INTEREST BASED GROUP MANAGEMENT MECHANISMS FOR E-LEARNING DESIGN USING THE PEER-TO-PEER TECHNOLOGIES

Surya Bahadur Kathayat
Information and Communication Technology(ICT) Program, Asian Institute of Technology (AIT), Thailand

Nandana Rajatheva
Telecommunication(TC) Program, Asian Institute of Technology (AIT), Thailand

Keywords: Application Layer Multicasting, e-Learning, Group Management, Peer-to-Peer (P2P) Technologies.

Abstract: Traditional client-server based e-Learning architecture has many limitations. There is the overhead in a single learning server and inefficient use of resources. There is a lack of real-time interactive ness among learning group members therefore learning is not effective. There is also difficulty in the collaborative work because learners may have different interest, may feel lonely and may leave the system as well. Synchronous, any time and any where, and interactive e-Learning architecture where every learner in the learning group can contribute their resources like in traditional class-room based system, is the requirement of next generation e-Learning architecture. In this paper we purpose novel mechanisms for next generation e-Learning architecture using alternative technologies, peer-to-peer technologies. The proposed framework is based on P2P architecture for scalability, robustness, efficient sharing of resources and interactivity. Purposed system also incorporate efficient and reliable interest based e-Learning grouping and management mechanisms in the top of application layer. Such Interest based interactive P2P based group management mechanisms for e-Learning will combine the tools that are already available, independent of the installed infrastructure and offer a great deal of potential for workgroup collaboration, communities of practice, and self-directed learning.

1 BACKGROUND

Various forms of e-Learning that have been deployed so far are based on client-server technology where learning management server plays a key role providing contents, connectivity and services to the members of the learning system, as shown in figure 1. Though client-server system in the e-Learning is easy to implement and cost beneficial there is wastage of resources in the system, less interactive and collaborative besides the possibility of single point overhead and failure.

Peer-to-Peer (Nowell et al, 2002), though is not a new technology, however only recently has been exploited throughout the Music and entertainment industry especially sharing content files (Lee et al., 2002) containing audio, video, data or anything in digital format, and real-time data. As the peers in the Peer-to-Peer computer network relies on the computing power and bandwidth of the participants in the network rather than concentrating these in a relatively few servers, P2P technology will also help many of the limitations of the traditional e-learning system.

Basic motive of this work is that P2P, in concept, can also be a natural tool for educators allowing the

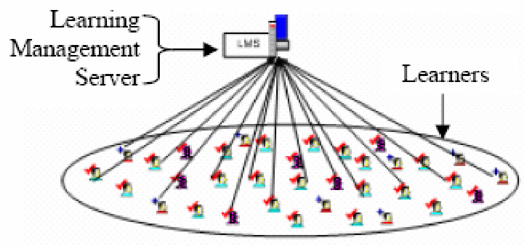

Figure 1: Traditional e-Learning Model.

Bahadur Kathayat S. and Rajatheva N. (2006).
INTEREST BASED GROUP MANAGEMENT MECHANISMS FOR E-LEARNING DESIGN USING THE PEER-TO-PEER TECHNOLOGIES.
In *Proceedings of WEBIST 2006 - Second International Conference on Web Information Systems and Technologies - Society, e-Business and e-Government / e-Learning*, pages 339-346
DOI: 10.5220/0001241603390346

group collaboration, management and sharing of resources for a constructivist approach to the learning. One scenario of such motive is briefly explained as below. In the traditional e-Learning system, single learning management server is responsible for handling large number of users which limits the problem of scalability, overhead, and inefficient use of resources. Besides that the interest of the different learner may be different and so that it is very difficult to do a collaborative work. If there is lack of collaboration then a user may feel lonely and there is chance of leaving the system and hence the effectiveness of the learning system will be significantly reduced. These limitations are inevitable even if it is assumed that if there is only a single common interest group, like learners of a single class room, in the client-server based e-Learning system. Therefore it is interesting to group the users of e-Learning system according to their interest and apply decentralized P2P technology within and among such groups. This approach will result better resource utilization because every peer can contribute their resources and more interactive ness because each peer can communicate in two ways either group mode via overlay multicasting or peer mode with other peer.

