

A Comparative Simulation Analysis of P2P System Architectures

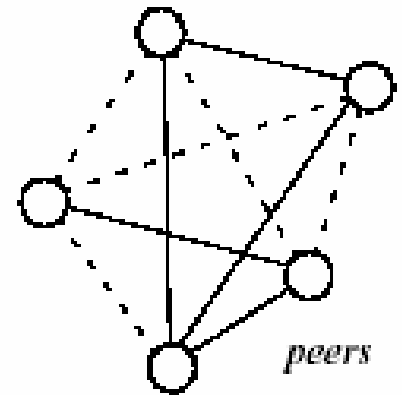


Anjan Mahanta

Sirinthon International Institute of
Technology, Thammasat University, Thailand

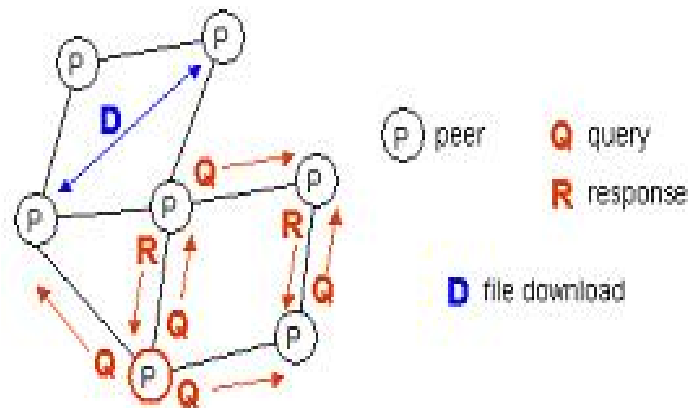
Overview of Peer-to-Peer(P2P) System

- The term “peer-to-peer” (P2P) refers to systems and applications that functions in a decentralized manner
- The benefits of a P2P approach include
 - improving scalability by avoiding dependency on centralized points
 - eliminating the need for costly infrastructure by enabling direct communication among clients
 - enabling resource aggregation



About Gnutella

- Gnutella provides a purely distributed file sharing solution
- Users can run software that implements the Gnutella protocol to share files and search for new files



Gnutella



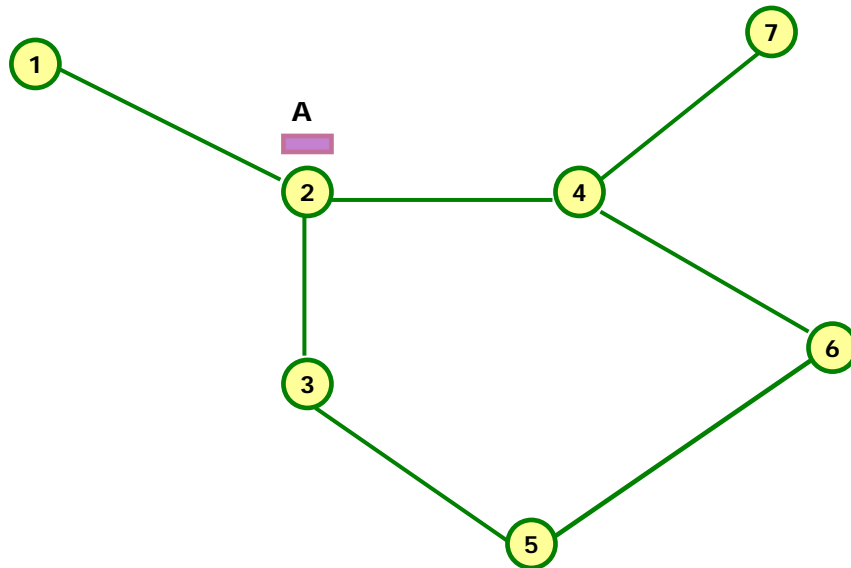


Gnutella Protocol Messages

- Broadcast Messages
 - **Ping**: initiating message (“I’m here”)
 - **Query**: search pattern and TTL (time-to-live)
- Back-Propagated Messages
 - **Pong**: reply to a ping, contains information about the peer
 - **Query response**: contains information about the computer that has the needed file
- Node-to-Node Messages
 - **PUSH**: file download



