Traffic Localization for DHT-based BitTorrent networks

Moritz Steiner and Matteo Varvello

Alcatel·Lucent

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

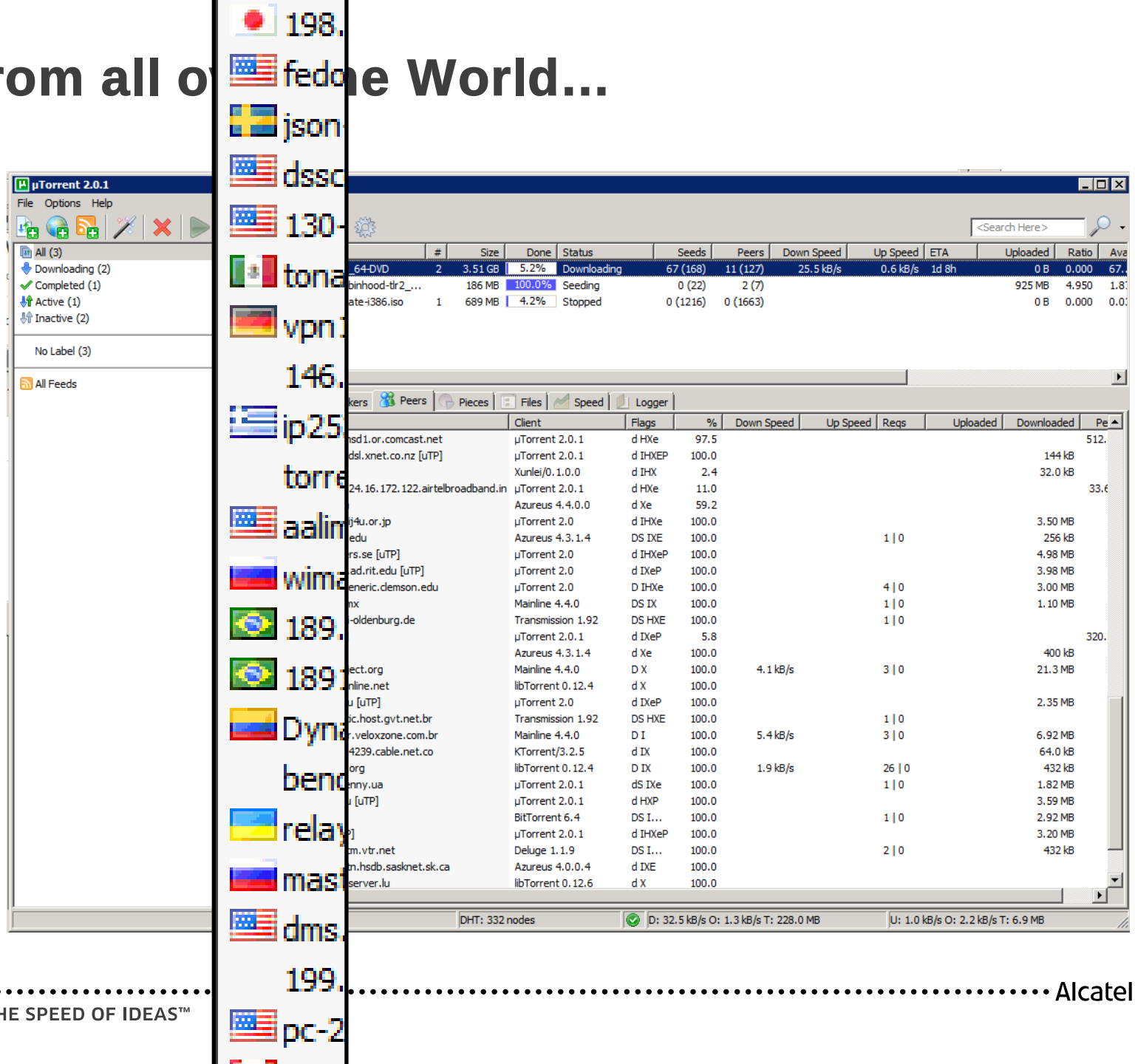
OUTLINE

- Motivation
- DHT Traffic Localization
- DHT Traffic Localization in Action!
- Evaluation
- Conclusion

Motivation

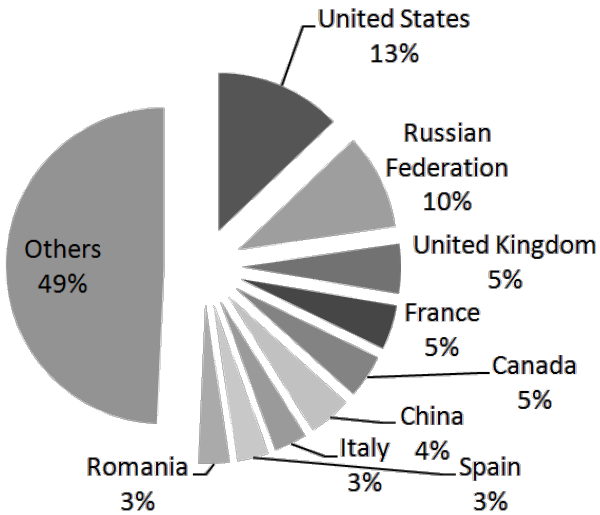
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

