International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 4
ISSN: 2454-4248
417 – 420

Big Data Gears:The Data Accelator For Large Data In Mobile Networks

Mr. Virat R Jani
M.E. Computer Engineering
Noble Group of Institutions
Junagadh, India
virat.jani@ngivbt.edu.in

Prof. Ashutosh A. Abhangi
Computer Engineering
Noble Group of Institutions
Junagadh, India
Ashutosh.abhangi@ngivbt.edu.in

***Abstract* —** BigData management is merely concerned with the large-scale data, generated from the cloud, mobile computing, media accessing and many more. Based on the research survey, such as BigCache, which emphasizes on efficiently serve the data to the end-user by provisioning the large cached data at local site, thereby to benefit effective bandwidth utilization, low server load and rapid data access at low cost. To extend this work, we would like to incorporate the notion of data gears in BigData management, to accelerate the large data access in mobile networks, which consist of the small size portable devices with mobility and low resources. These lead to the many challenges such as small cache, bandwidth management, power consumption, data sharing, secured access and so no. We mainly focus on the data access and bandwidth level in BigData through providing 'Gear Architecture' for the flexibility, portability and interchangeability of the various components in BigData over mobile networks. We will target to improve the utilization through proposing data access techniques in extended ranges using BigData analytic as a Gear.

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

*****

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

A. BigCache: A Cache-based BigData Management in Mobile Networks [1].

Now a days BigData is most trending concept and when we integrate BigData with mobile networks for mobile computing it will generate new concept called BigCache. BigCache is very much important system for addressing efficient caching of a data item in mobile networks.

In this paper author proposes a new cache base system for mobile networks with low Internet bandwidth and less mobile resource that constraints cause a frequent disconnection, thereby hard to access BigData ,thereby author describe three caching method schemes, namely *Request-based Caching scheme* (RC), *Priority-based Caching scheme* (PC) and *Scale based Caching scheme* (SC) along with cache decision system (CDS) and cache replacement process for cache based BigData management.

B. Cache Replacement Strategies for Mobile Data Caching[2]

In Mobile computing environment data caching is for improving data availability and access latencies particularly because these computing environments are characterized by narrow bandwidth wireless links and frequent disconnections. Cache replacement policy plays a vital role to improve the performance in a cached mobile environment, since the amount of data stored in a client cache is small.

In this paper author reviewed some cache replacement policies proposed for mobile data caches. We made a comparison between these policies after classifying them based on the criteria used for evicting documents. In addition, this paper suggests some alternative techniques for cache replacement.

C. Big Data Meets Telcos: A Proactive Caching Perspective [3]

Depending on the rating density and storage size, the numerical results showed that several caching gains are possible in terms of users' request satisfactions and backhaul off loadings. An interesting future direction of this work is to conduct a more detailed characterization of the traffic which captures different spatio-temporal content access patterns. In order to estimate the content access patterns for cache decision, the development of novel machine learning algorithms is yet another interesting direction. Finally, design of new deterministic/randomized cache decision algorithms are required and should not be purely based on content popularity and storing most popular contents, so that higher backhaul offloading can be achieved while satisfying users' requests.

D. Big Data Caching for Networking: Moving from Cloud to Edge [4]

In this article, author introduced a proactive caching architecture for 5G wireless networks by processing a huge amount of available data on a big data platform and leveraging machine learning tools for content popularity prediction. Additionally, relying on this prediction and using extracted traffic information from this data, the gains of caching have been investigated throughout numerical studies. One possible direction of this work is to investigate the proposed big data analysis framework in a real-time fashion. For this, recent frameworks that exist in the Eclipse eco-system, such as Apache Spark and its built-in libraries Spark Streaming for real-time data processing and MLlib for machine learning libraries, are of interest.

