Journal of Computer Applications ISSN: 0974 – 1925,
Volume-5, Issue EICA2012-4, February 10, 2012

A Review on Intelligent Deployment of Nodes for Improving the Performance of Overlay Networks

Muthuselvi S
Final Year Msc.,
Software Engineering,
Sathyabama University,
Chennai 119
E-Mail:muthuselvichitra@gmail.com

Prayla Shyry S
Faculty, Dept of CSE,
Sathyabama University,
Chennai 119

Ramachandran V
Professor,
Dept. of IST,
Anna University,
Guindy

Emalda Roslin S
Faculty,
Dept. of Electronics and
Telecommunication Engineering,
Sathyabama University,
Chennai 119

Nandhitha N.M
Faculty,
Dept. of Electronics and
Telecommunication Engineering,
Sathyabama University,
Chennai 119
E-Mail:nandhi_n_m@yahoo.co.in

Abstract - Overlay network is a network built on top of one or more existing networks. It adds an additional layer of virtualization to the network. Overlays enhance the performance of the network. They provide various services to the underlying network for efficient resource utilization. Overlay nodes present in the network can overcome path outages in a short duration and can also bypass congestion if necessary. However theoretically if overlay nodes are substituted for every node in the network, it complicates the performance of the network. Hence optimized number of overlay nodes must be deployed. This paper reviews various available techniques for deploying overlay nodes and the impact of deployment of overlay nodes in the network by various methods followed by different researchers. From the literatures surveyed, it is clear that many static approaches are already available for the deployment of overlay nodes. However the performance of the network can be improved by dynamic deployment of overlay nodes using intelligent methods.

Keywords: *overlay nodes, deployment, fuzzy, genetic algorithm.*

I. INTRODUCTION

Of the various issues that are to be considered while deploying the overlay services for large networks, the following real world conditions are more important. Connectivity management is a critical issue in managing overlay networks. It is because extending the existing overlays requires re-wiring of the network to cope with the changing conditions. Routing research aims at avoiding the inefficiencies in the network level routing. This leads to allowing the hosts to choose either source routing (the host choosing the network path themselves) or overlay routing networks (detouring the path). There are theoretical results that show selfish routing can result in suboptimal system behavior. Optimal routing has many constrains of deployment in terms of hardware, network connectivity and human effort. It is a very critical criterion about providing high level of path diversity while designing an overlay network. This diversity of the paths highly depends on how the relays (*i.e* Overlay nodes) are placed in the network.

Infrastructure overlay nodes is a good solution in order to satisfy the QOS requirements of applications like teleconferencing and chat room. It is a complicated job to correctly place the infrastructure overlay nodes to balance the tradeoff between performance and resource constraints.

All the limitations in the existing overlay nodes can be overcome by intelligent deployment of these nodes. This paper reviews various methods that are already available for the deployment of overlay nodes. Extensive research is carried out in this area, to find out the actual locations for the deployment of overlay nodes. In this paper, section II deals with various methods that are used for deployment of overlay nodes. Different techniques for improving the performance of overlay networks are summarized in section III and section IV concludes the entire survey.

II. OVERVIEW OF INTELLIGENT DEPLOYMENT OF OVERLAY NODES

Sabyasachi et al (2007) investigated efficient approaches to perform intelligent placement of overlay nodes for improving routing reliability using Single Overlay Source Routing (SOSR) and improving TCP throughput using Shortened Loop Overlay Transport (SLOT). To compare the reliability characteristics of overlay routes, they developed two objective functions for SOSR and SLOT. They developed incremental placement algorithms and found that a hybrid approach combined with a greedy and random approach gave a tradeoff between performance and computational complexity. They also observed that intelligent placement algorithms outperform a random heuristic by varying traffic load.

Jungher Han et al (2005) proposed a random architecture for topology aware overlay networks to maximize the path diversity without degrading the network performance. They first proposed various heuristics for the placement of overlay nodes based on extensive data collected from 232 points in 10 ISPs, and 100 Planet lab nodes. They deployed the overlay nodes using off-line topology analysis rather than deploying it randomly. It was believed that this off-line topology analysis gives good guidance about the underlying topology to the administrators to explicitly choose the overlay nodes. They showed that the clustering-based deployment reduced the number of overlays required in the network. They also kept a high level of availability and performance.

