Routing on Overlay Networks: Developments and Challenges

Adrian Popescu

Dept. of Telecommunication Systems
School of Engineering
Blekinge Institute of Technology
371 79 Karlskrona, Sweden
adrian.popescu@bth.se

Overlay networks are networks operating on the inter-domain level, where the edge hosts learn of each other and, based on knowledge of underlying network performance, they form loosely coupled neighboring relationships. These relationships can be used to induce a specific graph, where nodes are representing hosts and edges are representing neighboring relationships. Graph abstraction and the associated graph theory can be further used to formulate routing algorithms on overlay networks. The main advantage of overlay networks is that they offer the possibility to augment the IP routing as well as the QoS functionality offered by the Internet.

One can state that, generally, every Peer-to-Peer (P2P) network has an overlay network at the core, which is mostly based on TCP or HTTP connections. Because of the abstraction offered by the TCP/IP protocol stack at the application layer, the overlay and the physical network can be completely separated from each other as the overlay connections do not reflect the physical connections.

Overlay networks allow designers to develop own routing and packet management algorithms atop of the Internet. Overlay networks can therefore be used to deploy new protocols and functionality atop of IP routers without the need to upgrade the routers. New services can be easily developed, with their own routing algorithms and policies.

Actually, there are two general classes of overlay networks, i.e., routing overlays and storage and lookup overlays. Routing overlays operate on the inter-domain IP level and are used to enhance the Border Gateway Protocol (BGP) routing and to provide new functionality or improved service, e.g., as reported in [2]. However, the overlay nodes operate, with respect to each other, as if they are belonging to the same domain on the overlay level. QoS guarantees can be provided as well.

On the other hand, storage and lookup overlays have the focus on techniques to use the power of large, distributed collections of machines, e.g., Chord [15]. These overlays are actually used to support a number of projects on large distributed systems.

I. Overlay routing

A number of research activities are done today on overlay routing as well as on resource discovery, load balancing and security to find optimal solutions with reference to QoS provisioning [8], [11], [16].

Strategies for overlay routing describe the process of path computation to provide traffic forwarding with soft QoS guarantees at the application layer. There are typically three fundamental ways to do routing. These are source routing, flat (or distributed) routing and hierarchical routing. Source routing means that nodes are required to keep global state information and, based on that, a feasible path is computed at every source node. Distributed routing relies on a similar concept but with the difference that path computation is done in a distributed fashion. This may however create problems like distributed state snapshots, deadlock and loop occurrence. Better routing algorithms are using flooding but at the price of large volumes of traffic generated. Finally, hierarchical routing is based on aggregated state maintained at each node and the routing is done in a hierarchical way. The main problem in this case is related to imprecise states.

There are two main categories of routing protocols, i.e., proactive protocols and reactive protocols. Proactive protocols periodically update the routing tables, independent of traffic arrivals. On the other hand, reactive protocols update the routing tables on-demand, i.e., only when routes need to be created or adjusted due to changes in routing topology or other conditions (e.g., traffic must be delivered to an unknown destination). Proactive protocols are generally better at providing QoS guarantees for real-time traffic like multimedia. The drawback lies in the traffic volume overhead generated by the protocol itself. Reactive protocols scale better, but they experience higher latency when setting up a new route.

A specific difficulty with overlay routing is related to the presence of high churn rates in P2P networks [14]. The consequence is that the topology information is very dynamic, which makes it difficult to provide hard QoS guarantees.

QoS constraints associated with each route define an optimization problem. To solve this problem, the overlay nodes have dedicated algorithms associated with a traffic flow or with a group of flows sharing common characteristics (e.g., similar QoS constraints). To solve the optimization problem each algorithm can be connected, e.g., to a Random Neural Network (RNN) to continuously adapt the existing routes according to the quality experienced by traffic flows passing the node. This can be done by Reinforcement Learning (RL) [7]. Other methods to solve the optimization problem may be applied as well, e.g., swarm intelligence [4], and genetic algorithms [6].

