \text{ and } X_k = x_k) \quad (5.1)$$

$$= \begin{cases} \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}, & \text{When } \sum_{i=1}^n X_i = n \\ 0, & \text{Otherwise} \end{cases}$$

The above formula can be used to guess the probability of each user category ‘i’ get the resource.

**VII.ACCESS RIGHTS**

As earlier discussion it is clear that the different users need different services and resources. Their access to the resources could be restricted based on their category. The limitations or the restrictions can be illustrated with the matrimonial sites.

- Any category of the user should not have access to edit the profile of any other user except own.
- Male Seeking Female (Priority User) could access only the profiles of the females recommended based on his interest. The photographs will be shown only if both parties are agreed based on the profiles. Any male seeking female should not have access to view the profile of other males.
- Female Seeking Male (Priority User) could access only the profiles of the Males recommended based on her interest. The photographs will be shown only if both parties are agreed based on the profiles. Any Female seeking male should not have access to view the profile of other females.
- Male seeking Female (Free User) could access only the profiles of females recommended based on his interest limited by 5 views. The Photographs would be viewed only after his registration as a priority user to the site.
- Female seeking Male (Free User) could access only the profiles of males recommended based on his interest limited by 5 views. The Photographs would be viewed only after her registration as a priority user to the site.
- And so on.

Hence different access rights for the different category of the user is needed and the access rights are defined as view, modify and execute permission for a particular resource for a category of the user. This is modeled as a three dimensional access matrix in the proposed system.

**A. Access Matrix**

Access control matrix is normally a two dimensional matrix in which, rows represent different user categories and columns represents objects. Each entry in this matrix represents the set of access rights granted to user category for the corresponding object. It is an abstract security model. In reality, security is enforced either by Access control list (ACL) or capabilities, because access control matrix is sparse in nature, big in size and more space is required to store. The ACLs are columns of the access matrix. Each object is associated with a list of pairs for all user category and access permissions. Capability list is associated with each user category, indicates access permission to all objects. The row of access matrix represents the capability list. Both mechanisms require complex data structures and associated functions. It is hard to identify objects associated with a user category in ACL and user categories associated with an object in capabilities. Vincent et.al (2006) has made a detailed study on limitations of these access control mechanism. To overcome these limitations a three dimensional access matrix of data type binary, is proposed which reduces space requirements of access matrix.

The proposed access matrix is the three dimensional matrix in which first dimension is the resource, the second dimension is the user category and the third dimension is the access rights. Each row indicates the resources and each column indicates the category of user and the intersection of this is the access rights. If any website is having a ‘N’ resources, its clients could be classified as ‘M’ categories, and each category of the user can be provided ‘A’ access then the access matrix will be N x M x A of the form as shown in the Figure 1.

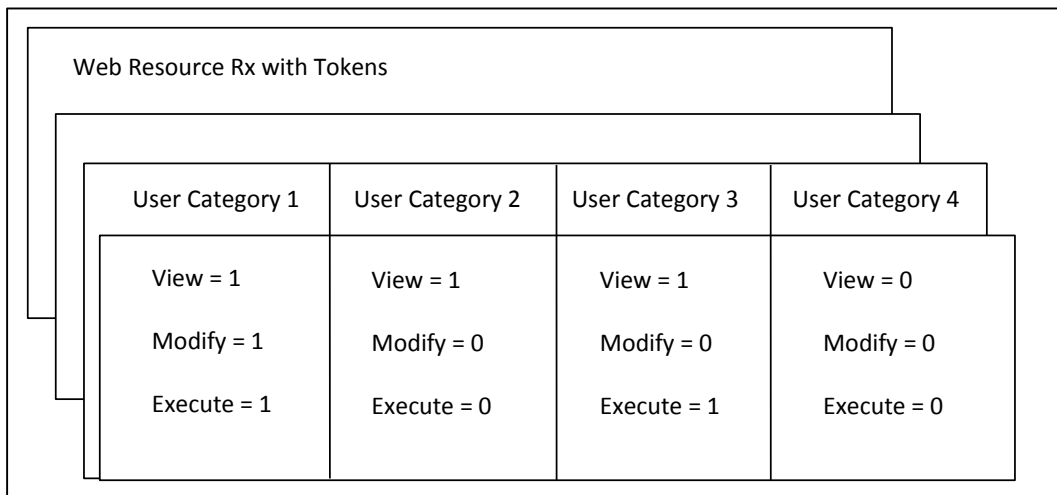


Fig. 1 Illustration of the web resource Rx and its access permission for different categories of the user

