Overlay (and P2P) Networks

Part II

- Complex Networks
- Applications of Overlay networks
- Advanced Topics
- Summary

Samu Varjonen

Ashwin Rao

Schedule

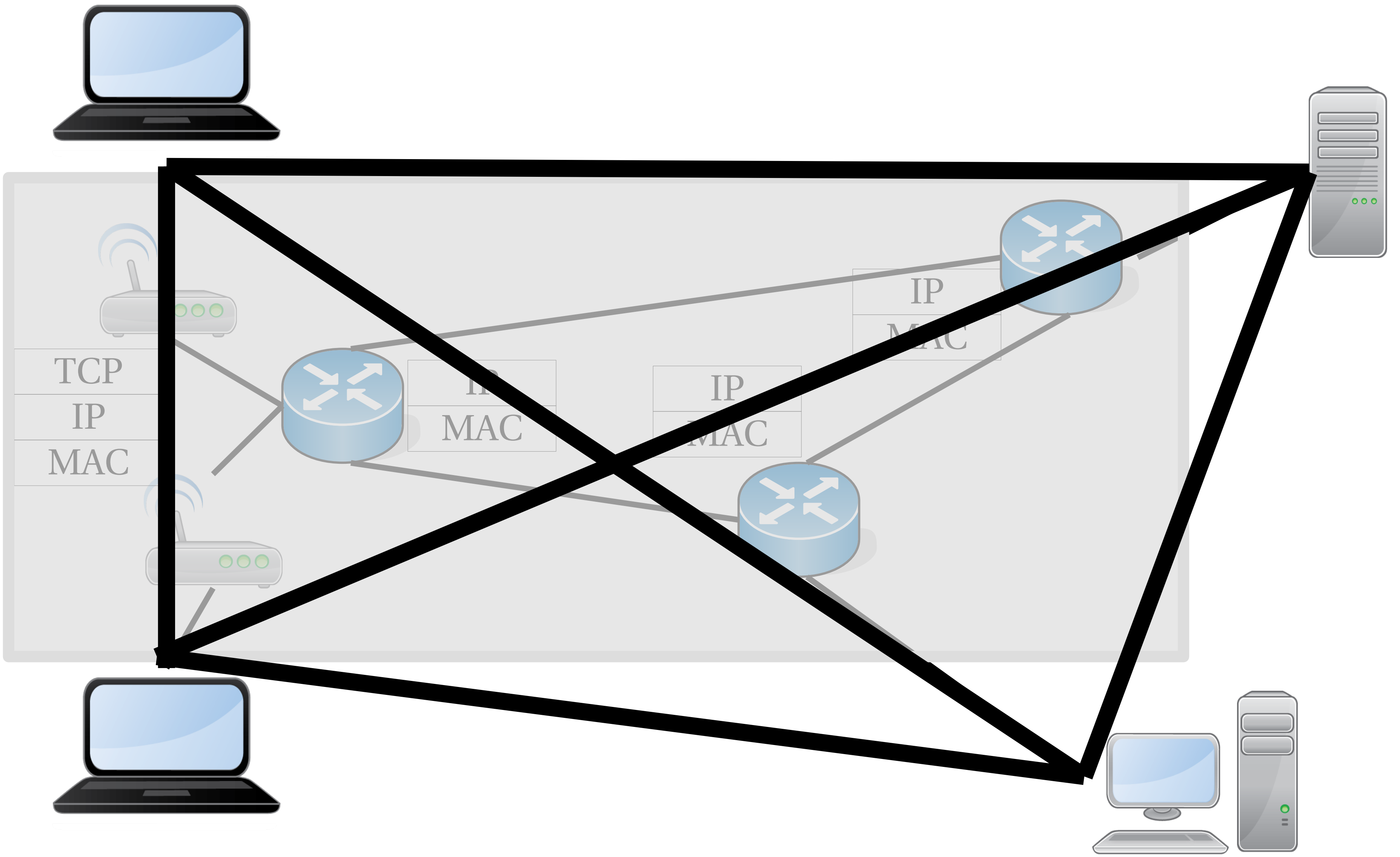
- Complex Networks 11.02
- Complex Networks 15.02
- Complex Net & Apps 18.02
- Applications 22.02
- Advanced Topics 25.02
- Conclusion Summary 29.02

- Why Complex Networks?
- Scale Free & Small World
- Zipf's law
- Power Law
- Search in Small World

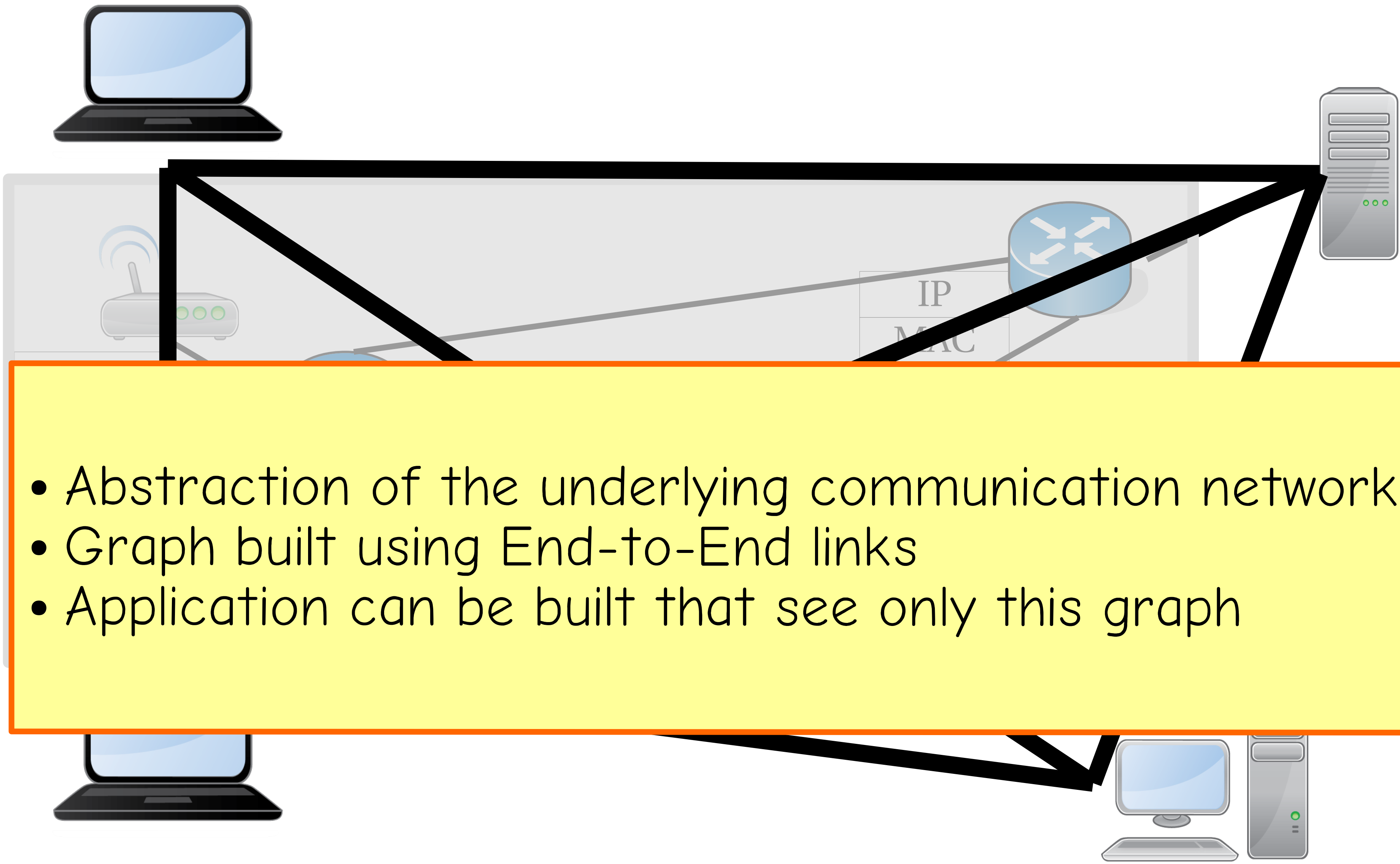
- Internet Indirection Infrastructure (I3)
- Content Delivery Networks
- Dynamo
- SDN and Clouds

Any suggestions for advanced topics?