From all over the World…

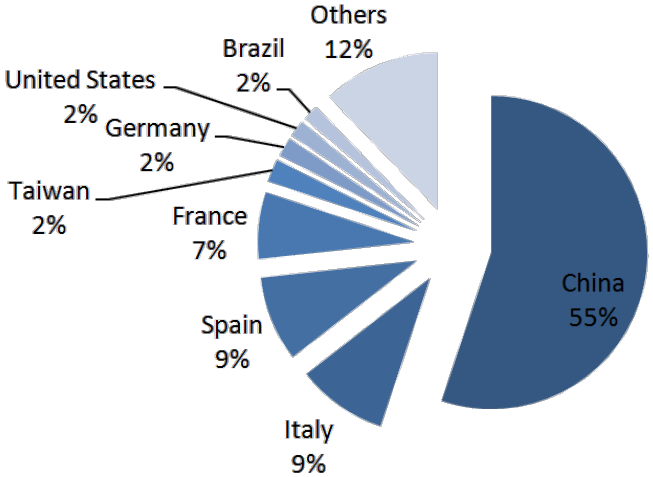


Geographic Distribution of Peers

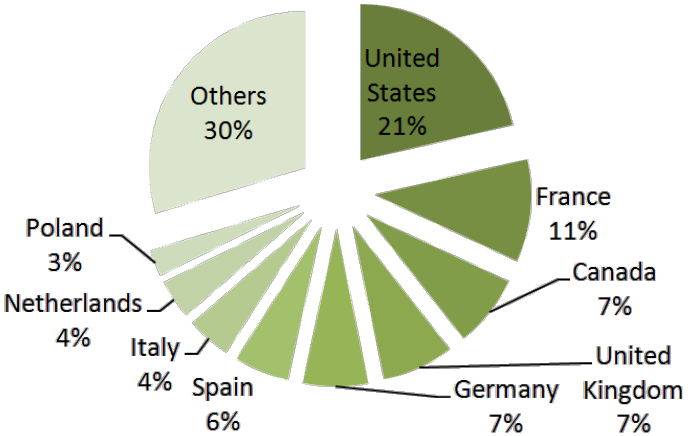
Mainline



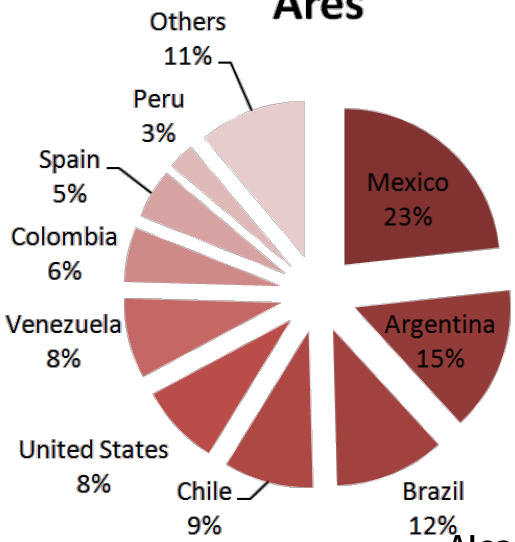
eMule



Azureus

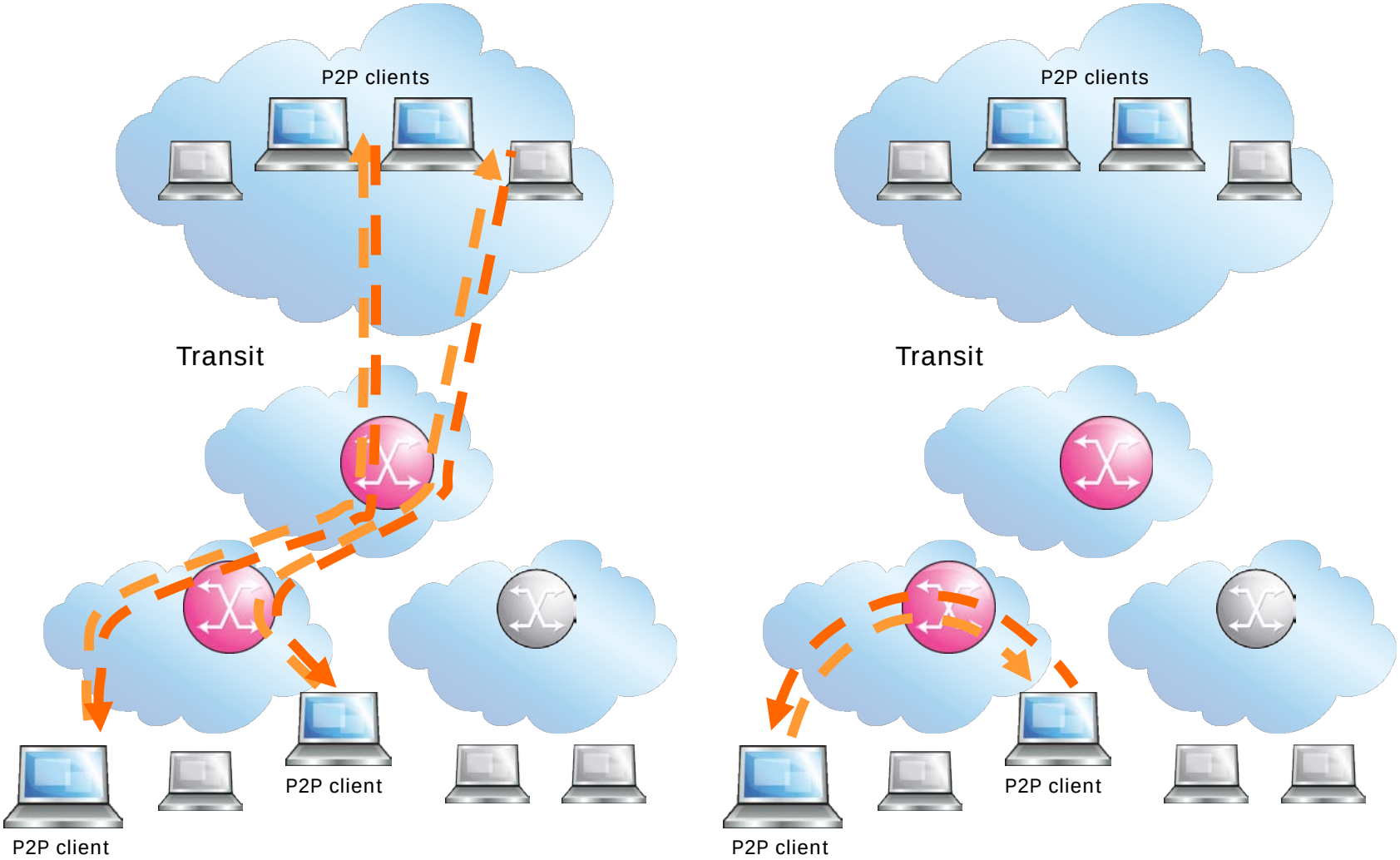


Ares



COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

P2P Traffic Localization



Random vs. Localized Peer Selection

Before: Random Peer Selection

After: “Local” Peer Selection

Approaches to Localization

- Client based
- Client – Network collaboration
- Network based

Trackers disappear – DHT takes over

German Police Shut Down Private Torrent Tracker

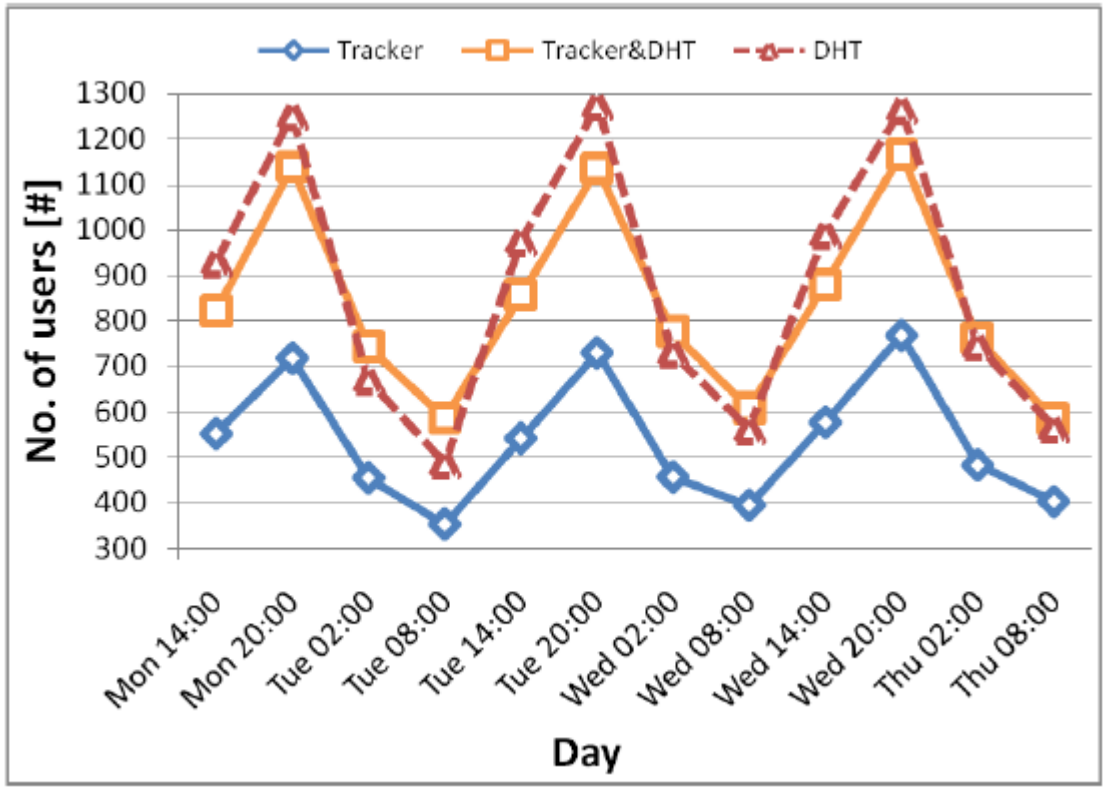
March 30, 2009 by Yonah
Filed under: Announcements & Events, File-Sharing Programs, Networks & Services, Legal P2P News & Issues

The battle against P2P file sharing has a lot of facets – w
on the abstract ground of law the repeat file sharers have
reasons to rejoice (as the European Parliament dismiss
the notion of 'three strikes' policy), on the concrete groun
the piracy combat authorities engage in rapid raids to ca
the pirates red-handed. In one such raid German authori
have terminated a private BitTorrent tracker this week an
confiscated lots of PCs, HDDs, DVDs, CDs and money.

30.7.07
Police shuts down Portuguese Trackers

This blog post talks about the news that the Portuguese Police shut down three Portuguese BitTorrent trackers. The whole issue arised many coments in the media and in the blogosphere. All my comments in this post are based on the press releases from four entities, PJ, FEVIP, ASAE and AFP, and the statements they've been giving about this issue (that I had access to, not everything). Also, I'm afraid I'm not using the correct English legal terms in some cases - sorry. Also, I'm not saying here that , specially at this level.

