

A Case for OneSwarm

Tom Anderson
University of Washington

With: Ivan Beschastnikh, Colin Dixon, Dan Halperin, John P. John,
Ethan Katz-Bassett, Tomas Isdal, Arvind Krishnamurthy, Ed Lazowska,
Harsha Madhyastha, Ratul Mahajan, Mike Piatek, Arun Venkataramani,
David Wetherall, Xiaowei Yang, John Zahorjan

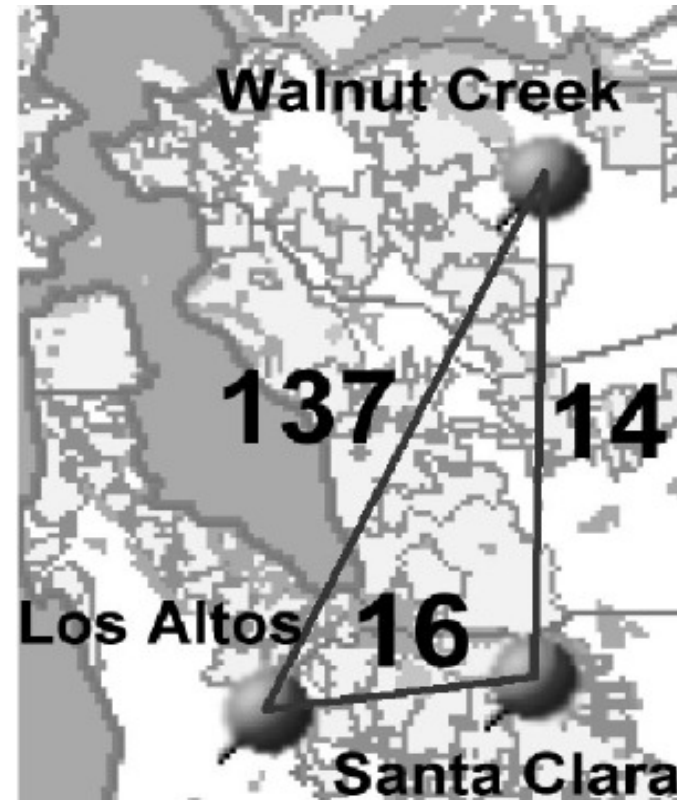
An Anecdote

In 1997, we observed that many routes in the Internet did not obey the triangle inequality

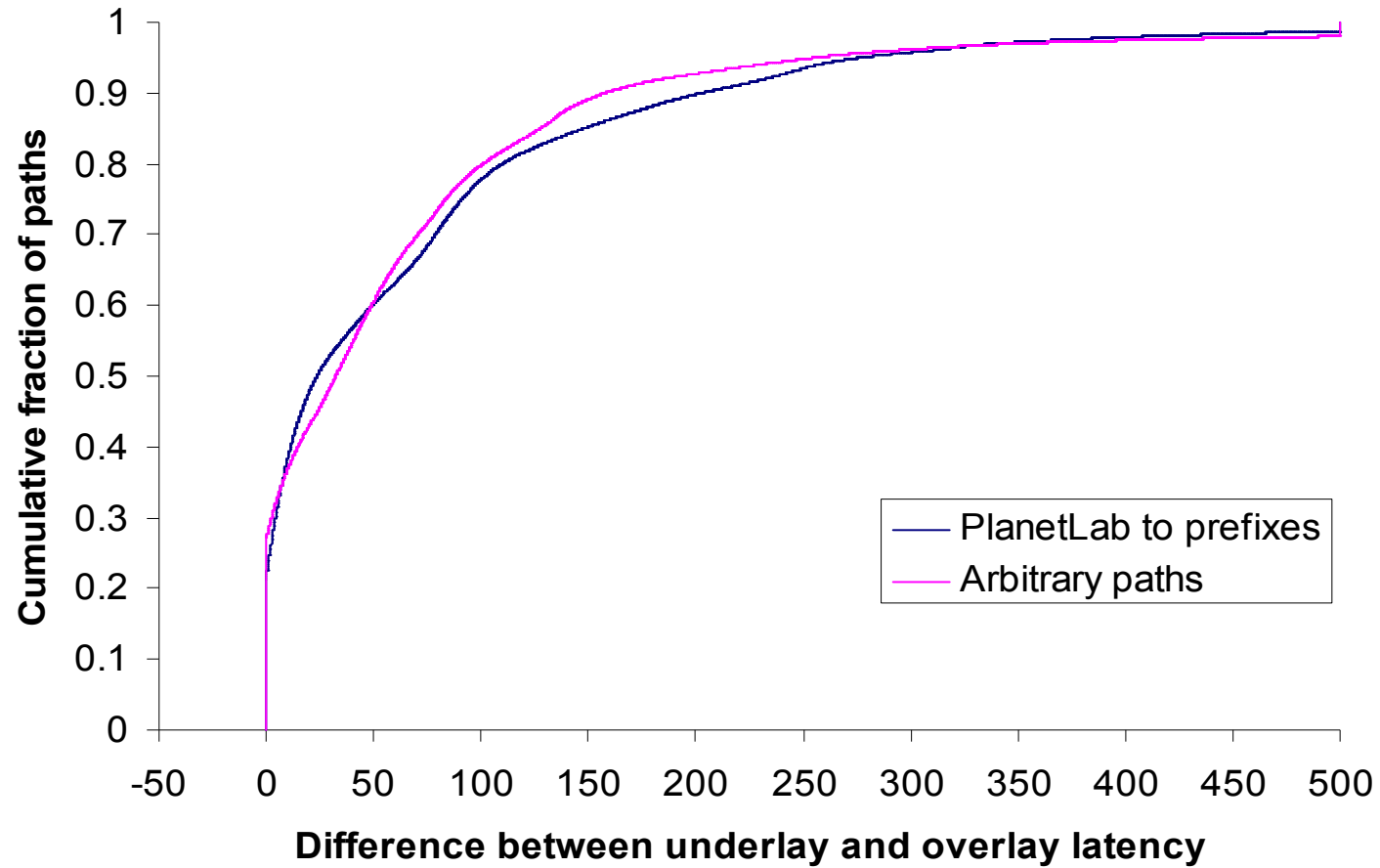
- 40% of all Internet routes
- 10% pathological

Fix via overlay routing?

