CS 554: Peer-to-Peer Systems

Mema Roussopoulou

Fall 2007

University of Crete

Class Logistics

- Professor: Mema Roussopoulou
  - Office: G 215
  - Office hours: Mondays/Wednesdays, 2:00-3:00 pm
- Teaching Assistants:
  - Dimitris Antoniades, Elias Athanasopoulos
  - Office: TBD
  - Office hours: TBD
- Class Mailing List: hy554-list@csd.uoc.gr
- Course info on readings, office hours, etc. at:

http://www.csd.uoc.gr/~hy554

Readings, Discussions, Reviews

- Read 1-2 papers per class
  - Write a review (in specified format)
  - E-mail to *hy554@csd.uoc.gr* **before** class begins
- Participate in class discussions on the papers
  - Be prepared to discuss questions/comments in detail
  - Send email before class with questions if shy asking about a particular detail
  - NO LAPTOPS ALLOWED (unless you are a speaker that day)
  - Print out papers and bring them to class

Paper Presentations

- You will present one or more papers in the class
  - 25-30 minute presentation
  - Explain overall contributions of the paper
    - What problem does the paper focus on?
    - Why is this problem important?
    - What solution do the authors propose?
    - How do the authors evaluate the solution?
  - Lead the class discussion on the paper
    - This is an interactive class – you are not meant to drone on
    - Solicit opinions from others
    - I will help you
  - Send presentation day preferences for October, November, and December to hy554@csd.uoc.gr be assigned a paper

Research Project

- Project logistics
  - Groups of two-three
  - Problem you work on will be your choice from a list of problems (based on papers we've read)
  - Evaluate solution and write a paper about what you learned
  - Groups formed, choices made: by November 14
  - Project proposals due: November 30
    - Short (3 pages max) description of problem (given to you), why you think it's worth studying, and how you will approach solving it (i.e., the work plan – who will do what by when?)
  - Status reports due: December 21
    - Short (1-2 pages) describing what you've done so far and your plan for completion

Research Project (cont'd)

- Final code + paper due: January 25
  - 10-12 pages long;
  - Should read like typical conference paper
  - Could form the basis for a future conference paper
- We will start early so you aren't panicking at the end! ☺
  - Ideal outcome: project efforts lead to a publication (I can help you write the paper)
  - Great outcome: we learn something from your efforts
- We will discuss progress in class to get feedback from others (short, very informal presentations)

GOAL: Gain some experience doing research

Grading

- Reviews: 10%
- Paper Presentations 20%
- Class Participation: 10%
- Project: 60%

What are Peer-to-Peer Networks?

- P2P concept: leverage idle resources
- Definition:
  - Self-organizing distributed system
  - Nodes provide and receive services in cooperative effort
  - Nodes have equal functionality
- Features:
  - Scalability
  - Availability
  - Fault tolerance
  - Decentralized administration, etc.