Gnutella Search Mechanism

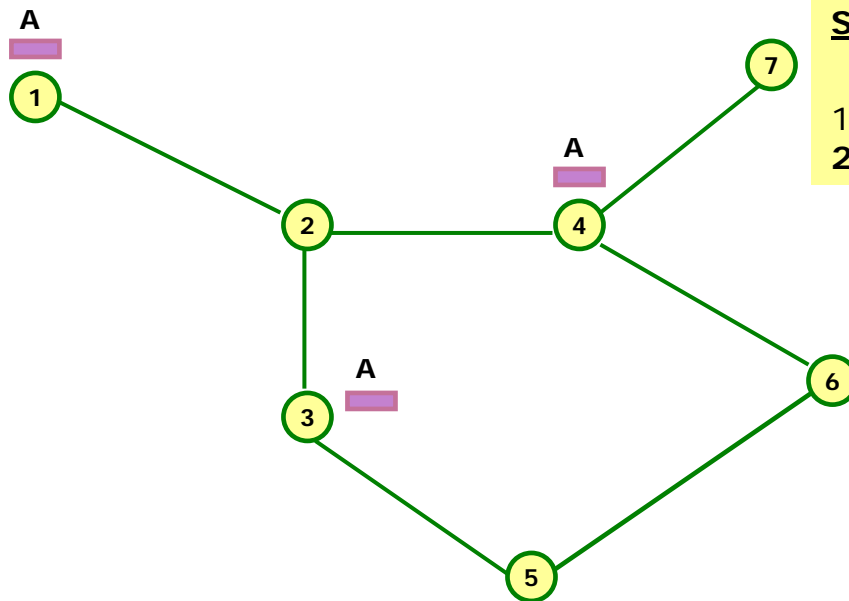


 Peer

Steps:

1. Peer 2 initiates search for file A.

Gnutella Search Mechanism

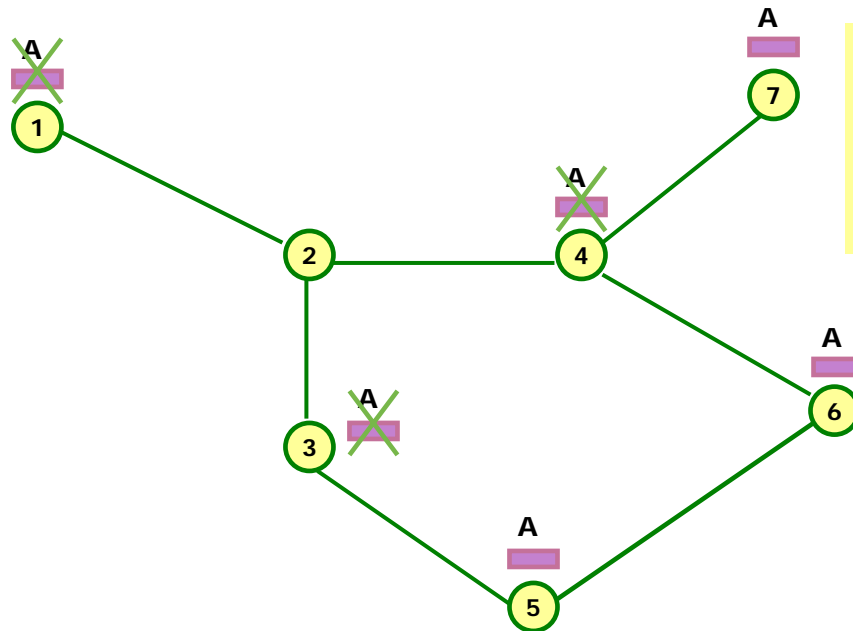


Steps:

1. Peer 2 initiates search for file A.
2. **Forwards Query request to all it's neighbors.**

 Peer

Gnutella Search Mechanism

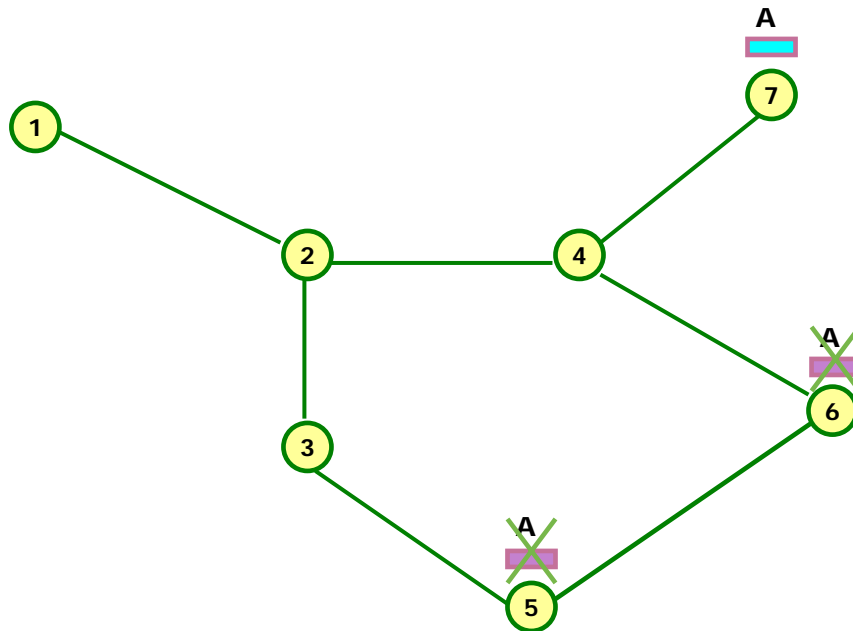


Steps:

1. Peer 2 initiates search for file A.
2. Forwards Query request to all it's neighbors.
3. **File A not found, forwards Query request.**

 Peer

Gnutella Search Mechanism



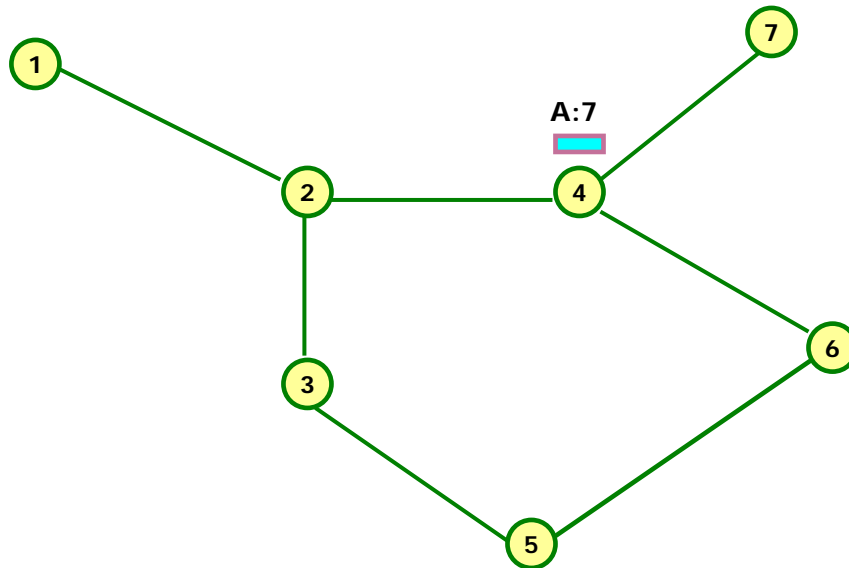
 Peer

Steps:

1. Peer 2 initiates search for file A.
2. Forwards Query request to all it's neighbors.
3. File A not found, forwards Query request.
- 4. File A found in Peer 7.**