- Embarrass ISPs into improving their routing



Detour Routing Today



Another Anecdote

In April 1997 and again in April 2001, a small mis-configuration at one ISP disrupted Internet connectivity on a global scale.

- Nothing prevented one ISP from announcing that it can deliver packets for any Internet prefix

2005: Internet is still vulnerable to this same problem

- Over half of all new Internet route announcements are misconfigurations!
- Until recently, Cisco's Internet prefix was hijacked on a regular basis

A Third Anecdote

In 2001, <anonymous internet founder>'s home machine was DoS'ed

- Packet floods rendered internet access unusable
- Despite repeated requests to his ISP, they were unable to stop the attack
- After a year, he gave up and got a new IP address (and made sure to always tunnel his packets through work)

Today: no effective DoS solutions for this case

How Do We Fix This Impasse?

GENI: Global Infrastructure for Network Innovation

- Testbed with programmability in every network device
- Architecture-neutral
- Necessary but not sufficient

Research to address fundamental limits of understanding in how to build secure, reliable, flexible, and manageable global scale systems

What would cause anyone to use these new solutions?

- Incentives...

UW Architecture Research

Enabling architectural change (PlanetLab, GENI)

- Make it easy for clients to sign up for new architectures (OASIS)
- Make it easier to develop/deploy new distributed services (cPlane)
- Leverage naming to make endpoint protocols evolvable (Active Names)

Measurement and diagnosis (Rocketfuel)

- Diagnose paths from end-points (Tulip, kPlane)
- Internet information as a service (iPlane, topology based geolocation)

Congestion Control

- Near optimal endpoint congestion control (PCP)

Security

- How do we design a network that is (nearly) DoS-free? (TVA)
- Secure controlled access to remote resources (CRISIS, GENI)

Routing

- How do we design routing that is maximally robust? (routing by consensus)
- Incentive compatible interdomain routing (Wiser)

Enabling new applications

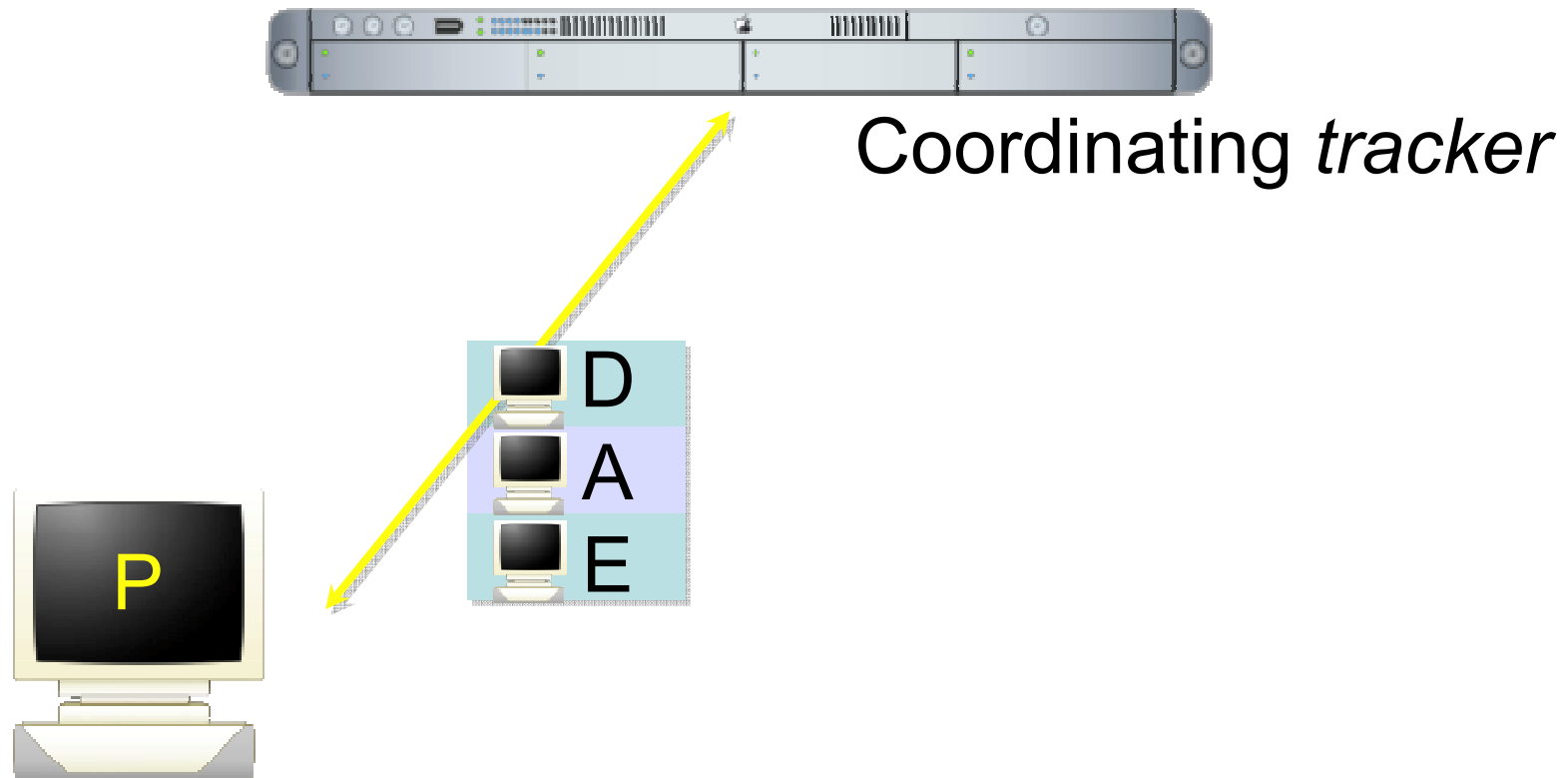
- Everyone a content producer (LiveSwarms, Digital StudyHall)

