



Peer organization in structured P2P network

Mohanjeet Singh¹, D.S Dhaliwal², Neeraj Garg³

¹Research Scholar

PTU Jalandhar, Punjab, INDIA

²Bharat Group of Colleges. Sardulgarh, Punjab, INDIA

³Haryana Engineering College Jagadhri, Haryana, INDIA

Abstract: One of the most popular P2P applications on the internet is P2P media streaming. Finding nearby peer over the internet is challenging job in live media streaming because it is time constrained application and nearest neighbors have to discover as quickly as possible. Objective of media streaming is low latency, QoS and scalability. For applications like Gnutella, finding the nearest peer can reduce network load for queries and responses. Nearest-peer finding techniques can also be used to efficiently construct overlay networks. There are different schemes which can be used to organize the peers. In this paper we will discuss some of them.

Keywords: P2p, Overlay network, round-trip-time, PingPong Msg

I. Introduction

Distributed data storage and lookup services like Chord [7] and CAN [6] is more efficient when the neighbors on the overlay are near to each other on the underlying topology. Super node selection is more complex than classic dominating set and p-centers from graph theory, known to be NP-hard problems, because it must respond to dynamic joins and leaves (churn), operate among potentially malicious nodes, and function in an environment that is highly heterogeneous. We broadly define the super node selection problems that of selecting some subset of the peers in a large scale peer-to-peer overlay network to take a special role, with the designated super nodes providing service to the non-super nodes. To provide quality of services, multimedia streaming requires contribution of various intermediate peers [5]. We need new simple system which quickly and accurately find the closest peer by simple probes or by using information obtain from network distance estimation. On way to construct P2P media streaming is Overlay layer and Data scheduler layer. Overlay layer provide partner node and organize the peers for data transferring. Data scheduler layer exchange the data between peers. Designers of peer-to-peer systems are usually concerned by the match between logical overlay and physical infrastructure. We face a usual challenge of peer-to-peer simulations: as peers are expected to be spread all over the world, the simulator should conform to a model of Internet, although this mapping is known to be a real issue [1]. Some previous research assume that it would be convenient to estimate the closest peer for newcomer accurately and quickly if exact route between each peer and central server can be store by the central server itself. The reasoning behind such assumption indeed refers to the statistical regularities observed in the large-scale structure of Internet [9].

II. P2P Classification

A. Unstructured

In this type of network as the name indicates they do not maintain any specific structure on the network topology. As such, there is not much systematic information to help index objects across nodes. So nodes often locally maintain objects they share, and search in the network is more or less a blind process, passing queries around nodes to check whether they have the desired objects. Since search is done locally, complex queries like wildcards can be supported as they were in a centralized system. Search space grows in proportion to recall rate because search is a basically a blind process. [13]

B. Structured

These networks are so named because they maintain a structure or topology to maintain a network. Most Structured peer to peer operate using distributed hash tables (DHT) where given a key, they can efficiently locate an object having the key by just maintaining neighboring tables entry. Nodes cooperatively maintain routing information about how to reach all nodes in the overlay [14]. Structured overlays provide a limit on the number of messages needed to find any object in the overlay which is particularly important when searching for infrequently occurring or low popularity objects. Local routing table which is there with each peer is used by the forwarding algorithm. When the peer joins the overlay using a specified bootstrap procedure, the peer's routing table is initialized. Peers periodically exchange routing table changes as part of overlay maintenance. Peers who are neighbors in the overlay can be distant in the underlying network because the address space is virtualized and peer addresses are typically randomly assigned.

III. Issues of peers

A. Managing peer dynamically

During streaming session dynamicity is difficult. Behavior of peers is unpredictable. They can join and leave the system at any time due to their dynamic nature. Other node doesn't find any prior notification regarding these changes. To manage the changes, strong and adaptive mechanism is needed because peer leaving and entrance interrupt the smooth working. A mechanism to handle these sudden changes which happened in network must be implemented. Perform streaming in smooth manner it is essential that when any peer leave the system it must be known as early as possible.

B. Peer heterogeneity

Accurate characterization of peer heterogeneity is difficult due to huge population and disparate properties of individuals [10]. Peers are not homogeneous because they have different capability. Heterogeneity may be due to different access networks which connecting the peers, or due to difference of peer's desire regarding contribution.