Gnutella Search Mechanism

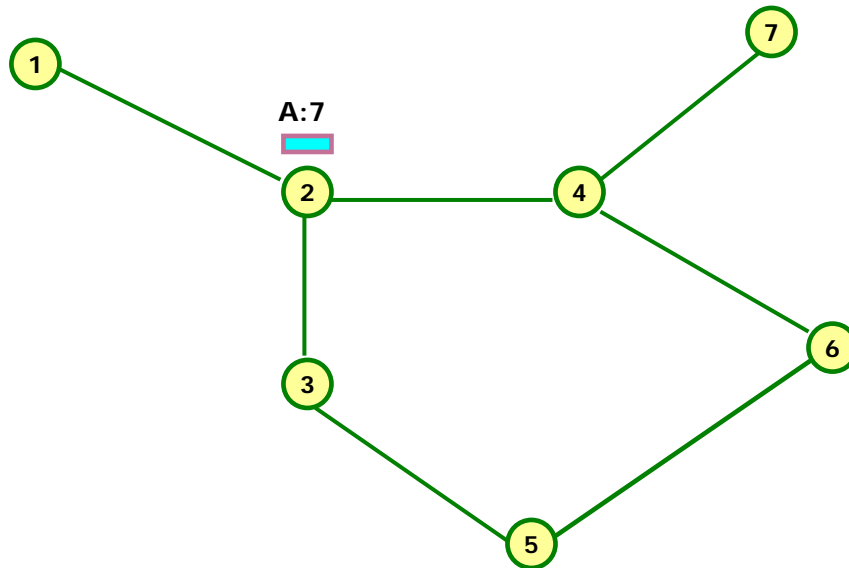


 Peer

Steps:

1. Peer 2 initiates search for file A.
2. Forwards Query request to all it's neighbors.
3. File A not found, forwards Query request.
4. File A found in Peer 7.
- 5. Query reply message is backpropogated.**

Gnutella Search Mechanism

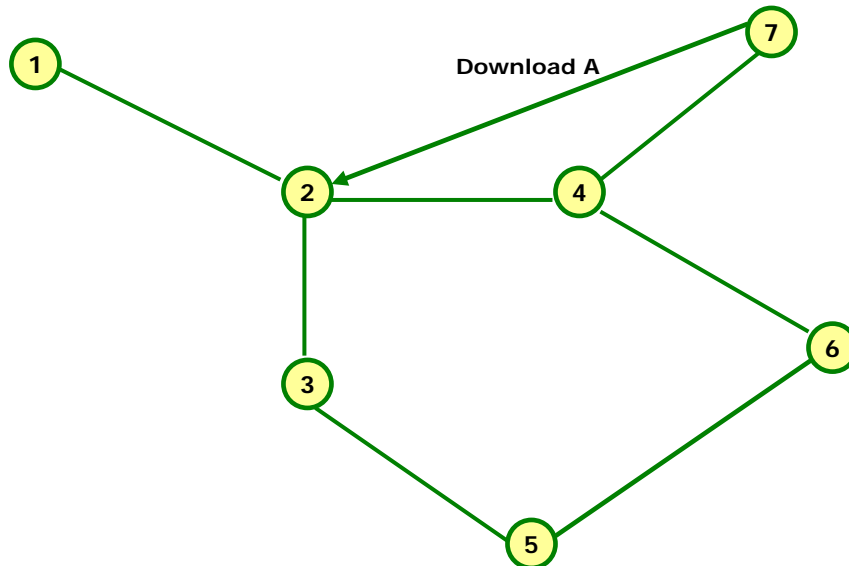


 Peer

Steps:

1. Peer 2 initiates search for file A.
2. Forwards Query request to all its neighbors.
3. File A not found, forwards Query request.
4. File A found in Peer 7.
- 5. Query reply message is backpropogated.**

Gnutella Search Mechanism



 Peer

Steps:

1. Peer 2 initiates search for file A.
2. Forwards Query request to all it's neighbors.
3. File A not found, forwards Query request.
4. File A found in Peer 7.
5. Query reply message is backpropogated.
- 6. File download.**



Gnutella Open Problems

- Scalability
 - *the number of queries and the number of potential responses increases exponentially with each hop*
- Download failure
 - *frequent download failure due to nodes crashing or disjoining*
- Security
 - *nothing to prevent tempering in a gnutella network*