Self-configuring secure wireless (Catch, Wit)

Examples of Incentives

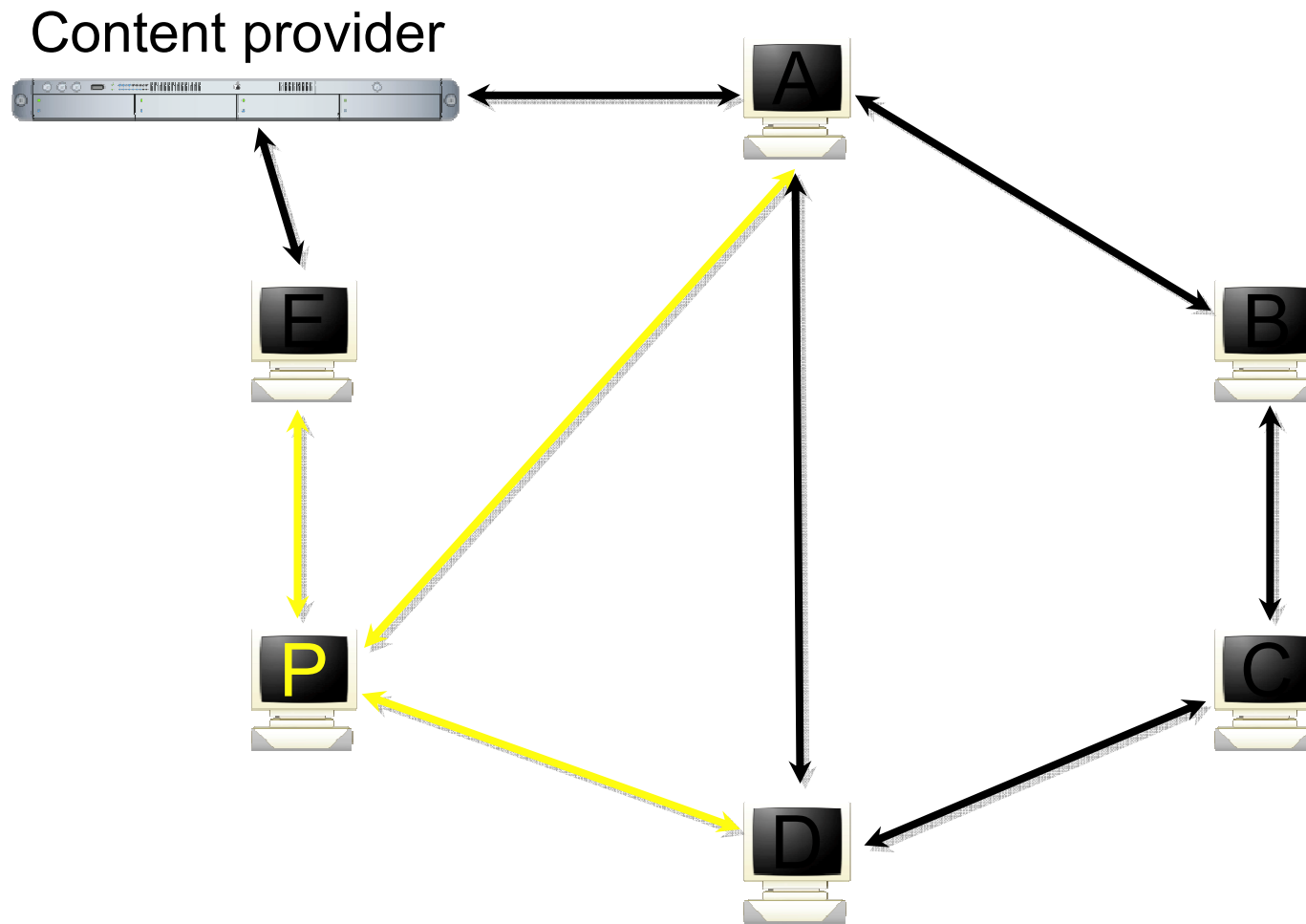
- ❑ Interdomain routing
 - Is selfishness the cause of poor interdomain routing behavior?
 - Or is it protocol design? (A: protocol design)
- ❑ BitTorrent file sharing: explicitly designed for incentives
 - Is BitTorrent incentive compatible?
 - What is the impact of selfishness on swarm performance?
 - Can swarming replace the Internet? (A: OneSwarm)

