CS314 2011 31

Peer-to-Peer (p2p) Networks

- Overview
- BitTorrent
- Skype (separate slide set)

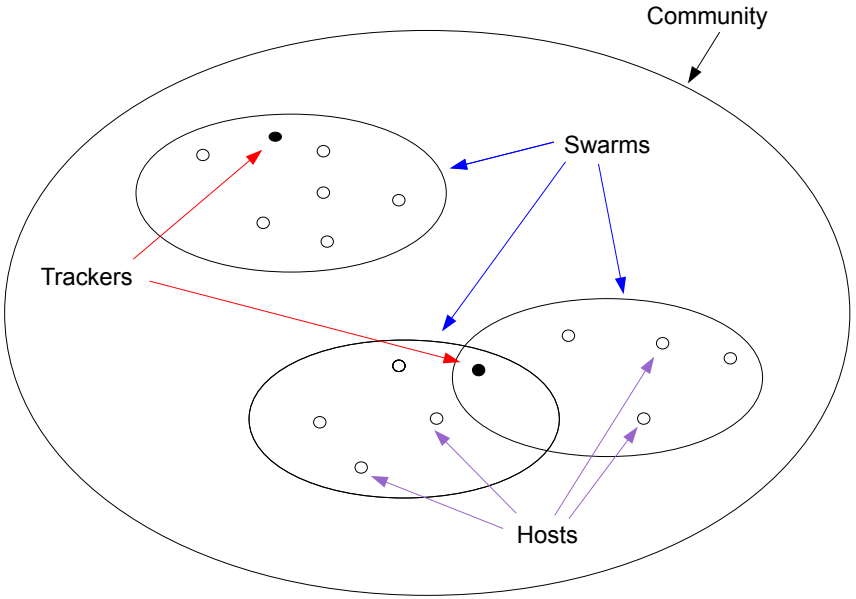
P2P Basics

- Peers cooperate to form an overlay network
  - At the *Application* layer
  - Using IP protocols to communicate (*network* layer)
- Their goal is to *share resources* between peers, e.g.
  - Storage (files or sets of files)
  - Transmission paths ('phone calls)
- They may have a shared directory system, e.g. using *distributed hash tables*
- We look at two examples
  - BitTorrent: file sharing
  - Skype: audio and video telephony

BitTorrent

- Shares files amongst *peers* that run implementations of the BitTorrent protocol (i.e. BitTorrent clients)
  - Files are broken into small pieces for transmission
  - The pieces can be downloaded in parallel from many peers
  - A peer that has the whole file is a *seed* peer
  - A *tracker* is a host that keeps track of other peers have that have pieces of one or more files
  - The peers a tracker is watching is called a *swarm*
  - *Communities* are based at sites that provide portals to one or more swarms

BitTorrent Universe



Sharing: 'tit-for-tat' principle

- A peer must share files with others so that it can go on downloading
  - Peers that don't are called leechers
  - Many client implementations try to enforce tit-for-tat, e.g. ranking peers by the amount they share
  - Ranking can be unfair to new users. Clients can be given some credit to get them started. That's called *opportunistic unchoking*

To share a file or set of files ...

- A user creates a *.torrent file* and makes it public
- A .torrent file contains
  - A unique identifier
  - Names and sizes of files to be shared
  - Piece hashing information
    - Each piece carries a cryptographic hash to protect its integrity
  - Address of one or more trackers
- The torrent is advertised on a community – usually via a web site

BitTorrent Summary

- Very widely used, e.g. for distributing software releases
- Many different clients available
  - Many use TCP
  - µTorrent uses UDP, and does its own congestion management
- Allows users to download large files faster than would be possible from a single server
- Should help ISPs by reducing the size of servers and amount of download bandwidth, but ..
- Means that ISPs need to provide symmetric link capacities