



EPiC Series in Computing

Volume 75, 2021, Pages 80–89

CAINE 2020. The 33rd International Conference on
Computer Applications in Industry and Engineering



Design of Efficient Broadcast Protocol for Pyramid Tree-based P2P Network Architecture

¹Indranil Roy, ²Nick Rahimi, ¹Swathi Kaluvakuri, ¹Koushik Maddali, ¹Bidyut Gupta and
³Narayan Debnath

¹School of Computing
Southern Illinois University
Carbondale, IL, USA

²Information Systems and Applied Technology
Southern Illinois University
Carbondale, IL, USA

³School of Computing and Information Technology
Eastern International University, Vietnam

{indranil.roy, shrahimi, swathi.kaluvakuri, koushik, }@siu.edu, bidyut@cs.siu.edu,
ndebnath@gmail.com

Abstract

In this paper, we have considered an existing non-DHT-based structured P2P network. It is known as pyramid tree. A node i in this tree represents a group (cluster) of peers that are interested in a particular resource of type i . It is not a conventional tree. In the present work, such a P2P architecture has been the choice because in a pyramid tree, search latency for its inter-group data lookup algorithm is bounded by the tree diameter and is independent of the total number of peers present in the system. In addition, any intra-group data lookup communication needs only one overlay hop. In the present work, we have explored some important structural properties of the tree and incorporated them to design an efficient inter cluster broadcast protocol with complexity $O(\log n)$ for complete pyramid tree architecture, where n denotes the number of nodes in the tree.

Index Terms — Structured P2P network, residue class, pyramid tree, broadcast

I. INTRODUCTION

Peer-to-Peer (P2P) overlay networks are widely used in distributed systems due to their ability to provide computational and data resource sharing capability in a scalable, self-organizing, distributed manner. P2P networks are classified into two classes: unstructured and structured ones. In unstructured systems [2] peers are organized into arbitrary topology. It takes help of flooding for data look up. Problem arising due to frequent peer joining and leaving the system, also known as churn, is handled effectively in unstructured systems. However, it compromises with the efficiency of data query and the much needed flexibility. In unstructured networks, lookups are not guaranteed. On the other hand, structured overlay networks provide deterministic bounds on data discovery. They provide scalable network overlays based on a distributed data structure which actually supports the deterministic behavior for data lookup. Recent trend in designing structured overlay architectures is the use of distributed hash tables (DHTs) [3] - [5]. Such overlay architectures can offer efficient, flexible, and robust service [3] - [5], [7], [8].

However, maintaining DHTs is a complex task and needs substantial amount of effort to handle the problem of churn. So, the major challenge facing such architectures is how to reduce this amount of effort while still providing an efficient data query service. In this direction, there exist several important works, which have considered designing hybrid systems [1], [6], [9] - [11]; these works attempt to include the advantages of both structured and unstructured architectures. However, these works have their own pros and cons.

Our Contribution

In this paper, we have considered designing an efficient broadcast protocol for the Pyramid tree architecture reported in [19]. The class of interest-based structured P2P systems has gained considerable attention [6], [12], [13], [15], [17], [18], [19], [20] from the viewpoint that users sharing common interests are likely to share similar contents, and therefore searches for a particular type of content is more efficient if peers likely to store that content type are neighbors [14]. The considered architecture is a non-DHT-based structured P2P system. We have earlier pointed out the disadvantages of DHT-based systems. A mathematical model based on modular arithmetic, specifically residue class (RC), has been used to design the Pyramid tree architecture [16].

Note that it is not a conventional tree. This tree structure found its applications in the past in the area of design for testability for VLSI chips [16]. In the present work, a node i in this tree represents a group of peers (also called as *cluster_i*) that possess a particular resource of type i . Hence, each node corresponds to a distinct resource type. In the present work, such a tree architecture has been the choice because in a complete pyramid tree, multiple paths exist between most of its nodes. Such a structural characteristic can be helpful from the viewpoint of designing communication protocols that are load balanced as well as robust. In this architecture, search latency for its inter-group data lookup algorithm is bounded by the tree diameter and is independent of the total number of the peers present in the system [19]. This is its most important advantage that makes it superior to some important existing DHT-based systems.

In the present work, we have explored some important structural properties of the tree and incorporated them to design an efficient inter cluster broadcast protocol with complexity $O(\log n)$ for complete pyramid tree architecture, where n denotes the number of nodes in the tree. The existing architecture has appeared in Section II on preliminaries. The architectural properties appear in Section III. In Section IV, a new concept ‘virtual neighbors’ has been introduced and in Section V, we present the proposed broadcast protocol.