BitTorrent overview

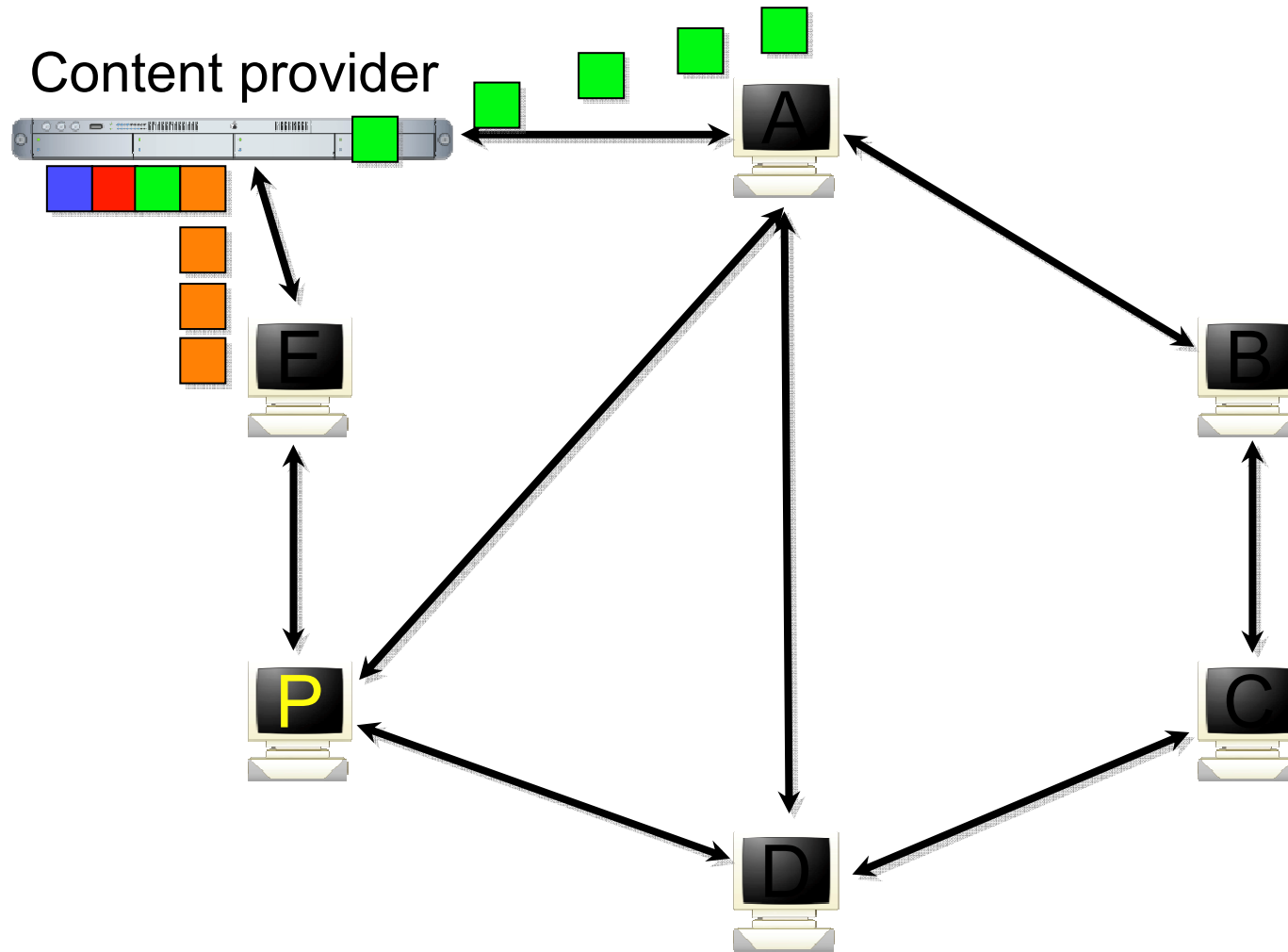


P joins the system by obtaining a **random** subset of current peers from a centralized coordinator