1) Hop-by-Hop -> Overlay



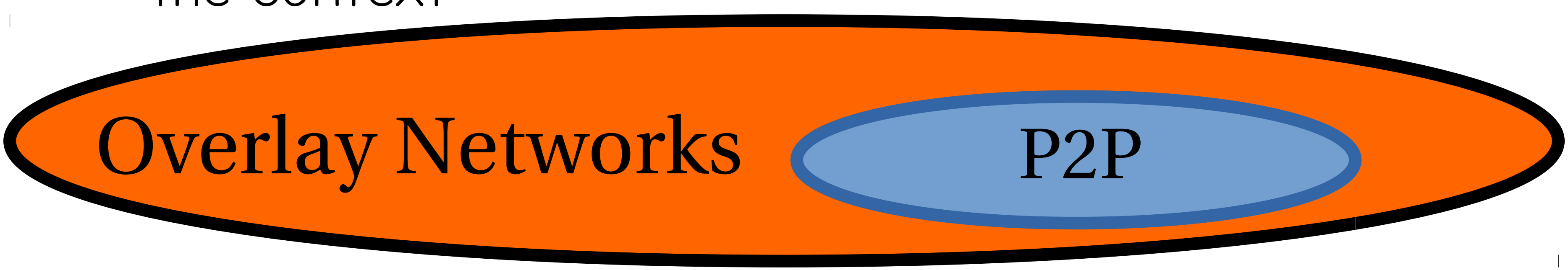
1) Hop-by-Hop -> Overlay



- Abstraction of the underlying communication network
- Graph built using End-to-End links
- Application can be built that see only this graph

2) Peer-to-Peer

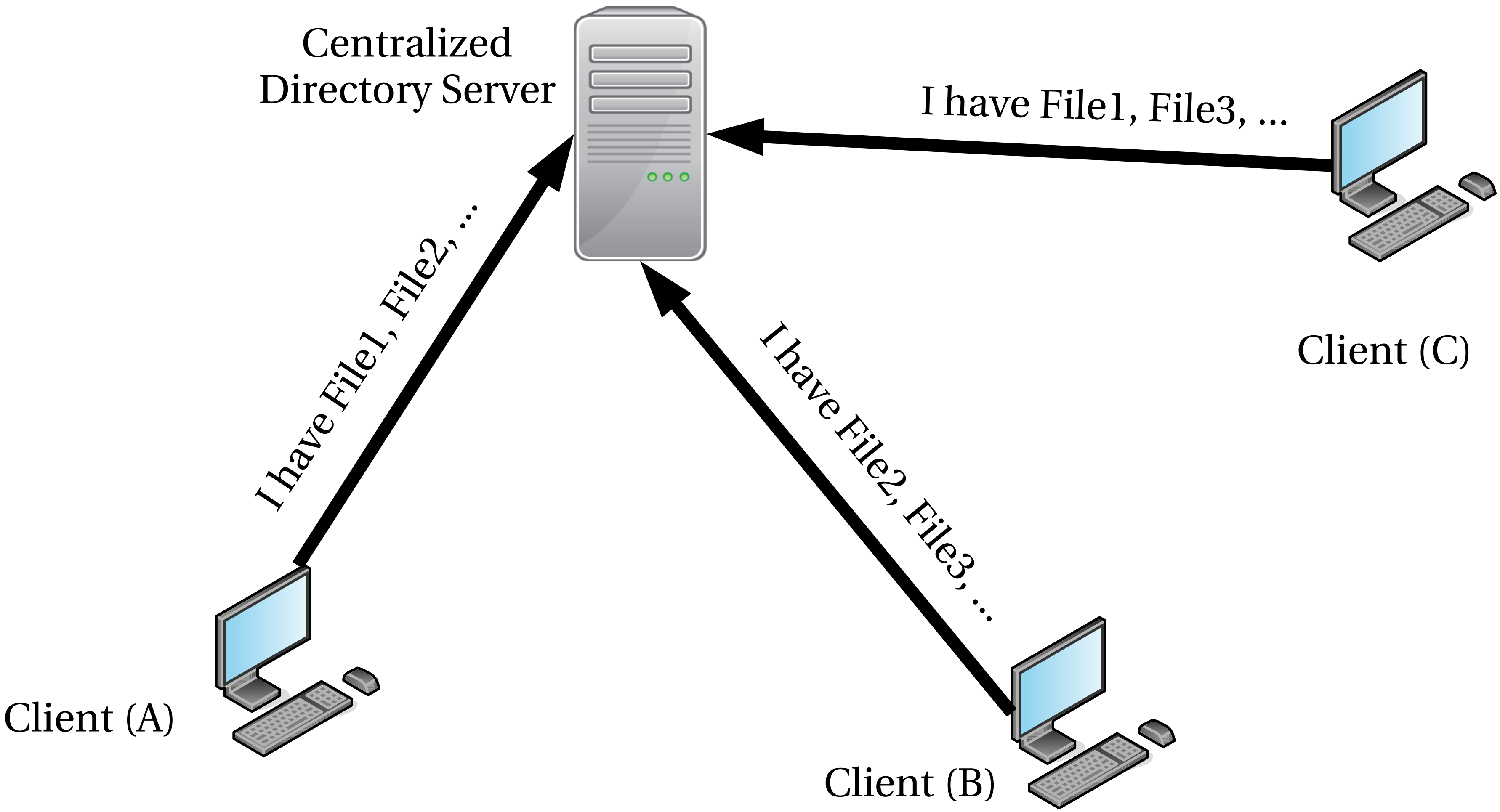
- Overlays (End-to-End Systems)
- **Peer-to-Peer: example of an overlay system**
- Peer
  - An **End** in an End-to-End system
    - All peers are "end-systems", but all "end-systems" are not peers
  - A computer, an end-user, an application depending on the context



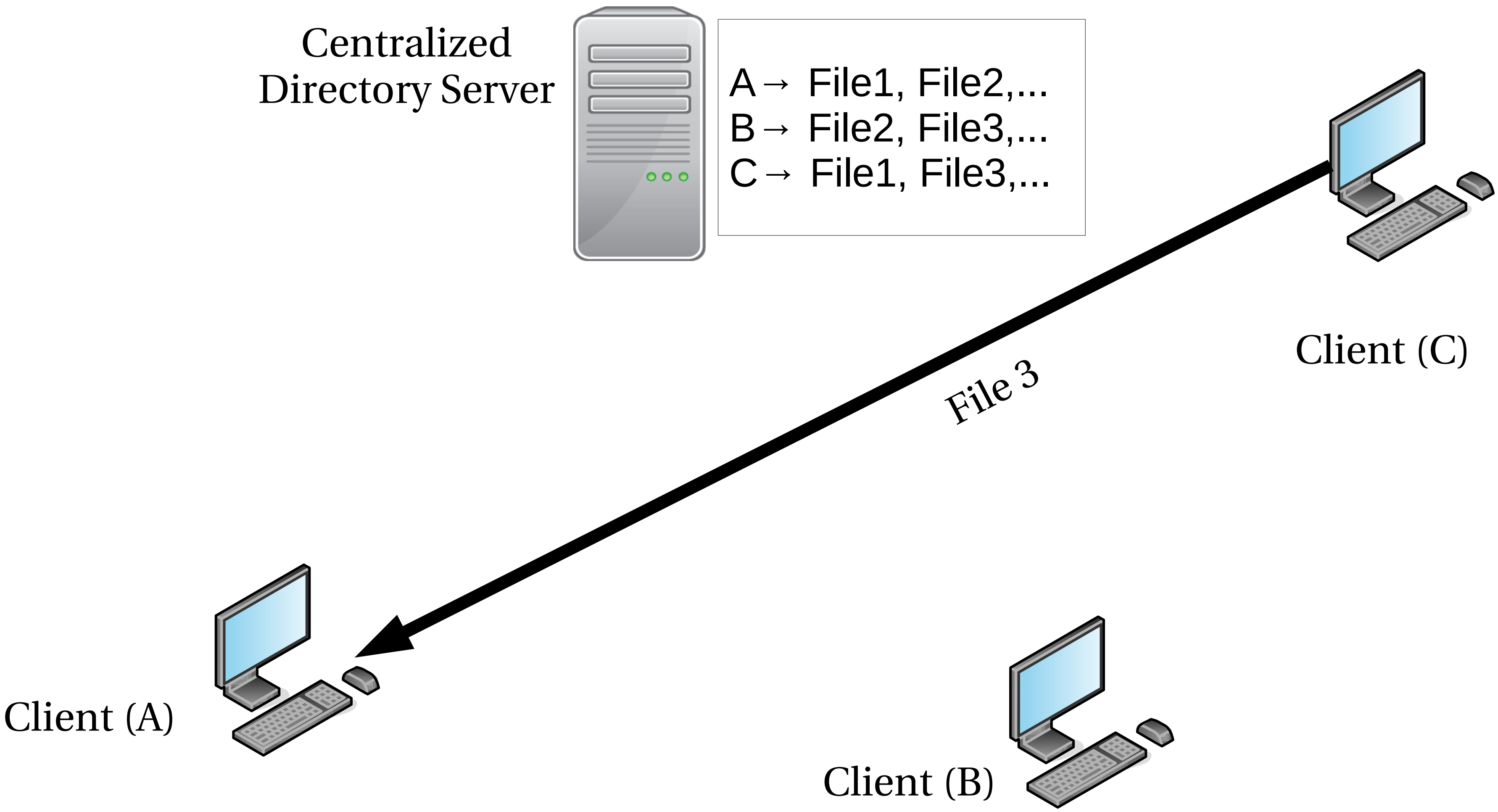
3) Unstructured Networks

- Flat or heirarchical organization
    - Napster
    - Gnutella
    - Skype
    - BitTorrent
    - Freenet