Bulgarian Torrent Tracke



tTorrent

orrent trackers, as two more
cal police. Whether the
l. File.lv, Latvia's largest
summer but has since

lv, Latvia's largest BitTorrent
by the Latvian Economic
sic Producers Association,
ite has since returned, local
t down further BitTorrent

IFPI, Italian po

By Jacqui Cheng | Last upd

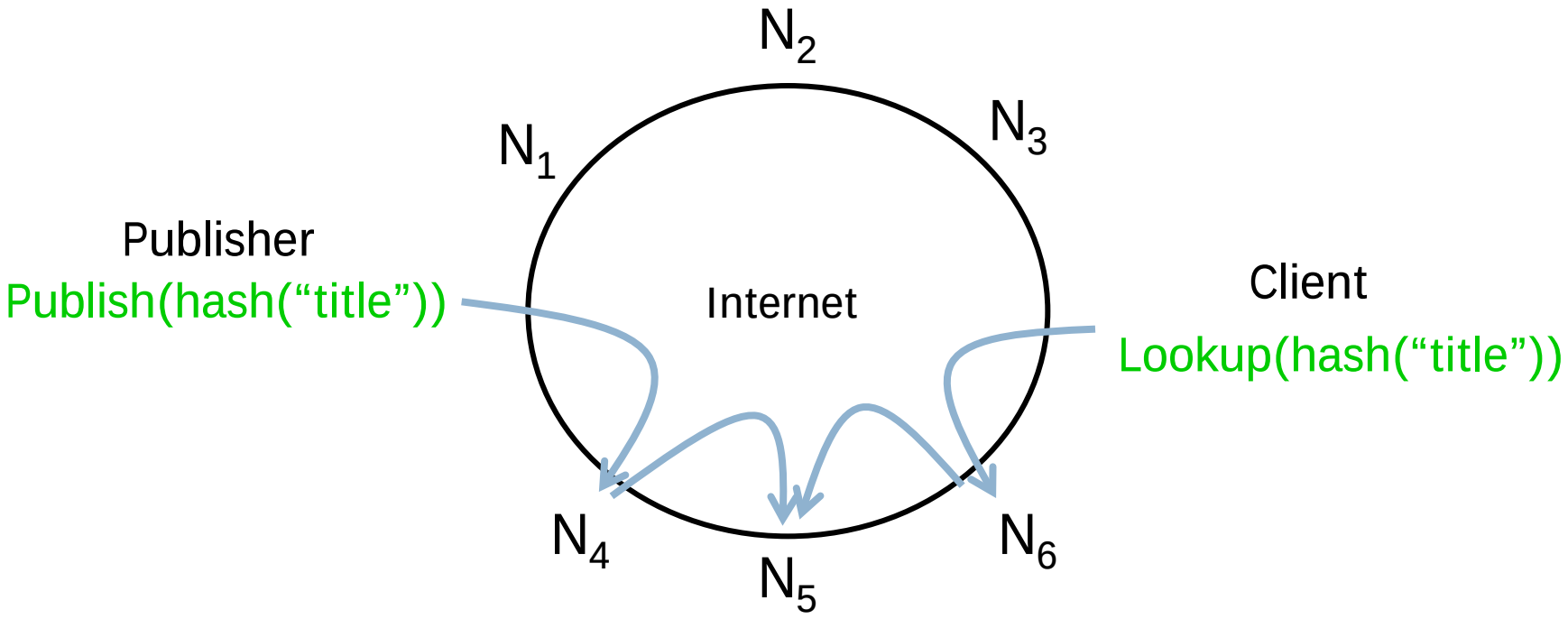
A popular torrent tracke
another victory against P2P. Colombo-BT.org now points to a page that says, "Access denied in execution of an Italian Court Authority injunction." The shutdown came after action was taken by the Guardia di Finanza, an Italian police force that operates under the Italian Armed Forces. Unsurprisingly, the IFPI applauded the move and took it as a sign of progress against piracy online.

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

DHT Traffic Localization

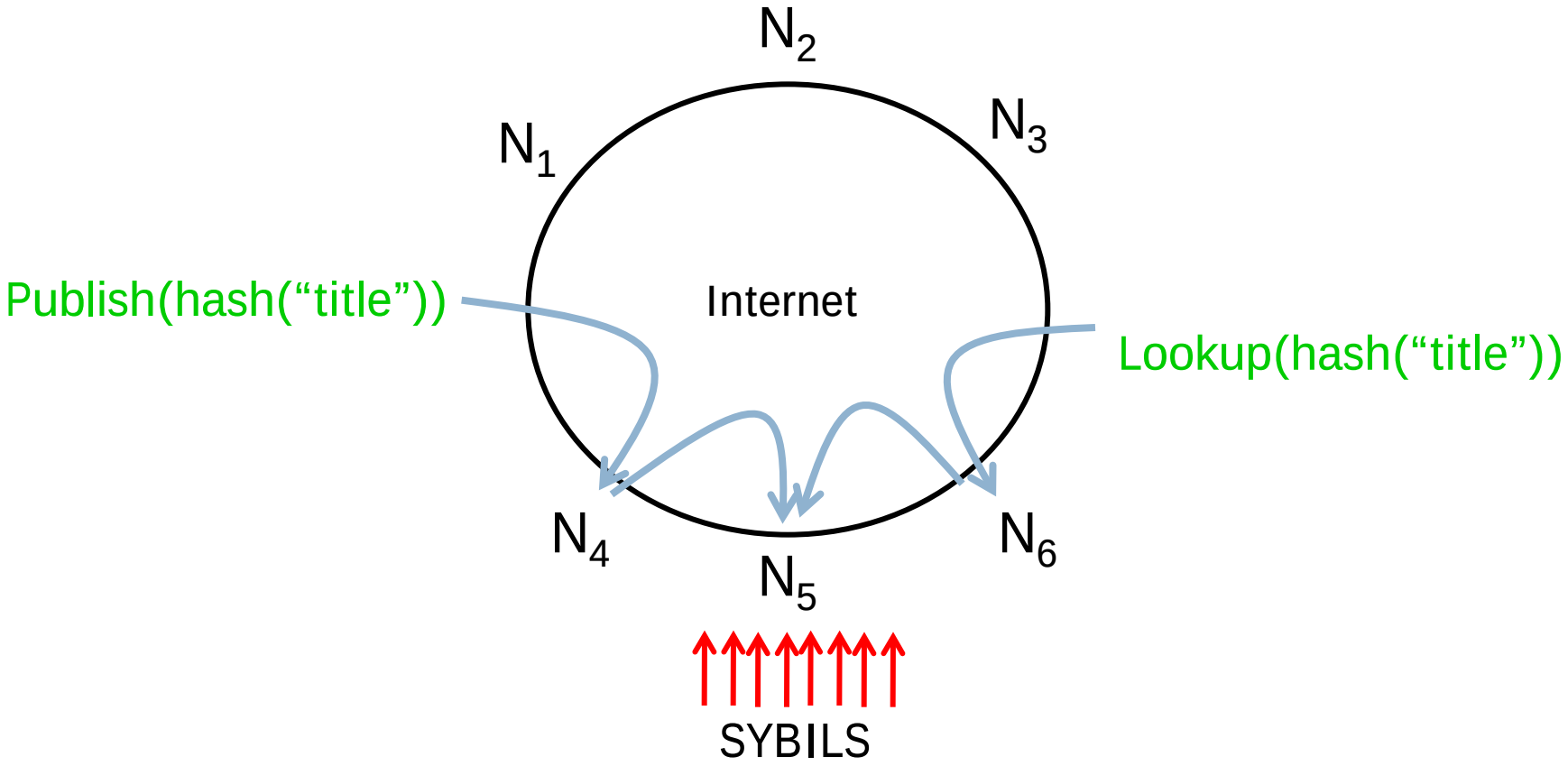
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

What is a DHT?