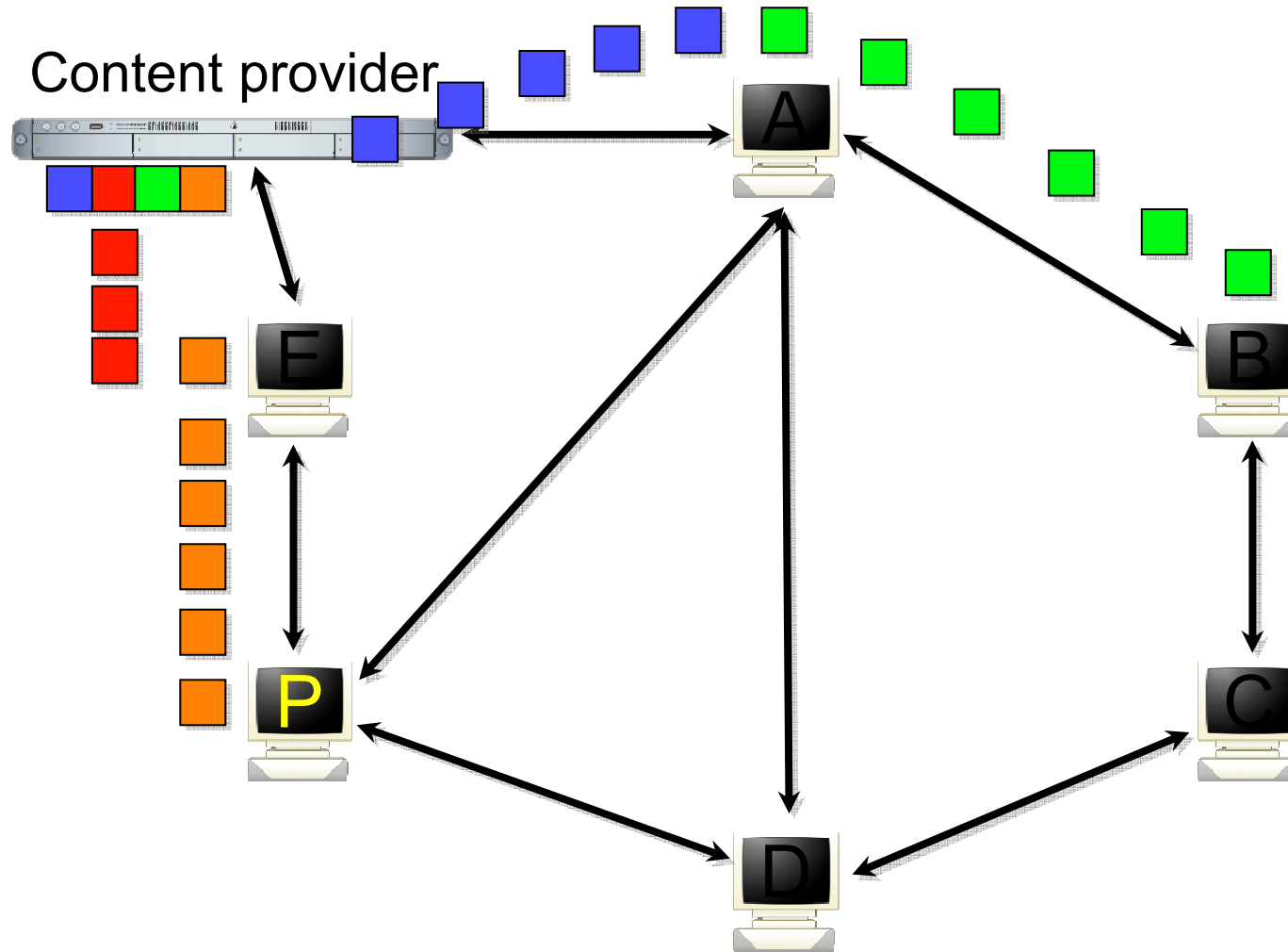
BitTorrent overview



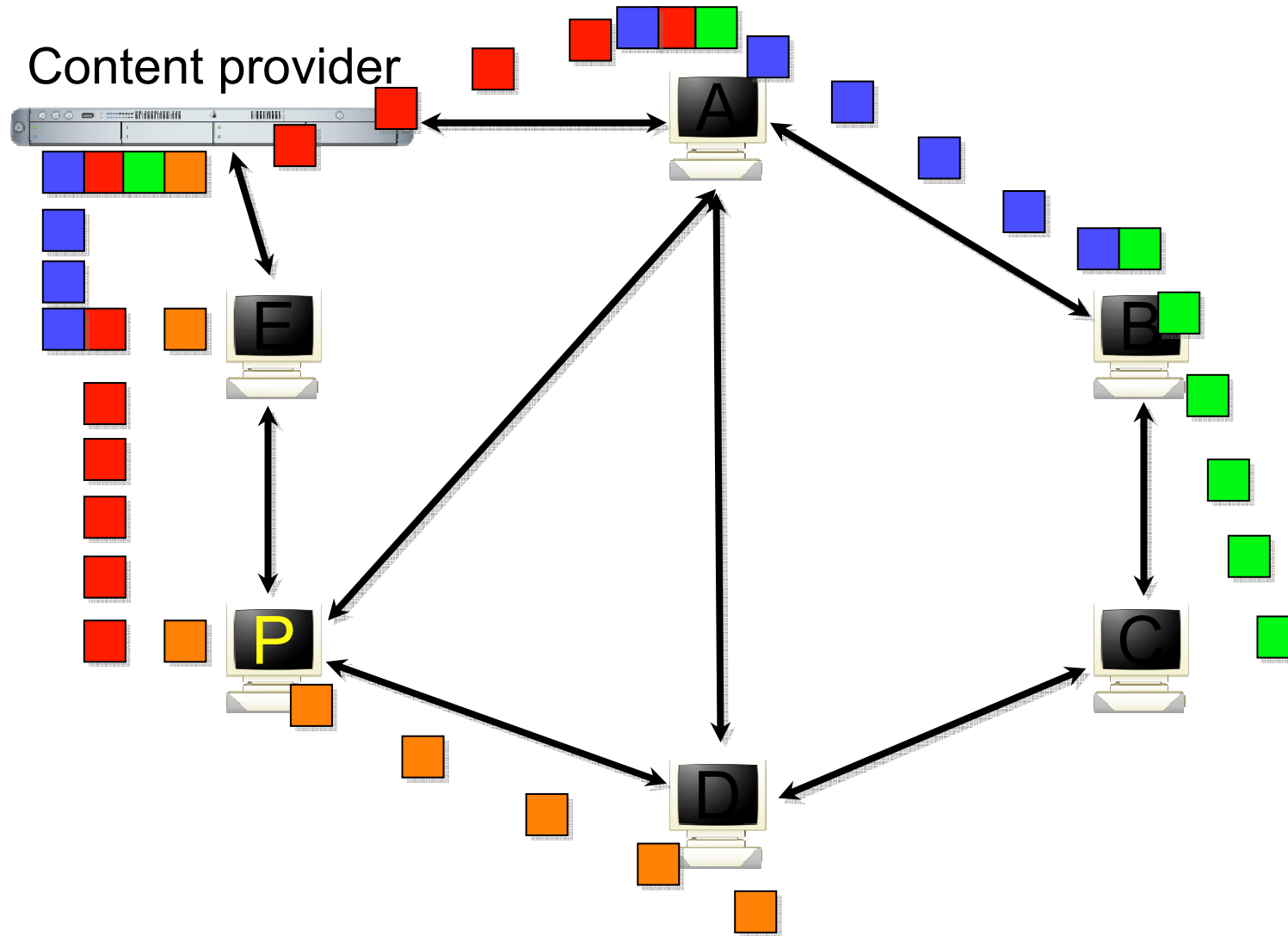
BitTorrent overview



BitTorrent overview



BitTorrent overview



Tit-for-tat in BitTorrent

Choosing *peers* and *rates*:

Sort *peers* by incoming data rate

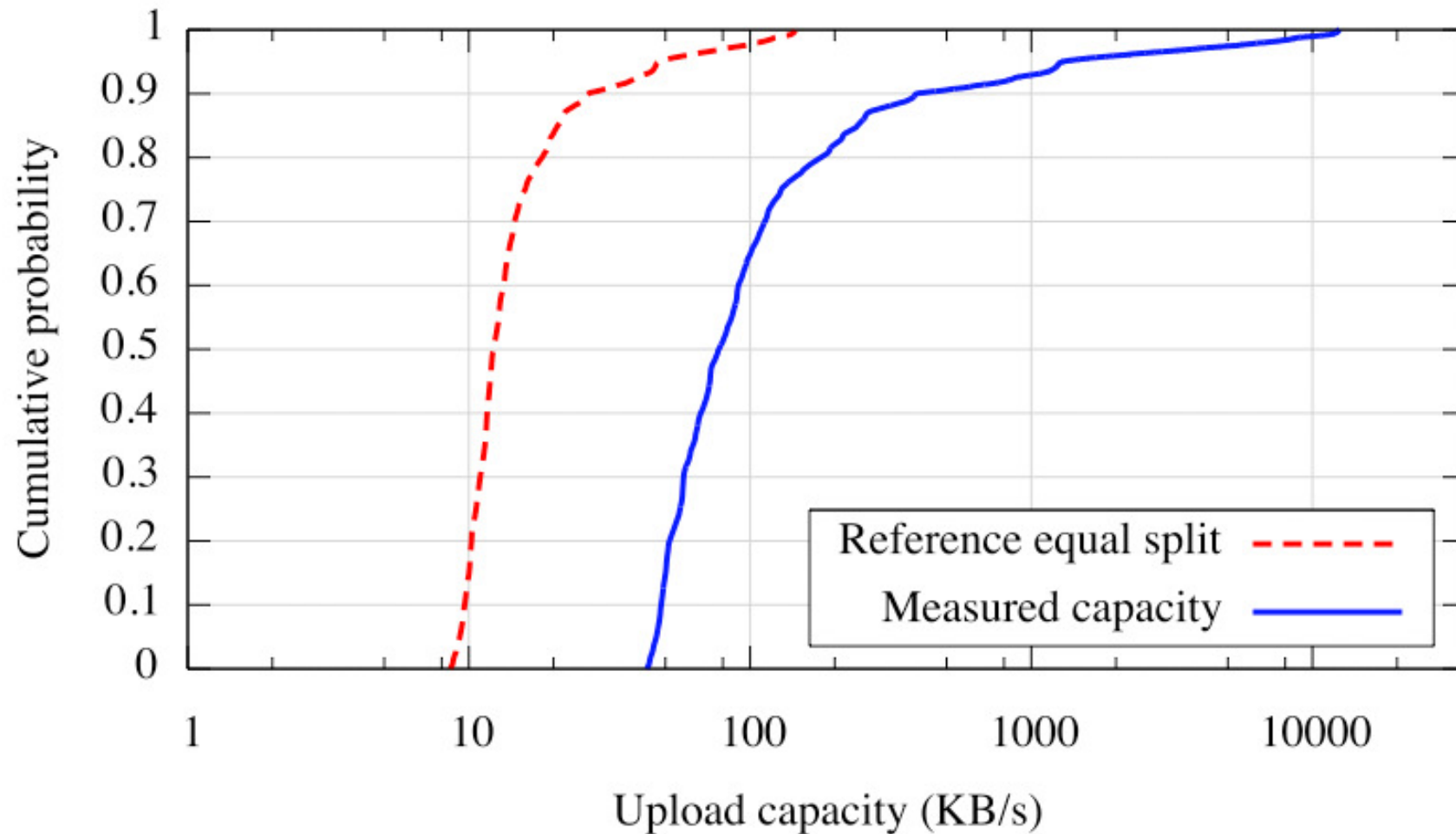
Reciprocate with *top k*

Send each peer selected in (2) an *equal split* of capacity

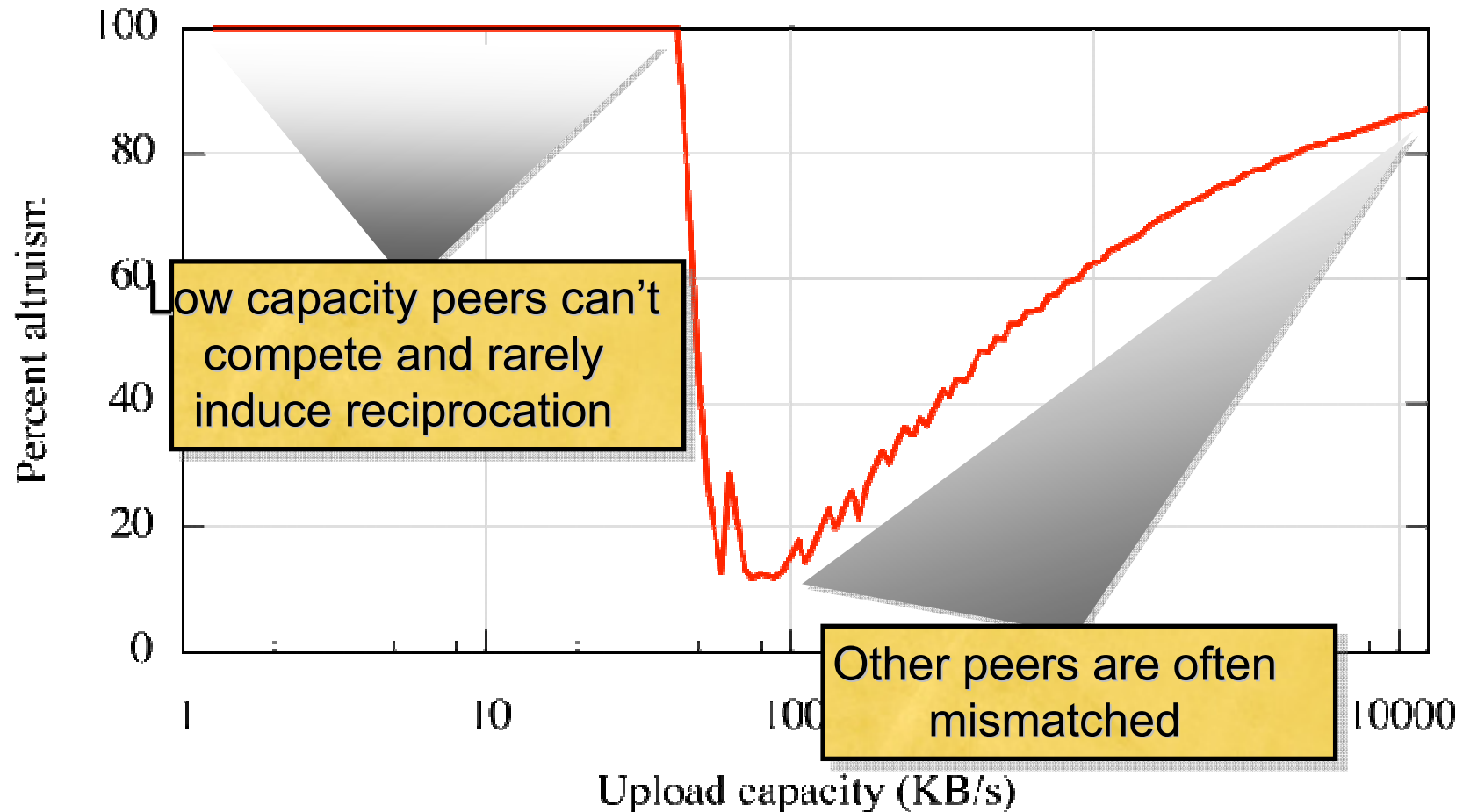
<i>Peer</i>	<i>Rate</i>
A	17
D	8
E	5

If $k=2$, P reciprocates with A and D, sending to each at an equal rate

Measured Upload Capacities



Churn creates altruism in BitTorrent



All peers make unnecessary contributions

BitTyrant

Can altruism in existing BitTorrent swarms be exploited by a selfish client?

