

Peer to Peer Applications: Napster, Gnutella

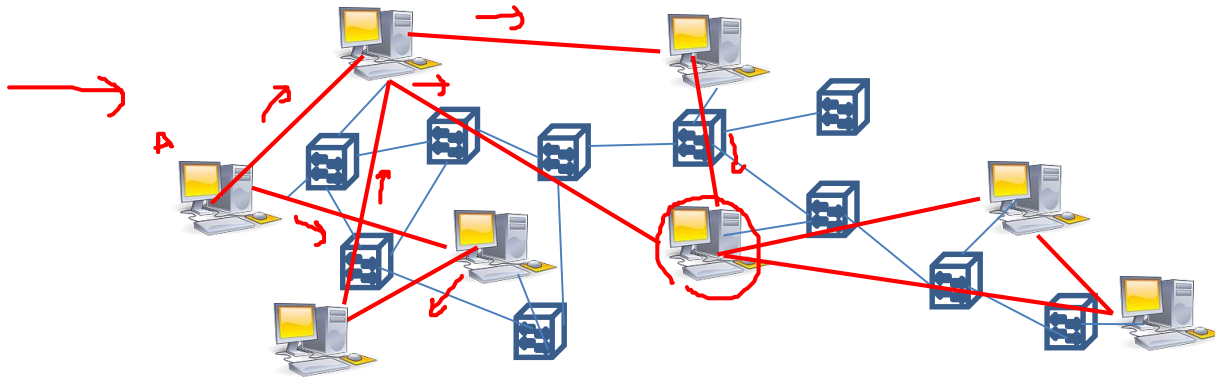
Kameswari Chebrolu

All the figures used as part of the slides are either self created or from the public domain with either 'creative commons' or 'public domain dedication' licensing. The public sites from which some of the figures have been picked include:

<http://commons.wikimedia.org> (Wikipedia, Wikimedia and workbooks); <http://www.sxc.hu> and <http://www.pixabay.com>

Gnutella

- Basic Idea: Eliminate central server; peers form an overlay network
 - Peers search and fetch themselves
 - Search performed via flooding the overlay

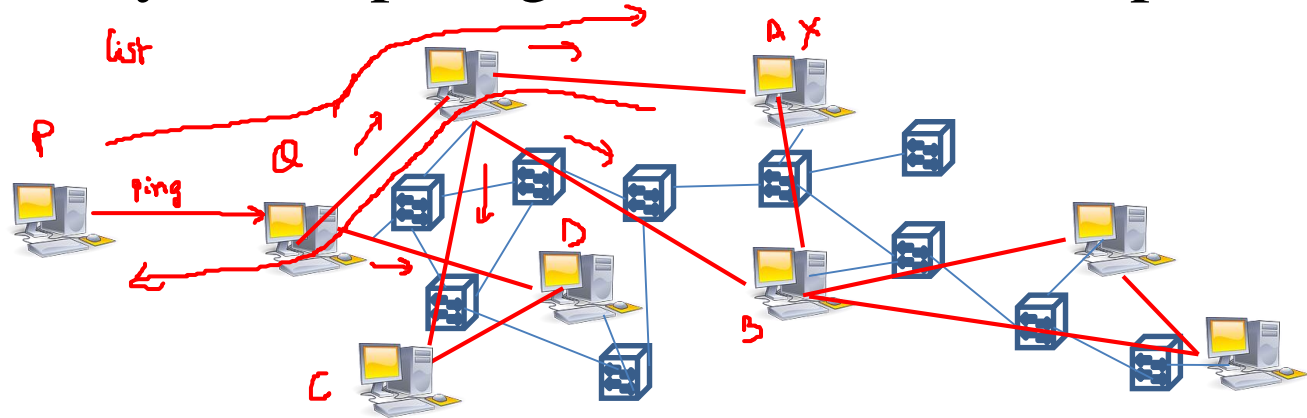


Join

- To join the overlay network, new peer P needs to find some other peers to connect to
 - Use a ^{server} bootstrap node to get IP address/port info of existing Gnutella nodes
- P attempts to connect with peers on the list using TCP until it succeeds (say it connects to Q)
- P sends a “Ping” message to Q; Q floods it into the overlay (TTL restricted)

flood TTL → no of hops
3

- All peers receiving Ping message respond with Pong message
 - Pong messages sent over reverse path
- P receives many Pong messages. It can select a subset and setup additional TCP connections
- Peers ping periodically to keep neighbour list fresh in spite of churn