There are many users in the realistic large learning domain and we assume that these users represent the node or peer in the overlay network (Zhang and Hu, 2003). Different peers may have different interest so peers having common interest will be organized together to form a group. There will be two possibilities; either the peer may join the already existing group or peer may create its own group and other peers may join it later. So there will be two categories of the users in the group creators or leaders and the normal users. For simplicity, if we assume that learning will be done by chatting (not limited to this) then interactive e-learning scenario in such particular case will be as follows:

First the creator, say c1, will create a group according to its interest, say computer network, and seeking for the other interested peers. If other peer, say p1, in the overlay network also have the same interest in the computer network, it will first find out the creator 'c1' and then pop-up chat window will appear for the learning by chatting. Similarly if other peers having same interest may find the group and join the group learning process. As shown in the Figure 2, there will be many possible cases. Most likely case is that one peer may have more than one interest and would like to participate in on the multiple groups. For example peer 'p1' may have common interest in the computer network and it is already the member of the group 'g1', peer p1 may also have interest in the database design so it may wish to join the database group, say g2, (at the same time) and get involved in the learning process.

Another likely case will be that there will be more than one groups having nearly common interest and either members or groups leaders may wish to merge these groups. It is also interesting to consider peer in a particular group may leave the group and or may wish to create another group having different interest than that of the current group and advertise its group members.

Briefly, framework of P2P groups' management mechanisms (interest based group formation, efficient group communication and groups management) will be proposed (potential use in collaborative learning) to incorporate interactive ness among the members, allow the efficient use of resources reducing the overhead in the server and single point of failure, and add scalability, decentralization and many more.

Rest of the paper is organized as below. Section 2 describes the statement of the problem and section 3 discusses about the objective and scope of this work. Section 4 explains briefly about the related literature on the P2P technology and grouping mechanisms. Finally section 5 of this paper discusses about the proposed system. Peer-to-Peer interest based grouping mechanism, efficient data delivery mechanism and management mechanisms, and learning environment model are also included in last section.

2 PROBLEM STATEMENT

Existing Client-Server (C/S) based e-Learning systems are facing many problems like inefficient use of resources, single point overhead and failure, limited interaction among the members, scalability etc. With these limitations, C/S based e-Learning could not be significant alternative to the traditional classroom-based learning. P2P technology, which is a hot technology recently for the online music and file sharing, has potential applications on the e-Learning as well. However, to date, there is very limited use of P2P in the e-Learning. From instructors' point of view, it is challenging and interesting to create interest based group, sub-group formation, and merging groups having similar interest. From the students' point of view, the challenge is to join into the group having specific interest and to get the multiple group membership. Common challenge for both is to efficiently distribute messages to other members.

To the best knowledge there are no existing e-Learning models for the collaborative learning using the structured P2P network especially for interest based group formation, for merging of two groups having nearly similar common interests and also group splitting or sub-grouping if the interest among the members of the group are in conflict. Therefore it is interesting to the design a semi-decentralized e-Learning framework that will provide efficient and more effective, collaborative and synchronous learning environment.

3 OBJECTIVES AND SCOPE

The major objectives of this research is to purposing a design of the group management mechanism for structured P2P network which will incorporate virtual ring interest based group formation, multi-virtual ring based data delivery mechanism, and group merging and group splitting.

This work focuses on more fundamental issues like peer organization, group communication and fundamental management issues for the collaborative synchronous e-Learning. Peer organization includes basically group formation based on an efficient multicast group ring; joining nodes, leaving nodes in the group are also handled with ring repair mechanism. Data forwarding will be based on the multi-virtual multicast-ring (multi-unicast and unicast based) and group leader plays an important role for the group communication. The potential scope of proposed mechanisms or algorithms is that these can be suitably applied for the synchronous, effective, and collaborative e-Learning system.

4 RELATED LITERATURES

4.1 P2p Technology

According to (Rowstronand and Druschel, 2001), P2P is a network architecture in which nodes are relatively equal, in the sense that each node is in principle capable of performing each of the functions necessary to support the network. In (Pandurangan, 2001), P2P systems are defined as the distributed systems without any centralized control or hierarchical organization. The software running at each node is equivalent in functionality.