Properties of P2P Systems

- Distributed over lots of sites
    - exploit existing (often free) resources
    - share costs among many
- Sites are autonomous (different owners)
    - they are both clients and servers
- Assumes peer cooperation
    - although people are not nice
- Sites have equal functionality
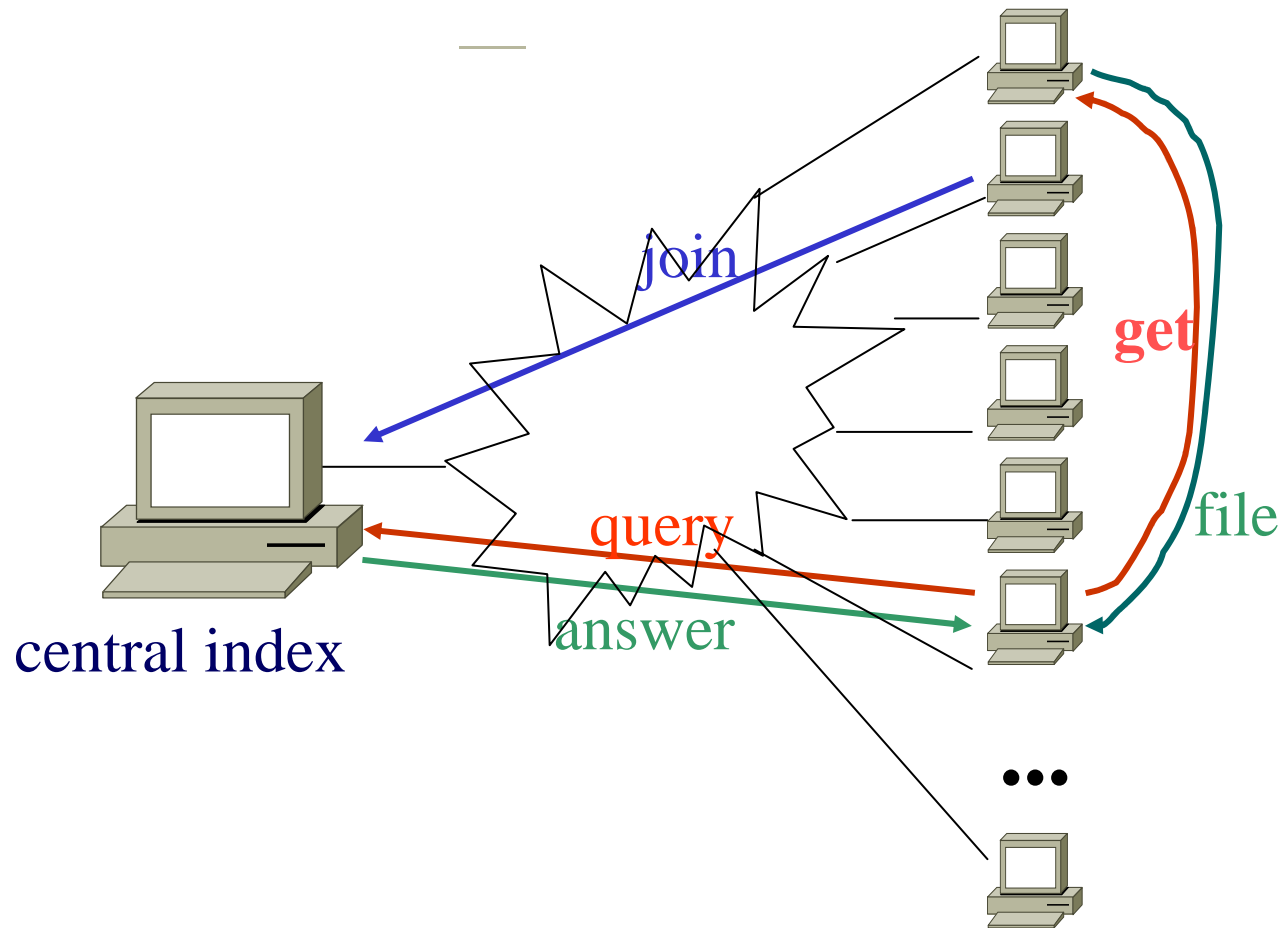- Anonymity and legal protection is desirable