E. Big Data Analytics in Mobile Cellular Networks [5]

Big data analytics will be an indispensable part of the mobile cellular operators' consideration of network operation, business deployment, and even the design of the next-generation mobile cellular network architectures. In this paper, the connection between big data analystics and mobile cellular networks has been systematically explored. Here author provided a broad overview of big data analytics based on radom matrix theory. Next, an architectural framework for the applications of big data analytics in cellular networks was presented. Moreover, several illustrative examples were provided.

In addition author also provide some illustrative case studies including different types of data like signaling data, traffic data, location data etc.

F. JeCache: Just-Enough Data Caching for Just-in-Time Prefetching in Big Data Applications [6]

In this paper, author propose a novel just-enough big data caching scheme for just-in-time block prefetching to improve the cache effectiveness of big data clusters. With just-in-time block prefetching, a block is cached in just before the task begins to process the block, rather than being cached in along with other blocks of the same dataset being processed. We monitor block accesses to measure the average processing time of data blocks, and then estimate the minimal number of blocks that should be kept in cache for a big dataset, so that the speed of data processing matches with that of data prefetching, and each upper level task can obtain its input blocks from cache just in time.

III. PROPOSED WORK

Here we proposed notion of big data gear in which we proposed four gear system for accurate large data in mobile networks.

A. Basic Architecture

Here we can present the basic architecture of the proposed system.

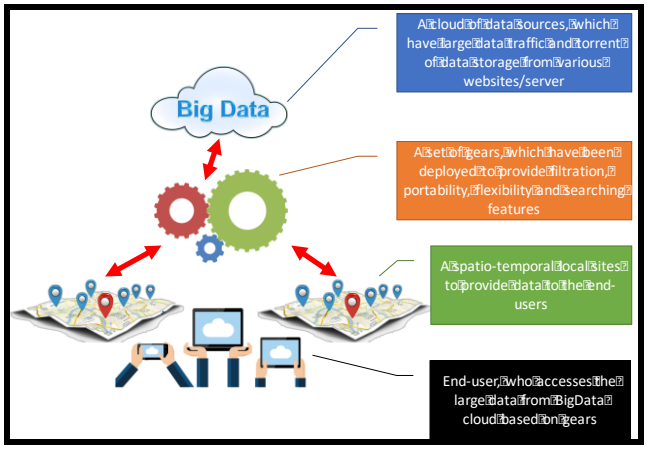

Fig 1. Basic Architecture of the system

B. Proposed Gears

1) Analytical Gear

This gear is responsible to perform analytics of the user-data retrieval based on frequency and data-size parameter. The highly-accessed data (i.e., frequent data) has been filtered and stored at the local-site, hence to reduce the Internet bandwidth and server load. For example, the recent videos related to the political event, which has high influence on the user-access. Such videos will be recognized by the G1

 ISSN: 2454-4248
417 – 420

and stored on the local site. Over a period of time, frequency analysis gives the decision that for how much time that video should be kept on local cache.

Analytical Gear also contain cache decision system (CDS) and cache replacement policy. The major work of cache decision system is to check that data should be cached or not.

a) CDS

The major work of cache decision system is to check that data should be cached or not. Indexing parameter n_d is indexing factor.

$$n_d = x1\left(\frac{s_d}{S}\right) + x2\left(\frac{T - t_d}{T}\right) + x3\left(\frac{f_d}{F}\right) + x4\left(\frac{r_d}{R}\right) + x5\left(\frac{e_d}{s_d}\right)$$

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

where, *s, t, f, r and e* are *d*'s size (in bytes), durability (i.e., freshness in seconds), access-frequency (i.e., number of accesses) , rating (i.e., number of likes) and engagement factor (i. e average data transferred to user) respectively.

While, *S, T, F* and *R* are the respective LS's largest-size data, oldest-time-stamp data, total access-frequency counts and total rating counts of all existing data in cache. Thus, the values of *S, T, F* and *R* are computed as follows:

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

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

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

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

Now, if $n_d \geq n_{LS}$, here $n_{LS} = \frac{1}{n}\sum_{i=1}^{n} n_i$ where n_i is decisive factor of *i* data item[1]

Then data item might be cached and if $n_d < n_{LS}$ then cache miss[1], If not then data will be cached into LS(Local Site).