Every client, customer or visitor for his requesting resource may be defined with three different access rights as view, modify and execute permissions in the model. The access matrix will be binary matrix where the entry as '1' if user could be granted the required access for the requested resource, else '0'. The access matrix is shown in the Figure 2.

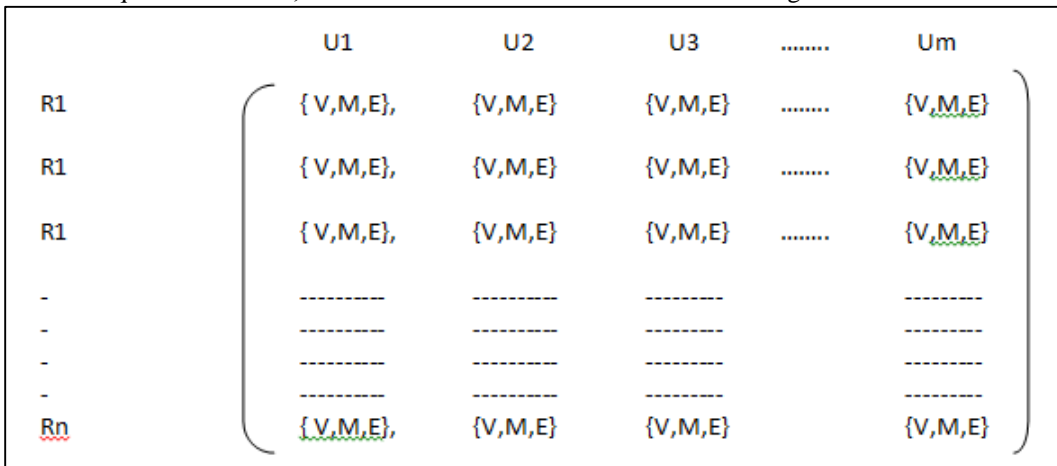


Fig. 2: Access Matrix

The access control list corresponds to rows and capability corresponds to columns in the proposed three dimensional access control matrix. The memory requirement for the three dimensional binary matrix is given by  $M \times N \times A$  which is relatively small when considering minimum number of user categories for websites and access rights. The metrics for comparison of various access control mechanism is studied by Vincent et.al. (2006). Based on their study, comparative analysis is made on different access control mechanisms like access control list, capabilities and three dimensional access matrix and tabulated in Table (5.1).

Table – 1  
Comparison of access control mechanism

Access control list	Capabilities	Three dimensional access matrix
Separate method required to pair objects and subjects	Built in method to pair objects and subjects	No method required to pair objects and subjects.
Hard to identify the related objects for a specific category of users.	Easy to identify the related objects for a specific category of users	Easy. The column corresponds to the user category gives the objects associated.
Easy to identify the related user category for an object.	Hard to identify the related user category for an object.	Easy. The row corresponds to related user category for an object.
Delegation is difficult	Delegation is easy.	Delegation is easy.
Less complex to implement.	More complex to implement.	Less complex to implement.
Updating user capabilities into system is difficult.	Updating user capabilities into system is easy.	Updating user capabilities into system is easy.
Updating objects access control entries into system is easy.	Updating objects access control entries into system is difficult.	Updating objects access control entries into system is easy.

## VIII. ALLOCATION OF THE RESOURCES

Allocation of resource is granting access to the web resource for the request raised by the client or customer. This allocation of the resource should be guaranteed in minimum response time, if the client is an authorized person to raise such request. Multiple clients may raise the request for the same web resource should be dealt with care because this may pull down the system performance. To avoid this problem a circulating token algorithm is adopted which solves the mutual exclusion problem and guarantees lockout-freedom.

