EPiC Series in Computing
Volume 98, 2024, Pages 69–76
Proceedings of 39th International Conference on Computers and Their Applications

On further improvement of data look-up latency in Residue Class (RC) based Peer-2-Peer networks

Indranil Roy[1], Reshmi Mitra[1], Swathi Kaluvakuri[2], Nick Rahimi[3] and Bidyut Gupta[4]
[1] Department of Computer Science, Southeast Missouri State University, Cape Girardeau, MO
iroy@semo.edu
[2]Enterprise Fleet Management, St. Louis, Missouri
Swathi.kaluvakuri@efleets.com
[3] School of Computing Sciences & Computer Engineering, University of Southern Mississippi, Hattiesburg, MS
nick.rahimi@usm.edu
[4] School of Computing, Southern Illinois University; Carbondale, IL
bidyut@cs.siu.edu

Abstract

In this work, the existing Residue-class based peer-to peer (P2P) network has been considered because of its manifold advantages. Two of the most prominent advantages relevant to the present work are: (1) all peers with the same interest (or possessing same resource type) structurally form a group of diameter one, and (2) the group heads are connected in the form of a ring and the ring always remains connected even in presence of any churn. However, data look-up latency is n/2 for an n group-network. It may become substantial if n is large. To improve the latency, some topological properties of Star inter-connection network have been used to modify the existing RC based network appropriately resulting in remarkable improvement of latency.

1 Introduction

P2P networks are classified into two classes: unstructured and structured ones. In unstructured systems [1] 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) [2]-[4]. Such overlay architectures can offer efficient, flexible, and robust service [2]-[6]. 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. There exists another important approach; it is interest-based P2P networks. It is a structured approach without using DHT. Therefore, it does not have the problems that DHT based architecture faces, while at the same time it supports the deterministic behavior for data lookup. There exist several works in the literature in this direction [7]-[9].

Our Contribution

The existing two level residue class-based (RC based) peer-to-peer (P2P) architecture [10] has been considered in this work because of its several advantages. It is an interest-based system. Two of the most prominent advantages relevant to the present work are: (1) all peers with the same interest (or possessing same resource type) structurally form a group of diameter one, and (2) the group heads are connected in the form of a ring and the ring always remains connected even in presence of any churn. These two advantages help in designing a very efficient data look-up algorithm [10]. However, we observe that there is scope to improve further the latency of the data look-up algorithm; all that is required for this purpose is to reduce the diameter of the ring, because this latency is dependent on the diameter (n/2) of the ring (with n group heads). As n increases, latency of the existing look up algorithm increases as well. We observe that a STAR topology with same number of nodes (group heads) has much smaller diameter compared to the RC based architecture. It has led us to borrow some architectural properties of STAR (or simply, Star) interconnection networks [11] and apply these appropriately to redesign the existing RC based architecture. Our objective is to make the diameter of the redesigned RC based architecture comparable with that of a STAR network with the same number of nodes (group heads), thereby ensuring remarkable improvement of the data look-up latency. In this work, we shall use the words, architecture and topology, interchangeably.

The present work is organized as follows. In Section 2, we reintroduced briefly the RC based topology followed by the Star topology. We compare the diameters of the two architectures. In Section 3, similarities and dissimilarities between a six-node basic RC component and a complete 3-Star are stated here. In Section 4, the proposed modified RC based topology has been stated. It is followed by a brief discussion about our future works in this area.

2 RC Based Topology and Star Topology

We first state briefly the architecture of RC based network followed STAR topology.

2.1 RC Based architecture [10]

Definition 1. *We define a resource as a tuple <R_i, V>, where R_i denotes the type of a resource and V is the value of the resource.*

A resource can have many values. For example, let R_i denote the resource type ‘movies and V' denote a particular actor. Thus <R_i, V'> represents movies (some or all) acted by a particular actor V'.

Definition 2. *Let S be the set of all peers in a peer-to-peer system. Then $S = \{P^{Ri}\}$, $0 \le i \le n\text{-}1$, where P^{Ri} 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^{Ri}, we assume that P_i is the first peer among the peers in P^{Ri} to join the system. We call P_i as the group-head of group G_i formed by the peers in the subset P^{Ri}*

A. Two level P2P architecture

It is a two-level overlay architecture and at each level structured networks of peers exist. It is explained below.