2) Storage Gear

This gear keeps track of the availably of storage at local sites and distribute effectively the cached data on the local sites. This distribution effectively provides the data storage and reduces the redundancy at local sites, thereby improving the data availability in mobile networks.

Eq^n

$$S = \left(\frac{w1}{di} + (w2 * si)\right)$$

,where $\sum_{i=1}^{n} wi = 1$

w1,w2 are weight coefficients,
d for distance
s for storage space

3) Bandwidth Gear

This gear effectively selects the global site of data source. For example, a YouTube video may be streamed from the various location as it has been stored at various locations across the globe. Hence, G3 selects the best local site which provides the high bandwidth to the end-user, who has requested the video from YouTube cloud.

Eq^n

$$S = \left(\frac{w1}{di} + (w2 * bi)\right)$$

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

w1,w2 are weight coefficients,
d for distance
b for available speed

4) Searching Gear

This gear is designed to provide the effective search mechanism to search the local copy of data across all the cached-data at local sites. This reduced the retrieval time as well as communication time among all the local sites. This gear stores the metadata of all the local cached data for effective searching mechanism.

IV. RESULT

Here we created sixteen thousand users, 500 local site(LS) And we focused on cache missed ratio[1] and we what to check its performance so we first apply on to existing CDS equation[1] And then we applied to proposed equation and we get improved result shown below.

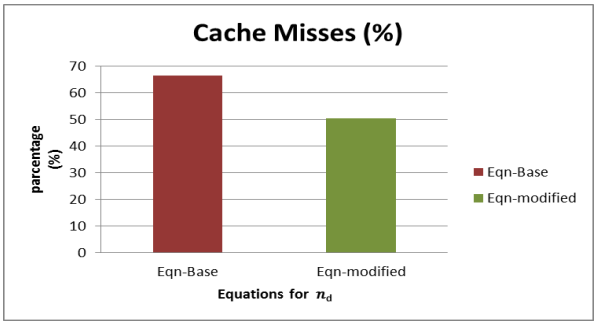

Chart-1: Cache Misses Ratio Analysis

 ISSN: 2454-4248
417 – 420

Here eqn-base[1] gives average 64% cache miss ratio where our proposed Equation gives 50% cache miss ratio , so that our proposed system gives better optimal output.

V. CONCLUSION

There are several Big Data management technique in mobile network. In this we propose the notion of gears (BigData Gears) in BigData management, to accelerate the large data access in mobile networks, which consist of the small size portable devices with mobility and low resources. Here we discussed about this new architecture and implement it ,we get optimal output. In future we can add more parameter in Cache decision system ,we can also try different caching schemas and also add more cache replacement policy.

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] 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.

[4] 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.

[5] 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.

[6] Yifeng Luo, Jia Shi, Shuigeng Zhou," JeCache: Just-Enough Data Caching for Just-in-Time Prefetching in Big Data Applications" © 2017 IEEE 37th International Conference on Distributed Computing Systems.

[7] Subramaniyaswamy , Vijayakumar , Logesh and Indragandhi ," Unstructured Data Analysis on Big Data using Map Reduce" © 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15).

[8] V.Vignesh," Predictive Analytics In Data Mining With Big Data: A Literature Survey" © International Journal of Research in Engineering and Applied Sciences (IJREAS)

[9] Julio C. L´opez, Tiankai Tu, David R. O'Hallaron," Big Wins with Small Applicationaware Caches",© 2004 0769521535/ IEEE-2004.

[10] Yinan Li ∗, Ing-Ray Chen , "Adaptive per-user per-object cache consistency management for mobile data access in wireless mesh networks" © J. Parallel Distrib. Comput. 71 (2011) 1034–1046.

[11] Minkeun Ha, Daeyoung Kim," On-demand Cache Placement Protocol for Content Delivery Sensor Networks" © 2017 International Conference on Computing, Networking and Communications (ICNC).

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