REVIEW PAPER ON CACHE BASED BIG DATA MANAGEMENT IN MOBILE NETWORKS

Virat R. Jani[1], Prof. Ashutosh A. Abhangi[2]

[1]Computer Engineering, Noble Group Of Institutions - Junagadh
[2]Computer Engineering, Noble Group Of Institutions - Junagadh

Abstract — *In this paper we reviewing the big data cache management in mobile networks in this we mainly focus on three terms that is caching schema, cache decision system and cache replacement policy ,In this we discuss major cachiong policies like request based caching, priority based caching, etc. we discuss about cache decision system respect to caching schemas and cache replacement policy for replace the cache in local storage.*

Keywords- *Big Data; Big Cache; Mobile Networks; Big data Management;*

I. INTRODUCTION

21st century is century of information and technology where each and every entity produce the torrent of data so called big data, So just because of that it is difficult to manage this type of large data so that we have to manage the big data by using various technique so that we can easily, securely and fast access these data.

Now a days the big challenge is to analyze[8] and process these torrent of data into the time so that data is transfer useful information to users. Further mobile networks is growing faster and faster and it will also grow the growth[10] of data like multimedia (youtube , vimo, etc..), image, audio, video etc. So that there is technique so called cached based approach for the frequently used data by mobile users, the main motive of using cache based approach is to effective bandwidth utilization , fast access on lower bandwidth devices such as mobile, tablet etc. So that caching based technique is mainly divided in to three part , first is caching schema, second is cache decision system and third is cache replacement policy.

II. LITERATURE SURVEY

Mainly we classify big data caching in three part Caching Schema ,Cache decision system and Cache replacement policy.

A. Caching Schema
Caching schema means technique to cache data so that here we discuss several caching techniques.

1) Request Based Caching (RC)
In Request based Caching[1] it is based on user request that means as user request for data it will directly go for caching. That means data will saved directly on user request based that will increase entropy and lead to frequent cache changes.

Request based caching is also maintains requested cache queuing in cache of storage is not available. Cache replacement in RC is FCFG(First Come First Go) or Pipelining that means the first cached data will remove from storage in case of storage full and new requested data is there.

2) Priority Based Caching (PC)
Priority based caching[1] is based on several parameters like event-related, access-frequency, rating, size etc. In this the concept is that user is more interested about current or new events rather then the old event or completed event, for example current sport event or live concert or any natural disaster etc. More over such event create spike on internet for period of time, Once event completed its access frequency decreased. So this is the idea behind PC.

In PC cache replacement is based on cache decisive factor. If storage is full and new data arrived for caching then the data with lowest cache decisive factor will removed.

3) Proactive Caching
Proactive caching[5][6] is impactful caching schema, In proactive caching some related data will pre fetched by the local storage and when user request for the data it is already in local storage, For example

Here big data platform is in charge of tracking or predicting user demand. Popularity matrix is used for cache decision system. In case of cache replacement data with lower popularity matrix will replace by new cache data. It is very complex system.

4) *Scale base Caching(SC)*

Scale base caching[1] is mainly based on decisive parameter that's whys it is called scale based. In SC it is dependent on mainly four parameter , access-frequency, rating, size and freshness, by using all for parameter it will generate cache decisive factor and on based on the factor data is cached.

Scale based caching most effective way of caching. It evaluate new caching data with the existing data and then cached. In cache of cache replacement the data with lower indexing replaced by the higher index.

SC is optimized way to caching the data. It has lowest cache miss ratio and give optimal throughput.

B. *Caching Desition System*

Caching decision systemis core part of Big data cache[7] in this we decide that on which bases data should be cache or not ,There are several cache decision system based on their caching schema, There is Request based cache decision system in this it is based on " First Come First Go"[1] method.

There are also some user based cache decision system which is based on user interaction and user behavior.