### A. Circulation Token Algorithm

The simplest mutual exclusion algorithms for the asynchronous send/receive network setting works when the network is a unidirectional ring. A token representing control of the resource circulates continuously around the ring. When process  $P_i$  receives the token, it checks whether there is an outstanding request from user  $U_i$ . If there is no such request,  $P_i$  passes the token to  $P_{i+1}$ . On the other hand, if there is an outstanding request,  $P_i$  grants the resource to  $U_i$  and holds the token until  $U_i$  returns the resource. When  $U_i$  returns the resource,  $P_i$  passes the token to  $P_{i+1}$  (Sedigheh et al 2009).

#### 1) Theorem 8.1

“The Circulating Token algorithm solves the mutual exclusion problem and guarantees lockout-freedom (Chandra et al 2003).”

### B. Proof:

The proof is straightforward. Mutual exclusion is guaranteed, because there is only one token, and only the user where the token is located can be in C. Progress is guaranteed, because the token keeps circulating until it finds a request. Lockout-freedom is guaranteed, because no process satisfies two consecutive requests without allowing the token to circulate around the ring in the interim, thus giving every other process a chance. Hence the circulating token algorithm is used in the proposed method for scheduling the multiple resources and the deadlock is completely avoided. With the help of access matrix, authenticity is verified.

In the method the concept of circulating the token is used to allocate the web resource to the request raised by the clients. Each web resource is holding ‘n’ tokens, where ‘n’ is a predetermined integer indicates that the web resource could be allocated to ‘n’ request simultaneously without degrading the performance metrics.

- A web resource is allocated to the request raised, and then the token associated with the web resource is reduced by one.
- A client finishes his job with the web resources; the token associated with the web resource is increased by one, with the maximum upper bound as ‘n’.
- If the token associated with the web resource reaches zero then the further request for that web resource is put on pending and after a random amount of time it check for the token availability.

## IX. PROPOSED MODEL

Every web resource is viewed as objects that are stored in a remote server. At any time

- $Token[i] = n$ , if the object ‘i’, can be simultaneously accessed by ‘n’ clients without sacrificing its performance. The  $token[i]$  is a shared variable can be accesses by all the threads, and it is a similar concept of N-ary semaphore, which enforces the mutual exclusion.
- Clients may request any object randomly at any point of time, if the token is available with that object then server checks for the access rights defined in three dimensional access matrix.
- The object will be granted to the requesting client only when the client is having an access rights and tokens. Hence authenticity of is enforced.

In this proposed method the clients’ access is restricted and server overloading is avoided completely. When the client is requesting an object and the object is not available then the client request is temporarily postponed for a random amount of time and retried again. This is being done until the client gained token. The concept is illustrated in the Figure 5.3. Figure 5.1 shows a web resource  $R_x$  with 4 tokens. The remote object  $R_x$  can be accessed by 4 clients simultaneously without any sacrifice in the performance. The clients fall under the user category 1 is having full access rights for the resource  $R_x$  hence the view, modify and Execute permission is set to ‘1’ whereas the clients falls under the user category could not access the resource  $R_x$  since the access permission is completely denied for this category of the user by setting the access rights as view, modify and execute to ‘0’. The clients of user category 2 can only view the resource and its modify and execute permission is restricted, whereas the clients comes under the user category 3 is granted only the view and execute permission and modify permission is restricted. Any client requesting resource  $R_x$ , for the access ‘a’, first its user category is identified and their access right for the access ‘a’ is checked with access matrix. The resource is allocated to the client only when the token is available and the access is granted for that category of the user.