Key idea: strategic selection of which peers and at what rates to send data, maximizing ROI

BitTyrant Algorithm

Each round, rank order each peer p by the ratio d_p/u_p , and choose those of top rank until the local upload capacity is reached.

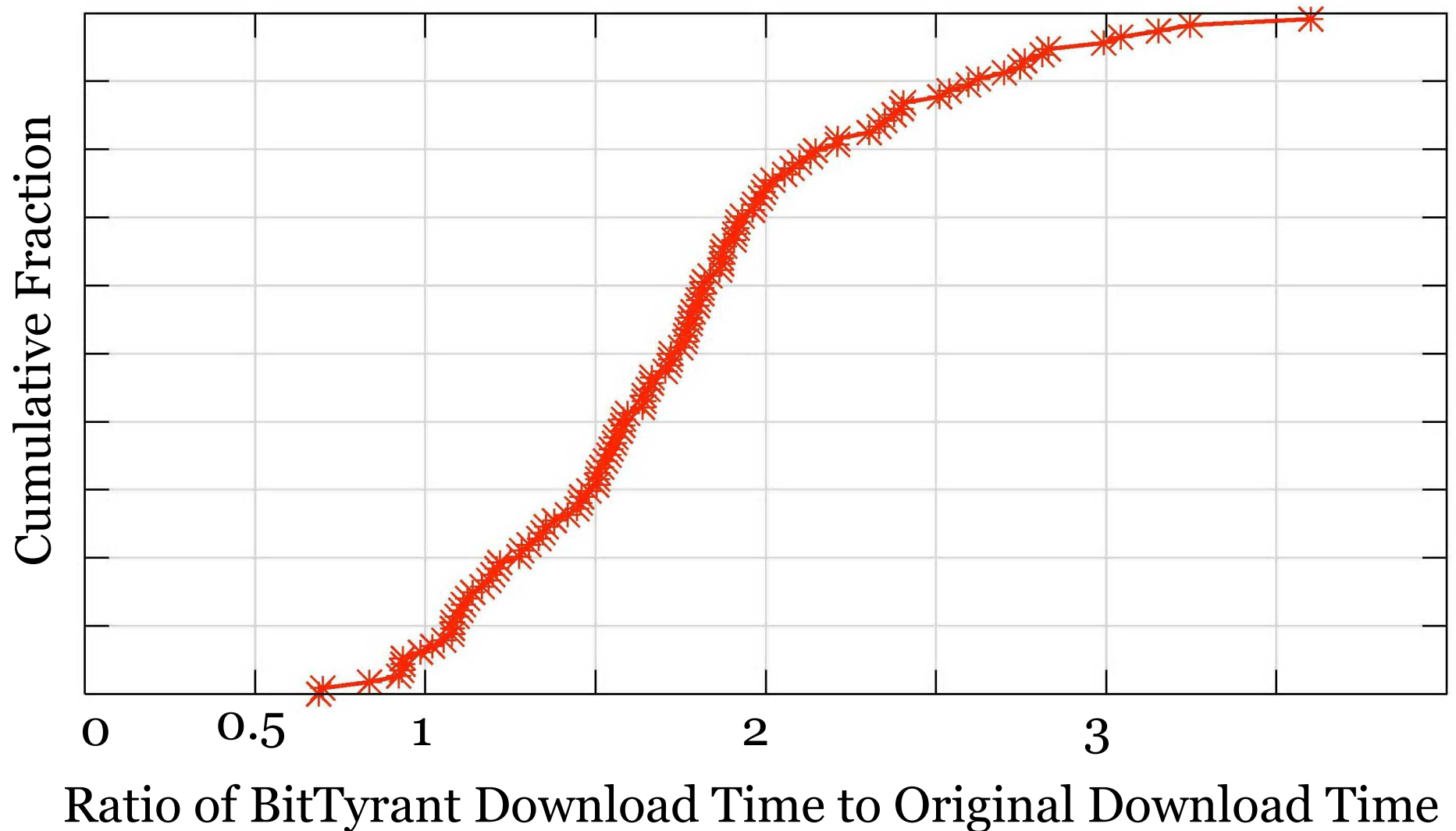
$$\underbrace{\frac{d_0}{u_0}, \frac{d_1}{u_1}, \frac{d_2}{u_2}, \frac{d_3}{u_3}, \frac{d_4}{u_4}}_{\text{choose } k \mid \sum_{i=0}^k u_i \leq \text{capacity}}, \dots$$