II. PRILIMINARIES

In this section, we present some relevant results from our recent work on the Pyramid tree architecture [19] for interest-based peer-to-peer system.

Definition 1. We define a resource as a tuple $\langle R_i, V \rangle$, where R_i denotes the type of a resource and V is the value of the resource.

Note that a resource can have many values. For example, let R_i denote the resource type ‘songs’ and V denote a particular singer. Thus $\langle R_i, V \rangle$ represents songs (some or all) sung by a particular singer V .

Definition 2. Let S be the set of all peers in a peer-to-peer system. Then $S = \{P^{R_i}\}$, $0 \leq i \leq n-1$, where P^{R_i} denotes the subset consisting of all peers with the same resource type R_i , and the number of distinct resource types present in the system is n . Also for each subset P^{R_i} , we assume that P_i is the first peer among the peers in P^{R_i} to join the system. We call P_i as the group-head of group G_i (cluster _{i}) formed by the peers in the subset P^{R_i} .

We have assumed that no peer can have more than one resource type. Generalization of the architecture is not considered in this paper.

A. Pyramid Tree

The following overlay architecture has been proposed in [19].

- 1) The root of the tree is at level 1 and there are n numbers of completely connected networks (groups/clusters) of peers. Each such group, say G_i is formed by the peers of the subset P^{R_i} , ($0 \leq i \leq n-1$), such that all peers ($\in P^{R_i}$) are directly connected (logically) to each other, resulting in the network diameter of 1. The group-heads of the n groups are connected to form the edges (links) of the pyramid tree.
- 2) Any communication between a peer $p_i \in G_i$ and a peer $p_j \in G_j$ takes place only via the respective group-heads P_i and P_j and with the help of tree traversal.

B. Residue Class

Consider the set S_n of nonnegative integers less than n , given as $S_n = \{0, 1, 2, \dots, (n-1)\}$. This is referred to

as the set of residues, or residue classes (mod n). That is, each integer in S_n represents a residue class (RC). These

residue classes can be labelled as $[0], [1], [2], \dots, [n-1]$, where $[r] = \{a: a \text{ is an integer, } a \equiv r \pmod{n}\}$.

For example, for $n = 3$, the classes are:

$$[0] = \{\dots, -6, -3, 0, 3, 6, \dots\}$$

$$[1] = \{\dots, -5, -2, 1, 4, 7, \dots\}$$

$$[2] = \{\dots, -4, -1, 2, 5, 8, \dots\}$$

In the P2P architecture, we use the numbers belonging to different classes as the logical addresses of the peers; therefore, for the sake of simplicity we shall use only the positive integer values. Before we present the mechanism of logical address assignments, we state the following relevant property of residue class [19].

Lemma 1. *Any two numbers of any class r of S_n are mutually congruent.*

C. Assignments of Overlay Addresses

Assume that in an interest-based P2P system there are n distinct resource types. Note that n can be set to an extremely large value a priori to accommodate large number of distinct resource types. Consider the set of all peers in the system given as $S = \{P^{R_i}\}$, $0 \leq i \leq n-1$. Also as mentioned earlier, for each subset P^{R_i} (i.e. group G_i) peer P_i is the first peer with resource type R_i to join the system.

In the overlay architecture, the positive numbers belonging to different classes are used to define the following:

- (a) Logical addresses of peers in a subnet P^{R_i} (i.e. group G_i). Use of these addresses can be shown to justify that all peers in G_i are directly connected to each other (logically) forming an overlay network of diameter 1. In graph theoretic term, each G_i is a complete graph.
- (b) Identifying the edges (links) that connect different group-heads of the n groups in the tree.

The assignment of logical addresses to the peers in the groups and the resources happen as follows:

- 1) Each group-head P_r of group G_r is assigned with the minimum nonnegative number (r) of residue class $r \pmod{n}$ of the residue system S_n .
- 2) All peers having the same resource type R_r will form the group G_r (i.e. the subset P^{R_r}) with group-head P_r . Each new peer joining group G_r is given the group membership address $(r + j.n)$, for $j = 0, 1, 2, \dots$
- 3) Resource type R_r possessed by peers in G_r is assigned the code r which is also the logical address of the group-head P_r of group G_r .