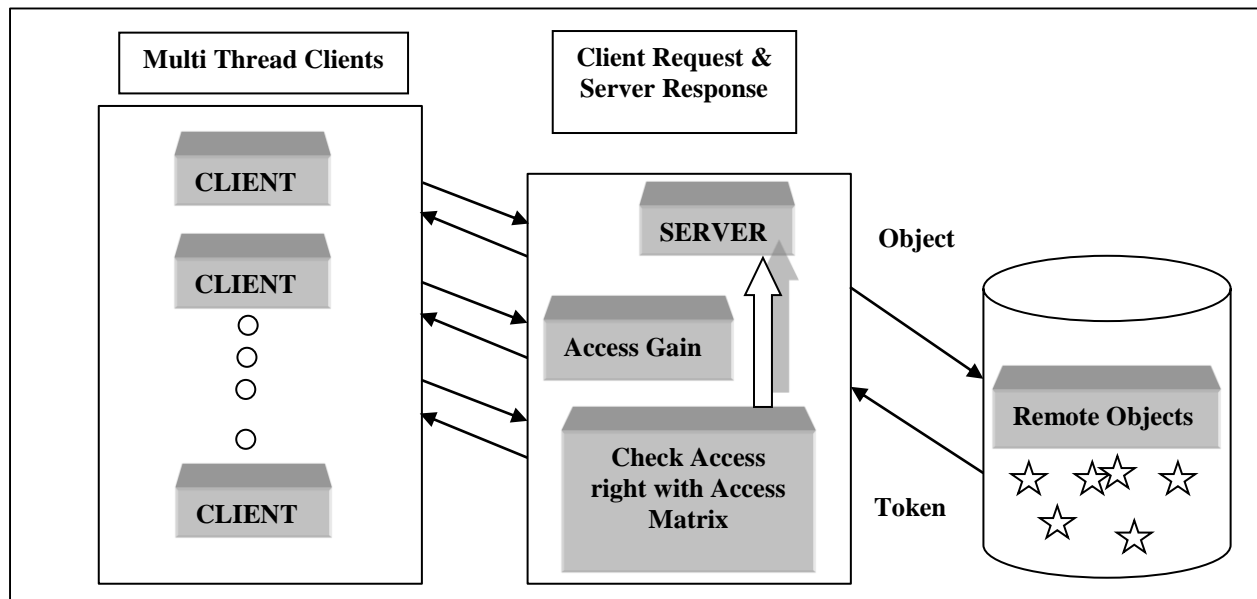


Fig. 3: Block diagram of the proposed algorithm

## X. ALGORITHM

- 1) Step1: Categories the users of the website into 'm' different category. This can be done by using any classification algorithm
- 2) Step 2: Obtain the access matrix based on the previous data collected from website log details, the constraints, value and price of the resource.
- 3) Step 3: For each resource Rx initializes the tokens for the resource such that the server performance would not be degraded.  
For  $R_x = 1$  to no of the resources  
token  $[R_x] = n$ ;  
/\*The resource Rx can be simultaneously accessed by  
'n' clients without sacrificing its performance \*/
- 4) Step 4: For Each client repeat the steps 4, 5 and 6.
- 5) Step 5: Client U1, raise a request for a web resource Rx for the access 'A'
- 6) Step 6: if token  $[R_x] < 1$  then
  - Step 6a: While (token  $[R_x]$ )
  - Retry to gain the token after a random amount of time
  - Else
  - Step 6b: token  $[R_x]--$ ;
  - Step 6c: Check the access for the Rx for the category of the user U1;
  - Step 6d: if((access = Access Matrix $[R_x][U1][A]$ ) == 1)Requested resource is granted for the user u1. Increase the positive count
  - Else
  - Increase the negative count
  - Step 6e: Process with the resource Rx with random amount of time.
  - Step 6f: If the process over update the token  $[R_x]$  by returning it to the resource.
  - token  $[R_x]++$ ;
- 7) Step 7: Repeat the steps 4 and 5 until the client U1 finishes the job.

## XI. MODEL ASSUMPTIONS

- The web site is assumed to have eight resources.
- Each resource is attached with four tokens.
- The clients of the website are categorized as four different groups.
- The access matrix for these resources is defined for all the category of the clients with three different accesses as view, modify and execute for all the web resources. Hence the order of the access matrix will be  $8 \times 4 \times 3$ .
- Simultaneous request raise from the clients is simulated by the multithreaded clients (Jain et al 2002).
- The categorization of users and request for specific resources are simulated by using random functions.

