International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

On Demand Cache Management and Cache Migration to Balance the Cache Load

Ramya S[1], Prof. Chaithra[2]
P. G Student[1], Associate Professor[2], Department of Computer Science and Engineering, SCE, Bangalore, Karnataka

Abstract: In Cloud Computing, balancing the workloads and managing the workloads across multiple servers is very important. The distribution of cache data among multiple servers according to their actual demand is also important. In existing system, cache management has a disadvantage where useful data is flushed out. In this project, an on demand cache management method is used where cache is allocated based on their demand and in order to store all important, a download count is used and a cache migration is used in order to balance cache load.
Keywords: Cache Management, Cache Migration, Load balancing, Cloud computing, Cloud Cache

I. INTRODUCTION

In cloud, load balancing is a process of distributing the workloads and the resources in the cloud environment. Multiple servers, network or computer is maintained where workload and resources are distributed. A cache is a small part of memory where it contains the recently used data or more frequently used data.

Cloud Cache is on-demand cache management where users used data is stored in cache of the cloud and is easily provided to other customers rather than getting the data from main server every time when it's been requested from other users .Cache of the cloud would have specific space to hold the data .When there is a condition wherein useful data of cache is flushed out from the cache, to hold the new data ,then some part of the data of the cache which is important will be migrated to some other location that is cache migration.

II. EXISTING SYSTEM

The Cloud cache of the cloud stores recently used data in the cache. All the recently accessed data is stored so that new user who wants to access same data will be redirected to cache rather than to main server where processing time will be decreased. But the issue with existing system is when there is more data which is frequently accessed to be placed on cache there is no space in cache to hold the data at that time the first placed data gets deleted from the cache.

III. METHODOLOGY

The Cloud cache of the cloud stores the recently used data in the cache but only if the data is used n number of times from many users, there is a count maintained for the data to be placed in the cache and it also overcomes the cache overload situation where the data will not be flushed out to place new data rather the earlier data will be migrated to other location.

The methodology used here is as shown in Fig.1.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