Various P2P

pastry can jxta fiorana

napster united devices freenet

aim ocean store open cola netmeeting

gnutella farsite skype

morpheus icq ebay

limewire seti@home

uddi bearshare

grove jabber popular power

kazaa folding@home tapestry

process tree mojo nation

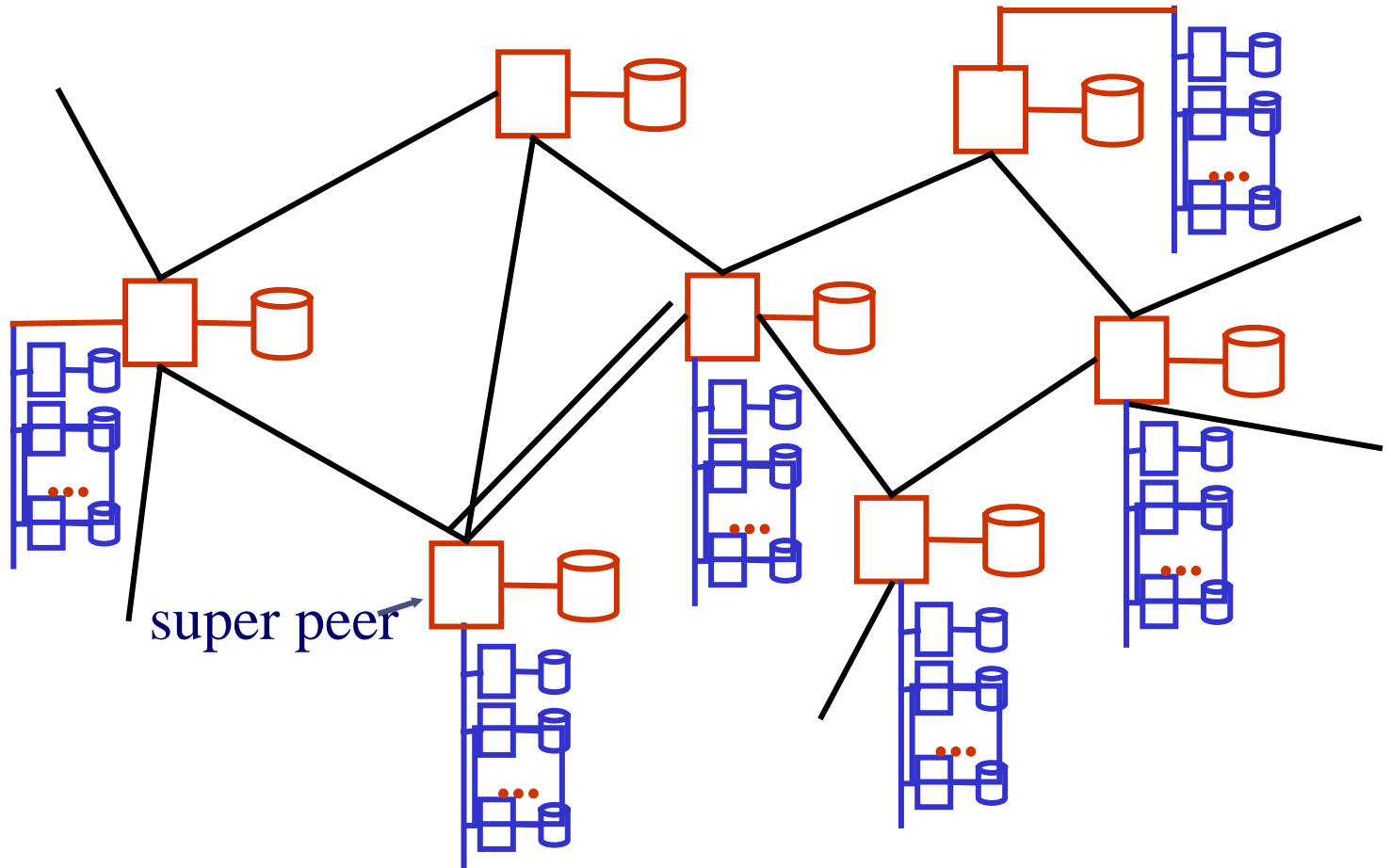chord bittorrent

edonkey

Taxonomy 1: Shared Resources

- Files
  - Napster, Morpheus, Kazaa, etc
- Indexes
