International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 17 Issue 2 – SEPTEMBER 2015.

BIG DATA PROCESSING WITH GEO-DISTRIBUTED DATA CENTERS

P.Elamathi[#1], M.Jothilakshmi[*2]

[#1]*M.phil, Research Scholar, Vivekananda Arts and Science College for women*
[*2]*Assistant professor, Department of Computer Science And Application, Vivekananda Arts and Science College for women*

***Abstract*--In this paper, we jointly study the data placement, task assignment, data center resizing and routing to minimize the overall operational cost in large-scale geo-distributed data centers for big data applications. We are motivated to study the cost minimization problem via a joint optimization of these three factors for big data services in geo-distributed data centers novel system for monitoring and analyzing large-scale network traffic data. To describe the task completion time with the consideration of both data transmission and computation, we propose a HADOOP two-dimensional Markov chain and derive the average task completion time in closed-form data security. The flexibility in the data block placement policy to increase energy efficiency in data centers and propose a scheduling algorithm, which takes into account energy efficiency in addition to fairness and data locality properties. Furthermore, we model the problem as a mixed-integer non-linear programming (MINLP) and propose an efficient solution to linearize it. The high efficiency of our proposal is validated by extensive simulation based studies.**
Keyword: Big data, Hadoop, Network Traffic analysis, Scheduling algorithm, Mixed-integer non-linear programming (MINLP)

I. INTRODUCTION

Disruption Tolerant Networks (DTNs) consists of mobile nodes. It exploits the intermittent connectivity between mobile nodes to transfer data. Two nodes exchange data when they move into the transmission range of each other, due to lack of consistent connectivity. Thus DTNs utilize the contact opportunity for data forwarding with "Store-Carry-and-Forward"; i.e when a node receives some packets it stores these packets into its buffer, it will carry those packets until it contact another node, and then forwards the packets to the node (fig 1). It provides hop-by-hop. If the next contacted node doesn't receive packets, it can easily retain those packets from its previous node. It is a main advantage of DTNs by using this method called Store-Carry-and-Forward. Because of mobility the duration of contact may be short. Due to wasted transmission (bandwidth), mobile node has limited buffer space.

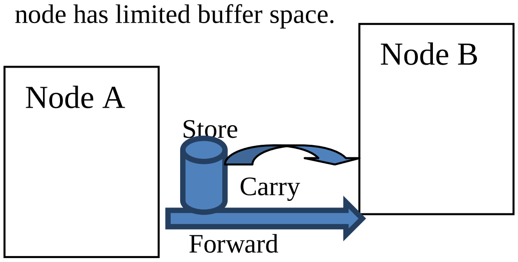

Fig 1: DTNs Store-Carry-and-Forward.

There is a limitation in bandwidth and buffer space. So DTNs are vulnerable to flood attacks. In order to collapse the network, the selfishly motivated attackers inject many packets or the replicas of the same packets as possible in to the networks. There are two types of flood attacks: Packet flood attacks and Replica flood attacks. Flooded packets and replicas wasted the bandwidth and buffer space and this attack will degrade the network service provided to good nodes.

Moreover mobile nodes spend much energy on transmitting/receiving flooded packets or replicas which may shorten their battery life. Therefore, it is important to secure DTNs against flood attacks. Rate Limit is used to prevent against flood attacks. In this approach each node has a limit over a number of packets that it has sent by a node, can send into the network in each interval time. Each node has a limit over a number of replicas that it can generate for each packet. If a node has break up its rate limit, it will be detected and data traffic will be detected. In this way the flood attacks can be avoided in DTNs. The main aim is to detect if a node has violated its rate limit. In the internet and telecommunication network has egress router and base station can account each user's traffic, so it is easy to find, if there is violation in its rate limit. Hence, it is challenging in DTNs due to lack of communication infrastructure and consistent (constant) connectivity. Even though it provides opportunistic contacts [nodes can communicate directly with each other]. Since a node moves around and may send data

To count the number of packets sent by a node, use a method called Claim-Carry-and-Check. Each node itself count the number of packets or replicas that it has sent out, and claims the count to another node, the receiving nodes carry the claims when they contact and cross-check if these claims are constant. If the attacker floods more packets or replicas than its rate limit, it has to use the same count in more than one

II. RELATED WORKS

This scheme bears some similarity with previous approaches [1] that detect node clone attack in sensor networks. To detect the attacker, both on the relay of identification of some kind of inconsistency. However, that approach has consistent connectivity which is unavailable in DTNs. Long delays of detections are also not handled. Wormhole attacks [2] are severe threat to normal network operation; it is detected by using forbidden topology. A malicious node records the packets at one place and channels into another colluding node, which replays them locally into the network. The blackhole attack [3] in which malicious node forge routing metrics to attract the packets and drop all received packets. Message delivery is in sparse Mobile Ad-hoc Networks (MANETs) is difficult due to the fact that the network is rarely connected, here ego networks can be used [4]. When the sending and receiving nodes have low connectivity and routing outperforms PROPHET routing. There are several other reasons to avoid authentication schemes [5] for DTNs. Such mechanisms imply administrative registration and key distribution ahead of deployment; however, DTNs can span hundreds of miles and many administrative domains, having a common or cooperative administrative authority for all users is unwieldy. *MobiCent*, [6] a credit-based incentive system for DTN. It allows the underlying routing protocol to discover the most efficient paths, it is also incentive compatible. Therefore, by using MobiCent, rational nodes will not purposely waste transfer opportunity or cheat by creating non-existing contacts to increase their rewards. Also introduced a scheme to detect resource misuse attack detection [7] in DTNs. If there any deviation in the expected behavior, it should noticed by the DTNs to detect an attack. A few recent works also address security issues in DTNs.