Proposed Architectures

- Towards these problems two new system architectures are proposed
 - *1-layer NGnu*
 - *Single layer is added into the original gnutella architecture*
 - *2-layer NGnu*
 - *Double layer is added into the original gnutella architecture*



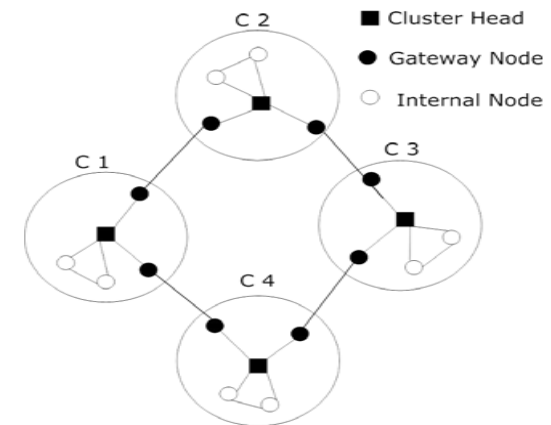
Objectives of the Proposed Architectures



- provide a *distributed file sharing solution*
- *increase scalability*
- *minimize download failures*
- *provide security*

1-layer NGnu Architecture

- The existing nodes (computers) are grouped into a set of clusters.
- In each cluster, there are three main types of nodes
 - the cluster head
 - gateway nodes
 - internal nodes
- The cluster head acts as a local coordinator of transmissions within the cluster.
- Gateway nodes located at the boundary of the cluster, communicates with nodes outside the cluster.
- Each node in the cluster has their unique identifier.
- When a new internal node joins the network, it will register itself to the cluster head.
- The cluster head keeps tracks of all internal nodes within its responsible cluster.



1-layer NGnu Search Mechanism

- A node sends out a query request to its neighboring nodes by a set TTL value.
- If the query file is not found or time out in the local cluster the request is forwarded to cluster head.
- The cluster head forwards the query to its neighboring cluster heads via gateway node.
- The cluster head sends out the query request to the local cluster nodes.
- If the requested file is found the request node receives a reply message via the cluster head & gateway node.
- A direct connection is established and the file is downloaded.

```

1: EXECQUERY(query, respond)
2: begin
3:   pid = GETPID();
4:   if( FINDOBJECT(query) ) then
5:     respond = {pid};
6:   else
7:     if (TTL < threshold) then
8:       nlist = GETNEIGHBOR(pid);
9:       foreach nid in nlist
10:        EXECQUERY(query, respond1)@nid;
11:        respond = respond ∪ respond1;
12:       else
13:         respond = ∅;
14:       end if
15:     end if
16:   if( ISCLUSTERHEAD(pid) AND
17:      respond == ∅ ) then
18:     clist = NEIGHBORCLUSTER(pid);
19:     foreach cid in clist
20:       EXECQUERY(query, respond1)@cid;
21:       respond = respond ∪ respond1;
22:     else
23:       respond = ∅;
24:     end if
25:   end

```

```

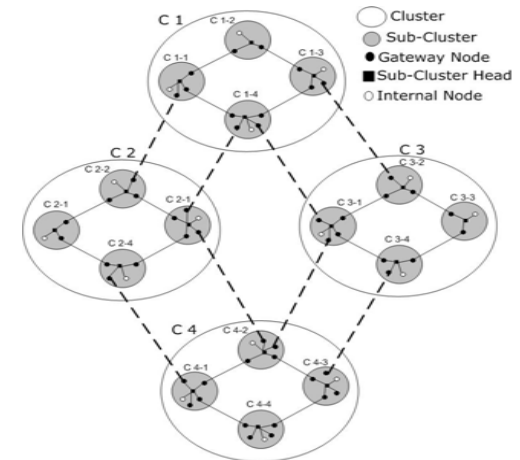
26: INITIALQUERY(query, respond)
27: begin
28:   pid = GETPID();
29:   if( ISCLUSTERHEAD(pid) ) then
30:     EXECQUERY(query, respond)@pid;
31:   else
32:     cid = CLUSTERHEAD();
33:     EXECQUERY(query, respond)@cid;
34:   end if
35: end