There is also parameter based cache decision system[1] where it is based on decisive parameter n_d

Here,

$$n_d = x1\left(\frac{s_d}{S}\right) + x2\left(\frac{T - t_d}{T}\right) + x3\left(\frac{f_d}{F}\right) + x4\left(\frac{r_d}{R}\right) \qquad (1)$$

where $\sum_{i=1}^{n} xi = 1$

$S = Max(s_i) , i \in N$

$T = Max(t_i) , i \in N$

$F = \sum_{i=1}^{|N|} f_i$

$R = \sum_{i=1}^{|N|} r_i$

Here, parameter *s, t, f, r* is respectively data's size, freshness, frequency, rating and *S, T, F, R* respectively Total storage size, total freshness time, total access frequency, total rating.

There are also some user based cache decision system which is based on user interaction and user behavior.

C. *Cache Replacement policy*

Caching decision system is core part of Big data cache in this we decide that on which bases data should be cache or not ,There are several cache decision system based on their caching schema, There is Request based cache decision system in this it is based on " First Come First Go" method.

1) *Scale based cache replacement policy*

In Scale based cache replacement policy[1] it is based on scale index so that cache with lower index will be replaced by the higher index cache. Scale based replacement policy is most effective and commonly used policy.

2) *LRU (Least Recently Used)*

LRU (Least Recently Used) [2] is based on the observation that data that have been heavily used recently will probably be heavily used again in the future. Conversely, data that have not been used for ages will probably remain unused for a long time. In LRU when cache is full the data item that has been unused for the longest time has been replaced.

3) *LRU Min*

LRU min[3] is as like LRU but the difference is LRU Min is try to minimize the number of cache replaced, It is basically depend on the data size and data access frequency.

LRU Min policy will increase the hit ratio of smaller sized data items.

4) *SXO*

SXO[4] is a local replacement policy[8] which is based on the two parameters data size and access frequency for replacement .In this larger sized data items are removed first because they occupy more space in storage.

Main benefit of using this policy is that the these two parameters are available easily.

5) *(LUV) Least Utility Value*

As name suggest it is based on least utility value[7], So that it is dependent on access portability, size of cache data and physical distance between mobile user and data source.

6) *LUV Min*

This replacement policy[10] it has two parts replacement and migration. The replacement decision is based on a utility value generate by combining the parameters access probability, distance, Size and Coherency. In the migration part the replaced data is stored in the neighboring nodes which have sufficient space.

7) *Count Vector*

In count vector policy[11] each data item maintains a count which gives the number of nodes having the same data. Whenever the cache is full data item with maximum count is removed first as this will be available in the nearby nodes. Whenever a data item is removed from the cache the access count will be decremented by one. Initially when the data is brought in to cache the count is set to zero.

III. PROPOSED WORK

We are proposing the notion of gear architecture for accelerate large data in mobile networks. In this we propose four gear system which contain Analytical Gear for user data analysis and track data activity like rating , access frequency etc., Storage Gear for effective search for local storage , Bandwidth Gear for effective bandwidth utilization and Searching gear for effective search mechanism.
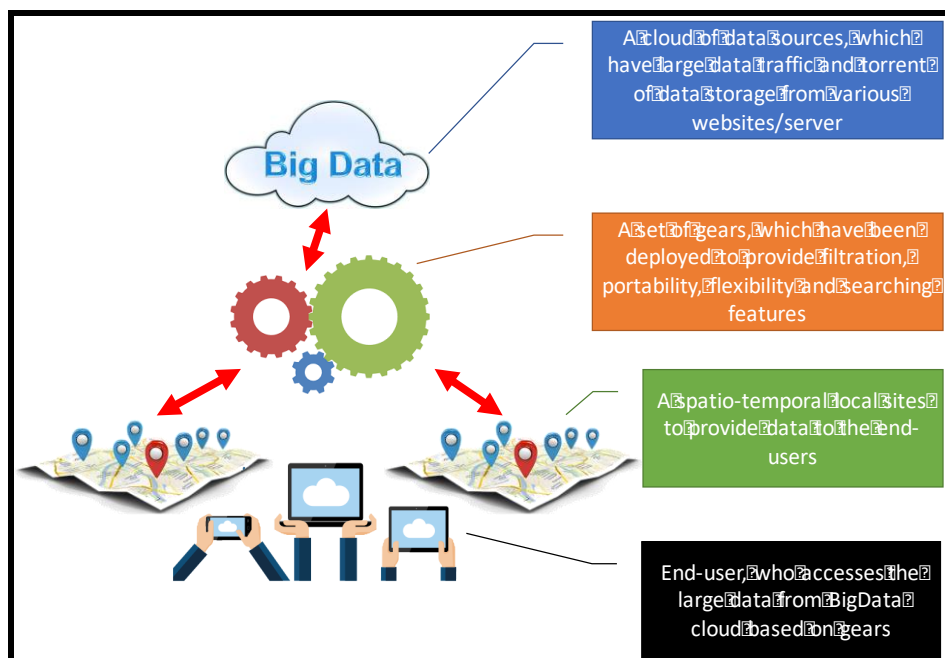

