Volume 4 Issue III, March 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Product Comparisons with Improved Reliability

P.Chitra[1], A.Bhargavi[2], C.H.Lohith[3], K.Pavithra[4]
Professor, Department of Information Technology, Prathyusha Engineering College, Thiruvallur-602025

***Abstract*— Big data computing is a new critical challenge for the ICT industry . It process the large quantity of information that cannot be analyzed with traditional computing techniques. In this paper, foremost we focus on investigating the Big Data Broadcasting problem for a single source node to broadcast a big chunk of data to a set of nodes. We model the Big-data Broadcasting to build a Scalable, Efficient and Precise system for service level comparison between products in Market.**

***Index Terms*—Big data computing, cloud computing ,big data management ,distributed computing.**

I. INTRODUCTION

Big Data is the process of generating value from storage and processing of very large quantities of digital information that cannot be analyzed with traditional computing techniques. It increases storage capacities, processing power and the availability of data..Big data broadcasting used to broadcast a big chunk data from a single source node to a set of nodes . In this we model the Big-data Broadcasting to build a Scalable, Efficient and Precise system for service level comparison between products in Market. The product comparisons for users request will b given using well known TSV (Tab Separated Value) Format is used in this paper. The proposal of TSV format gives service comparison, recommendations in a tabular format based on the users request. Here we model the service level comparisons and completes the banking process from a single gateway to all service providers. With this single gateway banking process the users need not to maintain multiple E-commerce accounts. Here we model Map Reduce Algorithm. The Map Reduce algorithm is a software framework that is ideal for big data because it enables developers to write programs that can process massive Amounts of data in parallel across a distributed group of processors. The Map Function performs “filtering” and “sorting” in whatever fashion the user desires. The Reduce Function takes the output of a map function and “reduces” the list and produces the final output and hence fast searching can be done. In this paper, our application connects to the several web applications and pull all the necessary data’s to our backend .A huge Amount of data got accumulated now .These data are then Reduced by Map reduce Framework and converted into a single object .This Reduced Object Contains all the necessary information for providing comparison and Recommendations. The users can register and can login to view various products available in market. The Users are provided with neat and clean indexes and Recommendations. The Recommendations were given based on the QOS, Availability, Delivery, Offers, Price and Specifications of the particular product. So that the user can pick a best provider for a particular product. The picked products were added in Cart and can be purchased later. . The Cart can be reviewed at any time and can be purchased later whenever the Shopper Wants the Product.

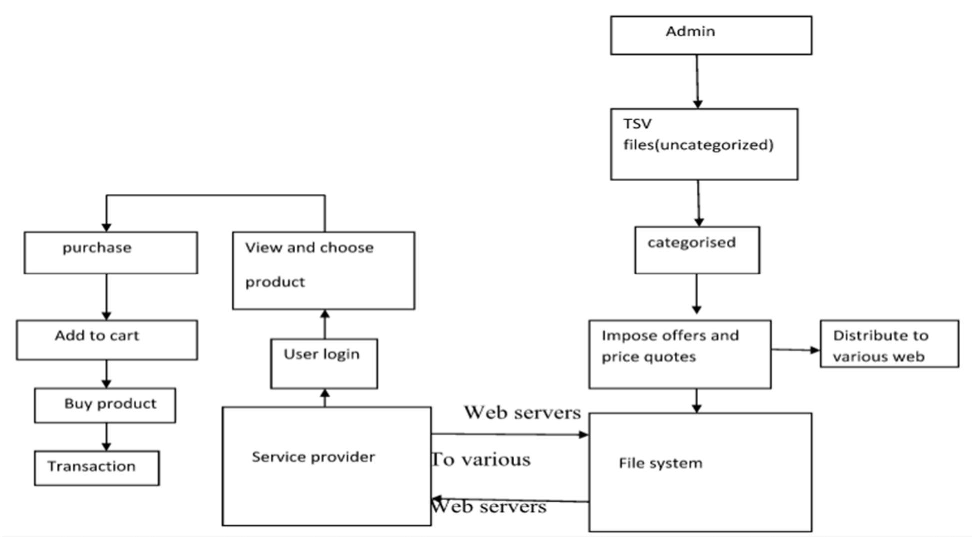

Fig I.1 Shows the architecture diagram of our application.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

II. VARIOUS WEB APPLICATIONS BUILDING AND BROADCASTING

In this module, Datasets were prepared from various web applications so that the users can compare their products with different service providers. The applications uses sample datasets that has been crawled in Amazon previously. Different datasets were prepared using

(Preprocessing(loads the data from different web servers and makes it ready for processing),clustering (Groups the data based on the category),classification(splits up the relevant resources-

Information, Features) and Distribution (distribute the data to various web servers-Broadcasting) **)** based on the users request.