Definition 3. *Two peers of a group G_r are logically linked together if their assigned logical addresses are mutually congruent.*

Lemma 2. *Each group G_r forms a complete graph [19].*

Observation 1. *Any intra-group data look up communication needs only one overlay hop [19].*

Observation 2. *Search latency for inter-group data lookup algorithm is bounded by the diameter of the tree [19].*

Observation 3: *A pyramid tree is a complete one if at any level j , number of nodes (i.e. clusters) is also j .*

Observation 4: *In an incomplete pyramid tree any level j has j nodes except the leaf level.*

Observation 5: *Joining of a new group (i.e. joining of a new cluster) always takes place at the leaf level.*

Observation 6: *A node that does not reside either on the left branch or on the right branch of the root node is an internal node.*

Observation 7: *Degree of an internal non-leaf node is 4.*

Observation 8: *Degree of an internal leaf node is 2.*

An example of a complete pyramid tree is shown in Fig. 1 [19]. In the figure, assignments of the overlay addresses of the group-heads are done in the following way. A group-head P_0 with resource type '0' is assigned the address '0' if it is the first peer to join the network, thereby becoming the root node of the tree. In general, suppose that there exist already i number of group-heads (i.e. P_0, P_1, \dots, P_{i-1}) in the tree. Then the next peer joining the system as the group-head with resource type i will be assigned with the address i . For example, the sixth group-head joining the system will have the logical address as 5 and the code of the resource type it possesses is 5. The edges are formed according to the pyramid tree structure [16]. Observe that no node can have a degree larger than 4 and only two nodes 6 and 9 at the leaf level have degree one each. Node 4 is an internal non-leaf node whereas nodes 7 and 8 are internal leaf nodes.

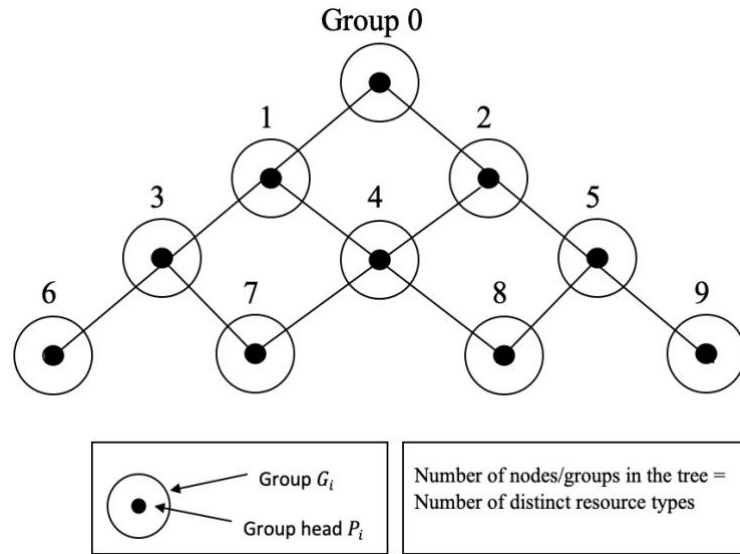


Fig. 1 A 4 level complete pyramid tree

III. SOME OTHER RELEVANT PROPERTIES

In this section, we state some other structural properties, which we shall incorporate in designing the proposed broadcast protocol. We shall state the properties using an example of a complete pyramid tree. The structural properties to be revealed are invariant, i.e. they do not change with the size (no. of levels) of the tree, or if the tree is a complete or an incomplete one. So, let us consider a complete pyramid tree of 5 levels as shown in Fig. 2. It means that it has 15 nodes/clusters (clusters 0 to 14, corresponding to 15 distinct resource types owned by the 15 distinct clusters). It also means that residue class with $\text{mod } 15$ has been used to build the tree. As pointed out above, the nodes' respective logical addresses are from 0 to 14 based on their sequence of joining the P2P system.

Each link that connects directly two nodes on a branch of the tree is termed as a *segment*. In Fig. 2, a bracketed integer on a segment denotes the difference of the logical addresses of the two nodes on the segment. It is termed as *increment*. This increment can be used to get the logical address of a node from its immediate predecessor node along a branch. For example, let X and Y be two such nodes connected via a segment with increment d, such that node X is the immediate predecessor of node Y along a branch of a tree which is created using *residue class with mod n*. Then, *logical address of Y = (logical address of X + d) mod n*.