There are three types of well-defined P2P architectures namely pure, hybrid and hierarchical architectures. In pure P2P architecture (Schollmeier, 2002), all functions and all relevant digital objects are distributed across many nodes, such that no node is critical to the network's operation and hence no node can exercise control over the network. Flooding and document routing algorithms are used for peer search and resource discovery in this P2P architecture. An example of pure P2P architecture is original Gnutella. In hybrid architecture, index is centralized like C/S system, therefore peers first contact the central peer to locate other peers and shared resources. Example of hybrid architecture is Napster (Thilliez et al., 2003) where the index is accessed in client-server mode, whereas the digital objects are transferred directly among peers. In hierarchical architecture, index is hierarchically structured and accordingly hierarchy of normal peers and super peers are maintained.

GRID computing architecture also seems like P2P but there are some differences. Grid computing is a means whereby available processing resources can be located, used and coordinated; whereas P2P encompasses both processing and data resources. Grid computing also differs from P2P in that it is largely pragmatic engineering effort, rather than scientifically designed architecture (Zhuge, 2005).

4.2 Grouping Mechanisms in P2P

Distribute hash tables (Eastlake and Jones, 2001; Sit and Morris, 2002) are the core for the routing in the P2P networks. Major structured P2P protocols are Pastry (Zhang and Hu, 2003), Tapestry (Zhao et al., 2004), Chord (Stoica et al., 2001), CAN (Ratnasamy et al., 2001) etc. All of them take, as input, a key and, in response, route a message to the node responsible for that key. The keys are strings of digits of some length. Nodes have identifiers, taken

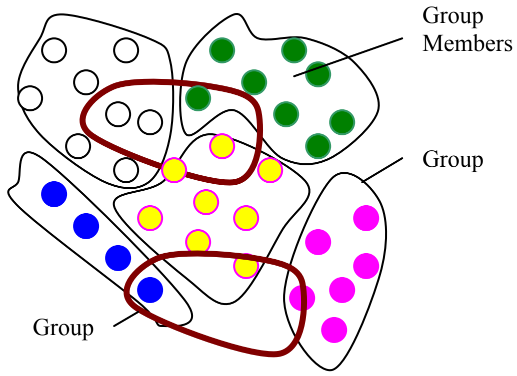

Figure 2: P2P Group Formations.

from the same space as the keys (i.e. same number of digits). Each node maintains a routing table consisting of a small subset of nodes in the system. When a node receives a query for a key for which it is not responsible, the node routes the query to the neighbor node that makes the most "progress" towards resolving the query. The notion of progress differs from algorithm to algorithm, but in general it is defined in terms of some distance between the identifier of the current node and the identifier of the queried key. Some group multicasting algorithms are pastry based SCRIBE (Castro and Rowstron, 2002) which is a reverse-path forwarding tree based publish/subscribe System, Tapestry based Bayeux (Zhao et al., 2001) which uses forward-path forwarding tree, and Brog (Zhang and Hu, 2003) which uses the same concept but the multicast is formed by the hybrid approach i.e. reverse path and forward path tree approach. Controlled and directed flooding concepts are also used for the mini-CAN and CHORD multicasting (Ratnasamy et al., 2001).

In distributed environment (Plaxton et al., 1997) group multicasting can be done either by mesh or multicast tree or ring. Mesh strategy provides more than one path between the group members and in tree case a single path is established between any pair of nodes. It is also feasible to apply a mesh first followed by tree construction algorithm to implement overlay multicast. Mesh provides routing stability and QoS but Tree approach have advantages in terms of link stress, no routing loops and simplicity. Traditional tree approaches use root based approaches for forwarding the messages which is well suited for the 1-to-m multicast. If the sender desires to send the message to the multicast group, it sends the message to the root which in turn forwards the message along the tree to all receivers. Network efficiency can be improved by using a source based tree algorithms in which each source builds an optimal tree from the source to all receivers in the group. However this approach introduces more overload as each node must run the routing algorithm and must maintain large amount of supporting information. So there are different alternatives for the overlay multicast protocols and existing initiatives tends to focus on the specific optimization parameters for the targeted application environment. Most Tree based approaches are proposed and implemented in the structured P2P overlay that has lower data delivery percentage with no back up path to each member (as in ring topology) but provide lower path stretch or link stresses.