Publish

- No Need to support

Search

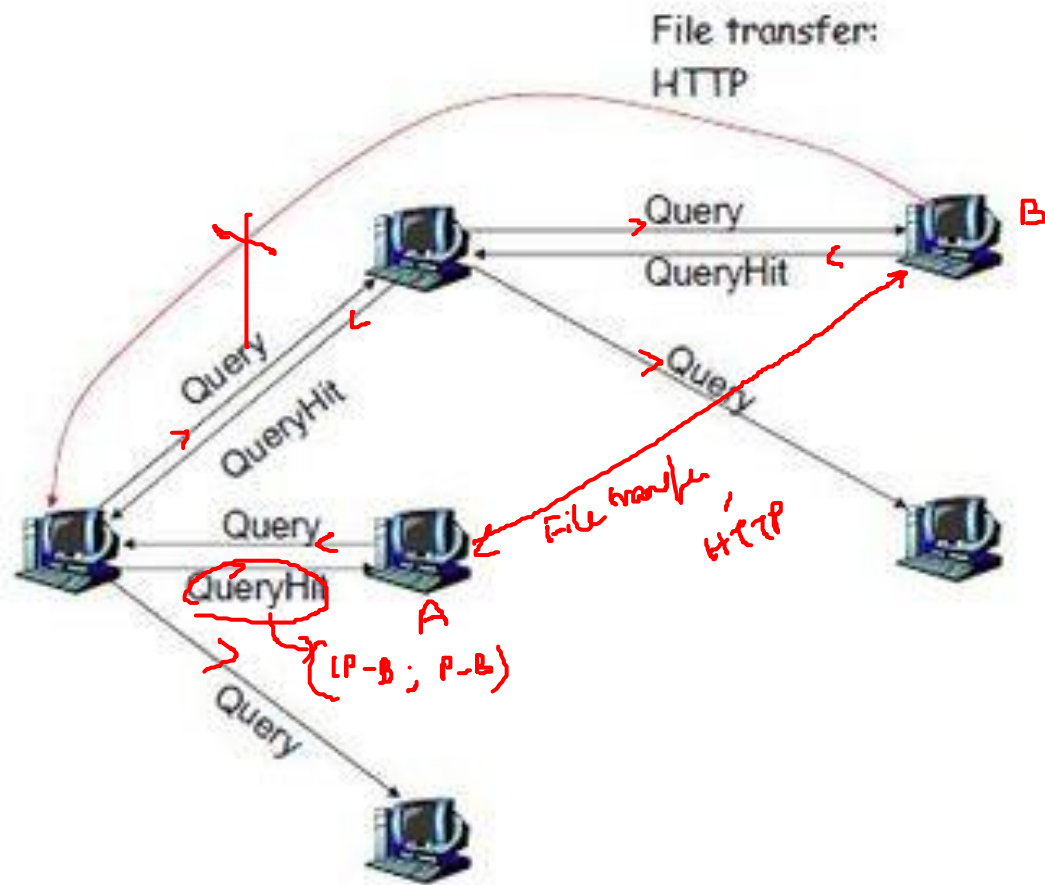
- Based on query flooding
 - Query messages are flooded in the overlay (much like ping messages) over the existing TCP connections
 - TTL restricted
 - TTL restricted (limited scope flooding) and forwarded only once
- Successful hit (QueryHit) sent on reverse path (much like pong messages)
 - QueryHit contains info (IP-addr, port, speed) of responder
 - 50Kbps
 - 1Mbps

Fetch

List

speed

- Searcher can choose “best” QueryHit responder and download file directly
- File download based on HTTP
 - HTTP is a widely used standard



Discussion

- Pros:

- Fully distributed (cost of search/processing shared by peers)
- Difficult to pull the plug

- Cons:

- Flooding can cause excessive traffic. *limited scope TTL*
- Unreliable search due to limited scope flooding and network churn
- No mechanism to handle freeloaders *→ download no sharing*

Summary

- Looked at two P2P systems: Napster and Gnutella
- Napster employs a centralized server to facilitate search while Gnutella is fully distributed and employs flooding
- Each system has its own set of pros and cons
- Both faced legal hurdles. Napster fully shut down while Gnutella lost popularity