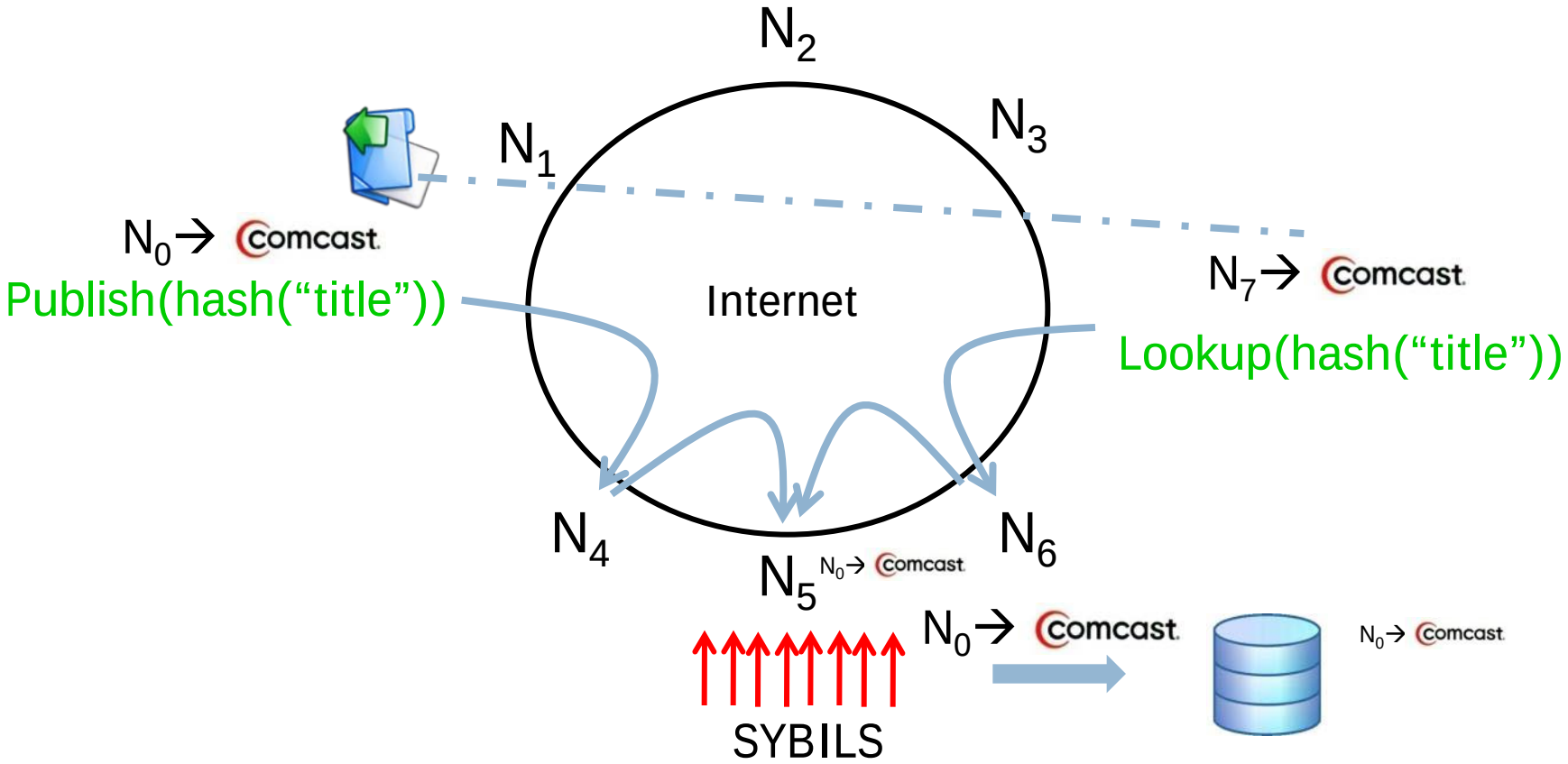
- A distributed database for publishing and searching information
  - Each peer and each object is identified by an info-hash
  - Use object info-hash to decide on which peer to store information

How to control a DHT?



- Intercept all the announces for a file
- Intercept all queries for a file

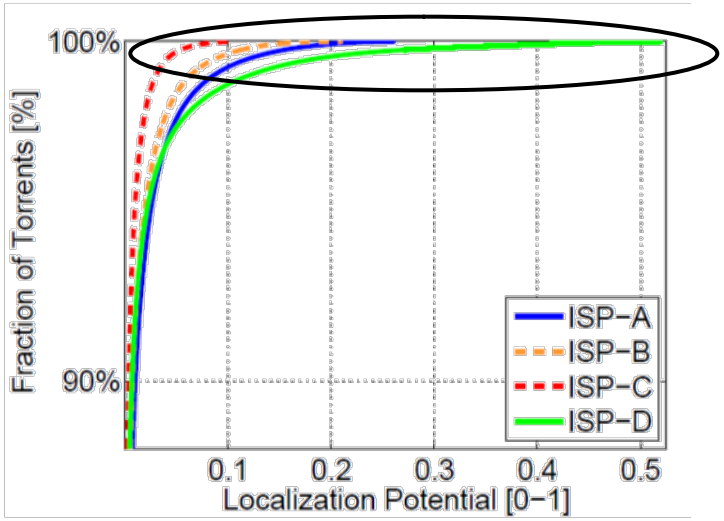
How to localize P2P traffic using the DHT?



- Build a database <peer, ISP> by intercepting all announces for a file
- Localized responses: request from Comcast response with peers from Comcast

Which files should we localize?

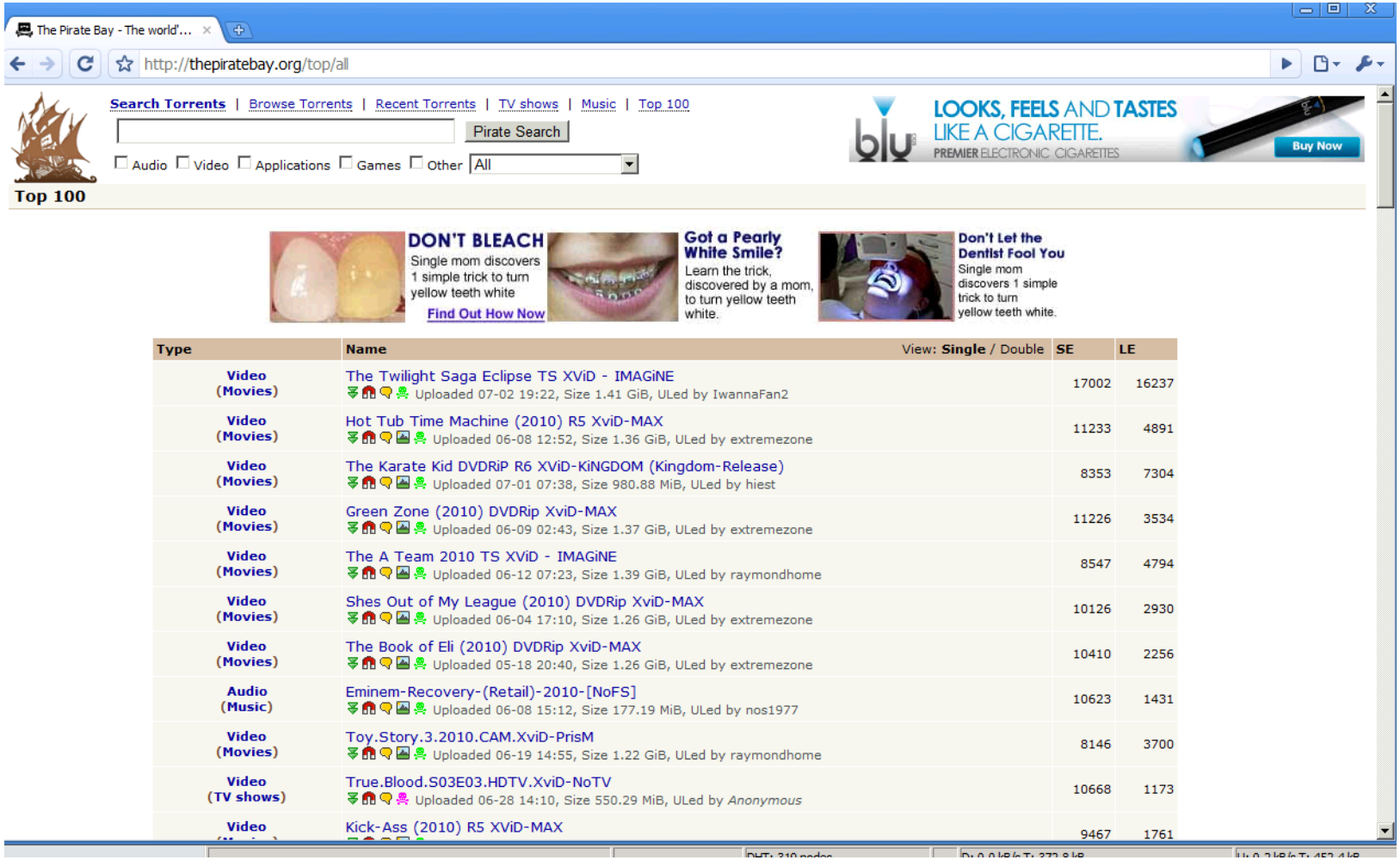
- Ideally, all files
  - Resource consumption scales linearly with the number of files to localize
- We only focus on **popular** files for two reasons:
  - It is infeasible to monitor all files available in a DHT
  - Localizing each file is unnecessary



Traffic Localization in Action!