5 PROPOSED SYSTEM

5.1 Technological Infrastructure

Both instructors and students in the e-Learning, like in the class room based learning, construct their own domain and it is at least somewhat different from others. These are self organizing and towards decentralization. Recent technological developments on the self-organizing and decentralized P2P network substrate, like Pastry (Zhang and Hu, 2003), Tapestry (Zhao et al., 2004), CHORD (Stoica et al., 2001), and CAN (Ratnasamy et al., 2001), point to a new paradigm for research and for building distributed applications. Each of these overlays implements a scalable, fault-tolerant distributed hash table for node ID, object ID representation and also for limiting the number of routing hops to locate them. So the new platform for e-Learning will be designed on such substrates where each instructor and/or learner node ID (based on IP address) and their class or group IDs (Based on Group Name) is uniquely obtained and uniformly distributed using the SHA-1 (Sit and Morris, 2002) hash function. Every node is identified by m-bits on the overlay network. All the nodes that are the members of all the particular groups will be the nodes in the common domain. Interest based groups like classes and common domain is like university where there are many mini-domains.

5.2 Peer Organization in Group

Here, Ring Based Group formation over the overlay network is proposed, such mechanisms to the best of knowledge, in structured P2P network, are not proposed and implemented yet. Main limitation of the ring based multicast group is the routing delay, but node degree is constant and they are suitable for secure, reliable and ordered delivery of messages, and effective against single node failure. If the routing delay is reduced in the ring topology, then it will be suitable for more cases, therefore a similar group formation mechanism (virtually multi-ring group multicast) is proposed here. Groups are assumed to be a medium sized classes having 10 - 100 peers and the group ID will determine or represent the group's interest.

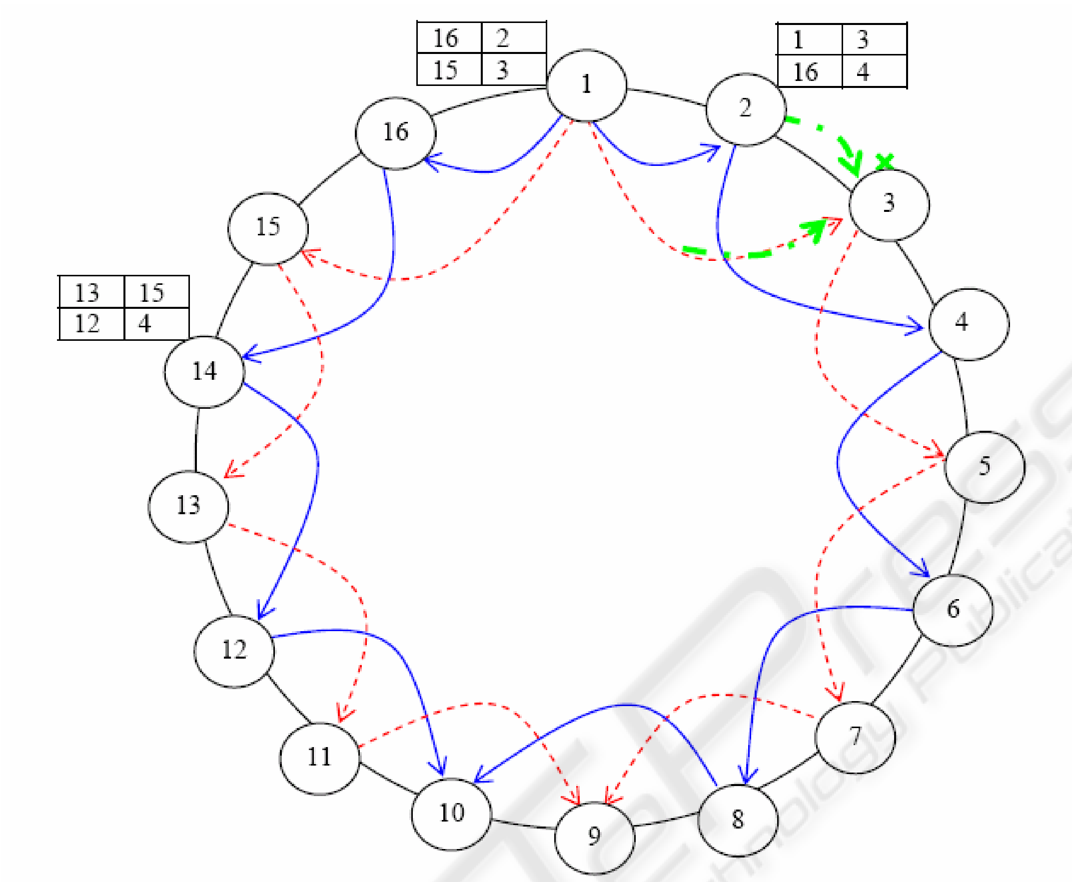

Figure 3: Group Ring and virtual multi-ring.