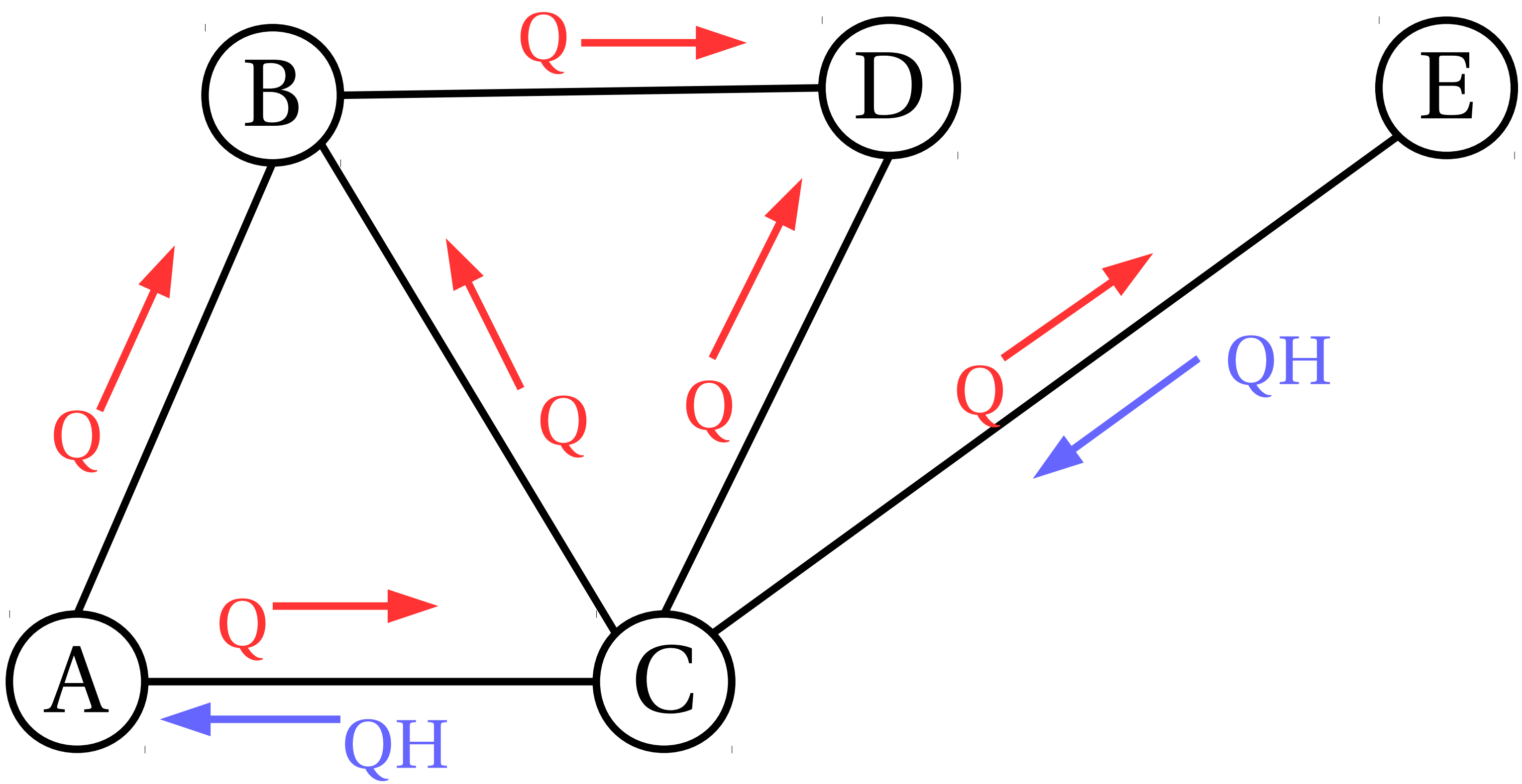
Napster (simple file sharing)



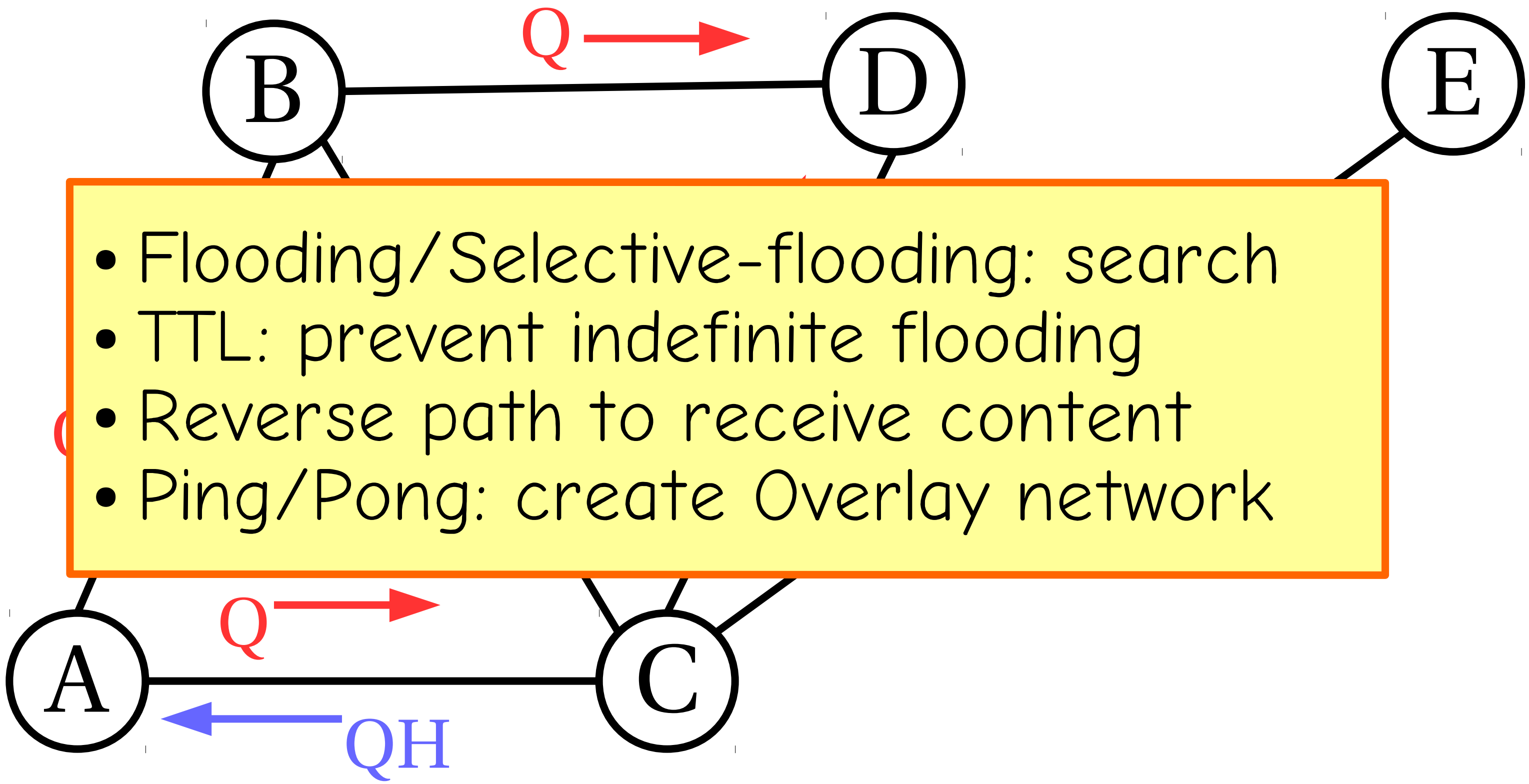
Napster (simple file sharing)