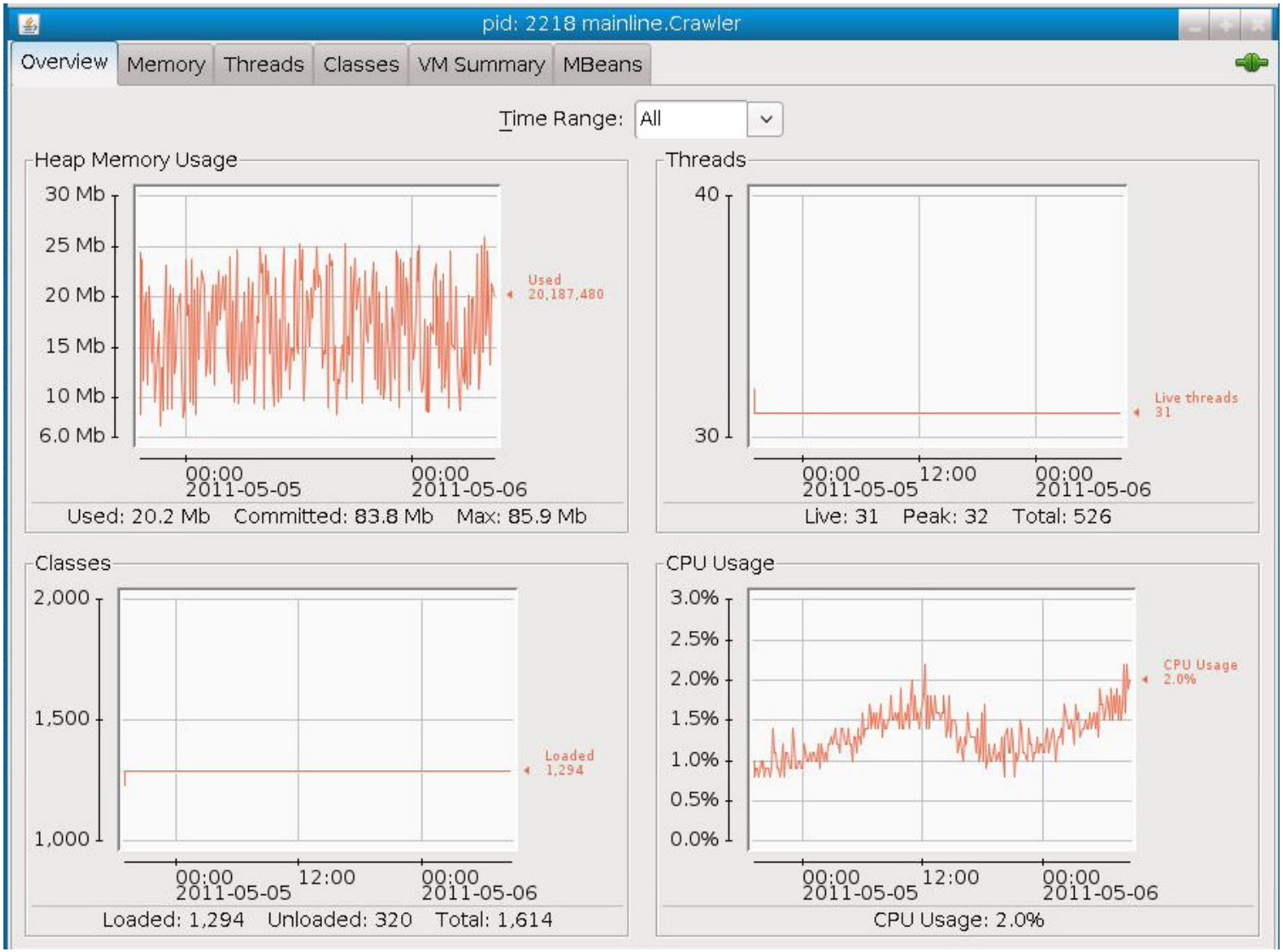
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

1) Choose a popular file - *thepiratebay.org/top/all*



COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

2) Start the Localization Mechanism



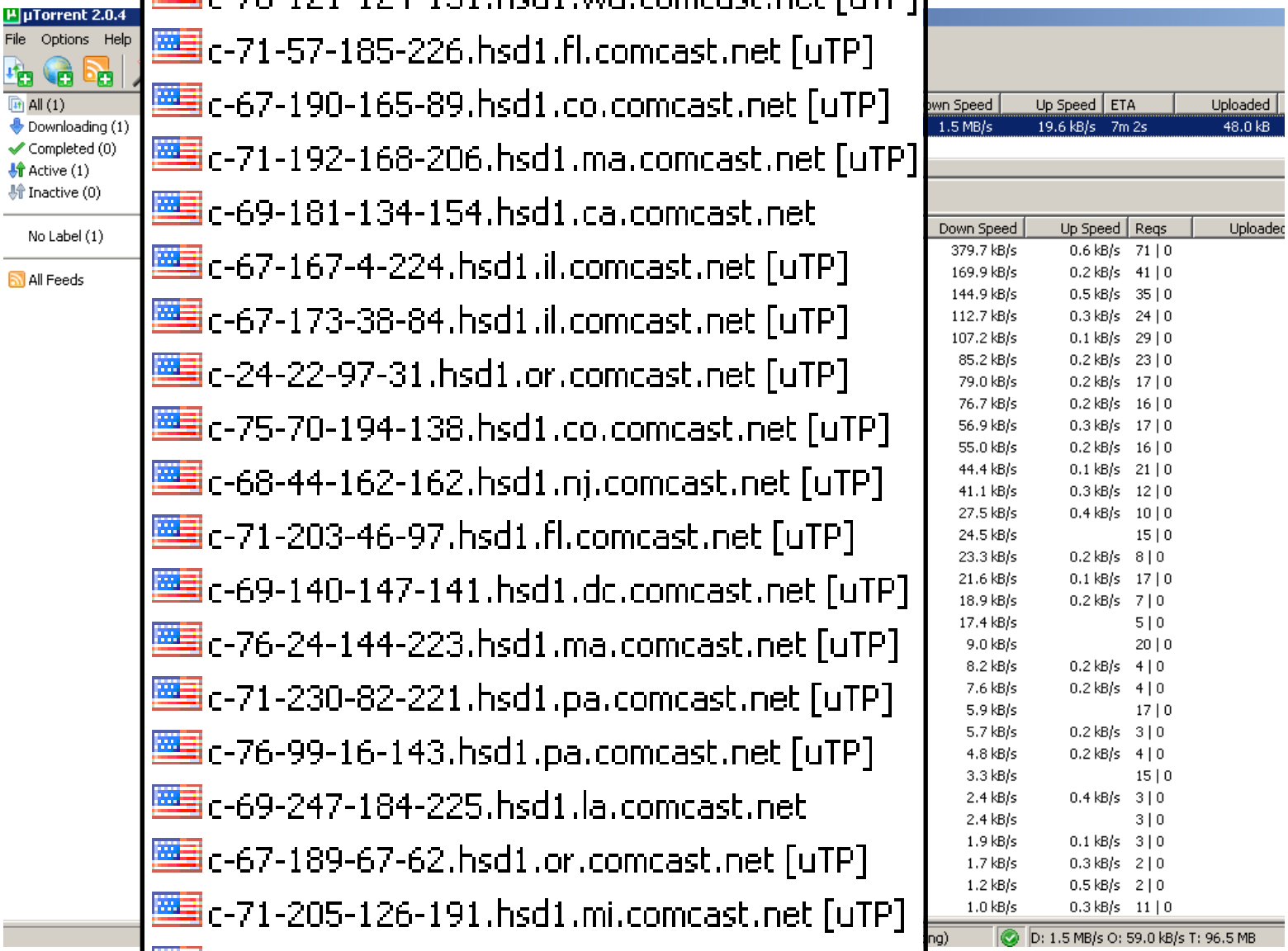
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