C. Efficient overlay network construction:

Constructing an efficient overlay network on top of which one can efficiently provide services such as broadcasting, information gathering, routing and data searching is a big problem. To understand the underlying topology peers must be organize into logical topology. Overlay topology if not suitable can decrease the system performance and can generate extra overheads. There for overlay network construction should be scalable.

D. Best peer selection:

The selection of peers is a key issue for improving resource usage and end users Quality of Experience (QoE) [11]. For selection of sender and intermediate peers a flexible and efficient strategy must be there. In fact Overlay network is a group of connected intermediate nodes and multimedia contents go through these intermediate nodes which cause end to end delay. For providing more live multimedia contents, this delay should be less.

E. Encouragement to participating peers:

In literature, it is found that mainly peers join the network for their own benefit. They never share their own resources but share other's resources. In 2000, near about 70% of Gnutella users shared no file but they only download contents [12]. This issue let the P2P network work like client-server because no peer shares its resources but want to get the resources from others. This model will fail when number of client peer increase. To overcome this issue there must be some encouragement for peers who participate in streaming. Encouragement can be in economical form.

IV. Different scheme to select nearby peer

A. Hierarchy based scheme

The hierarchy is based on topological clustering of these peers, where nearby peers are grouped into the same cluster. The querying member (termed query-host) finds its closest peer by successively refining its search in a top down manner over this hierarchy. Tiers create a hierarchy of the application peers. In this scheme, each application peer dynamically discovers a few other application peers, and is required to make distance measurements to a subset of them. Arrange the set of application peers into a hierarchy. Logically, each peer keeps detailed state about other peers that are near in the hierarchy, and only has limited knowledge about other peers in the group. The hierarchical structure is also important for localizing the effect of peer failures. Peers that are "close" with respect to the distance metric are mapped to the same part of the hierarchy. The closest peer-finding operation proceeds stop-down on the hierarchy thus successively refining the search at each step, till the appropriate peer is identified. Hierarchical routing in the Internet offers several benefits over non-hierarchical routing, including scalability and administrative autonomy (e. g., at the level of a university, a corporate campus, or even the coverage area of a base station in a mobile network).

B. Coordinate-based schemes

Its aim is to fix the location of any host on Internet, to allow two peers to estimate their latency by a basic distance computation [9, 15, 2]. Previous works show that these virtual coordinates can be obtained by active probing, i.e. by collecting round-trip-time (RTT) measurements between peers and a small set of landmarks [3]. Unfortunately, network coordinate systems require a substantial amount of time before to deliver accurate information and do not take physical network topology into consideration. It is assume that if a central server

can store the exact route between each peer and itself, it would be able to estimate accurately and quickly the closest peers of any newcomer. The requirement of coordinate based schemes is to install a handful of globally distributed “landmark” servers. Peers are made to open TCP connections with each landmark by adding the addresses of landmarks to the first list of peers returned by the tracker. The measurements obtained from the TCP connection are used to compute the location of the peer in a centralized network coordinate system (GNP) [3]. The tracker can thus reply to any subsequent request from that peer with a list of carefully selected peers in place of the usual random choice. Each time a peer P connects to the tracker, the tracker returns a set of other peers which exhibit low latency paths to P . The client remains completely unaware of this bias and behaves as usual.

C. Physical Topology based scheme

Most traditional P2P media streaming systems don not consider the physical network among the peers, while selecting partner candidates according to physical network can reduce the latency of the data transferring and improve the quality of service. How to decide the peers which are actually close to each other in physical network? Several schemes have been proposed to estimate the internet distance. Internet Distance Maps (IDMaps) [2] places tracers at key locations in the Internet. These tracers measure the latency among themselves and advertise the measured information to the clients. Clients get the latency with the closest tracer, and can compute the distance between them without extra probing in particular way. GNP (Global Network Positioning) [2] and its sequel NPS (Network Positioning Systems) [4] are another methods for estimating the network distance. In physical topology based scheme there will be a Media Server which is the data server which pushes the data to clients as bootstrap. Index Server stores the landmark list. The client with a tree icon is landmark node who manage sends probing packets (PingPong Msg) and records the RRTs value from the landmark nodes. Once p receives the response from the landmark nodes, the latency from landmark nodes of p make up the vector. The vector is stored on landmark node. Landmark node manages the entire vector in the cluster and returns partner candidates to p according to specific algorithm