```



2-layer NGnu Architecture

- The 2L-NGnu architecture aims to provide same functionality similar to that of the 1L-NGnu but more scalable with one more layer added.
- A cluster is organized in the form of a 1L-NGnu architecture.
- There are a sub-cluster head, gateway nodes and the internal nodes for each sub-cluster.
- In the outer layer (layer 2), a system is virtually formed by a number of clusters.
- Sub-clusters and clusters are organized by taking geographical locations into account.
- In each sub-cluster, the sub-cluster head node acts as a local coordinator of transmissions within the sub-cluster.



2-layer NGnu Search Mechanism

- A node sends out a query request to its neighboring nodes by a set TTL value.
- When the requested file is not found or time out, the request is sent to the sub cluster header.
- The sub cluster header forwards the request to its neighboring sub cluster headers.
- The neighboring sub cluster headers execute the query request in its local sub cluster.
- If the requested file is not found or time out, the requested file is sent to the neighboring sub cluster header belonging to a different cluster.
- If the object is found in the neighboring cluster the object is shared by a *virtual link* between the two internal nodes.

```

1: EXECQUERY(query, respond)
2: begin
3:   pid = GETPID();
4:   if ( FINDOBJECT(query) ) then
5:     respond = {pid};
6:   else
7:     if (TTL < threshold) then
8:       nlist = GETNEIGHBOR(pid);
9:       foreach nid in nlist
10:        EXECQUERY(query, respond1)@nid;
11:        respond = respond ∪ respond1;
12:       else
13:         respond = ∅;
14:       end if
15:     end if
16:   if( ISSUBCLUSTERHEAD(pid) AND
17:      respond == ∅ ) then
18:     clist = NEIGHBORCLUSTER(pid);
19:     foreach cid in clist
20:       EXECQUERY(query, respond1)@cid;
21:       respond = respond ∪ respond1;
22:     else
23:       if ( ISBOUNDARYNODE(pid) ) then
24:         glist = GATEWAYNODE(pid);
25:         foreach gid in glist
26:           EXECQUERY(query, respond1)@gid;
27:           respond = respond ∪ respond1;
28:         else
29:           respond = ∅;
30:         endif
31:       end if
32:     end

```

```

33: INITIALQUERY(query, respond)
34: begin
35:   pid = GETPID();
36:   if( ISSUBCLUSTERHEAD(pid) ) then
37:     EXECQUERY(query, respond)@pid;
38:   else
39:     cid = SUBCLUSTERHEAD();
40:     EXECQUERY(query, respond)@cid;
41:   end if
42: end

```





Simulation Model

- Each system Gnutella, 1L-NGnu and 2L-NGnu are simulated as a separate process.
- We created a process that emulates the behavior of peers leaving and joining the network.
- A peer continuously services incoming requests and performs appropriate actions based on the request type.
- Each peer has an unique identifier and contains about 20 files for sharing.
- All the peers are connected randomly with atleast 2 connections.
- One or more peers can send a query request at the same time, so it is capable of handling multiple query sessions.





Simulation Settings

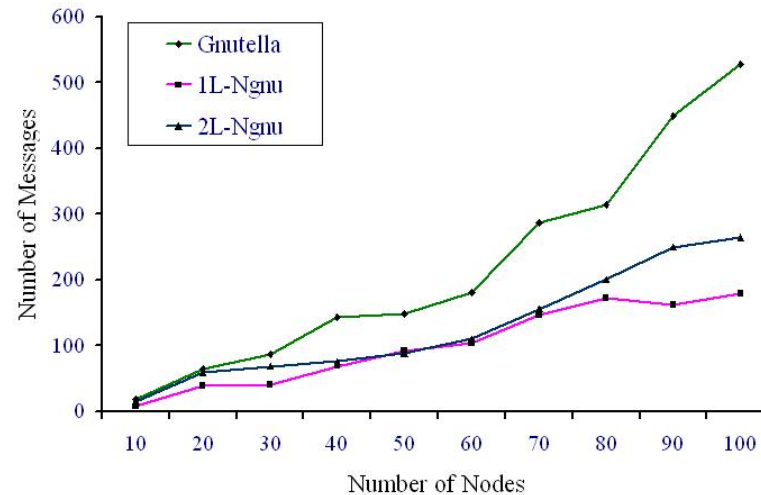
- There are 10000 nodes with identical capabilities participating in the network.
- The Time-To-Live (TTL) value can be set to any value from (1 – 10).
- Two alternative query methods are used
 - random query
 - static query
- The simulation duration can be set to any length