Bui et al (2008) proposed a novel approach to maximize the diversity of the overlay paths to reduce the number of relays required. They applied a novel Genetic algorithm for optimal relay placement. This algorithm was developed by hybridizing the standard GA and k-Shortest path algorithms. The network path diversity was defined as follows

Where ||K|| denotes the total number of paths *i.e* the size of K and is the scaling constant chosen to regulate the objective function to balance the utilization of the network. was maximized with a intend to obtain an overlay configuration, in which each link is shared by a minimum number of paths in average and a fair link utilization is obtained. Roy et al (2009) proposed approaches to perform intelligent placement of overlays nodes in order to achieve two main goals. The first was to attain resilient routing using SOSR and the second was to improve the performance of the TCP using SLOT. They developed objective functions to accurately capture the behavior of the application in terms of reliability and TCP performance.

Then they developed incremental placement algorithms for the applications and demonstrated their effectiveness using Simulation and Planet lab experiments. They observed that intelligent placement algorithms of infrastructure overlay nodes significantly outperform the random heuristics. They also observed a hybrid approach that the combination of a greedy and a random approach gave a good tradeoff between the performance and the computational complexity.

Srinivasan et al (2006) proposed an algorithm called "RouteSeer". This algorithm was to solve the overlay node placement problem. They initially split the problem into two parts. The first part was to place some overlay nodes called Client Proxies "close" to the clients of the overlay service. In the second part, the local routing tables available at the client proxies was used to decide the locations of the intermediate overlay nodes in order to provide non-overlapping network paths between the client proxies. This placement of intermediate overlay nodes were only based on the routing tables at the client proxies and do not require any global information.

This "RouteSeer" algorithm was able to better results than the previous schemes by 50-100%.
HE et al (2010) used k-core decomposition for intelligent placement of overlay nodes. This decomposition was based on recursive pruning of the least connected vertices, to obtain the node coreness of the overlay placement. Using that strategy, they developed a heuristic algorithm that can achieve intelligent placement of overlay nodes which in turn facilitates the TCP performance improvement. They have also demonstrated the effectiveness of the placement algorithm.

Yang et al (2007) formulated the problem of optimal proxy placement using Integer-Linear Programming (ILP) which was NP-Hard when the number of proxies was limited. After this, they quantified the benefits from using proxies in six-real world networks. They observed that out of six networks, only two got benefited from the proxies. Then they discovered a diminishing gain from the proxy usage for reducing end-to-end delays. In particular, only two to three proxies were sufficient to realize most of the performance gain.

Kamal et al (2011) introduced intelligent deployment of overlay based on network coding protection scheme in contrast to the conventional single and multiple link failures. They made sure that in a connection, each node receives two copies of the same data unit. One copy was given to the working circuit and the second copy was extracted from the linear combination of data units transmitted on a shared protection path. This guaranteed instantaneous recovery of data units upon a failure of working circuit. This strategy was implemented in the overlay layer, which makes its deployment simple and scalable. The sharing of the protection circuit by a number of connections was the key to the reduction of the cost of protection.

Fujita et al (2005) proposed a scalable overlay network deployment scheme in order to minimize the impact of membership changes in collaborative groups. In their scheme the IPsec policy required for delivering packets to a destination node was resolved by an on-demand basis to eliminate the advertisement based updates of the membership changes. They also provided two modes of overlay topology operation that addresses the dynamic changes in the number of nodes. The first mode called "mesh mode" eliminates the initiation/teardown behavior for membership changes where as the second mode known as "graph mode" creates a graph structured topology reconfigurable with a constant number of initiated/torn-down tunnels for node joins/leaves. An efficient performance increasing the number of nodes was shown. They also showed that their topology reconfiguration algorithm providing a smaller number of initiated/torn-down tunnels for changes in the number of nodes when compared with the previous approaches.

Enachescu et al (2008) addressed the placement of relay service agents (routelets) in the internet in order to assist multipath transport protocols for efficient data transmission. They identified three routelet deployment scenarios providing Linear Programming (LP) formulations for each of the scenarios. With that they proved that placement problem was Nondeterministic Polynomial (NP) hard in two of the scenarios. They then provided "rounding algorithms" for a subset of the scenarios and compared their network performance using simulations on several Boston University Representative Internet topology gEnerator (BRITE) topologies for varying values. It was observed that their rounding algorithm lead to 2-5 fold increase in bandwidth utilization when compared to single path routing.