DEMO TIME

1. Connect to a windows machine at my place
   - Verizon FIOS
   - Utorrent
2. Download a content item being localized using the presented approach.
3. Observe
   - from which peers the client downloads from
   - the interaction with the sybils

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

3) Download locally!



c-76-121-124-131.hsd1.wa.comcast.net [uTP]
c-71-57-185-226.hsd1.fl.comcast.net [uTP]
c-67-190-165-89.hsd1.co.comcast.net [uTP]
c-71-192-168-206.hsd1.ma.comcast.net [uTP]
c-69-181-134-154.hsd1.ca.comcast.net
c-67-167-4-224.hsd1.il.comcast.net [uTP]
c-67-173-38-84.hsd1.il.comcast.net [uTP]
c-24-22-97-31.hsd1.or.comcast.net [uTP]
c-75-70-194-138.hsd1.co.comcast.net [uTP]
c-68-44-162-162.hsd1.nj.comcast.net [uTP]
c-71-203-46-97.hsd1.fl.comcast.net [uTP]
c-69-140-147-141.hsd1.dc.comcast.net [uTP]
c-76-24-144-223.hsd1.ma.comcast.net [uTP]
c-71-230-82-221.hsd1.pa.comcast.net [uTP]
c-76-99-16-143.hsd1.pa.comcast.net [uTP]
c-69-247-184-225.hsd1.la.comcast.net
c-67-189-67-62.hsd1.or.comcast.net [uTP]
c-71-205-126-191.hsd1.mi.comcast.net [uTP]
c-68-82-165-106.hsd1.pa.comcast.net [uTP]

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

The Sybils control the content

No.	Time	Source	Destination	Protocol	Info
3	5.504641	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49403
4	5.552394	192.11.19.2	192.168.1.10	UDP	Source port: 49403 Destination port: 25834
5	5.552656	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49404
6	5.602966	192.11.19.2	192.168.1.10	UDP	Source port: 49404 Destination port: 25834
7	5.603153	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49408
8	5.650895	192.11.19.2	192.168.1.10	UDP	Source port: 49408 Destination port: 25834
9	5.651141	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49411
10	5.697484	192.11.19.2	192.168.1.10	UDP	Source port: 49411 Destination port: 25834
11	5.697728	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49410
12	5.746730	192.11.19.2	192.168.1.10	UDP	Source port: 49410 Destination port: 25834
13	5.746937	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49413
14	5.801294	192.11.19.2	192.168.1.10	UDP	Source port: 49413 Destination port: 25834
15	5.801546	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49412
16	5.851896	192.11.19.2	192.168.1.10	UDP	Source port: 49412 Destination port: 25834
17	35.015617	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49404
18	35.015677	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49407
19	35.015708	192.168.1.10	192.11.19.2	UDP	Source port: 25834 Destination port: 49408
20	35.062017	192.11.19.2	192.168.1.10	UDP	Source port: 49404 Destination port: 25834
21	35.063339	192.11.19.2	192.168.1.10	UDP	Source port: 49407 Destination port: 25834
22	35.063355	192.11.19.2	192.168.1.10	UDP	Source port: 49408 Destination port: 25834

```
⊞ Frame 17 (196 bytes on wire, 196 bytes captured)
⊞ Ethernet II, Src: Wistron_23:a8:47 (00:16:d3:23:a8:47), Dst: Netgear_bb:10:dc (00:1f:33:bb:10:dc)
⊞ Internet Protocol, Src: 192.168.1.10 (192.168.1.10), Dst: 192.11.19.2 (192.11.19.2)
⊞ User Datagram Protocol, Src Port: 25834 (25834), Dst Port: 49404 (49404)
⊞ Data (154 bytes)

0000  00 1f 33 bb 10 dc 00 16  d3 23
0010  00 b6 cf de 00 00 80 11  d5 98
0020  13 02 64 ea c0 fc 00 a2  d5 8b
0030  3a 69 64 32 30 3a 8d 9b  89 34
0040  8a cc cb f2 fe 03 da a0  55 27
0050  5f 68 61 73 68 32 30 3a  88 74
0060  f0 58 5a 42 4d 23 9f 32  4b 43
0070  72 74 69 32 35 38 33 34  65 35
0080  32 30 3a c7 00 19 9e 87  33 f3
0090  83 77 59 35 cf fe a3 65  31 3a
00a0  6e 6f 75 6e 63 65 5f 70  65 65
00b0  81 2a 00 00 31 3a 76 34  3a 55
00c0  31 3a 71 65
```

```
6f    .XZBM#.2 KCB.4:po
6e    rti25834 e5:token
cd    20:..... 3..aC2~.
6e    .wY5...e 1:q13:an
3a    nounce_p eer1:t4:
79    .*..1:v4 :UTH81:y
      1:qe
```


COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation

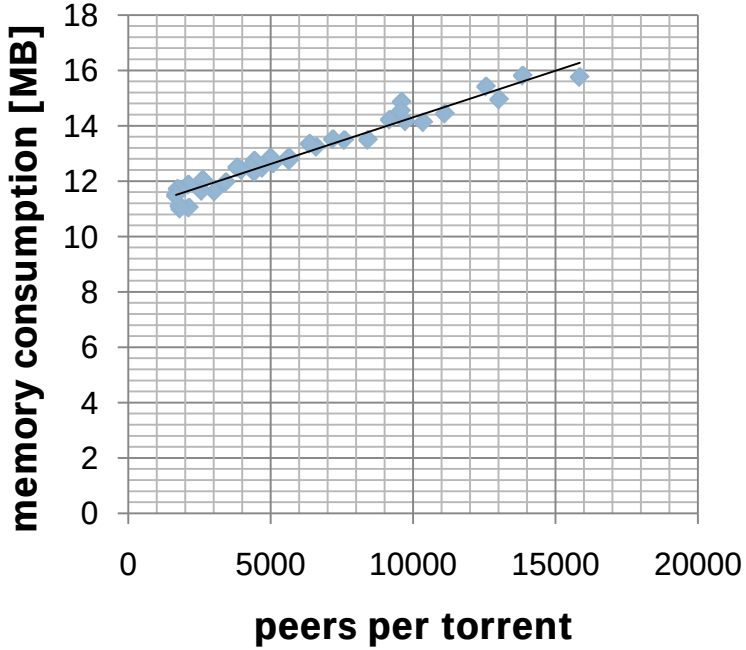
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