Simulation Comparison Case 1

The total number of network messages is examined

TTL=7, Time= 10 ms



It can be observed that as the number of nodes increases the network traffic in Gnutella is much higher than that of 1L-Ngnu or 2L-Ngnu throughout the entire simulation.

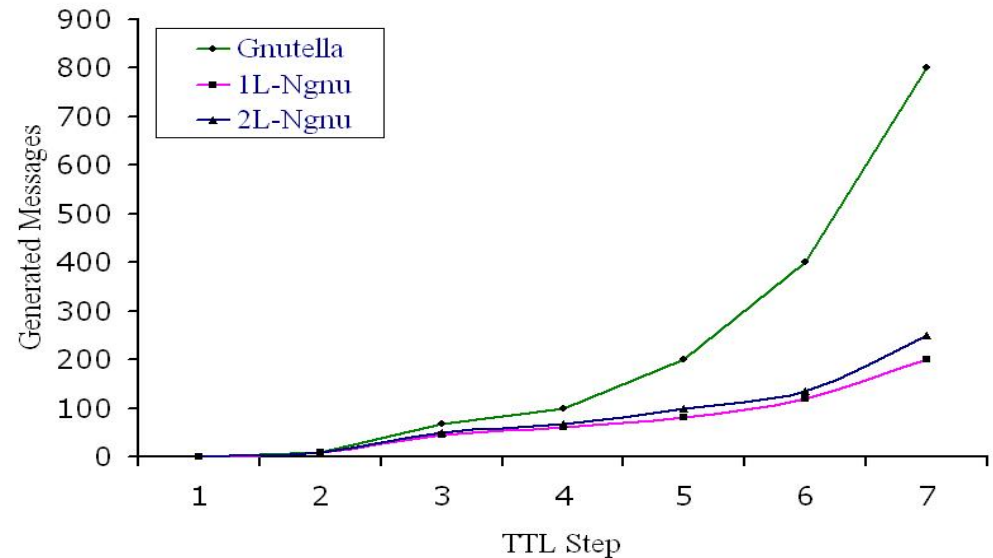
Simulation Comparison

Case 2

The total number of network messages is examined by varying the TTL

Time = 10 ms

- All the three systems have almost the same number of messages when the TTL is set to low.
- However, when the TTL value is set to a larger number, it is observed that the Gnutella network generates highest number of messages.