Gnutella
(file sharing-old version)

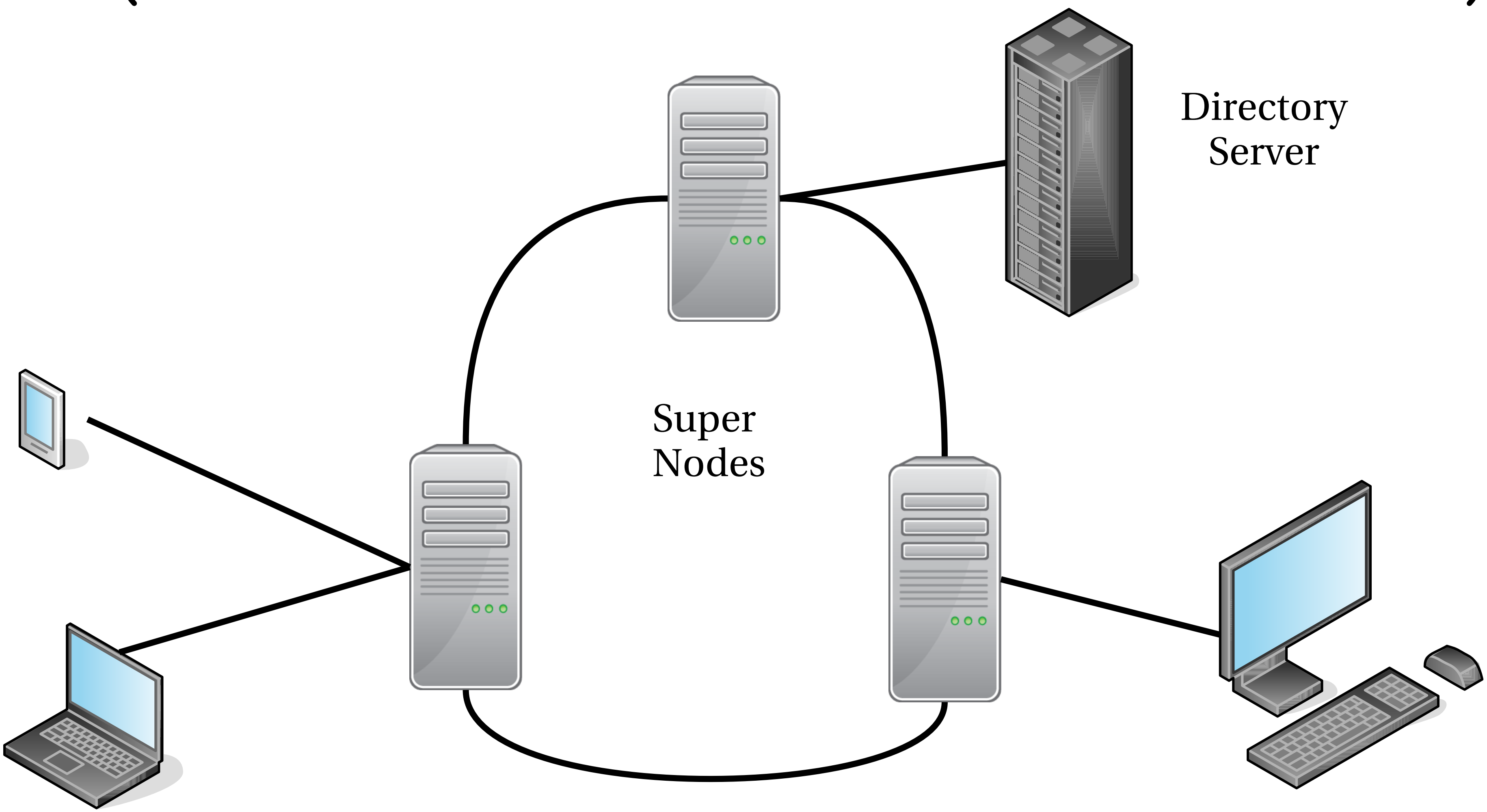


Gnutella
(file sharing-old version)

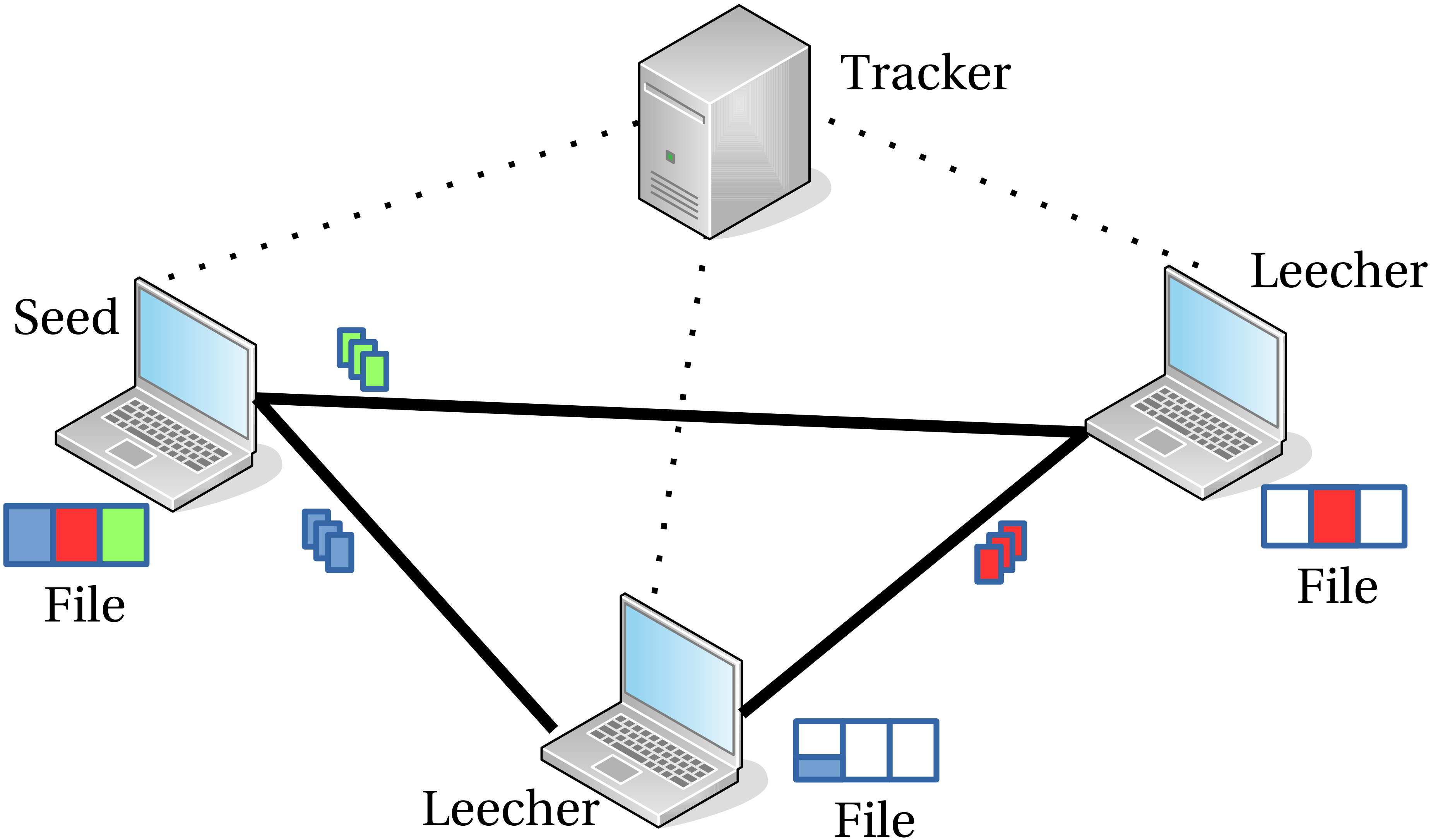


- Flooding/Selective-flooding: search
- TTL: prevent indefinite flooding
- Reverse path to receive content
- Ping/Pong: create Overlay network

Skype (P2P voice/video communication)



BitTorrent
(P2P content dissemination)



Freenet

- Self-organizing and Self-contained P2P network
- Collaborative (distributed) virtual file system
- Strong anonymity and censorship resistance

More Resources

- Video: Freenet at C3 Conference:
- B. Cohen.
  Incentives for building robustness in BitTorrent
- Stefan Saroiu, Krishna Gummadi, and Steven Gribble.
  "A Measurement Study of Peer-to-Peer File Sharing Sy stems."