5.2.1 Group Formation Mechanism

One node, having sufficient resources like bandwidth (BW), CPU, memory etc and willing to contribute more resources, can create a group specifying its interest in the structured P2P substrate and wait for other nodes. Other nodes will join according to their own interest; the virtual ring topology will be maintained in the overlay network. If any node wants to join the pre-existing ring, it will request a found first peer (bootstrap node) on the ring and the peer on the ring replies to the requesting node and forward the request towards the root i.e. leader of the group. After getting an acknowledgement from the leader, bootstrap node will reply to the requesting node along with its neighbor information (IP address, other existing group information) so that requesting node can join the ring. Then all these nodes (bootstrap node, neighboring node and requesting node) will update the neighbor list and the leader will update group information. Group Leader will send periodic live signal, root information, number of nodes in the system. When the particular node leaves the group, then neighbor nodes will know about it from the regular neighbor update information and accordingly maintain their new neighbor list and inform to the group leader to maintain updated correct group information.

5.2.2 Efficient Data Delivery Mechanism

Each learning peer will contain more than one (say $\sqrt{N}/2$) successor and predecessor list, so virtually there will be more than one ring (say $\sqrt{N}/2$ rings) for multicasting the group message. Each peer will get N (number of nodes in the group) from the group information circulated by the instructor (root node). Each node will also issue special request signal (node address, SUB-COUNT) in each direction to maintain neighbor list. The initial setting for SUB-COUNT value is $\sqrt{N}/2$. When the ring node get that special request signal, there will be two possibilities at the node, (i) node will reply (node address) to the requester if SUB-COUNT is greater than zero and then forward that request signal by decrementing the

SUB-COUNT value by 1, (ii) node will discard and terminate the special request signal if SUB-COUNT is equal to or less than zero.

Root node can control and manage the token for ordered data-delivery within the group. Each node willing to send data has to send data to all of its neighbors (say √N/2) using the multi-unicast mechanism, other peers will forward the message to the highest successor/predecessor (formation of multiple virtual rings). Each node in the group can forward message to the original ring and corresponding virtual ring. This mechanism will also reduce the overhead in the node and routing delay (in terms of hops) at most will be improved by √N/2 times than normal ring which can be mathematically expanded as

i. Source node can send data to $\sqrt{N}/2$ number of nodes in once in both direction of ring.

ii. Since $\sqrt{N}/2$ nodes get message in one hop, 1 node will get message in average $2/\sqrt{N}$ hop (unitary method) with reference to the original ring.

iii. $N/2$ nodes (half of the nodes in symmetrical ring) will get message in $2/\sqrt{N}$ * $N/2$ $=\sqrt{N}$ Hops

iv. Without multiple-virtual mechanism, $\sqrt{N}/2$ nodes in the ring will get message in $N/2$ hops.

v. Therefore routing delay improved will be improved as $\frac{N/2}{\sqrt{N}} = \sqrt{N}/2$ times

As an example shown in the Figure 3, there are 16 nodes in a ring i.e. in a group. According to the proposed data delivery mechanism, there will be √16/2 = 2 virtual-rings in the group ring. The routing delay will be improved (optimum case) by 2 times. Similarly if there are 100 nodes in the system, there will be 5 virtual rings within the group ring and routing delay will be reduced by 5 times. Besides reducing the routing delay, concept of the virtual ring will be useful as the backup link to the normal ring in the case of the failure of the particular node in the normal ring. Each node will keep the source information and maximum sequence number of the packet that it received from that source. Each node will then forward the received packet if that packet is not already received from the corresponding source; otherwise it simply discards. Suppose at time t, node '3' get the packet 'n' from node '1' in one hop using virtual ring. At time t' (t' > t), if node '3' get the same packet from node '2', node '3' will simply ignore it which is shown by thick line in Figure 3.

5.2.3 Group Merging and Splitting

As the interest of the peer or the learner may change from time to time, it should be able to participate in different groups having corresponding interests accordingly, so group merging and splitting have significant importance in the e-Learning.

For the sub-grouping, a peer having different interest than the current group first create a new group and deliver the message to the existing group members so that interested peers join it later. This sub-grouping is not be limited to the existing group members; rather other group peers having same interest may join the newly formed group. After negotiation between the two group leaders, leader for the newly formed common group is selected and that manages the groups.