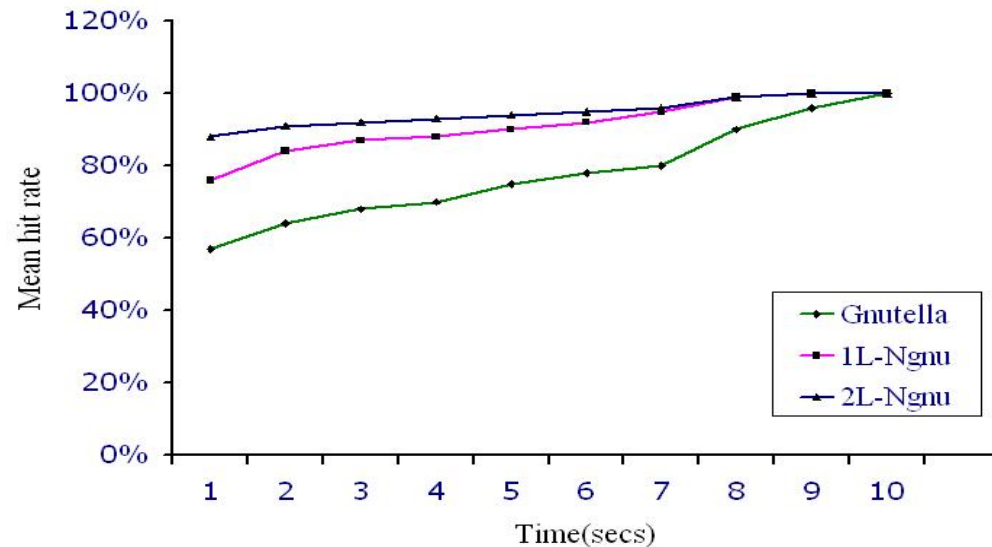
Simulation Comparison

Case 3

Comparison of query hit rate percentages

TTL = 7, No. of nodes=10000, Queries=300

- The query hit rate is recorded by varying the time to run the simulation.
- Initially, the Gnutella gains the lowest query hit rate of 60%, 1L-Ngnu has query hit rate of 80% and 2L-Ngnu has a query hit rate of nearly 90%.
- When the simulation duration is set longer, the query hit rate increases for all the systems and finally their query-hit rates are saturated to 100%, especially for 10 secs.



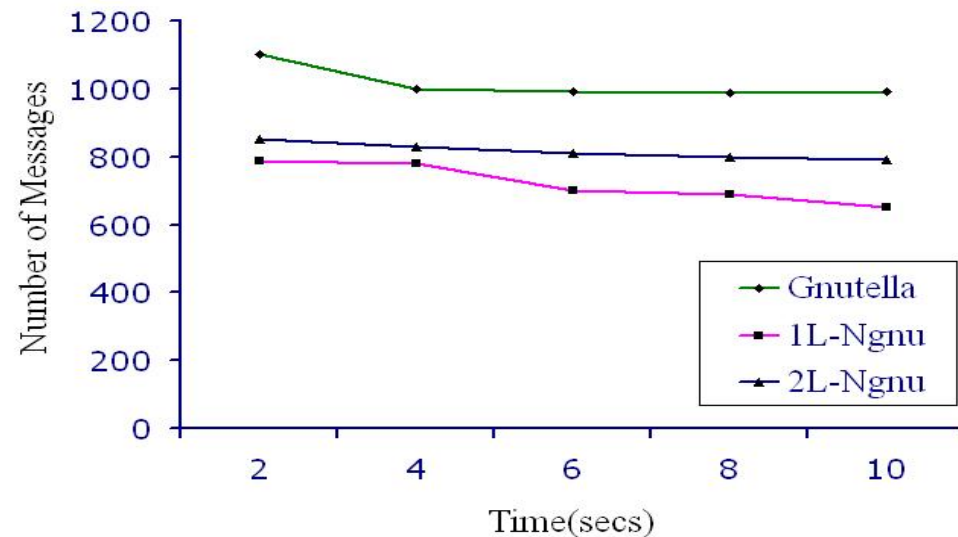
Simulation Comparison

Case 4

Message Comparison of a stable network

TTL = 7, No. of nodes=100, Queries=300

- It is observed that except for the initial state all the systems have almost constant number of messages when the simulation is recorded with various runtime durations.
- This shows that the traffic is almost stable and independent of the time passed.





Conclusions

- Two peer-to-peer (P2P) system architectures for file sharing mechanism, called 1L-Ngnu and 2L-Ngnu, are presented.
- In the architectures, multi-layer mechanism was introduced to extend a well-known P2P system named Gnutella.
- These system architectures provide more scalable alternatives to existing Gnutella algorithms, focusing mainly on the search and replication aspects.
- By simulation experiments, it was found that the total number of network messages in 1L-Ngnu and 2L-Ngnu is considerably less than Gnutella.
- The systems provides security by registering the users and secure key technology.
- The 1L-Ngnu and 2L-Ngnu architectures can be implemented as a part of file-sharing applications such as e-government and e-library.



Questions & Suggestions