1. At level-1, we have a ring network consisting of the peers P_i ($0 \leq i \leq n-1$). The number of peers (i.e., group heads) on the ring is n which is also the number of distinct resource types. This ring network is used for efficient data lookup and so, it is named as transit ring network.
2. At level-2, there are n numbers of completely connected networks (groups) of peers. Each such group, say G_i is formed by the peers of the subset P^{Ri}, ($0 \leq i \leq n-1$), such that all peers (ϵ P^{Ri}) are directly connected (logically) to each other, resulting in the network diameter of 1. Each G_i is connected to the transit ring network via its group-head P_i.
3. Any communication between a peer $p_i' \epsilon G_i$ and $p_j' \epsilon G_j$ takes place only via the respective group-heads P_i and P_j.

The architecture is shown in Figure 1.

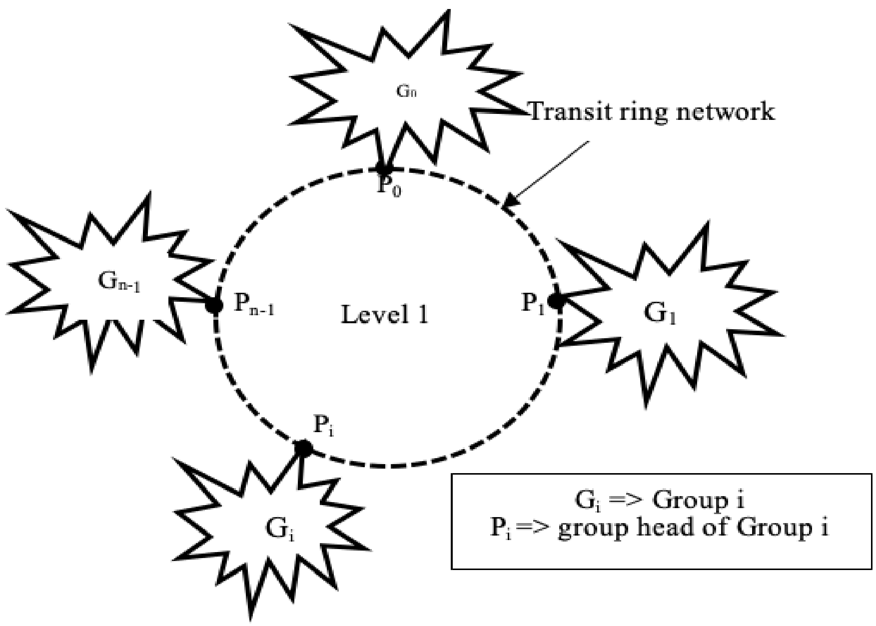

Figure 1: An RC-based P2P network

B. Assignments of overlay addresses

Consider the set S_n of nonnegative integers less than n, given as $S_n = \{0, 1, 2, \ldots (n-1)\}$. 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], …, [n – 1], where $[r] = \{a: a \text{ is an integer}, a \equiv r \pmod{n}\}$.

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

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

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

[2] = {…., − 4, − 1, 2, 5, 8, …}

A relevant property of residue class is stated below.

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

Assume that in an interest-based P2P system there are n distinct resource types. Consider the set of all peers in the system given as S = {P^{Ri}}, 0 ≤ i ≤ n-1. Also, as mentioned earlier, for each subset P^{Ri} (i.e., group G_i) peer P_i is the first peer with resource type R_i to join the system.

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

1. Logical addresses of peers in a subnet P^{Ri} (i.e. group G_i). Use of these addresses has been 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.
2. Identifying peers that are neighbors to each other on the transit ring network.
3. Identifying each distinct resource type with unique code.