Fig.1 Basic Architecture of the System

IV. CONCLUSION

In this paper we reviewed of Caching Schema, Cache Decision system and major replacement policies in big data and mobile networks and summarized the key schemas and general overview of Big cache. In cache decision system there are several techniques are available but the most effective schema is scale based caching, we can also merge the two caching schema . In Cache decision system it is mostly based on the caching schema so that it is changed as caching schema changed, the effective cache decision system is scale based system because it will give most effective way of decision system , Further in Big Cache the third entity is cache replacement policy ,In these there are so many schema most of them are dependent on size, access frequency and location.

In future more parameters can be considered in cache decision system as well as improvement in cache decision system.

REFERENCES

[1] Nilesh Padhariya , Kshama Raichura , "BigCache: A Cache-based BigData Management in Mobile Networks" © International Journal of Mobile Communication - 2017.

[2] Preetha Theresa Joy and K. Polouse Jacob," Cache Replacement Strategies for Mobile Data Caching" © International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.3, No.4, August 2012.

[3] Denko, M.K., Tian, J.,Cross-Layer Design for Cooperative Caching in Mobile Ad Hoc Networks, Proc .of IEEE Consumer Communications and Networking Conf(2008). Cooperative Caching in Mobile Ad Hoc Networks, Proc .of IEEE Consumer Communications and Networking Conf(2008).

[4] L. Yin, G. Cao: Supporting cooperative caching in ad hoc networks, IEEE Transactions on Mobile Computing, 5(1):77-89(2006).

[5] Ejder Bastug, Mehdi Bennis, Engin Zeydan, Manhal Abdel Kader, Ilyas Alper Karatepe, Ahmet Salih Er, and Merouane Debbah" Big Data Meets Telcos: A Proactive Caching Perspective" © Journal of Communications and networks, vol. 17, no. 6, December 2015.

[6] Ejder Bastug, Mehdi Bennis, Engin Zeydan, Manhal Abdel Kader, Ilyas Alper Karatepe, Ahmet Salih Er, and Merouane Debbah "Big Data Caching for Networking: Moving from Cloud to" © IEEE Communication magazine – 2016.

[7] Chand, N. Joshi R.C., and Misra, M., Efficient Cooperative Caching in Ad Hoc Networks Communication System Software and Middleware.(2006).

[8] Ying He, Fei Richard Yu, (Senior Member, IEEE), Nan Zhao1, (Member, Ieee),Hongxi Yin1, Haipeng Yao, And Robert C. Qiu, (Fellow, IEEE) "Big Data Analytics in Mobile Cellular Networks" © 2016 (IEEE) Digital Object Identifier 10.1109/ACCESS.2016.2540520.

[9] Michal Angelo , Yiki Xu, "BigCache for BigData Systems " IEEE confarance on Big Data- 2014

[10] Narottam Chand, R.C. Joshi and Manoj Misra, "Cooperative Caching Strategy in Mobile Ad Hoc Networks Based on Clusters," International Journal of Wireless Personal Communications special issue on Cooperation in Wireless Networks, Vol. 43, Issue 1, pp. 41-63, Oct 2007

[11] B. Z heng, J. Xu, and D. L ee. Cache invalidation and replacement strategies for location dependent data in mobile environments,. IEEE Transactions on Computers, 51(10) : 1141–1153, October 2002.