D. Landmark Nodes based scheme

Landmark nodes measure RTTs among themselves and use this information to compute the coordinates in a Cartesian space for each landmark node. New peer join the system to measure the distance with landmark nodes and compute the coordinate for itself according to the coordinates of landmark nodes. The Euclidean distance between nodes in the Cartesian space is directly used as an estimation of the network distance. Since any client who joins the system should contact Index Server firstly, Index Server records the peer information like online time, bandwidth and so on. The server selects the peer with long online time and high bandwidth as landmark nodes. In case the landmark nodes quit, there is a backup mechanism. Before landmark node quits, the node should inform Index Server to update landmark list. In order to increase the system stability, it is necessary to deploy some landmark nodes in advance. As explained in the former section, landmark node is also the leader of the cluster. Every landmark maintains a one hop route table between them in order to search the neighboring peers from one cluster to another. The route table is also stored on Index Server. If one route table changes, it should be updated to other landmark nodes.

V. Conclusion

Designers of peer-to-peer systems are usually concerned by the match between logical overlay and physical infrastructure: peers should preferentially be connected with the peers that are the closest in Internet. However determining these closest neighbors in a wide population spread on a large-scale network as Internet is still a challenge. The super node selection problem for peer-to-peer applications lies in its need to fulfill dispersal and access criteria in a way that is scalable, highly adaptive, fault tolerant, and that respects local node autonomy. Super node selection protocols must cope with an unknown, large scale, dynamically changing, and heterogeneous network. To provide quality of services, contribution of various intermediate peers is required. We need simple system which quickly and accurately find the closest peer by simple probes or by using information obtain from network distance estimation.

Acknowledgment

I would like to thanks Punjab Technical University Jalandhar, for providing me opportunity and technical support to complete this research work.

Reference

- [1] L. Dall'Asta, J. I. Alvarez-Hamelin, A. Barrat, A. Vazquez, and A. Vespignani. Exploring networks with traceroute-like probes: Theory and simulations. *Theor. Comput. Sci.*, 355(1):6{24, 2006.
- [2] T. S. E. Ng and H. Zhang, "Predicting Internet network distance with coordinates-based approaches," in *Proceedings of the IEEE INFOCOM 2002*, 2002.
- [3] Ng, T.S.E., Zhang, H.: Predicting Internet Network Distance with CoordinatesBased Approaches. In: *Proc. INFOCOM (2002)*

- [4] T. S. E. Ng and H. Zhang, "A network positioning system for the internet," in Proceedings of USENIX Annual Technical Conference, 2004.
- [5] M.Singh, D.S Dhaliwal, N.Garg. "Searching and streaming of multimedia content in p2p overlay network", International Journal of Computer Engineering & Technology Volume 3, Issue 2, August (2012), pp. 433-438.
- [6] S. Ratnaswamy, P. Francis, M. Handley, K. Karp, and S. Shenker. A scalable content-addressable network. In Proceedings of SIGCOMM, August 2001.
- [7] I. Stoica, R. Morris, D. Karger, M. Frans Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In Proceedings of SIGCOMM, August 2001.
- [8] C. Kommareddy, N. Shankar, and B. Bhattacharjee. Finding close friends on the Internet. In Proceedings of ICNP, November 2001.
- [9] T. S. E. Ng and H. Zhang, "Predicting Internet network distance with coordinates-based approaches," in Proceedings of the IEEE INFOCOM 2002, 2002.
- [10] Zhi Yang, "Exploring peer heterogeneity: Towards understanding and application", In IEEE International Conference on Peer-to-Peer Computing (P2P), 2011.
- [11] Fiorese, A," An approach to peer selection in service overlays", In IEEE 7th International Conference on Network and Service Management (CNSM) 2011
- [12] Eytan Adar and Bernardo Huberman, "Free Riding on Gnutella, First Monday". October 2000 available at http://www.firstmonday.dk/issue5_10/adar/
- [13] Yuh-JzerJoung(2007), "Wildcard Search in Structured Peer-to-Peer Networks", IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 11, pp. 1524-1540.
- [14] J. Buford, H. Yu, and E. K. Lua (2008), "P2P Networking and Applications" Morgan Kaufmann.