## XII. EXPERIMENTAL SETUP

The user request for a web resource is not deterministic and random in nature. More than one client may raise the request for the same resource or different resource simultaneously. These simultaneous raise of the request for the web resources could be simulated by making the clients request as multithreaded. Every thread is raising a client request for a specific web resource, checking the availability of that resource, checking the access rights for the resource to the clients, granting the resource for the client based on the availability and access rights, processing the resources, updating the tokens of the web resources after finishing the process. These steps are repeatedly done until the client finishes its job (Figure 5.4). Tokens and access matrix are synchronized shared resources used by all the threads. The java program used for simulation is given in appendix A1.5. The access matrix and the web resource tokens are made as common and shared among all the threads. For each client the positive hits that is the requested resource for the required access is granted and the negative hits are recorded. The positive hits matrix is a three dimensional matrix whose entry (Prua) records the number of times the resource(r) is allocated to a user category (u) with the specified access (a). Similarly the negative hits matrix entry Pure records the number of times the request is denied for a user category (u) for access (a) to a specific resource (r) due to enforce the authorization. The availability of the resource is ensured because of the clients thirsty for the tokens. That is, the clients are waiting and retrying again and again for the resource until the token is available. The cumulative positive and negative hits give the measure of the correctness of the access matrix. More number of negative hits implies that more number of clients requesting for a web resources and its access is denied for the clients requesting the resource. It becomes a measure for redefining the access matrix or assigns more value to the resource.

## XIII. RESULTS

The result is produced by simulating the simultaneous request by multithreading clients. These multithreading clients are raising requests randomly to any one of the available 8 resources with specific access permission defined by three dimensional binary access matrix. The request raised is simulated for 999 users. Cumulative total number of the request raised is 4537 out of it, 2854 was granted and 1683 was restricted. The possibility of the request restriction is approximately 37%. To reduce this number of restrictions the access matrix should be updated properly. The details of granting and denying access of individual resources for specific resources are captured in positive hits matrix and negative hits matrix respectively. The modification of the access matrix can be done based on the negative hits matrix. Higher the number in the negative hits matrix entry indicates that the many users who don't have access are willing to have access to that resource. The service to all requests is ensured since 155 requests are in waiting state to get the resource availability and the service is successfully completed by obtaining the same number of tokens.

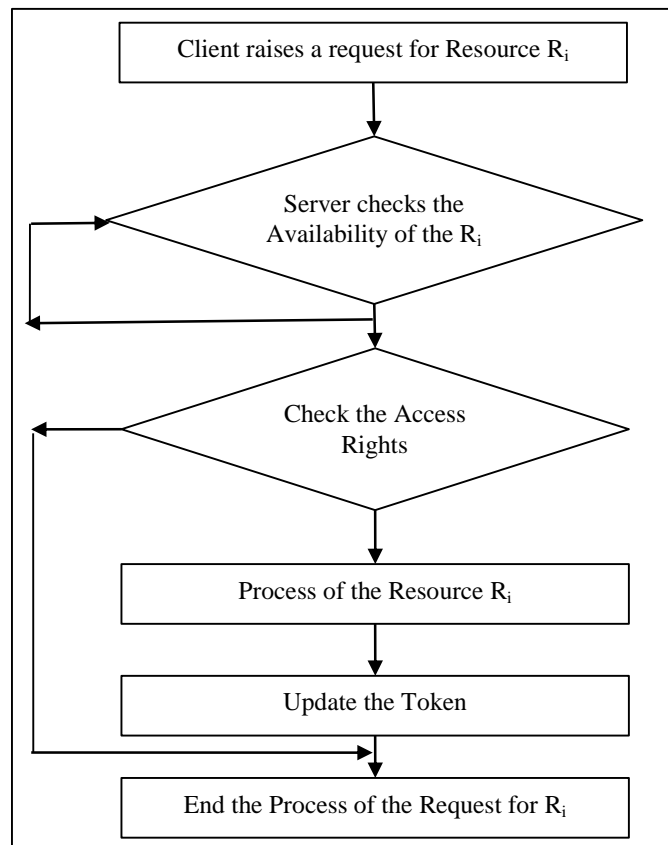


Fig. 4: Process involved in granting a resource to a request

Number of Users : 999  
Total hits : 4537  
Cumulative positive: 2854  
Cumulative negative : 1683  
No token negative : 155  
Token obtained : 155