For the group merging, there are different possible cases such as (i) one particular peer may be the member of two groups and may know that two groups are engaging in the similar activities, it will then inform its leaders and two leaders can communicate and exchange the information to merge the group (ii) one leader may be the member of the another group and these two can share the information to merge the group.

5.3 Implementation Model

Learners are in application layers, internet based overlay network. Each user run the standalone application software (P2P software developed using Jdk1.4.2) specifying its interests. Learners may have different interests and there may be more than one learner in the system having common interest. A peer first tries to find out the existing groups with its interest and if such groups are not found it creates a new group (we assume that group creator have sufficient resources) and wait for other peers to join it. Once there are two or more members in the group, they communicate with each other and discussion goes on (currently only messaging). Also in different cases as mentioned in the earlier section, two groups can merge together and be involved in the collaborative learning.

6 PRELIMINARY RESULTS

Experiments are going on parallel in two different scenarios. First is the deployment of these algorithms in the Internet. Using the FreePastry-1.4.2 structured peer-to-peer platform, algorithms are implemented (some in implementing phase) varying the control variable (SUBCOUNT) to control the number virtual rings. The results shown in Figure 4 are some results with data measured using 10 nodes (physically in the same laboratory) in the internet and running the developed software on each. Software is written in Java Jdk-1.4.2 version. These preliminary results clearly show that there is significant reduction in the delay in MVRing case compared to RING case i.e. amount of time to multicast the message in the group. Now, experiments are towards increasing the number of nodes to 100s of numbers and physically from different locations.

Besides that we are conducting research to calculate the node stress on each nodes and link stress between the nodes, fault tolerance of the proposed data delivery mechanism to compare its efficiency with that of existing tree based group communication mechanisms.

Second scenario of the experiment we are conducting is the modeling of the internet in transit-stub topology using GT-ITM topology generator and simulating the performance evaluation of the proposed algorithms using Network Simulator (NS-2). In this case the nodes in the group are chosen randomly and hence it is obviously not necessary that neighbor node is the nearest node in terms of time. Results as shown in Figure 4 show the latency profile for the scenarios first where P2P code is run on 10 machines and RING and MVRING algorithms are compared with exactly implemented SCRIBE. Figure 5 shows the result of second scenario. The average latency that each node experience from its predecessor in the case of the multiple-virtual ring cases is about 2000 ms, 1700ms and 1500ms for number of nodes (n) 50, 150 and 500 respectively, while these values are 2600ms, 2700ms and 2900ms in case of pure ring based grouping and data delivery mechanism.

From the results in Figure 4 and Figure 5, it is shown that nodes clear that latency in case of the MVRing is significantly improved than in the RING case and quite better compared to SCRIBE as well. Experiments are going on to measure node stress, link stress, fault tolerance of systems and efficiency of the data delivery mechanisms.

7 CONCLUSIONS

Decentralization addresses the overhead in a particular machine and all members of the learning group can share resources among each other. Grouping of the learners according to their interest in P2P technology increases interactive ness and effective collaboration in the e-learning system. The virtual multi-ring based data delivery mechanism for the application layer group multicasting adds the reliable communication among group members. Finally, the instructors and learners having variable interest with time can be handled by the interest based group merging and group splitting.

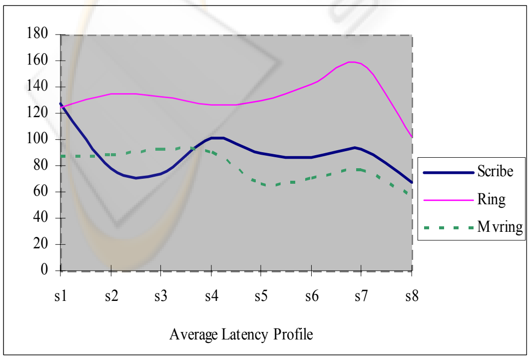

Figure 4: Average Latency profile.