III. OUR GATEWAY APPLICATION AND BATCH PROCESSING OVER THE TSV DATA

Now our gateway application is built which gives users with Recommendations and Comparisons between the Products in the Market. Generally the Resources provided by Various Web Servers are in TSV (Tab Separated Values) Format. In this module, the users login to our own application and select a product type which can be converted of text to object for information and conversion of text to object for features.

IV. WEB CRAWLING FOR RESOURCES AND MAP REDUCE

In this module, the users can register and can the view the various products available in the market using Map Reduce algorithm which pull down all the similar data to our application as a product list so that the user can pick the product whichever they desire.

V. CONCLUSION

Thus the Broadcasted data's are received using different web services. Map Reduce collects data from different web servers and gives service level comparisons and recommendations for the selected products. Recommendations for purchasing a product is given using CBR (case based recommendations) and the transaction process can be done. Hence, our applications stands unique and does not rely on the single service provider. We propose a scalable, efficient and precise system from various providers through broadcasting technique.

REFERENCES

[1] R. E. Bryant, R. H. Katz, and E. D. Lazowska, "Big-data computing: Creating revolutionary break throughs in commerce, science, and soci-ety," In Computing Research Initiatives for the 21st Century, 2008.
[2] A. Szalay and J. Gray, "2020 computing: Science in an exponential world," Nature 440, 413-414, March, 2006.
[3] G. Brumfiel, "High-energy physics: Down the petabyte highway," Nature 469, 282-283 January, 2011.
[4] J. Dean and S. Ghemawat, "Map reduce: Simplified data processing on large clusters," Proc. of Operating Systems Design and Implementation(OSDI), 2004.
[5] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Bur-rows, T. Chandra, A. Fikes, and R. E. Gruber, "Bigtable: A distributedstorage system for structured data," Proc. of Operating Systems Design and Implementation (OSDI), 2006.
[6] W. D. Hillis and G. L. Steele, Jr., "Data parallel algorithms," Communications of the ACM, vol. 29, pp. 1170–1183, December 1986.
[7] U. Rencuzogullari and S. Dwarkadas, "Dynamic adaptation to available resources for parallel computing in an autonomous network of workstations," Proc. of ACM SIGPLAN PPoPP, 2001.
[8] M. Chowdhury, M. Zaharia, J. Ma, M. I. Jordan, and I. Stoica, "Managing data transfers in computer clusters with orchestra," Proc. of ACM SIGCOMM, pp. 98–109, 2011.
[9] D. Nukarapu, B. Tang, L. Wang, and S. Lu, "Data replication in data intensive scientific applications with performance guarantee," IEEE Transactions on Parallel and Distributed Systems, aug. 2011.
[10] C. Peng, M. Kim, Z. Zhang, and H. Lei, "Vdn: Virtual machine image distribution network for cloud data centers," Proc. of IEEE International Conference on Computer Communications (INFOCOM), 2012.
[11] S.Khuller and Y.-A. Kim, "Broadcasting in heterogeneous networks,"Algorithmica, vol. 48, no. 1, Mar. 2007.
[12] J. Mundinger, R. Weber, and G. Weiss, "Optimal scheduling of peer-to-peer file dissemination," Journal of Scheduling, vol. 11, no. 2, 2008.
[13] L. Massoulie, A. Twigg, C. Gkantsidis, and P. Rodriguez, "P2p streaming capacity under node degree bound," Proc. of IEEE InternationalConference on Computer Communications (INFOCOM), 2007.
[14] O. Beaumont, L. Eyraud-Dubois, and S. K. Agrawal, "Broadcasting on large scale heterogeneous platforms under the bounded multi-port model," Proc. of IEEE International Parallel and Distributed Processing Symposium (IPDPS), 2011.
[15] S. M. Hedetniemi, S. T. Hedetniemi, and A. Liestman, "A survey of gossiping and broadcasting in communication networks," Networks,1988.
[16] P. Liu, "Broadcast scheduling optimization for heterogeneous cluster systems," J. Algorithms, vol. 42, no. 1, Jan. 2002.
[17] K. Wang, J. Li, and L. Pan, "Fast file dissemination in peer-to-peer networks with upstream bandwidth constraint," Future Generation Computer Systems, vol. 26, July 2010.
[18] K.-S. Goetzmann, T. Harks, M. Klimm, and K. Miller, "Optimal file distribution in peer-to-peer networks," Proc. of the 22nd International Symposium on Algorithms and Computation (ISAAC), 2011.
[19] M. Deshpande, N. Venkatasubramanian, and S. Mehrotra, "Heuristics for flash-dissemination in heterogenous networks," Proc. of the 13th international conference on High Performance Computing, 2006.