Properties of Unstructured Overlay Networks

- Unstructured Link Creation
- Random Arrivals and Churn
- Flat or Hierarchical organization
- and many more ...

Which graph abstracts unstructured (and other) overlay networks?

Why is modeling important?

- What is the expected latency to the peek the content?
  - Average distance between nodes
- What is the impact of a peer leaving the network?
  - Robustness to faults/down-times
- What are the factors that determine the evolution of the graph structure?
  - Impact of proposed optimizations such as localization of content
- …

Small World

- Stanley Milgram. "**The small world problem.**" Psychology today 2.1 (1967): 60-67

  ”Given any two persons in the world, person X and person Z, how many intermediate acquaintance links are needed before X and Z are connected?”

- Mathematical structure of society

Milgrams Approach

A random person X was selected and given a document.

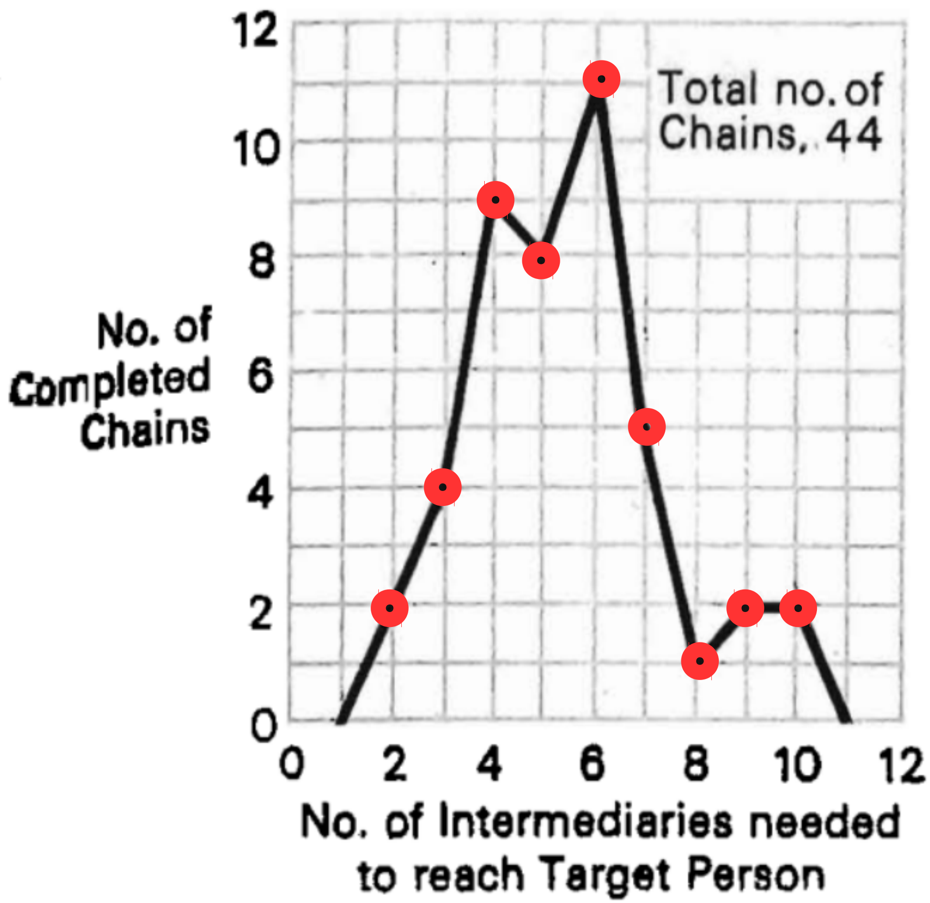
The document contained

(1) The name of person Z and some information such as profession and city name

(2) If X did not know person Z then X mails it to a person Y whom X knows on a first name basis and who is more likely than X to know Z. (Y is the new X)

(3) A roster on which each person to whom document arrives writes his/her name. Roster also prevents endless looping.

Key Result



Stanley Milgram. "**The small world problem.**" Psychology today 2.1 (1967): 60-67

- Median: 5 intermediaries
  - How many degrees of separation?
- Decay in active chains
  - 126 of 160 chains dropped out!
- Common pathway
  - Stock broker, clothing merchant
- Misc
  - Role of gender, ethnicity
- ***When will two sub-populations be isolated from each other?***

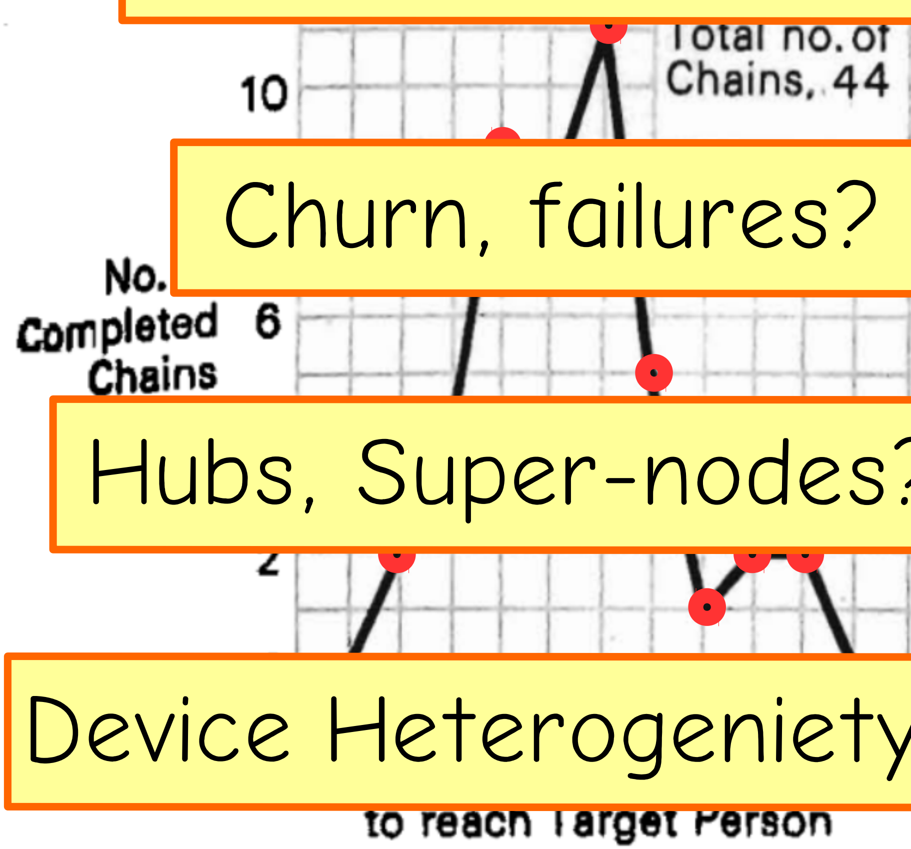
Key Result

Average diameter?

Churn, failures?

Hubs, Super-nodes?

Device Heterogeneity?