II. Multi-path overlay routing

IP routing protocols are forwarding data on a single-path between source and destination nodes. Single-path routing has the drawback that the achievable throughput could be limited due to many policy routing decisions existing today. BGP is primarily a policy-based routing protocol, which means that it may route a specific data flow on a path with lower bandwidth even if alternate paths with higher bandwidth are available. Furthermore, single-path routing is not performing well in wireless (ad-hoc) networks either. This is because of the relatively high route failures, due to mobility or to false failures created by interference effects.

An interesting solution is to develop multi-path overlay routers [13]. Multi-path overlay routing is an algorithm that can be deployed at the source node to stably and optimally split the data flow sent to a specific destination node. The algorithm may increase throughput, reduce latency and balance traffic loads. It may also provide robustness to link failures due to mobility and false failures that occur as a consequence of, e.g., IEEE 802.11 MAC protocol.

There are several fundamental questions that must answered about multi-path overlay routing and the associated algorithms. Some of the most important questions are as follows: how many paths are needed for the transfer of a specific amount of data?; given a specific topology, how to select the paths such as to provide the requested QoS and to balance traffic loads?; where to place the multi-path overlay routers given an existing network topology?; what is the effect of multi-path overlay routing on TCP stability and performance?; given a specific topology and a specific multi-path routing algorithm, how does one design stable TCP congestion control mechanism that exploits the multi-path routing capability?

III. Overlay routing vs BGP

Today, BGP suffers from performance problems created by increasing size and complexity of the Internet backbone [3], [9]. Increasing number of Autonomous Systems and the associated advertisements has the consequence that the routing tables are increased. Further, increased inter-domain connectivity means that the Internet topology is becoming less hierarchical due to multi-homing. Increased demand for policy-based routing has also a serious consequence in that the amount of reachability advertisements further increases. All together these factors create the situation that BGP routers need longer time (e.g., at least several minutes) to converge to a new valid route after a link or path failure. There are studies showing that inter-domain routers may even need tens of minutes to come to a consistent view of the network topology after a fault [10]. This further increases the risk for routing flaps (i.e., routing table oscillations) and instability. One can therefore state that, although the Internet is actually performing well, it is also inherently unreliable. Today's Internet is quite sensitive to router and link faults, configuration errors and malice and this has a direct impact on performance.

It is actually extremely hard to understand the dynamics of inter-domain routing and to debug routing problems [3]. It is therefore important to also focus the interest on alternative solutions like using overlay routing to bypass BGP's path selection and to improve performance and fault tolerance. Furthermore, it is important to compare the relative benefits of overlay routing with inter-domain routing as well. A number of key metrics can be considered, e.g., achievable throughput, end-to-end and round-trip delays and availability.

Two key elements can be considered for comparison, namely route availability and route selection algorithm. Route availability refers to the number of available routes whereas route selection algorithm refers to protocol complexity, performance and resilience.

It is important to study the performance of inter-domain routing, and compare it with that of overlay routing protocols. Furthermore, a very interesting question is related to what is the best architectural solution (with reference to performance) regarding the route selection algorithm itself. Does the solution with two route selection algorithms existing today (i.e., BGP and overlay routing protocol) offer acceptable good performance or maybe a solution with only one route selection algorithm at the overlay (as suggested by [5]) offers better performance? The first alternative raises also the question of coordination of routing mechanisms existing in two parallel overlays (i.e., BGP and overlay routing) to obtain the best performance like for instance in the case of IP and MPLS.

IV. Security issues

Unstructured and unadministered P2P networks like Gnutella present serious security challenges. There are generally three categories of threats that are acting at different levels, i.e., threats on the individual user, threats within the P2P network itself and threats to the Internet. For instance, one of the most serious threats acting at individual level is free riding [1]. Free riding is when users download documents and use so network resources but they do not share files and do not answer other P2P searches. This is mainly a fairness problem, as users with selfish behavior are consuming resources and deteriorating the network performance for own profit only.