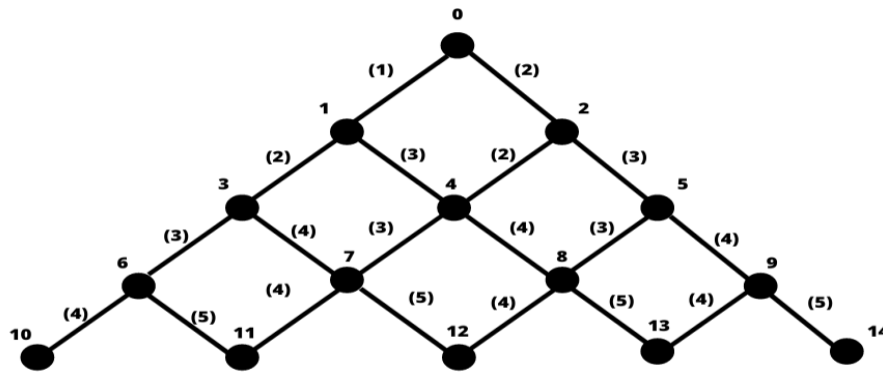


Fig. 2 A complete pyramid tree with root 0

Thus in the example of Fig. 2,

Logical address of the leftmost leaf node = (logical address of its immediate predecessor along the left branch of the root + increment) mod 15 = (6 + 4) mod 15 = 10.

We also use the following terminology. Any left branch originating at a node on the right branch of the root is termed a *left sub-branch*. For example, one such *left sub-branch* originating at node 2 on the right branch of the root is $2 \rightarrow 4 \rightarrow 7 \rightarrow 11$. Similarly, we can define right sub-branches as well.

Remark 1: *The sequence of increments on the segments along the left branch of the root, appears to form an AP series with 1st term as 1 and common difference as 1.*

Remark 2: *The sequence of increments on the segments along the right branch of the root, appears to form an AP series with 1st term as 2 and common difference as 1.*

Remark 3: *Along the 1st left sub-branch originating at node 2, the sequence of increments appears to form an AP series with 1st term as 2 and common difference as 1. Note that the 1st term is the increment on the segment $0 \rightarrow 2$.*

Remark 4: *Along the 2nd left sub-branch originating at node 5, the sequence of increments is an AP series with 1st term as 3 and common difference as 1. Note that the 1st term is the increment on the segment $2 \rightarrow 5$.*

In a similar way, the left most sub-branch (3rd one) can be interpreted. Observe that all such properties are invariant and do not vary with size of the tree. Also, similar kinds of remarks can be drawn in case of the right sub-branches.

IV. VIRTUAL NEIGHBORS

In this section, we shall present some important structural properties of the pyramid tree P2P system. To the best of our knowledge, no structured P2P system, either DHT or non-DHT based, possesses this property. It is stated below.

Let S_Y be the set of logical links that connect a node Y to its neighbors in a complete pyramid tree T_R with root R . Assume that the tree has n nodes (i.e. n group heads / n clusters). Let another tree T'_R be created with the same n nodes but with a different root R' . Let S'_Y be the set of logical links connecting Y to its neighbors in the tree T'_R .

Property 1. $S_Y \neq S'_Y$

Property 2. *Diameter of $T_R = \text{Diameter of } T'_R$*

Property 3. *Number of levels of $T_R = \text{Number of levels of } T'_R$*

Property 4. *Complexity of broadcasting in T_R with root R as the source of broadcast is the same for T'_R with root R' as the source of broadcasting. It is $O(\log n)$.*

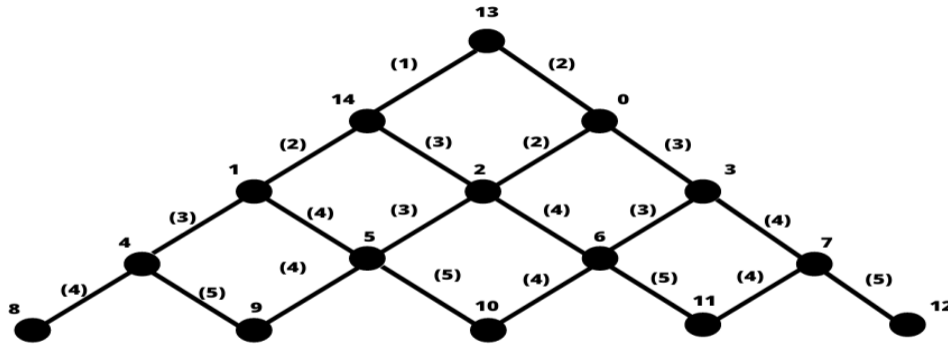


Fig. 3 A complete pyramid tree with root 13

An example: Consider the complete pyramid tree of 5 levels as shown in Fig. 3. Note that root of this tree is node 13, whereas root of the tree of Fig. 2 is 0.