  - Distribution entries: CAN, Chord, Tapestry, Pastry
  - DNS
- CPU cycles
  - Seti@home
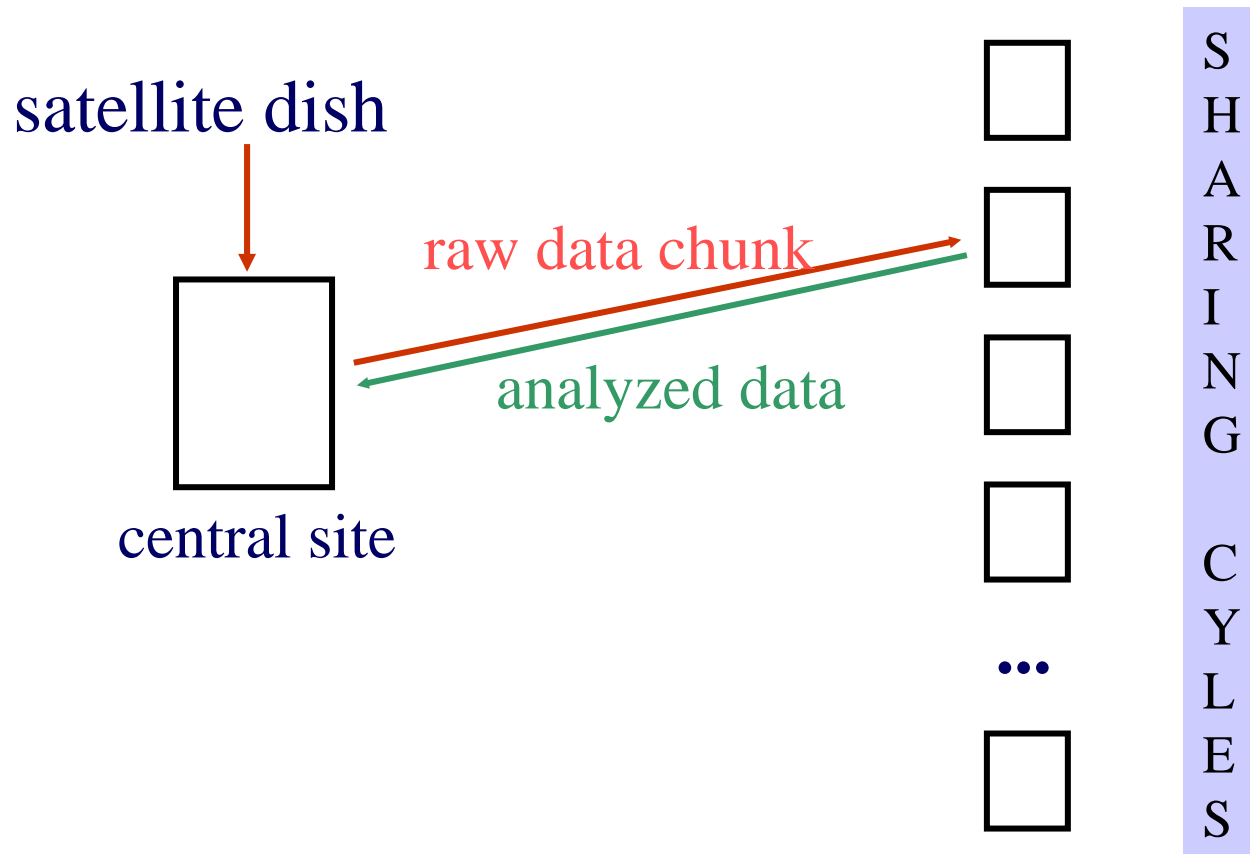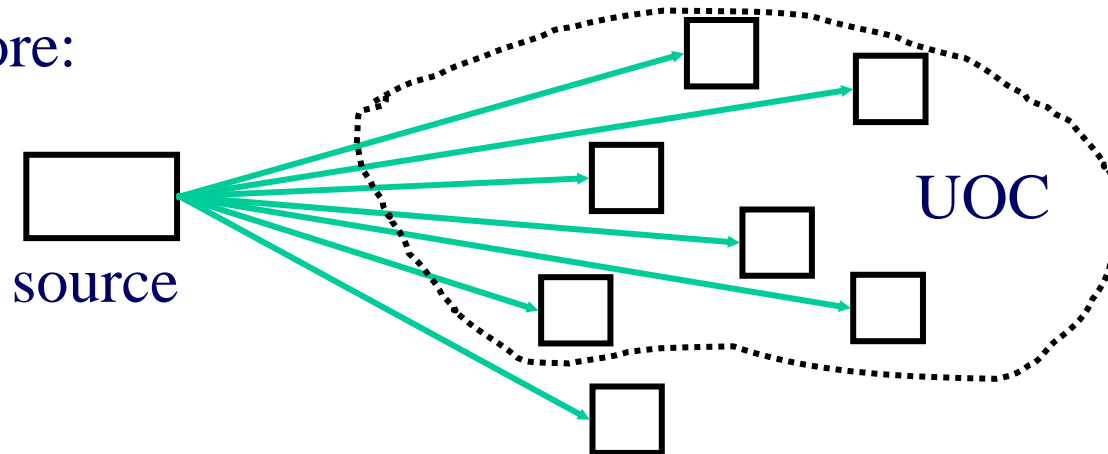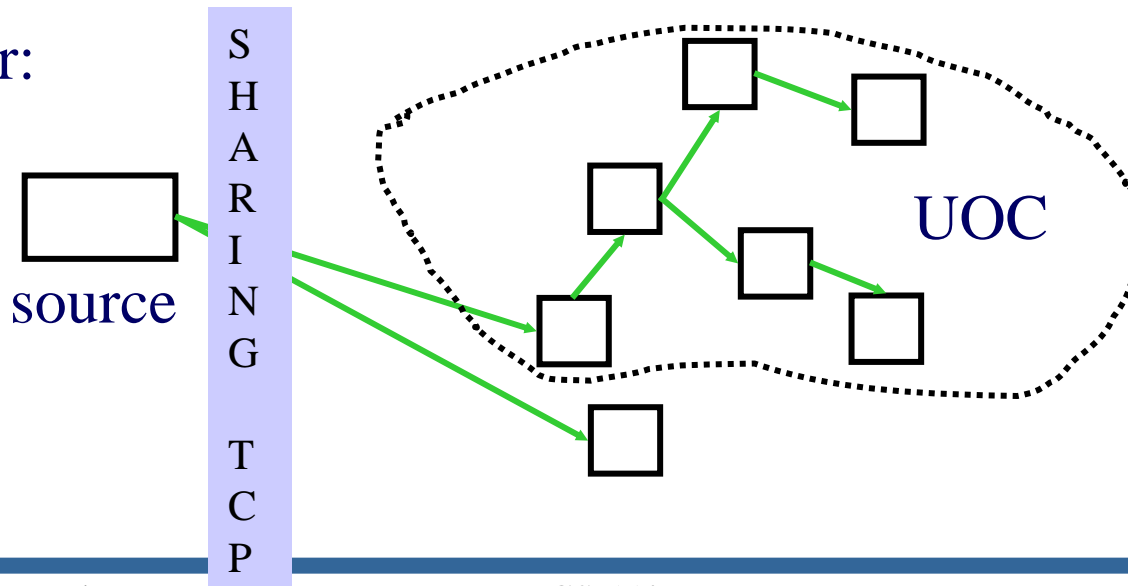- TCP Connections
  - PeerCast