The assignment of logical addresses to the peers at the two levels and the resources happen as follows:

1. At level-1, each group-head P_r of group G_r is assigned with the minimum nonnegative number *(r)* of *residue class r (mod n)* of the residue system S_n.
2. At level-2, all peers having the same resource type R_r will form the group G_r (i.e., the subset P^{Rr}) with the group-head P_r connected to the transit ring network. Each new peer joining group G_r is given the group membership address (r + j.n), for j = 1, 2, …
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.
4. Each time a new group-head joins, a corresponding tuple <Resource Type, Resource Code, Group Head Logical Address> is entered in the global resource table (GRT).

Definition 3: Two peers P_i and P_j on the ring network are logically linked together if (i + 1) mod n = j.

Remark 2: The last group-head P_{n-1} and the first group-head P_0 are neighbors based on Definition 3. It justifies that the transit network is a ring.

Definition 4: Two peers of a group G_r are logically linked together if their assigned logical addresses are mutually congruent.

Lemma 2: *Diameter of the transit ring network is n/2.*

Lemma 3: *Each group G_r forms a complete graph.*

2.2 STAR Architecture [11]

The address of a node in an n-star S_n is identified by a unique permutation of the digits {1, 2, 3, … , n}. Let f_i, (2 ≤ i ≤ n-1) be a function that maps permutation P_k of the digits {1, 2, 3, …, n} to another permutation P'_k, i.e. $f_i(P_k) = P'_k$ where the first and the i^{th} digits in P_k are interchanged to generate P'_k.

The mapping function f_1 on P_k is such that the first and the nth digits in P_k are interchanged to generate P_m. That is, $f_1(P_k) = P_m$. In other words, the last digit of P_m is identical to the first digit of P_k and vice-versa; meaning thereby that P_k and P_m belong to two different S_{n-1} stars which are the components of S_n and the nodes represented by these addresses, i.e., P_k and P_m are directly connected to each other. For example, in a 4-star, any two of the component 3-stars have two links connecting them. Thus, we see that number of nodes in a complete n-star is n!

An example of a complete 4-star, S_4 is shown in Figure 2. Note that it has 24 (4!) nodes and each component 3-star, S_3 is also a complete one and has 6 (3!) nodes. That is, the S_4 has four component complete 3-stars. However, in general, a star may be an incomplete one as well. Later when we use the topological properties of a star to design the modified RC based network, a node in a star means a group head as in the RC based one.

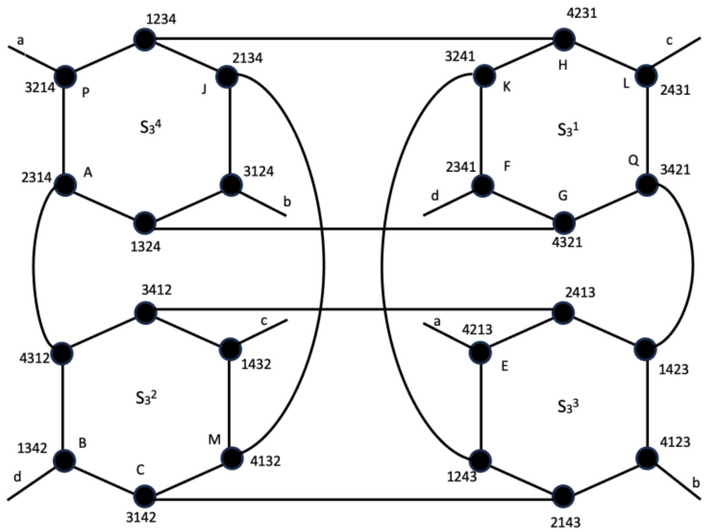

Figure 2: A complete 4-star architecture

2.3 Comparison of the diameters of the two architectures

We shall consider number of overlay hops between nodes as the parameter to determine diameters. For this purpose, in RC based one, we consider only the group heads located on the ring. The peers inside a group form a network of diameter 1 and so, we do not need to count it while determining the diameter of the (transit) ring because a group head may not have any other member present at a given time. For Star architecture [11] we consider only the nodes that form the architecture. (Later in this paper, we shall propose a variant of Star architecture in which a node represents a group head of peers with identical interest and each component (n-1)-star of an n-star is formed based on the RC based idea.). The comparison is stated below.