Another important threat issue at the individual level is regarding the copyright-infringement concerns as well as the drive of media industry to protect proprietary content and to constrain file copying. A consequence of this could be an eventual persecution of Gnutella supernodes, which are generating the bulk of data content. This further limits the technical development of P2P and overlay networking. A possible solution could be to develop and build up anonymity on top of Gnutella, but this further raises the question of interoperability among anonymized and non-anonymized users.

Another important research issue is regarding the protection of P2P and overlay networks that are facing security attacks, e.g., Denial-of-Service (DoS) attacks. This problem is further complicated because actually many standard security mechanisms are not effective for P2P and overlay networking. This is because P2P and overlay communication protocols are more sophisticated, communication patterns are more dynamic and port selection is more random than with other applications. Moreover, accountability and privacy are not yet solved in a satisfactory manner [12].

References

[1] Adar E. and Huberman B. A., *Free Riding on Gnutella*, First Monday, Vol. 5(10), October 2000, http://firstmonday.org/issues/issue5.10/adar/index.html.

[2] Andersen D., Balakrishnan H., Kaashoek F. and Morris R., *Resilient Overlay Networks*, 18th ACM Symposium on Operating Systems Principles (SOSP), Banff, Alberta, Canada, October 2001.

[3] Chandrashekar J., Zhang Z. and Peterson H., *Fixing BGP, One AS at a Time*, ACM SIGCOMM Workshop 2004, Portland, Oregon, USA, August 2004.

[4] Di Caro G., Ducatelle F. and Gambardella L. M., *Anthocnet: An Ant-Based Hybrid Routing Algorithm for Mobile Ad-Hoc Networks*, 8th International Conference on Parallel Problem Solving from Nature, Birmingham, UK, September 2004

[5] Feamster N., Balakrishnan H., Rexford J., Shaikh A. and van der Merwe J., *The Case for Separating Routing from Routers*, ACM SIGCOMM Workshop 2004, Portland, Oregon, USA, August 2004.

[6] Gelenbe E., Gellman M., Lent R., Lei P. and Su P., *Autonomous Smart Routing for Network QoS*, First International Conference on Autonomic Computing, New York, USA, July 2004.

[7] Gelenbe E. and Lent R., *Power-Aware Ad-Hoc Cognitive Packet Networks*, Ad-Hoc Networks Journal, Vol. 2, pp. 205-216, July 2004.

[8] Gelenbe E., Lent R. Xu Z., *Design and Performance of Cognitive Packet Networks*, Performance Evaluation, No. 46, pp. 155-176, 2001.

[9] Griffin T. G. and Wilfong G., *An Analysis of BGP Convergence Properties*, ACM SIGCOMM 1999, Cambridge, MA, USA, August 1999.

[10] Labovitz C., Ahuja A., Bose A. and Jahanian F., *Delayed Internet Routing Convergence*, ACM SIGCOMM 2000, Stockholm, Sweden, September 2000.

[11] Li Z. and Mohapatra, P., *QRON: QoS-Aware Routing in Overlay Networks*, IEEE Journal on Selected Areas in Communications, Vol. 22, No. 1, January 2004.

[12] Peer-to-PeerWG, http://www.peer-to-peerwg.org

[13] Pi R. and Song J., *Multi-Path Transmission Based on Overlay Network*, 18th International Conference on Advanced Information Networking and Application AINA'04, Fukuoka, Japan, 2004.

[14] Saroiu S., Gummadi P. K. and Gribble S. D., *Measuring and Analyzing the Characteristics of Napster and Gnutella Hosts*, Multimedia Systems, Vol. 9, No. 2, pp. 170-184, August 2003.

[15] Stoica I., Morris R., Liben-Nowell D., Karger D. R., Kaashock M. E., Dabek E. and Balakrishnan H., *Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications*, IEEE/ACM Transactions on Networking, Volume 11, Number 1, February 2003.

[16] Subramanian L., Stoica I., Balakrishnan H. and Katz R. H., *OverQoS: Offering QoS Using Overlays*, ACM SIGCOMM Computer Communications Review, Vol. 33, No. 1, January 2003.