At the end of each round for each unchoked peer:

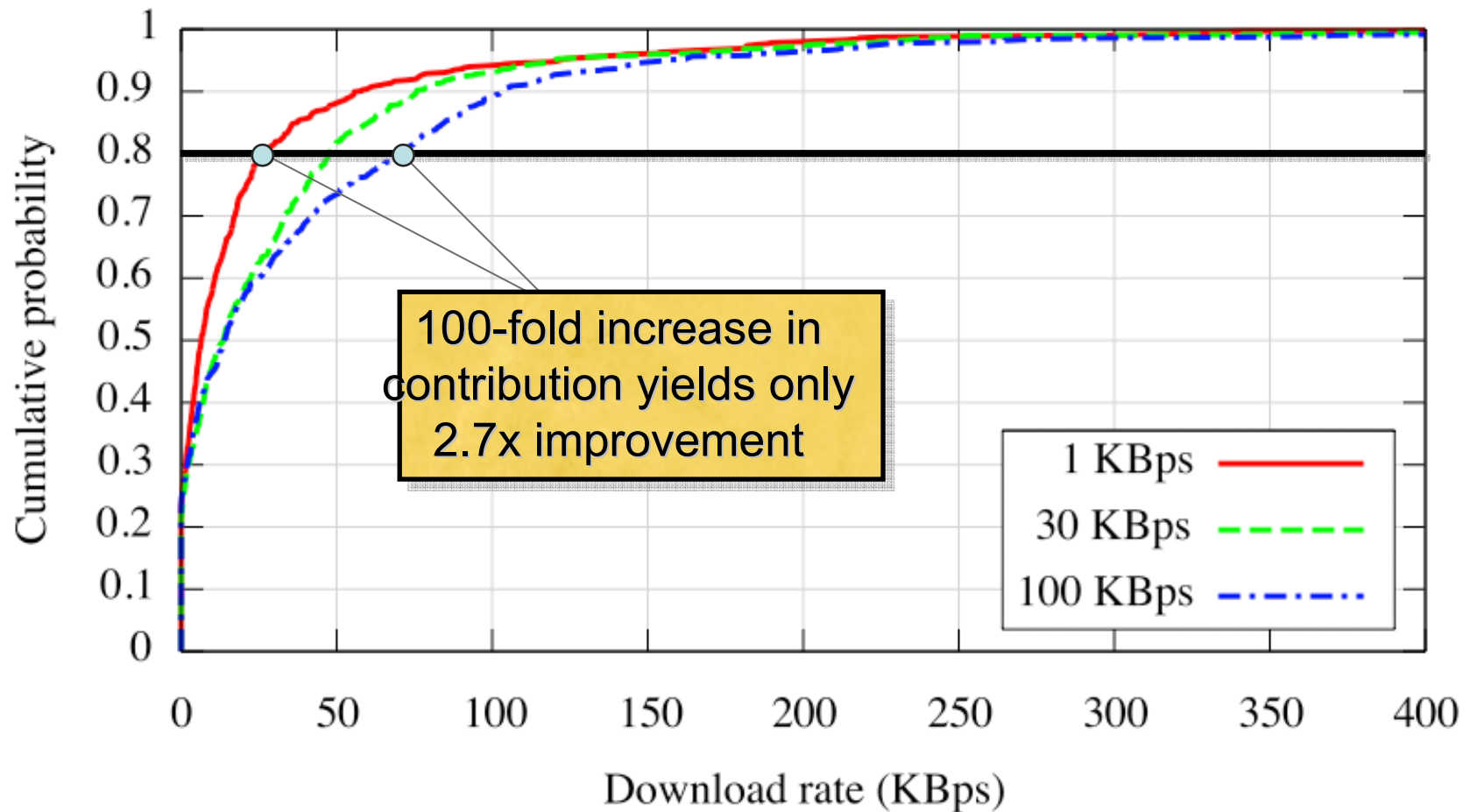
If peer p does not send data: increase cost estimate, u_p .

If peer p has unchoked us for the last minute:
reduce cost estimate, u_p .

BitTyrant on Popular Swarms



But most swarms are small...



For most swarms, incentives don't control performance

OneSwarm: Swarming as a Universal Data Layer?

Rewards for *high availability*, not churn

Incentives at the *granularity of users*, not objects

Bootstrapping via an *incentivized control plane*
(DHT)

Swarming distribution of metadata

Designed to exchange bandwidth *across services*
(bulk data, web, CDN, streaming video)