It is seen that $S'_4 = \{1,8,9\}$ and $S_4 = \{1,2,7,8\}$. Therefore, Property 1 holds.

Diameters of both trees are same; it is 8 in terms of number of hops. Besides, both trees use the same 15 nodes and have the same total number of levels. Finally, broadcasting from either root 0 in the tree

of Fig.2 or from root 13 in the tree of Fig. 3 can be completed in 4 hops. Thus all the properties as mentioned above hold.

Remark 5. *Set of neighbors of a given node Z varies as the root of the tree varies. Hence, it is termed as virtual.*

V. BROADCAST

It may be noted that internal nodes are present starting from level 3 in a pyramid tree and at this level there is only one such node. At any other level i , the number of internal nodes is $(i-2)$. Therefore, total number of such nodes in an n level pyramid tree is $(n-1)(n-2)/2$, that includes both internal non-leaf and internal leaf nodes. Therefore, for flooding based inter cluster broadcasting, each internal non-leaf node at level i will receive one extra copy from its neighbors at level $(i-1)$ and will generate one extra copy to its neighbors at level $(i+1)$. So effectively, there are two duplicate packets per internal non-leaf node. Each internal leaf node will receive an extra packet from its neighbors. Therefore, for flooding based broadcast, total number of duplicates per packet broadcast will be

$$N = 2((n-1)(n-2)/2) - (n-2) = (n-2)^2; \text{ Therefore, as } n \text{ increases (more clusters are formed), } N \propto n^2.$$

Therefore, we avoid flooding and design an inter cluster broadcast protocol that does not generate any duplicate packet. The protocol takes help of the properties mentioned in the previous section. Let a node X be the source of broadcast. The following steps are followed in the protocol.

Protocol Inter Cluster Broadcast

Step 1: Root X sends packets to its neighbors on left and right branches.

Step 2: Each receiving node on the left branch sends packets to its neighbor on this branch till a receiving node is a leaf node.

Step 3a: The i^{th} receiving node on the right branch sends packets to its neighbor on the i^{th} left sub-branch originating at the i^{th} node until the i^{th} receiving node is a leaf node.

Step 3b: The i^{th} receiving node sends packets to its neighbor, the $(i+1)^{\text{th}}$ node on the right branch until it is a leaf node.

Step 4: Propagation along the i^{th} left sub-branch continues as in Step 2.

Remark 6. *No internal non-leaf node at a level i receives any duplicate packet as it receives only one packet from one of its neighbors at level $(i-1)$.*

Remark 7. *The inter cluster broadcast protocol does not generate any duplicate packet because any receiving node gets a packet via only one link.*

IV. CONCLUSION

In this paper, we have considered pyramid tree-based P2P network architecture. We have explored some important structural properties of the tree; besides, we have introduced a new structural property, *virtual neighbors* which none of the existing structured P2P systems possesses. We have incorporated all these structural properties in the design of an efficient inter cluster broadcast protocol with complexity $O(\log n)$ for complete pyramid tree architecture, where n denotes the number of nodes (clusters) in the tree. The present work is part of an ongoing research project; future work is directed at developing broadcast protocol for incomplete pyramid tree.