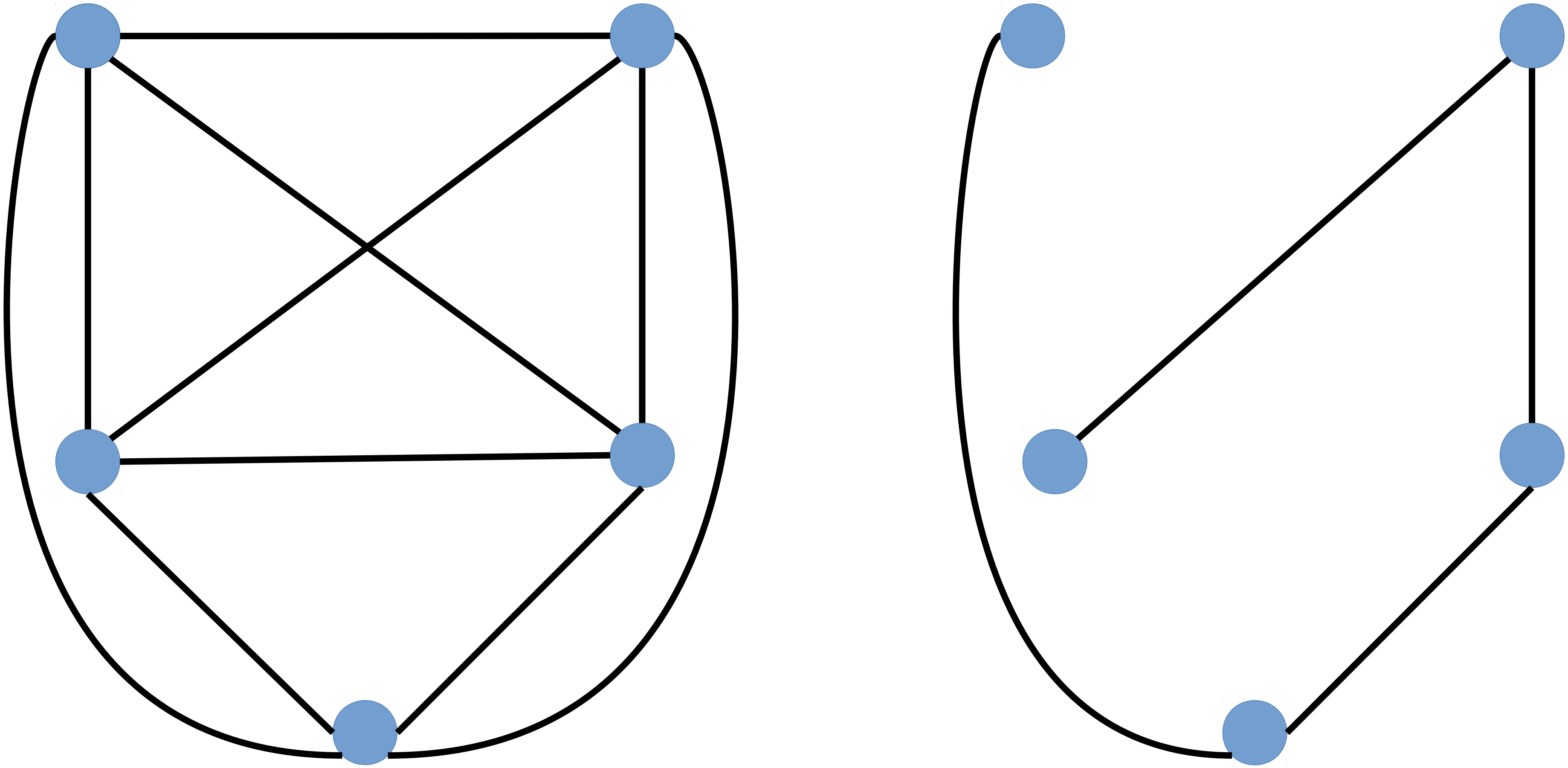


Stanley Milgram. "**The small world problem.**" Psychology today 2.1 (1967): 60-67

- Median: 5 intermediaries
  - How many degrees of separation?
- Decay in active chains
  - 126 of 160 chains dropped out!
- Common pathway
  - Stock broker, clothing merchant
- Misc
  - Role of gender, ethnicity
- ***When will two sub-populations be isolated from each other?***

Random Graph

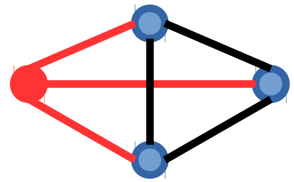
(Erdős–Rényi)/Gilbert Model



Small diameter but not robust to link/node failures?

Modeling the "Small World"

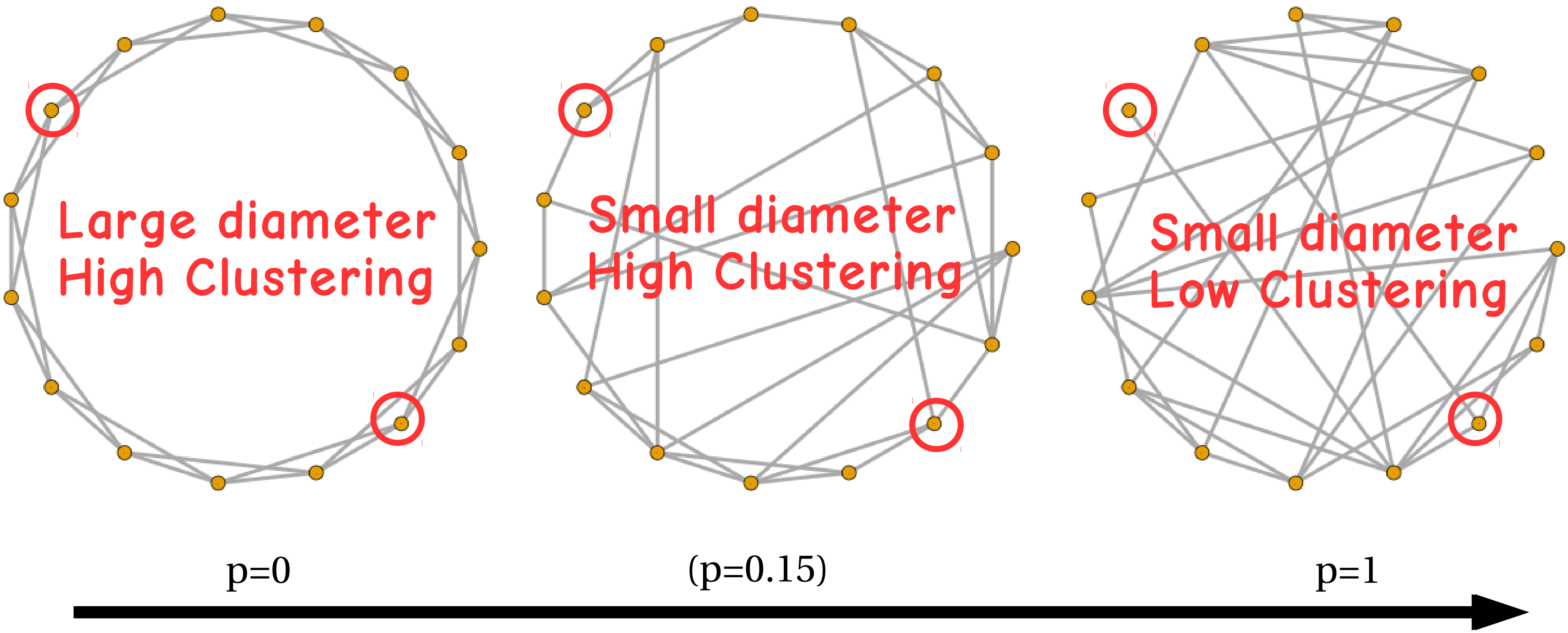
- Duncan Watts and Steven H. Strogatz. "Collective dynamics of 'small-world' networks." Nature 393.6684 (1998): 440-442.
  - Diameter: length of the longest path
  - Clustering Coefficient: how close are the adjacent vertices of a vertex to a complete graph (clique) if the vertex is removed.
  - Desirable properties: Small Diameter and High Clustering Coefficient



Random Rewiring Procedure

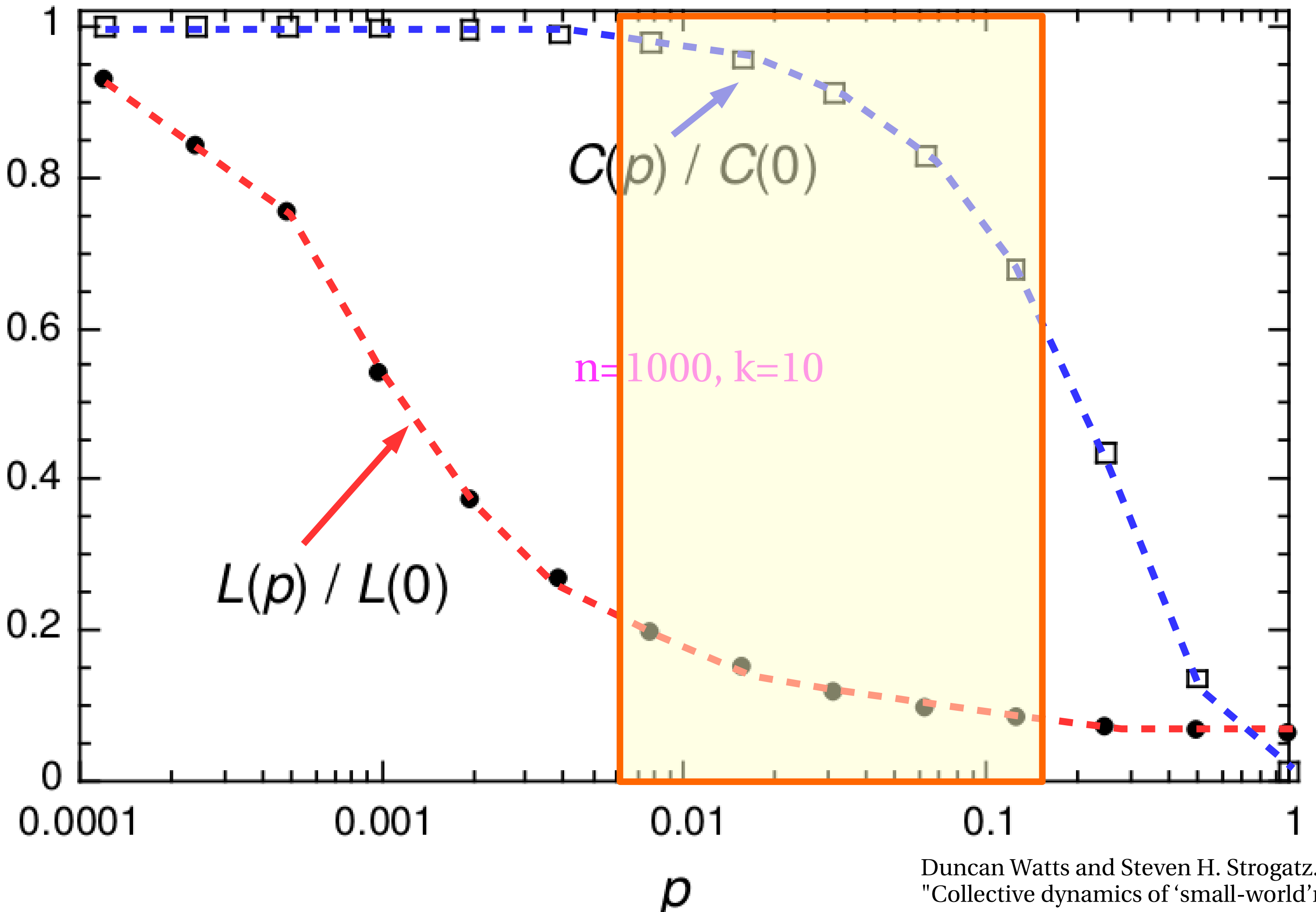
What is the diameter?
What happens if I remove a neighbour?