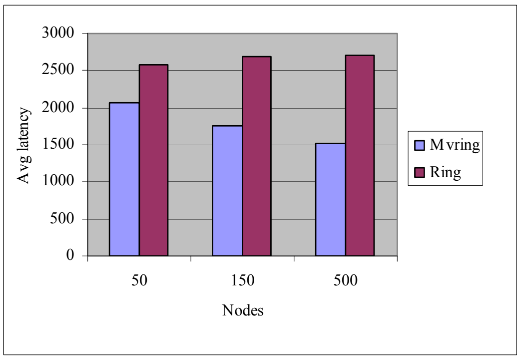

Figure 5: Latency Profile for n=50,150,500.

REFERENCES

Castro, M., and Rowstron, A. (2002). SCRIBE: A large – scale and decentralized application-level multicast infrastructure. *IEEE Journal on Selected Areas in Communications 20*(8), 1489-1499.

Eastlake, D., and Jones, P. (2001). US Secure Hash Algorithm 1 (SHA-1). *Published in RFC3174*. RFC Editor: United States

Lee, Y., Oh, C., and Park, E. K. (2002). Intelligent knowledge discovery in peer-to-peer file sharing. In *Proceeding of 11th International Conference on Information and knowledge management* (pp 308-315). Virginia : ACM Press.

Nowell, D., Balakrishnan, H., and Karger, D. (2002). Analysis of the Evolution of Peer-to-Peer Systems. In *ACM Conf. on Principles of Distributed Computing* (pp 233-242). Monterey: ACM Press.

Pandurangan, G. (2001). Building Low-Diameter P2P networks. In *Proceedings of the 42nd IEEE symposium on Foundations of Computer Science* (pp 492). Washington: IEEE Computer Society.

Plaxton, C.,Rajaram, R., and Richa, A. W. (1997). Accessing nearby copies of replicated objects in a distributed environment. In *Proceedings of the Ninth Annual ACM Symposium on Parallel Algorithms and Architectures* (pp 311-320). New York: ACM Press.

Ratnasamy, S., Handley, M., Karp, R., and Shenker, S. (2001). Application-level Multicasting using Content-Addressable Networks. In *Proceedings of the Third International COST264 Workshop on Networked Group Communication* (pp 14-29). London: Springer-Verlag.

Rowstronand, A., and Druschel, P. (2001). Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *Proceedings of the 18th IFIP/ACM International Conference on Distributed Systems Platforms*(pp 329-350) Heidelberg Germany.

Schollmeier, R. (2002). A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications. In *International Conference on P2P Computing* (pp 202-209). Munchen: IEEE Computer Society.

Sit, E., and Morris, R. (2002). Security Considerations for Peer-to-Peer Distributed Hash Tables. In *1st International Workshop on Peer-to-Peer System* (pp 129-138). Cambridge: Springer-Verlag.

Stoica, I., Morris, R., Karger, D., Kaashoek, M.F., and Balakrishnan, H. (2001). Chord: A scalable peer-to-peer lookup service for internet applications. *IEEE Journal on IEEE/ACM Transactions on Networking (TON), 11*(1), 17-32.

Thilliez. M., Delot. T., Lecomte. S., and Bennani. N., (2003). Hybrid Peer-To-Peer Model in Proximity Applications. In *17 th International Conference on Advanced Information Networking and Applications (AINA'03)* p. 306

Zhang, R., and Hu, C. (2003). Anycast in Locality Aware Peer-to-Peer Overlay Networks. In *Lecture Notes in Computer Science, 2816*, 34-46.

Zhang, R., and Hu, Y.C. (2003). Brog: A Hybrid Protocol for Scalable Application-Level Multicast in Peer-to-Peer Networks. In *Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video* (pp 172-179). New Work: ACM Press

Zhao, B.Y., Joseph, A.D., Katz, and R.H., Kubiatowicz, J. (2001). Bayeux: An Architecture for Scalable and Fault-tolerant Wide-Area Data Dissemination. In *Proceedings of the Eleventh International Workshop on Network and Operating System Support for Digital Audio and Video* (pp 11-20). New Work: ACM Press.

Zhao, B.Y., Huang, L., Stribling, J., Rhea, S.C., Anthony, D. Kubiatowicz, D. (2004). Tapestry: A Resilient global-Scale Overlay for Service Deployment. *IEEE Journal on Selected Areas in Communications, 22*(1), 41-53.

Zhuge, H., Sun, X., Liu, J., Yao, E., and Chen. X. (2005). A Scalable P2P Platform for the Knowledge Grid. *IEEE Transactions on Knowledge and Data Engineering 17*(12), 1721-1736.