REFERENCES

- [1] P. Ganesan, Q. Sun, and H. Garcia-Molina, "Yappers: A peer-to-peer lookup service over arbitrary topology," Proc. IEEE Infocom 2003, San Francisco, USA, March 30 - April 1 2003.
- [2] Y. Chawathe, S. Ratnasamy, L. Breslau, N. Lanham, and S. Shenker, "Making gnutella-like p2p systems scalable," Proc. ACM SIGCOMM, Karlsruhe, Germany, August 25-29 2003.
- [3] B. Y. Zhao, L. Huang, S. C. Rhea, J. Stribling, A. Zoseph, and J. D. Kubiatowicz, "Tapestry: A Global-Scale Overlay for Rapid Service Deployment", IEEE J-SAC, vol. 22, no. 1, pp. 41-53, Jan. 2004.
- [4] A. Rowstron and P. Druschel, "Pastry: Scalable, Distributed Object Location and Routing for Large Scale Peer-to-Peer Systems", Proc. FIP/ACM Intl. Conf. Distributed Systems Platforms (Middleware), pp. 329-350, 2001.
- [5] I. Stocia, R. Morris, D. Liben-Nowell, D. R. Karger, M. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications", IEEE/ACM Tran. Networking, vol. 11, No. 1, pp. 17-32, Feb. 2003.
- [6] M. Yang and Y. Yang, "An Efficient Hybrid Peer-to-Peer System for Distributed Data Sharing", IEEE Trans. Computers, vol. 59, no. 9, pp. 1158-1171, Sep. 2010.
- [7] M. Xu, S. Zhou, and J. Guan, "A New and Effective Hierarchical Overlay Structure for Peer-to-Peer Networks", Computer Communications, vol. 34, pp. 862-874, 2011.
- [8] D. Korzun and A. Gurtov, "Hierarchical Architectures in Structured Peer-to-Peer Overlay Networks", Peer-to-Peer Networking and Applications, Springer, pp. 1-37, March 2013
- [9] Z. Peng, Z. Duan, J. Jun Qi, Y. Cao, and E. Lv, "HP2P: a hybrid hierarchical p2p network," Proc. Intl. Conf. Digital Society, 2007.
- [10] K. Shuang, P. Zhang, and S. Su, "Comb: a resilient and efficient two-hop lookup service for distributed communication system," Security and Communication Networks, vol. 8(10), pp. 1890-1903, 2015.
- [11] M. Kleis, E. K. Lua, and X. Zhou, "Hierarchical Peer-to-Peer Networks using Lightweight SuperPeer Topologies," Proc. IEEE Symp. Computers and Communications, 2005.
- [12] R. Zhang and Y. C. Hu, "Assisted peer-to-peer search with partial indexing," IEEE Trans. Parallel and Distributed Systems, vol. 18(8), pp. 1146-1158, 2007.
- [13] E. Cohen, A. Fiat, H. Kaplan, "Associative search in peer-to-peer networks: harnessing latent semantics," vol. 2, pp. 1261-1271, 2003.
- [14] Andrea Passarella, "A survey on content-centric technologies for the current internet: cdn and p2p Solutions," Computer Communications, vol. 35, pp. 1-32, 2012.
- [15] Bidyut Gupta, Nick Rahimi, Shahram Rahimi, and Ashraf Alyanbaawi, "Efficient Data Lookup in Non-DHT Based Low Diameter Structured P2P Network," Proc. IEEE 15th Int. Conf. Industrial Informatics (IEEE INDIN), July 2017, Emden, Germany.
- [16] Bidyut Gupta and Mohammad Mohsin, "Fault-Tolerance in Pyramid Tree Network Architecture," J. Computer Systems Science and Engineering, Vol. 10, No. 3, pp. 164-172, July, 1995.

[17] N. Rahimi, K. Sinha, B. Gupta, and S. Rahimi, LDEPTH: A Low Diameter Hierarchical P2P Network Architecture, Proc. IEEE 14th Int. Conf. on Industrial Informatics (IEEE INDIN), Poitiers, France, July 2016.

[18] Indranil Roy, Koushik Maddali, Swathi Kaluvakuri, Banafsheh Rekabdar, Ziping Liu, Bidyut Gupta, Narayan Debnath, *Efficient Any Source Overlay Multicast In CRT-Based P2P Networks – A Capacity - Constrained Approach*, Proc. IEEE 17th Int. Conf. Industrial Informatics (IEEE INDIN), July 2019, Helsinki, Finland.

[19] Indranil Roy, Bidyut Gupta, Banafsheh Rekabdar, and Henry Hexmoor, *A Novel Approach Toward Designing A Non-DHT Based Structured P2P Network Architecture*, EPiC Series in Computing, Volume 63, 2019, pages 121-129, (Proceedings of 32nd Int. Conf. Computer Applications in Industry and Engineering).

[20] Swathi Kaluvakuri, Koushik Maddali, Bidyut Gupta and Narayan Debnath “Design of RC based Low Diameter Hierarchical Structured P2P Network Architecture ”, EMENA-ISTL, 2019; LAIS 7 (Learning and Analytics in Intelligent Systems 7, pp 312-320, 2020) Springer.