n=10, k=4



n: number of vertices
k: degree of each node, i.e., neighbours of a vertex (n >> k >> ln(n) >> 1)
p: probability of rewiring

Path Length L(.) and Clustering Coefficient C(.) in Rewired Graphs



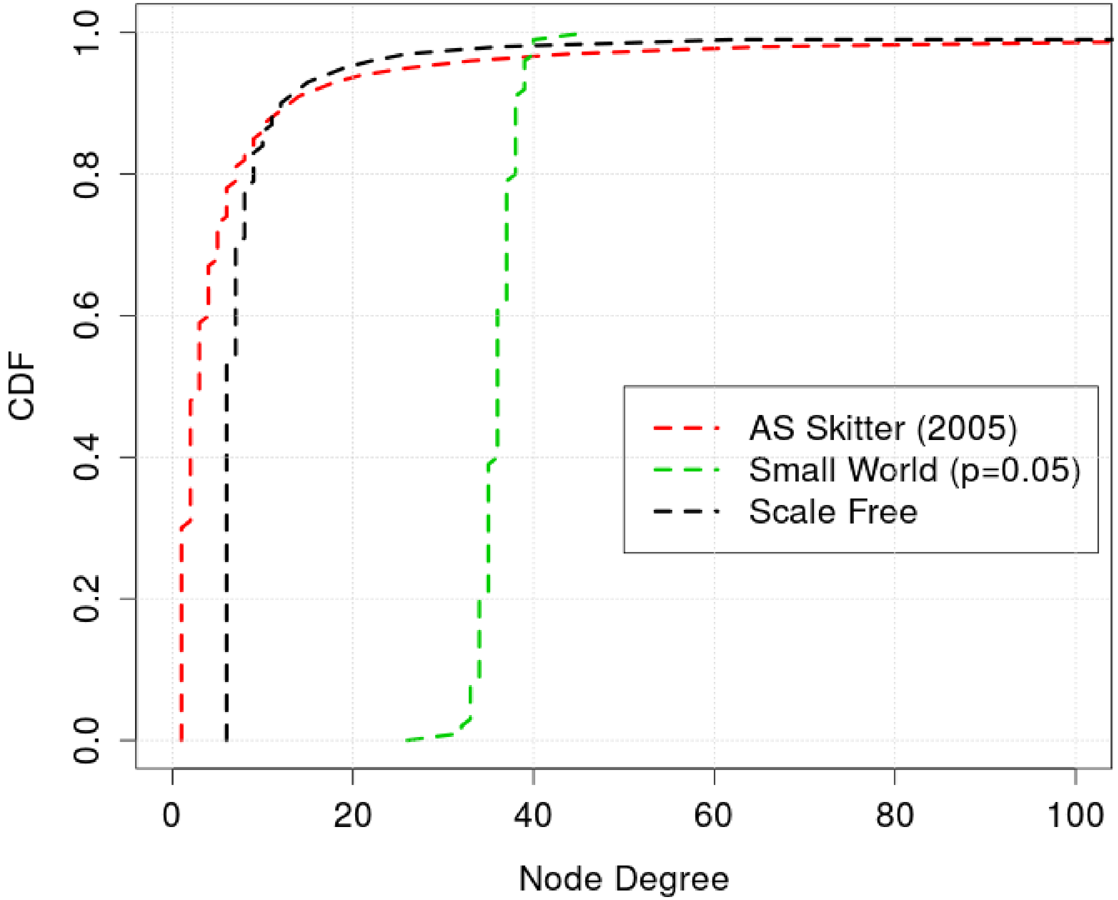
Duncan Watts and Steven H. Strogatz. "Collective dynamics of 'small-world' networks." nature 393, no. 6684 (1998): 440-442.

Is Small-World Networks the right model?

- What happens in dynamic networks whose members/vertices change with time?
- With which vertices of a graph do new arrivals prefer to attach?
- Is the the development of large networks governed by a robust self-organizing phenomena?
- What about heterogeniety in type of elements in and the degree distribution?
- ...

What about degree distribution?

- Do small-world/random-graph models capture the presence of hubs?



AS Topology of skitter dataset parsed by SNAP team
http://snap.stanford.edu/data/as-skitter.html

Assumptions of Random Graphs and Small-World

- Fixed number of nodes that are either randomly connected (Random Graphs) or randomly rewired (Small-World)
- Probability that two vertices are connected is random and uniform

Barabási, Albert-László, and Réka Albert. "Emergence of scaling in random networks." Science 286, no. 5439 (1999): 509-512.

Scale Free Networks

- Probability that a node has $\mathbf{k}$ links is $P(k) \propto k^{-\gamma}$
- γ exponent for the degree distribution
- Why the name Scale-Free?

$$P(x) \propto x^{-\gamma}$$

$$P(ax) \propto a^{-\gamma} P(x)$$

  – Same functional form across all scales
    - e.g. γ = 1 → P(x) = 1/x
    - - γ is the slope on log log scale

Preferential Attachment (Growth in Scale-Free Network)