Number of Group heads	Diameter	Complete Star	Diameter
6	3	3-Star	3
24	12	4-Star	5
120	60	5-Star	7
720	360	6-Star	9

From the information stated above, we observe that as the number of nodes (group heads) increases, diameter of RC based architecture increases much faster compared to the diameter of a star architecture.

Therefore, latency of any data look-up protocol will be much smaller in a Star network compared to that in an RC based network.

3 A Six-node Basic RC Component and A Complete 3-Star

We highlight below the similarities and dissimilarities between a six-node basic RC-based component and a complete 3-Star.

Let us consider the two architectures as shown in Figure 3 and Figure 4. We name the first one as a basic RC-based component. The second one shows a complete 3-star. One main reason to consider such an RC based component is that it is very unlikely that there will exist less than six distinct types of resources at any time. Even if it is, as will be clear later that the architecture will simply consist of a RC-based ring with less than six nodes. The structural similarities and dissimilarities of the two architectures are stated below. As is seen in the figures, the overlay addresses in the two architectures are based on RC based idea and star-based idea respectively. The two architectures have the same diameter (3 in terms of hop). So, data look-up latency in both can be at most 3. This is the similarity. Now, consider that node f has left. In Figure 3, the topology remains connected, i.e., it remains a ring. So, performance of any data look-up protocol does not degrade. However, in Figure 4, the same is not true because the topology is a disconnected one resulting in an incomplete 3-star. This is the dissimilarity. This simple and yet important observation will be used in constructing the proposed architecture.

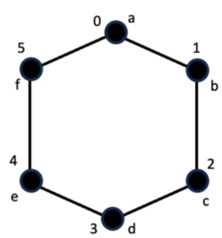

Figure 3: RC-based component

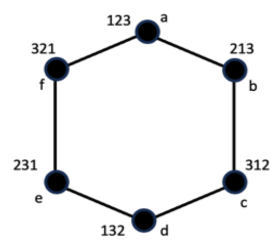

Figure 4: A complete 3-star

4 The Proposed RC-Based Architecture

In the proposed architecture we assume that an RC -based architecture will consist of a number of basic RC-based components and the components will be connected to form a complete n-star. Data look-up inside an RC-based component will follow the existing data look-up protocol reported in [10] using the RC-based overlay addresses of the nodes and inter-component routing (i.e., to send any query from component to component) will use the star-based overlay addresses. Note that in the architecture, every node in the star will be a group-head of peers with the same resource type. The reason for consideration of such intra-component and inter-component query protocols is two-fold: *1) under no circumstances an RC-based component can be disconnected, so there will be no degradation of performance of the look-up process, and the latency of intra-component query look-up will be limited by 3 hops, 2) data look-up latency between any two nodes in the architecture will at most be equal to the diameter of the n-star plus 3, resulting in a much less latency required for the query propagation compared to that in an RC based ring topology with the same number of nodes n!.*

In the above architectural design, a node needs to have a tuple of overlay addresses (a1, a2), in which a1 represents the node's RC-based overlay address and a2 represents its star-based overlay address. For example, in Figure 3, the address tuple of node a will now become (0, 123). The first component is used for intra-query propagation and the second component is used for the inter-component query propagation.

We now illustrate briefly how the proposed architecture is constructed. Let us start with a single node a. The address tuple of the node a is (0, 123). Now consider that a second node b joins (Figure 5). The address of this second node is (1, 213). Note that it is now a ring of two nodes (see the dotted link between the two nodes). Now consider that a third node c joins. Its address tuple is (2, 312). The Basic component structure is still a ring (see the dotted link in Figure 6). However, if we consider a 3-star with these three nodes, it is just incomplete and definitely the performance of any query in it will degrade. This will go on unless the component contains all six nodes. This is the reason which we mentioned earlier, why we each component will be constructed using RC based idea. However, when a new node, say g, (seventh one) joins, it will form the second basic component (of a 4-star) with one node only and the overlay address tuple of g will be 4231 while at the same time the second component address of each of the nodes in the first basic component will be updated as xxx4 (see Section 2.2). In addition, based on star topology, node a and node g will have a direct overlay link for inter-component query propagation. Note that the topology now becomes an incomplete 4-star. and yet there is no performance degradation of query propagation protocol (as discussed earlier in the previous section). The structure is shown in Figure 7. In this way, using RC based components, star topologies of higher dimensions can be constructed. As stated earlier the maximum look-up latency for inter-component query propagation in the new architecture can be equal to the diameter of the n-star plus 3. In this context, it may be mentioned that while the number of the distinct resource types is limited [12], the number of members in a group can be enormously large (theoretically infinite). This is one of the main characteristics (advantages) of the residue class-based design and it has been incorporated in the Star architecture (of the proposed design) consisting of basic RC based components.

