To Improve the Searching Efficiency on P2p Network System

By Shou-Fon Wu
06/03 2004
Outline

• P2p searching mechanism.
• The problem of p2p searching.
• To solve such problem.
• The concept of group.
• Selectively searching.
• Discussion.
• Conclusion.
P2p Searching Mechanism

- Poor p2p(Unstructured):
  - Flooding
  - Broadcast
P2p Searching Mechanism

• Hybrid p2p:
  - Need a central server to receive query message.
  - The central server helps the requesting peers to find supplying peers.
The Problem of P2p Searching

• The number of message may be too large.
• Too many redundant messages will cause the inefficiency of the system.
To Solve Such Problem

- Limit the searching range.
- Reduce the number of redundant message.
- Two kinds of method:
  - Group
  - Selectively searching
The Concept of Group

- Limit the searching range in a group.
- Assign each group a server (super peer).
- Construct another overlay called “super peer overlay” to perform the communication between different groups.
- How to form peer groups?
  - Different methods result in different “group characteristic”.
(1) Limit the Number of Hops

• Define an upper bound $k$ which indicates the number of hops.
• In a group, each peer can communicate with others by no more than $k$ hops.
• It is convenient for peers in the same group to communicate.
(1) Limit the Number of Hops
(2) Hash Function

- Each peer provide a “key” value.
- Design a hash function and use the key value as its input.
- The output of the function decide which group a peer would join.
- The function maps a peer into a group.
(2) Hash Function

A peer join

Hash Function

Output: 1..k

Group 1

Group 2

Group k
(3) Interest Group

• Peers with the same interest (file type) form a group.
• A peer may belong to multi-group.
• We can do searching in some specific groups depend on the information type we want.
• Each group has its own “characteristic”.
(3) Interest Group
Comparison

• Limit the number of hops:
  Easily connect within a group.
• Hash function:
  No special idea but easily compute.
• Interest group:
  It is convenient for a peer to search the information of the same type but is a little hard to implement.
Selectively Searching

- (1) Ranking
  - To reduce the connection complexity.
- (2) Reputation
  - To provide peer information that indicates which peer is better than others.
Ranking Mechanism

• Each peer compute the ratio of query hits and query message.
• By the ratio computed above, each peer can change its connection.
Reputation System

• Compute a reputation score for each peer.
• The higher score of a peer means that we can find the data we want in this peer with higher probability.
• The reputation system is dynamical (The reputation score changes rapidly).
Reputation System

• To compute the reputation score:
  U: Upload credit
      Increase the reputation score.
  D: Download credit
      Decrease the reputation score.
  S: Share credit
      Increase the reputation score.
Discussion

- Comparison:
  - Ranking: for some peer
  - Reputation: for all peers

- Both ranking and reputation reduce the connection complexity.
Conclusion

• To improve the searching efficiency:
  - Group:
    Limit the number of hops
    Hash function
    Interest-based
  - Selectively connect:
    Ranking
    Reputation system
Reference