III. DEFENDING SCHEMES AGA -NST FLOOD ATTACK DETEC -TION IN DTNS

3.1 Network Model

to any contacted node, so it is very difficult to count the number of packets or replicas sent out by this node.

claim according to pigeonhole principle [the n number of items are put into m number of boxes with n>m] and this inconsistent may lead to detection. When the attackers send the packets or replicas within the rate limit private key should be generated by Trusted Authority (TA). Depends on the packet count key should be generated. TA generates both private key and rate limit certificate.

The contact duration will be short in DTNs, so a large data item is usually splits into smaller number of packets to facilitate data transfer. All packets have predefined size. It is impractical to allow unlimited delays in DTNs because the allowed delay of packet delivery is usually long. Each packet has a lifetime, the packet become meaningless after the lifetime ends and it will be discarded.

3.2 Setting the Rate Limit (L)

Request approve Style is used to set the rate limit. When the user is ready transfer their packets into the network, request for rate limit to a Trusted Authority (**TA**). Network operator is acted as TA. In this request the user specifies the appropriate value of L based on the prediction of the traffic demand. If TA approves the user request, depending upon the packet size rate limit certificate should be issued which can be used by the user to prove the legitimacy of the rate limit. TA send RL certificate to each node. This certificate includes node ID, its approved rate limit L, validation time of certificate and TA's signature.

3.3 Basic idea: Claim-Carry-and-Check

3.3.1 Packet Flood Detection

To detect the attacker violates the rate limit L, so count the unique packets that each node as a source has generated and sent out in current interval. Since the node can contact any time and place the packets, no other way is to monitor all sending activities. To address this challenge, the node itself counts the packets that it has transferred through the network by using this method. So it is easy to identify if the packets violates its rate limit. If it is greater than the real value there is a clear indication of an attack. The claimed count must have been used by the attacker in another claim which is guaranteed by pigeonhole principle.

3.3.2 Replica Flood Detection

It is used to detect that the attacker forwards a buffered packet more times than its limit L. When the source node or intermediate hop transmits the packet to its next hop, it claims the transmission count which means the number of times the packet has been transmitted; it includes the current transmission count.

If the node is a source, the next hop can know the nodes rate limit L for the packet to ensure that the claimed count is in correct range. Thus, if an attacker transmits more than l times, it must claim a false count and clear indication of an attack in DTNs as used in packet flood.

3.4 Claim Construction

Two pieces of metadata are added to each packet. Packet count claim (P-Claim) and Transmission count claim (T-Claim) are used to detect packet and replica flood attacks

3.4.1 P-Claim

P-Claim is added by the source and transmitted to later hops along with the packet. When the contacted node receives the packet, it verifies the signature of P-Claim and checks the value of packet count (**C_p**). If C_p is larger than the rate limit it discards the packet, otherwise it stores as P-Claim

3.4.2 T-Claim

It is generated and processed hop-by-hop. There is a sequence increment in T-Claim (1,2..) it also includes the current transmission count. When the packets transmitted from one hop to another hop (node) it will increase its T-Claim's count.

If there is any inconsistency in both claims, there is a clear indication of an attack. Here sampling is used to reduce the communication cost by exchanging both claims and also to increase the probability of attack detection redirection is used. Both sampling and redirection is used for P-Claim and T-Claim to detect probabilistically in the network. T-Claim should count from starting node and increment continuously.

3.5 Private Key Generation

If the attacker send the packets within the rate limit there is no indication of attack. If the packets transmits within the rate limit in the network, private key should be generated by Trusted Authority (TA).Depending upon the packet count TA will generate the private key. If the node transmits the packet in the network, it can able to verify and match the key value; because each node has a private key value. The attacker cannot able to identify the private key. For additional security purpose the private key should be generated

IV. CONCLUSION

In this paper, rate limiting to mitigate flood attacks in DTNs and also identify the attackers who send the packets within the rate limit by generating private key. And also use Claim-Carry-and-Check to probabilistically detect the violation of rate limit in DTNs environments. Efficient constructions to keep communication, computation, storage cost low. These schemes are effective to detect flood attacks and it achieves such effectiveness in an efficient way. Thus this schemes works in distributed manner, which well fits the environments of DTNs. Besides, it can tolerate a small number of attackers to colluded.

REFERENCES

[1] B. Parno, A. Perrig, and V. Gligor, "Distributed Detection of Node Replication Attacks in Sensor Networks," Proc. IEEE Symp. Security and Privacy, 2005.

[2] Y. Ren, M.C. Chuah, J. Yang, and Y. Chen, "Detecting Wormhole Attacks in Delay Tolerant Networks," IEEE Wireless Comm. Magazine, vol. 17, no. 5, pp. 36-42, Oct. 2010.

[3] F. Li, A. Srinivasan, and J. Wu, "Thwarting Blackhole Attacks in Disruption -Tolerant Networks Using Encounter Tickets," Proc. IEEE INFO COM, 2009.

[4] E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," Proc. MobiHoc, pp. 32- 40, 2007.

[5] Disruption-Tolerant Networks without Authentication," Proc. ACM MobiHoc J. Burgess, G.D. Bissias, M. Corner, and B.N. Levine, "Surviving Attacks on, 2007.

[6] B. Chen and C. Choon, "Mobicent: A Credit-Based Incentive System for Disruption Tolerant Network," Proc. IEEE INFOCOM, 2010.

[7] V. Natarajan, Y. Yang, and S. Zhu, "Resource-Misuse Attack Detection in Delay-Tolerant Networks," Proc. Int'l Performance Computing and Comm. Conf. (IPCCC), 2011.