Napster†

join

get

query

answer

file

central index

• • •

Morpheus

super peer

Seti@Home

satellite dish

central site

raw data chunk

analyzed data

S H A R I N G   C Y L E S

PeerCast

before:

source

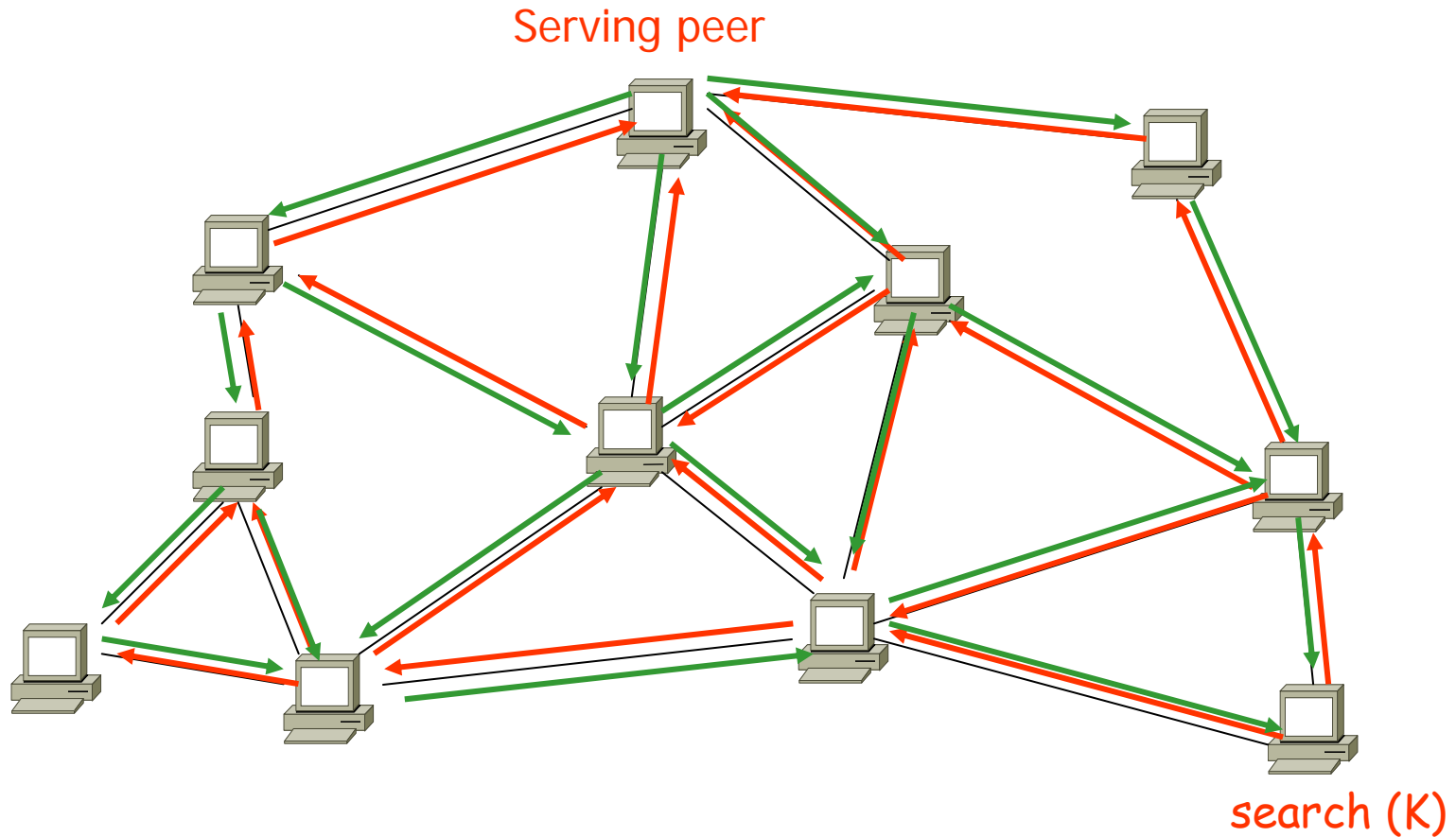UOC

after:

SHARING TCP

source

UOC

Taxonomy 2: Types of Search

- Index entry:
  - (key, value) pair
  - Key = name of content
  - Value = IP address of serving peer
- Unstructured search
  - Gnutella, FreeNet
- Structured search
  - CAN, Chord, Pastry, Tapestry