#### A. Total hits Matrix

{45 40 44}, {61 62 52}, {68 59 69}, {57 62 55},  
{42 36 43}, {51 52 33}, {57 56 59}, {52 46 51},  
{41 61 38}, {37 35 38}, {52 54 74}, {48 61 53},  
{39 31 48}, {41 57 53}, {42 57 49}, {42 56 47},  
{45 44 50}, {51 55 40}, {52 53 59}, {38 38 41},  
{45 45 41}, {57 49 44}, {52 50 69}, {36 39 54},  
{53 45 38}, {44 60 33}, {54 53 44}, {38 54 40},  
{51 41 58}, {54 44 38}, {56 50 50}, {43 40 58},

#### B. Positive hits Matrix

{45 40 44}, {61 62 52}, {68 59 69}, {57 6 5},  
{42 36 43}, {3 0 1}, {57 56 59}, {52 0 1},  
{41 61 38}, {37 35 38}, {2 1 0}, {48 2 0},  
{39 31 48}, {1 0 0}, {0 1 1}, {42 0 1},  
{45 44 50}, {51 55 40}, {52 53 59}, {38 0 0},  
{45 45 41}, {57 49 44}, {2 0 0}, {36 0 0},  
{53 45 38}, {1 4 0}, {54 53 44}, {38 0 0},  
{51 41 58}, {54 44 38}, {0 1 1}, {43 1 1},

#### C. Negative hits Matrix

{0 0 0}, {0 0 0}, {0 0 0}, {0 56 50},  
{0 0 0}, {48 52 32}, {0 0 0}, {0 46 50},  
{0 0 0}, {0 0 0}, {50 53 74}, {0 59 53},  
{0 0 0}, {40 57 53}, {42 56 48}, {0 56 46},  
{0 0 0}, {0 0 0}, {0 0 0}, {0 38 41},  
{0 0 0}, {0 0 0}, {50 50 69}, {0 39 54},  
{0 0 0}, {43 56 33}, {0 0 0}, {0 54 40},  
{0 0 0}, {0 0 0}, {56 49 49}, {0 39 57},

## XIV. CONCLUSION AND FUTURE DIRECTIONS

The three dimensional access matrixes are used to allocate the resource only to the authorized persons only. The N-ary semaphore token is used to avoid the mutual exclusion problem for shared resource allocation. This algorithm is very useful in cloud computing where the pricing of the web resources is essential and based on the website popularity and utilization. It also helps in increasing the Quality of service by identifying the most used and famous resources.

## REFERENCES

- [1] Bailey, J., Bry, F., Furche, T., and Schaffert, S. "Web and Semantic Web Query Languages: A Survey". In Proc. of Reasoning Web, First International Summer School. LNCS 3564, Heidelberg, Germany, pp. 35-133, 2005.
- [2] 10. Barry, D. K., "Web Services and Service-Oriented Architecture: Your Road Map to Emerging IT", Morgan Kaufmann, 2003.
- [3] Carlos Castillo, "Effective web crawler" Thesis submitted to University of Chile, November 2004. 12. Chakrabarti, S. "Mining the Web", Morgan Kaufmann, 2003.
- [4] Gottschalk, K., Graham, S., Kreger, H. and Snell, J. "Introduction to Web Services Architecture," IBM Systems Journal, Vol. 41, No. 2, pp.170-177, 2002.
- [5] Gupta, A. Stahl, D. O. and Whinston, A. B. "Managing Computing Resources in Intranets: an Electronic Commerce Perspective," Decision Support Systems, Vol. 24, pp. 55-69, 1998.
- [6] Menasce, "QoS issues in Web services", IEEE Internet Computing, Vol.6, No. 4, pp.72-75, November 2002.