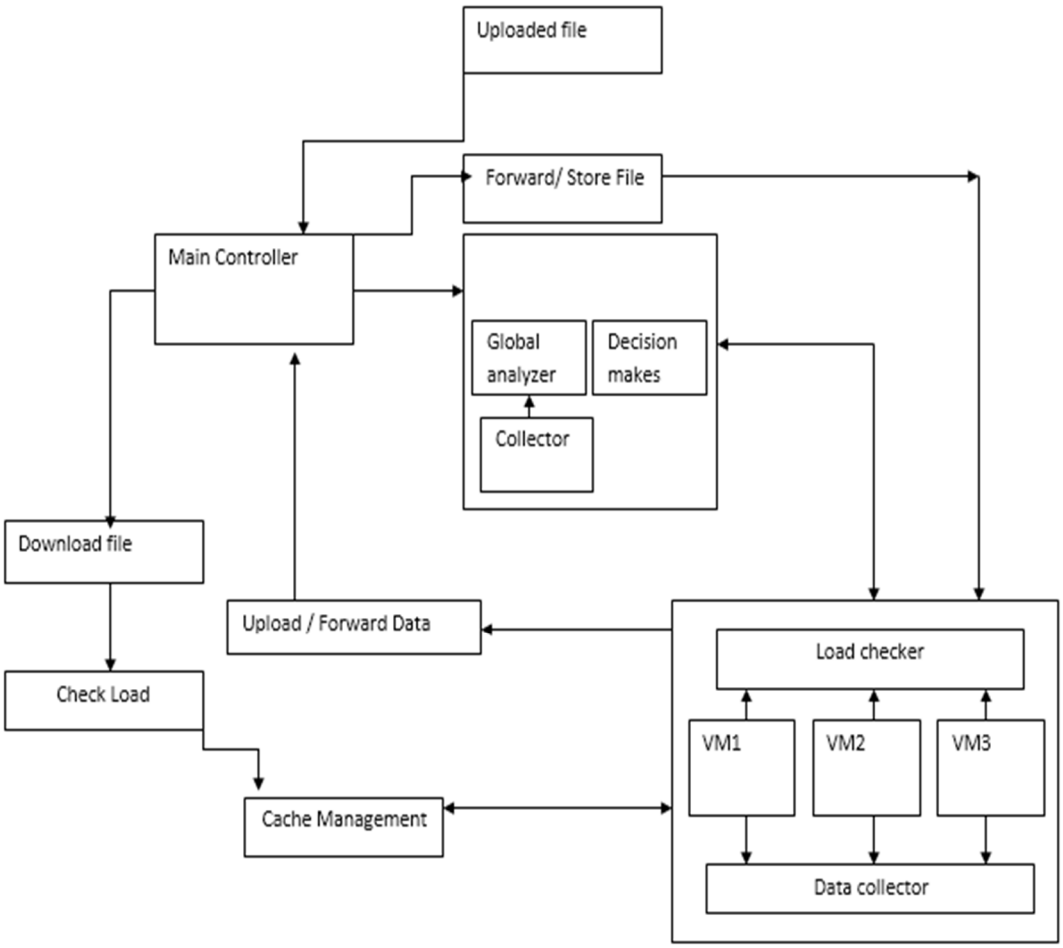

Fig.1 System Architecture

A. Load balancing by servers

In this module, a load balancer sits between the client and the server farm accepting incoming network and application traffic and distributing the traffic across multiple backend servers using various methods. By balancing application requests across multiple servers, a load balancer reduces individual server load and prevents any one application server from becoming a single point of failure, thus improving overall application availability and responsiveness A load balancer acts as the "traffic cop" sitting in front of your servers and routing client requests across all servers capable of fulfilling those requests in a manner that maximizes speed and capacity utilization and ensures that no one server is overworked, which could degrade performance. If a single server goes down, the load balancer redirects traffic to the remaining online servers.

B. Cache Management

In this module, how the data is placed in cache is observed. The admin logs in and uploads data into cache at that time data gets uploaded to all those servers which is maintained by main server. When the user downloads some file the control first go to main server from there the server which is least loaded gets into action and then a copy of the data is placed in cache of server which is been selected.

C. Cache Migration

In this module, When the cache gets full with all the data which is requested by user then cache gets overloaded, then some amount of data is deleted from cache so to overcome this instead of placing all of the requested data a download count is maintained for the data .The download count is 3 then data is taken out from cache of that server and data is placed into main server next time if that data is accessed then data is given to user from main server.

IV. RESULTS

Proposed system overcomes the cache problem where useful data is not flushed out, important data which is recently accessed by n number of times say three times is been measured by download count and when that condition is satisfied that data is moved to other location where it becomes very important data, so that other data could be placed in cache.

V. CONCLUSION AND FUTURE WORK

The future work of this project could be taken in a way there would be data in the cache which are all important which are all having same download count, when a new data encounters into cache that time cache would be filled with other data, there is cache insufficient problem so that, the present files of cache which is having more download count will be transferred into other location so new data will be placed.

REFERENCES

[1] Mingming Zhang, Songyun Wang, "Dynamic Load Balancing for Physical Servers in Virtualized Environment ", 2016 17th International Conference on Parallel and Distributed Computing, Applications and Technologies

[2] Peter J. Denning "The Working Set Model for Program", Communications of the AMC, Volume 11, Number 5, May 1968

[3] Dulcardo Arteaga and Jorge Cabrera, Jing Xu, Swaminathan Sundararaman, Ming Zhao, "Cloud Cache: On-demand Flash Cache Management for Cloud Computing", Proceedings of the 14th USENIX Conference on File and Storage Technologies (FAST '16), 2016

[4] Chaturong Sriwiroj, Thepparit Banditwattanawon "An Economic Model for Client Cloud Caching Service", Knowledge and Smart Technology (KST), 2015 7th International Conference