[20] B. Cohen, "Incentives build robustness in bittorrent," Proc. of ACM P2PECON, 2003.
[21] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "Splitstream: High-bandwidth multicast in a cooperative Environment," Proc. of ACM SOSP, 2003.
[22] D. Kosti, A. Rodriguez, J. Albrecht, and A. Vahdat, "Bullet: High bandwidth data dissemination using an overlay mesh," Proc. of ACM SOSP, 2003.
[23] C.-J. Wu, C.-Y. Li, K.-H. Yang, J.-M. Ho, and M.-S. Chen, "Time-critical data dissemination in cooperative peer-to-peer systems," Proc. of IEEE Global Telecommunications (GLOBECOM), 2009.
[24] A. Passarella, "A survey on content-centric technologies for the current internet: Cdn and p2p solutions," Computer Communications, 2012.
[25] M. A. Brown, "Traffic control how to. chapter 6. classless queuing disciplines," http://tldp.org/HOWTO/Traffic-Control-HOWTO/classless-qdiscs.html, 2006.
[26] A. Cayley, "A theorem on trees," Quarterly Journal of Mathematics,1889.
[27] A. R. Bharambe, C. Herley, and V. N. Padmanabhan, "Analyzing and improving a bittorrent networks performance mechanisms," Proc. of IEEE
[28] International Conference on Computer Communications (INFOCOM),2006.
[29] R. Thommes and M. Coates, "Bittorrent fairness: Analysis and improvements," Proc. of WITSP, December 2005.
[30] S. ul Islam, K. Stamos, J.-M. Pierson, and A. Vakali, "Utilization-aware redirection policy in cdn: A case for energy conservation," Proc. Of Information and Communication on Technology for the Fight against Global Warming, 2011.
[31] S. Saroiu, K. P. Gummadi, and S. D. Gribble, "A measurement study of peer-to-peer file sharing systems," Proc. of Multimedia Computing and Networking (MMCN), 2002.
[32] M. Banikazemi, V. Moorthy, and D. Panda, "Efficient collective communication on heterogeneous networks of workstations," Proc. Of IEEE International Conference on Parallel Processing, pp. 460–467, 1998.
[33] J. Edmonds, "Edge-disjoint branchings, in combinatorial algorithms," Algorithmics Press, 1972.
[34] G. M. Ezovski, A. Tang, and L. L. H. Andrew, "Minimizing average finish time in p2p networks," Proc. of IEEE International Conference on Computer Communications (INFOCOM), 2009.
[35] C. Chang, T. Ho, M. Effros, M. Medard, and B. Leong, "Issues in peer-to-peer networking: a coding optimization approach," Proc. of IEEE International Symposium on Network Coding (NetCod), 2010.
[36] S. Liu, R. Zhang-Shen, W. Jiang, J. Rexford, and M. Chiang, "Performance bounds for peer-assisted live streaming," Proc. of ACM SIGMET- RICS, 2008.
[37] S. Liu, M. Chen, S. Sengupta, M. Chiang, J. Li, and P. A. Chou, "P2p streaming capacity under node degree bound," Proc. of IEEE ICDCS,2010.