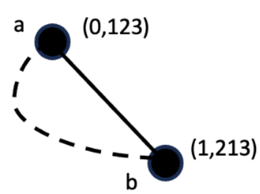

Figure 5:
A two-node component

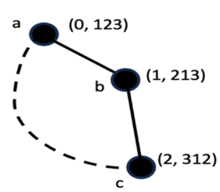

Figure 6:
A three-node component

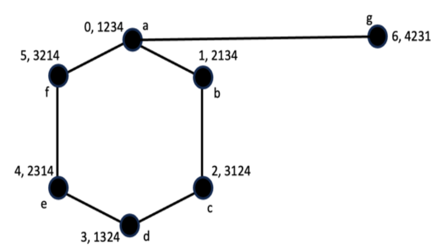

Figure 7:
An incomplete 4-star

In our future work, we plan to design the data look-up algorithm in the proposed architecture. In addition, we plan to incorporate security in it.

5 Conclusion

In this work we have used some topological properties of Star inter-connection networks to redesign an already existing Residue class-based peer-to-peer architecture. The existing RC-based architecture has been the choice in this work because of its manifold advantages. Specifically, the following two structural properties are behind the choice: these are (1) all peers with the same interest (or possessing same resource type) structurally form a group of diameter one, and (2) the group heads are connected in the form of a ring and the ring always remains connected even in presence of any churn. However, data look-up latency is n/2 for an n group RC-based network. It may appear substantial if n is large. In order to reduce this latency, some pertinent topological properties of Star network have been used to modify the existing RC based design. The proposed RC based-design has much less diameter than in its original design. It reduces the data look-up latency remarkably. Future works aim at designing secured data look-up process in the proposed design.

References

[1] 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).

[2] 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).

[3] 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).

[4] R. 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).

[5] 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).

[6] 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).

[7] Bidyut Gupta, Nick Rahimi, Henry Hexmoor, Shahram Rahimi, Koushik Maddali, and Gongzhu Hu, "Design of Very Efficient Lookup Algorithms for a Low Diameter Hierarchical Structured Peer-to-Peer Network", Proc. IEEE 16th Int. Conf. Industrial Informatics (IEEE INDIN), July 2018, Porto, Portugal.

[8] S.K. Khan and L.N. Tokarchuk, Interest-based Self Organization in Cluster-Structured P2P Networks", Proc. 6th IEEE Consumer Communications and Networking Conf, Las Vegas, pp. 1-5, 2009.

[9] L. Badis, M. Amad, D. Aissani, K. Bedjguelal, and A. Benkerrou, "ROUTIL: P2P Routing Protocol Based on Interest Links", Proc. 2016 Int. Conf. Advanced Aspects of Software Eng., Constantine, pp. 1-5, 2016.

[10] Swathi Kaluvakuri, Koushik Maddali, Nick Rahimi, Bidyut Gupta, and Narayan Debnath, "Generalization of RC-based Low Diameter Hierarchical Structured P2P Network Architecture", IJCA, vol. 27, no. 2, June 2020.

[11] S. B. Akers and B. Krisnamurthy, "A Group Theoretic Model for Symmetric Interconnection Networks", IEEE Trans. Computers, vol. 38, no. 4, pp. 555-566, 1989.

[12] Cheng, J., & Donahue, R. (2013). The pirate bay torrent analysis and visualization. International Journal of Science, Engineering and Computer Technology, 3(2), 38.