Gnutella Search

Unstructured Search
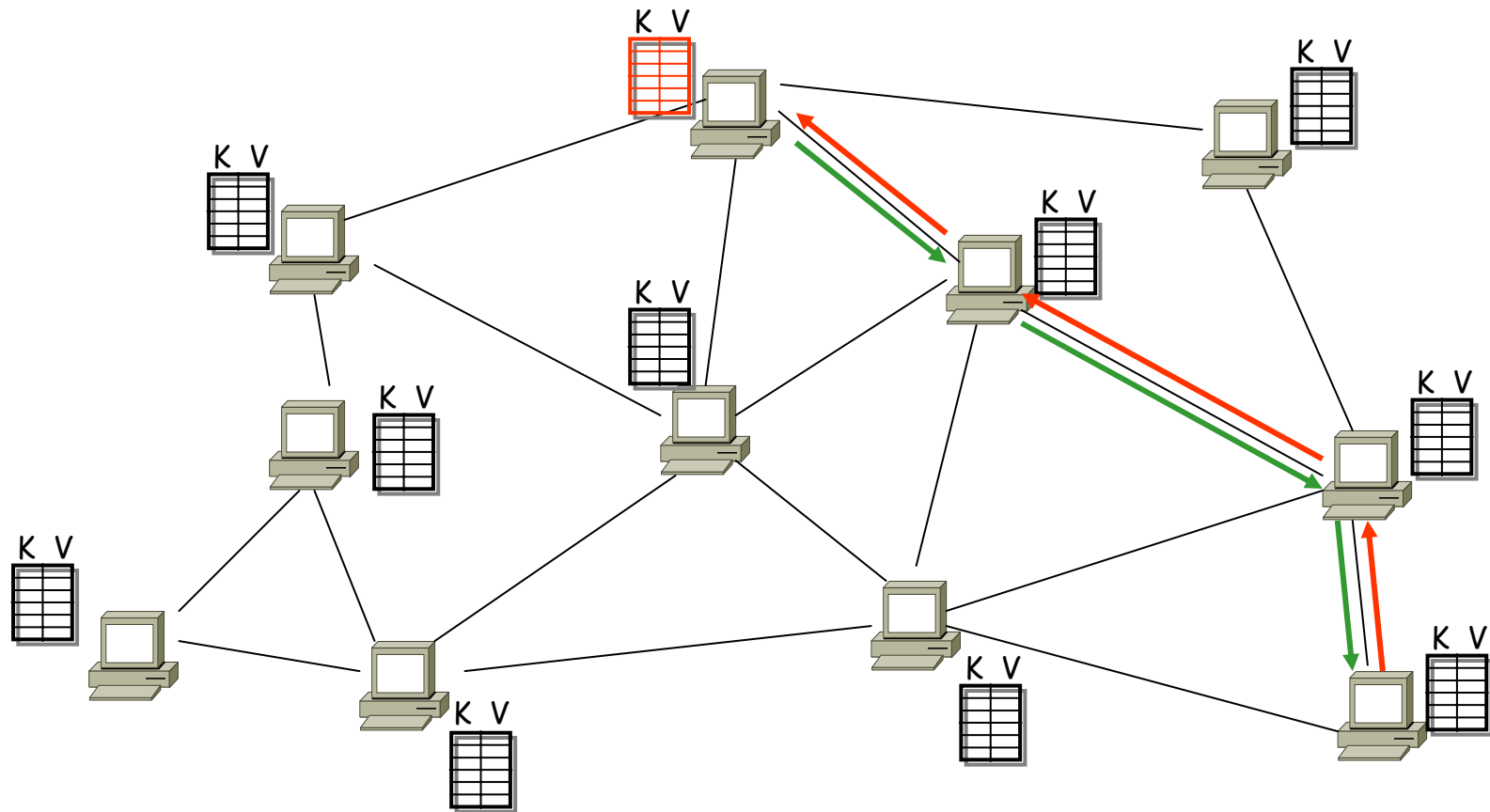
Serving peer

search (K)

Advantages & Disadvantages

- Allows search on multiple keywords +++
- NO global index information is needed ++
- Causes network flooding, wastes network ----
- Inefficient search (wander around until you find the content, repeated everywhere) ---

Content Addressable Network (CAN)

Structured Search

search (K)

Advantages & Disadvantages

- No network waste (no flooding) ++++

- Efficient single key (file name) search +++

- Overhead in maintaining not only network but distributed hashing consistent -

- Inefficient multiple keyword search --

P2P Challenges

- Routing
- Resource Management
- Security & Privacy
  - Slap down malign peers
  - Work out legal issues
- Providing incentives
  - Don't rely on peers being cooperative
- Identifying when P2P is a good solution, and when not

Writing a Paper Review

- Provide a short summary of the paper (1-2 paragraphs)
- List the strengths of this paper (in bullet form)
- List the weaknesses of this paper
  - Justify why you think these are weaknesses
    - E.g., "Your assumption that as long as a host remains up, it has the same IP address is incorrect because… This assumption can thus color the results of your measurement of uptimes of peers."
  - Provide concrete suggestions for improving the paper
    - E.g., "You should run your protocol for a longer period of time to adequately test the XYZ aspect of your protocol."
    - E.g., "You are missing citation XYZ from your related work section"
    - E.g., "You should report some statistics for your metrics (e.g., standard deviation) to give a better sense of the distribution of your results."

Sample Reviews

#1 "This paper reads like you are auditioning to be George W's science advisors! It is sensationalist and riddled with grammatical errors and malapropisms. (" ...results in such a dramatically condensed store that…") what language were you trying for with that little repeated sound bite?? Is reminds me a long lost friend of dopey, bashful and doc?)

#2 "PLEASE! Have some respect for your peers. You are not the be-all, end-all researchers in XXX XXX."

Avoid being emotional. It's one step close to insulting.

Sample Review

#3 "The paper considers the problem of (extremely!) long storage of objects in the face of machine failures, bit-rot, and malicious peers trying to convince you that valid copies of objects are invalid. The motivation is quite compelling (library archival storage) and a prototype is apparently run a limited production network.The intellectual meat of the paper is a secure voting protocol wherein a node can convince itself of the validity of its copy of an object through opinion polls with its peers. The protocol addresses confederations of malicious peers that conspire to vote against valid copies."

A good review benefits both the PC AND the authors.

Sample Review

#4 "In terms of originality, this paper doesn't propose significant new techniques. I believe that Algorithm XXX already appears in the distributed systems literature."

Provide citations when claiming "this is old stuff."

Next Classes

- Wed 3/10:
  - "A Scalable Content Addressable Network"
- Mon 8/10:
  - "Chord: A Scalable Peer-to-Peer Lookup Service"
  - "Serving DNS Using a Peer-to-Peer Lookup Service"

Before Next Class

- Read the CAN paper

- Write and send to us your review of the paper BEFORE class

- Send us by Friday, Oct 5, three days on which you would like to present papers (one in October, one in November, and one in December)

- Join the class mailing list (see web page to learn how)