Barford et al (2002) proposed methods for generating Autonomous System (AS) clusters based on information obtained from Border Gateway Protocol (BGP) routing tables. This method lead to iterative addition of small clusters into large clusters which in turn was based on minimizing the hamming distance between the neighbor sets of the clusters. This was done in order to identify regions of client demand, with a motto off placing caches close to groups of clients and then routing client requests to the nearest cache. The algorithm for creating a forest of AS numbers objectively discovered AS's that formed highly interconnected backbone for internet. This forest representation was considered as the idealization of the internet's true structure. The major advantage of AS clustering method was that it naturally lended itself for cache placement problem. They addressed two cache placement algorithm based on tree graph of demand. They then compared the effectiveness of their algorithm to two incremental techniques using a commercial web log.

Gast et al (2002) addressed the problem of effective resource deployment that use client clustering at the Autonomous System (AS) level. They argue that the deployment at the AS level was sufficient and appropriate for most of the web services. The major advantage of AS level resource deployment is that it was very efficient and consideration of that problem at route level made intractable topology and resource placement algorithms. The mechanism that was used starts with an algorithm which finds the centroid of highly connected nodes at the backbone of the internet. It then proceeds with trees being grown based on BGP information. Those links are learned from the trace route results. The entire mechanism was suitable for use operationally. The resulting AS forest was computationally tractable for use in service placement.

Cohen et al (2011) proposed an optimization problem to find a minimum set of overlay nodes in order to satisfy the routing properties. They showed that it was NP-hard and derived a non-trivial approximation algorithm for it. They examined practical aspects by evaluating the gain and got over two real scenarios. The first one was BGP routing. They showed that less number of relays (*i.e* less than 100) servers was sufficient to enable routing. They also demonstrated that their scheme for TCP Performance improvement resulted in an almost optimal placement of overlay nodes.

Cronin et al (2002) studied the performance improvements in client download time and server load in large networks. Mirrors of popular content are usually replicated on every site in order to maximize reach ability to clients. This was done by increased number of mirrors under different placement algorithm. Their results showed that increased number of mirrors under the constraint was effective in reducing the client download time and server load only for a small range of values regardless of the mirror placement algorithm.

III. OVERVIEW OF TECHNIQUES FOR IMPROVING THE PERFORMANCE OF OVERLAY NETWORKS

Smaragdakis et al (2008) designed a distributed overlay routing system that were implemented, deployed and evaluated on Planet Lab to solve the problem in two different perspectives in a unified manner. This aims at providing practical heuristics for specific applications designed for working well in the real deployments, and providing an idea about the underlying problem that are tractable, particularly via game-theoretic analysis. They demonstrated that EGOIST's neighbor selection primitives done by SNS (Selfish Neighbor Selection) which in turn used Best Response (BR) neighbor selection strategy significantly outperformed the existing heuristics in terms of delay, available bandwidth and node utilization. They have also demonstrated that EGOIST showed minimal overhead and how it could be used as a building block for efficient routing in overlay applications.

Qui et al (2003) disproved the theoretical worst cases of selfish routing using a game-theoretic approach to investigate the performance of selfish routing in Internet-like environment based on realistic topologies and traffic demands. In contrast to the worst cases of selfish routing, it achieved minimal latency that is close to optimal average latency in Internet like environments. This performance improvement in selfish routing comes out when there is an increased congestion on certain links.

Rowstron et al (2004) proposed a novel defense that prevents eclipse attacks. In an eclipse attack, a modest number of malicious nodes make a secret agreement in order to fool the correct nodes into adopting the malicious nodes as their peers, with the goal of dominating all the correct nodes of the neighbor sets. If successful, an eclipse attack enables the attacker to effectively "eclipse" correct nodes from each other's view mediating most of the overlay traffic. To the extreme an eclipse attack allows the attacker to control all the overlay traffic, resulting in arbitrary denial of service. They use anonymous auditing in order to bound the degree of overlay nodes. The proposed defense as compatible with both structured and unstructured overlays and also it allowed optimizations like proximity neighbor selection.

IV. CONCLUSION

Since overlay nodes present in the network provide considerable services to all the peers of the network, it is very much important to concentrate in the correct positioning of the relays in the network. From the various literatures, it is also clear that there is a huge impact of overlay node deployment in the performance of the network. Though there are many techniques that are carried out in order to solve this issue, each of which outperforms the existing method in an efficient way, it is found that usage of soft computing tool for intelligent dynamic deployment of overlay nodes will result in satisfying performance.