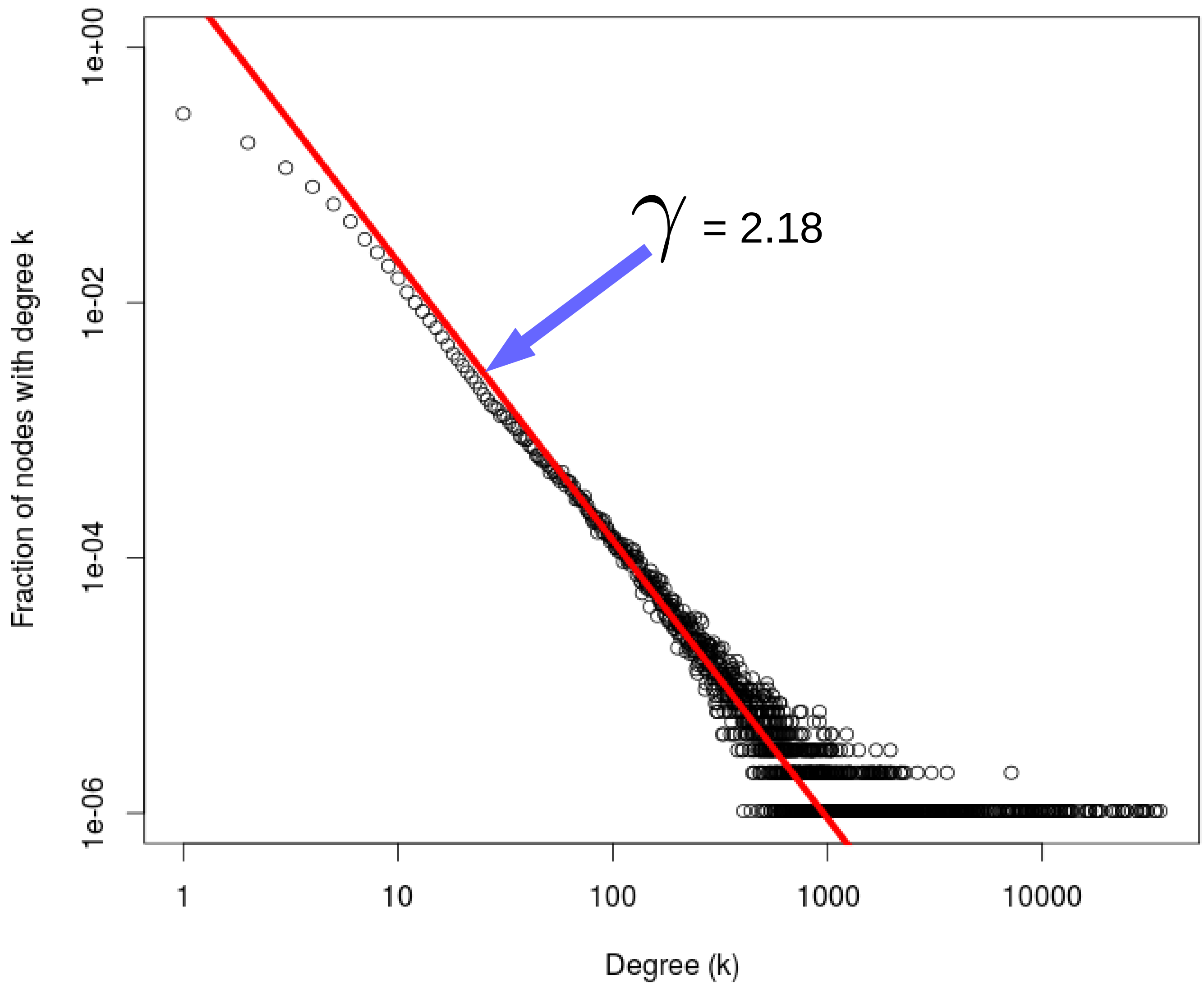
- Initial number of nodes $\mathbf{m_0}$
- At time **'t'** add a new vertex **j** with degree **m (< $\mathbf{m_0}$)**
- The probability that vertex **j** creates a link with vertex **i** is the probability **π** which depends on $\mathbf{k_i}$ (the degree of vertex **i)**

$$\pi(k_i) = \frac{k_i}{\sum_l k_l}$$

- Probability that a node has **k** links is $P(k) \propto k^{-\gamma}$
- γ exponent for the degree distribution

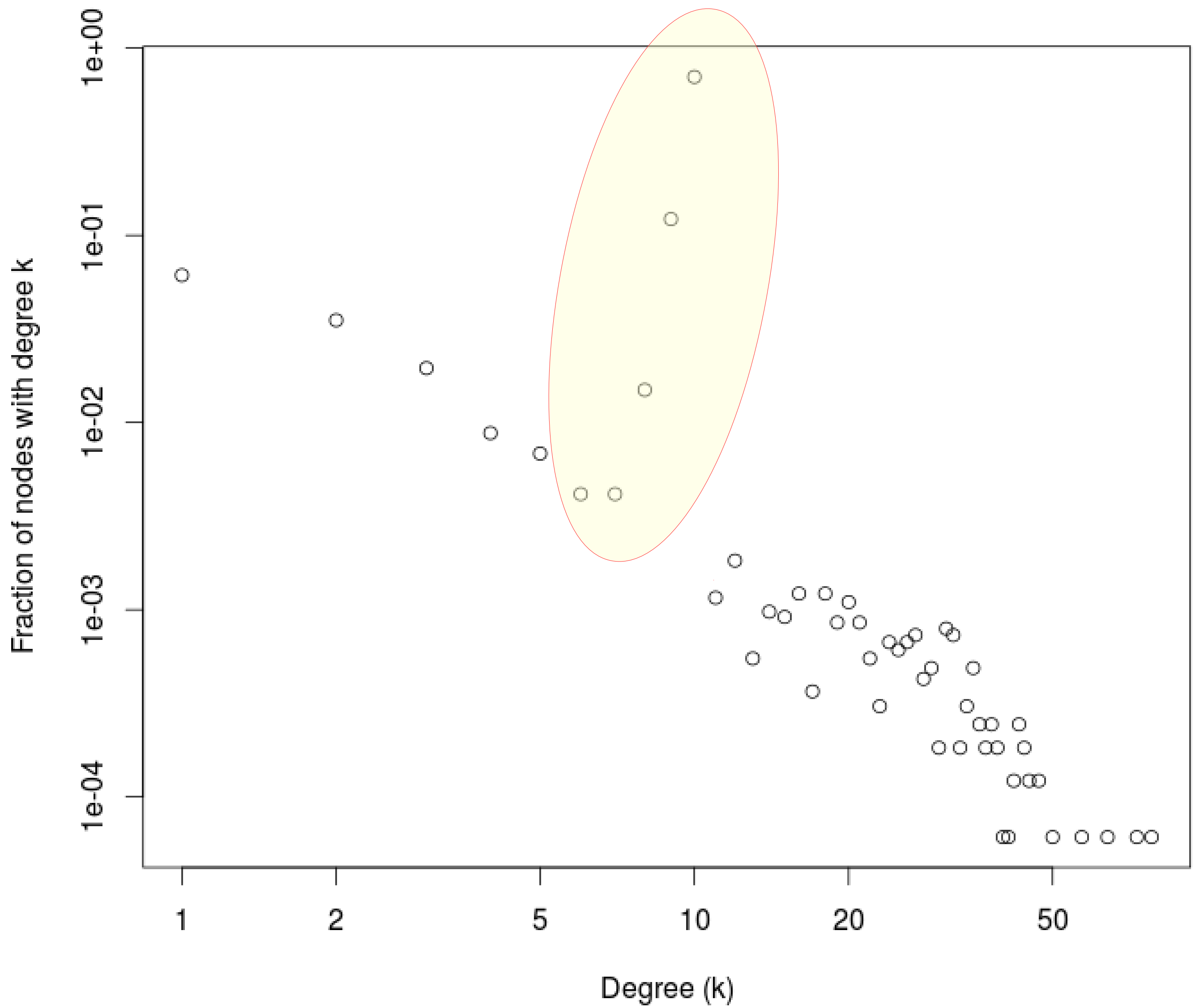
Barabási, Albert-László, and Réka Albert. "Emergence of scaling in random networks." Science 286, no. 5439 (1999): 509-512.

Scale-Free Model for AS-Graph



AS Topology of skitter dataset parsed by SNAP team
http://snap.stanford.edu/data/as-skitter.html

Scale Free for Gnutella?



Gnutella August 2002 dataset parsed by SNAP team
http://snap.stanford.edu/data/p2p-Gnutella31.html

Summary

- Large number of interacting elements
- System adapts and evolves
- Although elements follow simple rules, the system behaves in a *complex* manner
- System behavior is an outcome of the interaction between the elements, so studying elements in isolation will not provide detailed insights on system behavior
- Similar processes shape networks therefore many networks share similar characteristics

Methodology

1) Make observations (conduct measurement studies)
2) Build model to explain observations
   – Choose the right level of granularity (zoom level)
   – Strip the problem to a simple form
   – Attempt to formulate the problem and model the system
3) Validate model
   – Reproduce observations/measurements
   – Explain observations
4) Revisit step 2 (and 1) to improve understanding

Complex Networks

Important Papers

- Stanley Milgram. "**The small world problem.**" Psychology today 2.1 (1967): 60-67
- Duncan Watts and Steven H. Strogatz. "**Collective dynamics of 'small-world' networks.**" Nature 393.6684 (1998): 440-442.
- Barabási, Albert-László, and Réka Albert. "**Emergence of scaling in random networks.**" Science 286, no. 5439 (1999): 509-512.
- Albert, Réka, and Albert-László Barabási. "**Statistical mechanics of complex networks.**" Reviews of modern physics 74.1 (2002): 47.
- Mitzenmacher, M. (2004). "**A brief history of generative models for power law and lognormal distributions**." Internet mathematics, 1(2), 226-251.

Sources for these slides

- Sasu Tarkoma ”Overlay and P2P Networks”, 2015
- Arnaud Legout. ”Peer-to-Peer Applications : From BitTorrent to Privacy”, January 2012.
- Jari Saramäki ”Introduction to Complex Networks”, 2010.
- Datasets from Stanford Network Analysis Project (SNAP)