REFERENCES

1. Georgios Smaragdakis, Nikolaos Laoutaris, Vassilis Lekakis, Azer Bestavros, John W.Byers, Mema Roussopoulos,"EGOIST: Overlay Routing using Selfish Neighbor Selection",CoNEXT'08.
2. Lili Qui, Yang Richard Yang, Yin Zhang and Scott Shenkar-"On Selfish Routing in Internet Like Environments"- ACMSIGCOMM 2003,pp-151-162.

3. S.Prayla Shyry and V.Ramachandran,"Minimizing Link Utilization in selfish Overlay Routing",IJCCT 2010,pp.314-318.
4. Sabyasachi Roy, Himabindu Pucha, Zheng Zhang1, Y. Charlie Hu, and Lili Qiu,"Overlay Node Placement: Analysis, Algorithms and Impact on Applications",IEEE, ICDCS'07,pp.43-53.
5. Junghee Han, David Watson and Faenam Jahanian,"Topology Aware Overlay Networks",IEEE 2005- pp-2554-2565.
6. Vinh Bui, Weiping Zher and Lam Thu Bai,"Optimal Relay Placement for Maximizing Path Diversity",IEEE GLOBECOM 2008,pp.1-6.
7. Sabyasachi Roy,Himabindu Pucha,Zheng Zhang,Y.Charlie Hu, Lili Qiu,"On the Placement of Infrastructure Overlay Nodes"IEEE/ACM 2009,pp. 1298 – 1311.
8. Sridhar Srinivasan and Ellen Zegura,"Routeseer: Topological Placement of Nodes in Service",GIT-CC-06-03 2006.
9. ZhengHe,Yu Heng cao,Xiao,YanMa,"Overlay node deployment based on node coreness", The journal of chinese universities of posts & telecommunication 2010,vol 17, pp.99-103.
10. Shaoyu Yang, Yoo-Ah Kim, Bing Wang,"Designing Infrastructure Based Overlay Networks for Delay Sensitive Group Communication",GLOBECOM'07,pp.567-570.
11. Kamal A.E, Ramamoorthy.A, Long Long, Shizheng Li, "Overlay Protection against Link Failures using Network Coding",IEEE/ACM 2011,pp.1071-1084.
12. Fujita.N, Ishikawa.y, Koide.T,Tsukamoto,"Scalable Overlay Network Deployment for Dynamic Collaborative Groups",TechPark'05,pp.102-109.
13. Yong Zhu, Constantinos Davrolis, Mostafa Ammar-"Dynamic Overlay Routing Based on Available Bandwidth Estimation: A Simulation Study",International Journal of Computer and Telecommunications Networking,Overlay Distribution Structures and their Applications 2006.
14. Fraiwan.M and G.Manimaran,"On the Design of Overlay Networks for IP Links for Fault Verification",GLOBECOM'08,pp.1-5.
15. Frey.D, Murphy.A.L,"Failure-Tolerant Overlay Trees for Large Scale Dynamic Networks",Peer-To-Peer Computing'08-pp-351-361.
16. Enachescu.M, Kokku.R,"Routlet Placement for Multipath Transport",BROADNETS 2008,pp.657-664.
17. Paul Barford , Jin-yi Cai , Jim Gast," Cache Placement Methods Based on Client Demand Clustering",IEEE INFOCOM 2002.
18. Paul Barford , Jin-yi Cai , Jim Gast," Cache Placement Methods Based on Client Demand Clustering",IEEE INFOCOM 2002.
19. Jim Gast , Paul Barford," Resource deployment based on autonomous system clustering",IEEE GLOBE INTERNET 2002,vol.3,pp. 2240 – 2244.
20. Atul Singh, Miguel Castro, Peter Druschel, Antony,"Rowstron Defending against Eclipse attacks on overlay networks",ACM SIGOPS.
21. Cronin.E, Jamin.S,Cheng Jin,Kurc, A.R.Raz.D, Shavitt.Y,"Constrained mirror placement on the Internet",IEEE 2002, vol 1, pp.31-40.
22. Cohen. R, Raz.D,"Cost Effective Resource Allocation of Overlay Routing Relay Nodes" , IEEE INFOCOM -2011, pp. 3236 - 3244.