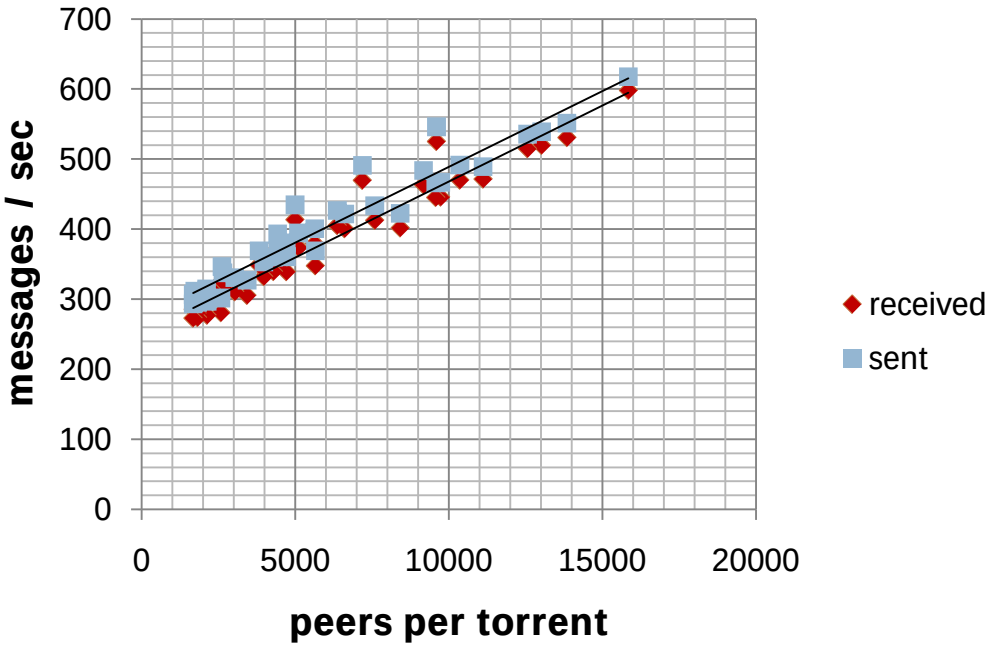
Evaluation – Methodology

- Select torrents to localize (use PirateBay)
  - Run the sybils
- Instrument a Transmission, Utorrent and BitComet client to repetitively download the selected torrents
  - Download time is bound
  - One day, one week, one month periods
  - No client-to-tracker communication
- We run the instrumented clients on Private machines connected to the Internet with
  - Cable
  - Fiber
  - DSL

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation – *Resource Consumption*

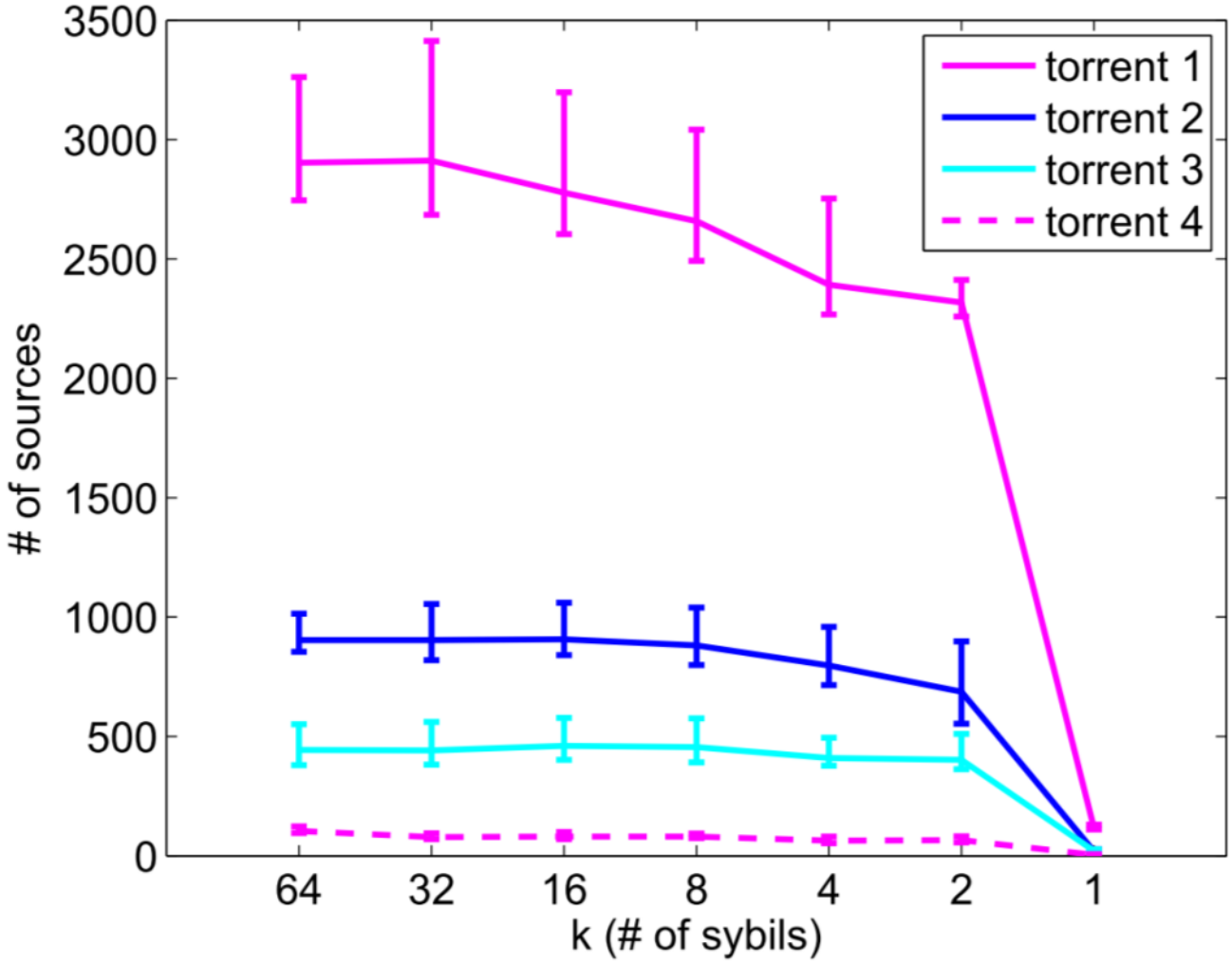




- Resource consumption linearly grows with the number of peers per torrent
- Number of message per second is the critical resource
  - ca. 20k / sec using a small linux server in our lab
  - ca. 10k / sec using a “high IO” amazon ec2 VM

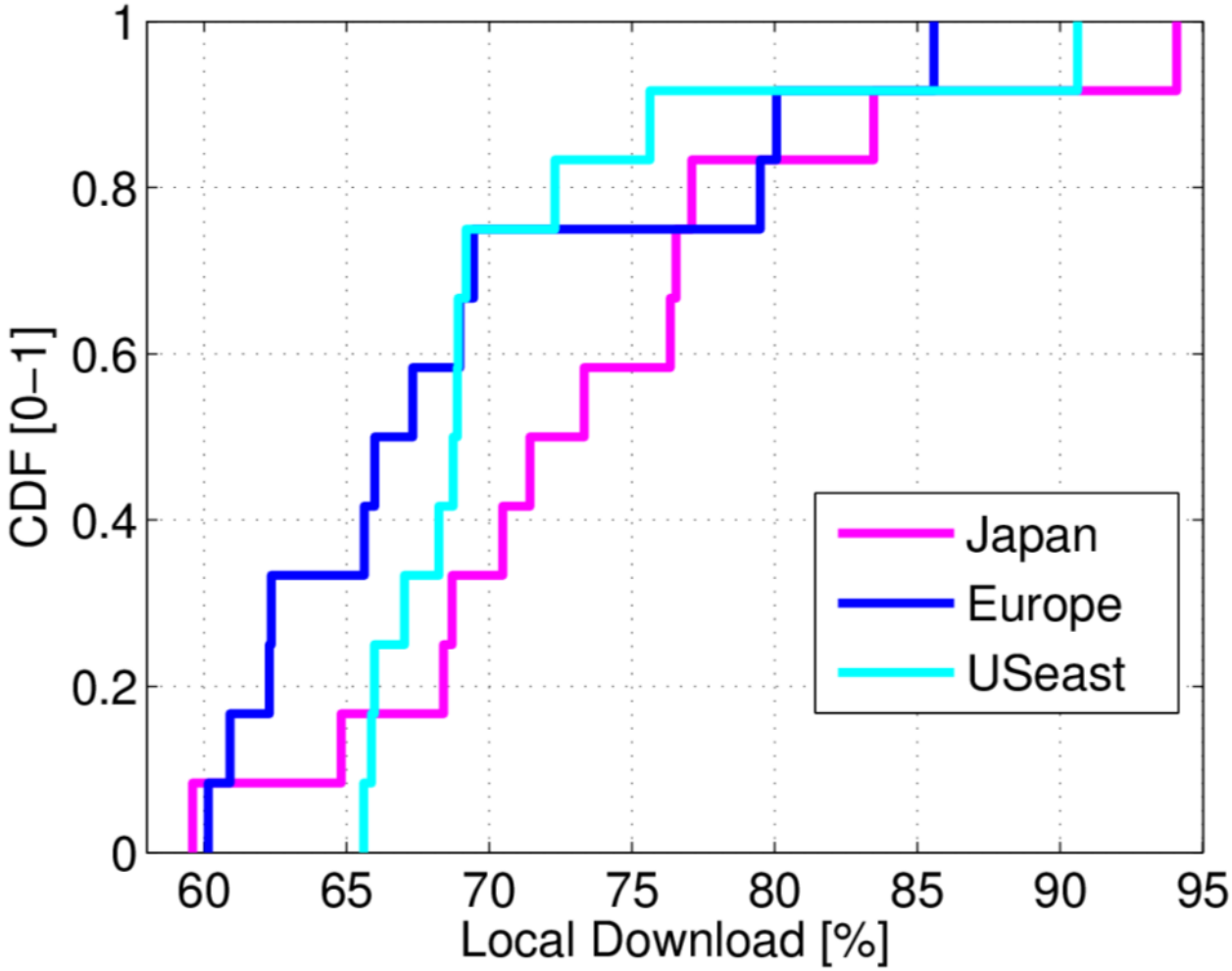
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Server Side Evaluation – *Nr of Sybils*



COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

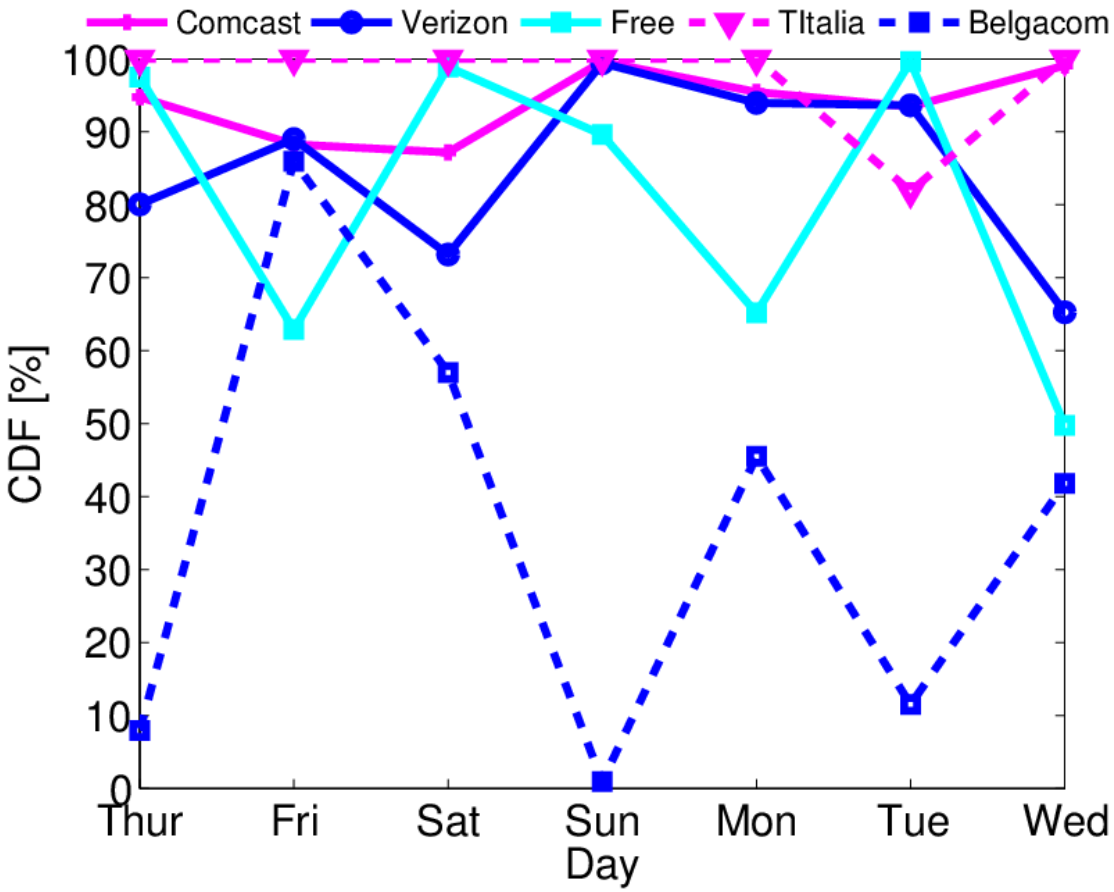
Server Side Evaluation – *Running from the Cloud*



COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

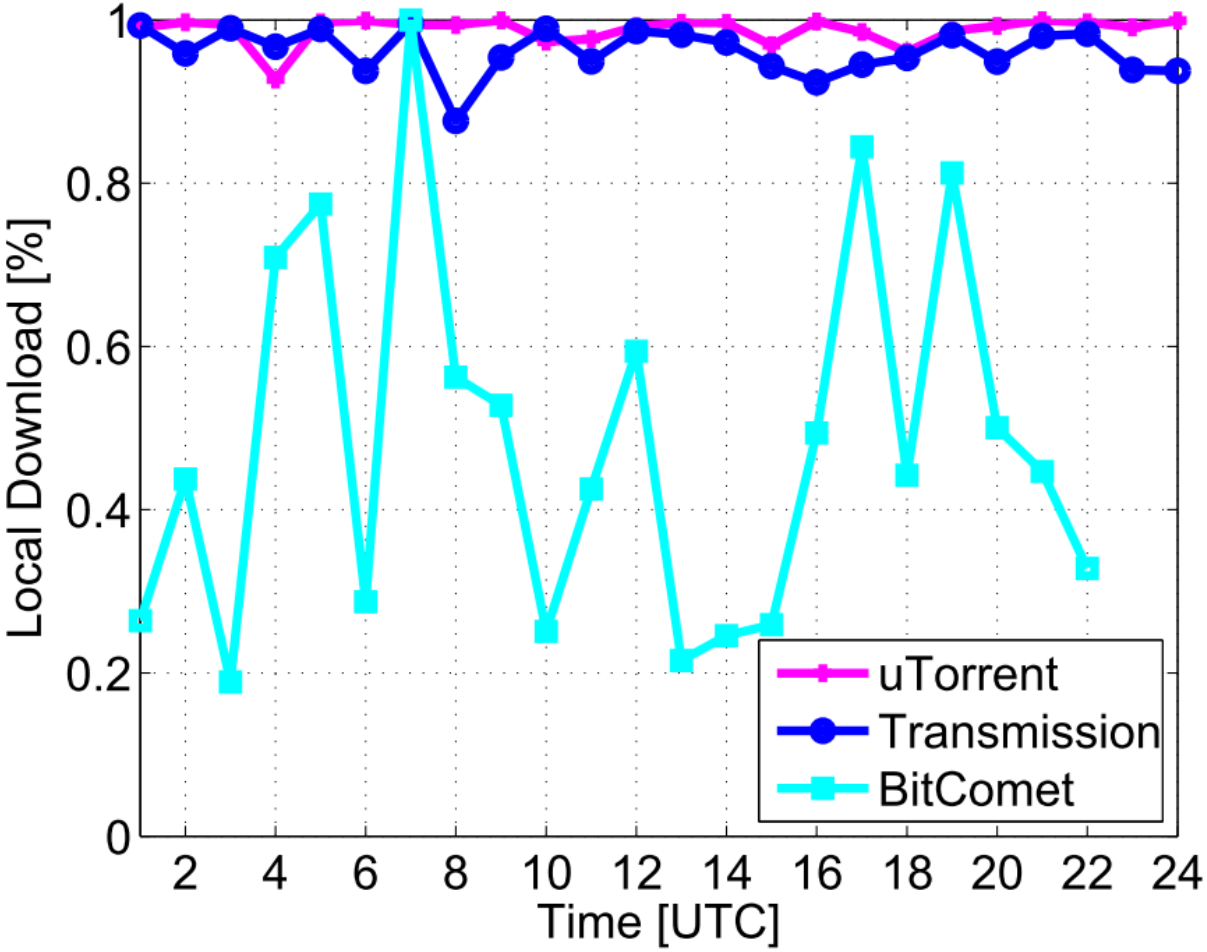
Evaluation – *Traffic Localization*

BitComet, Ktorrent …



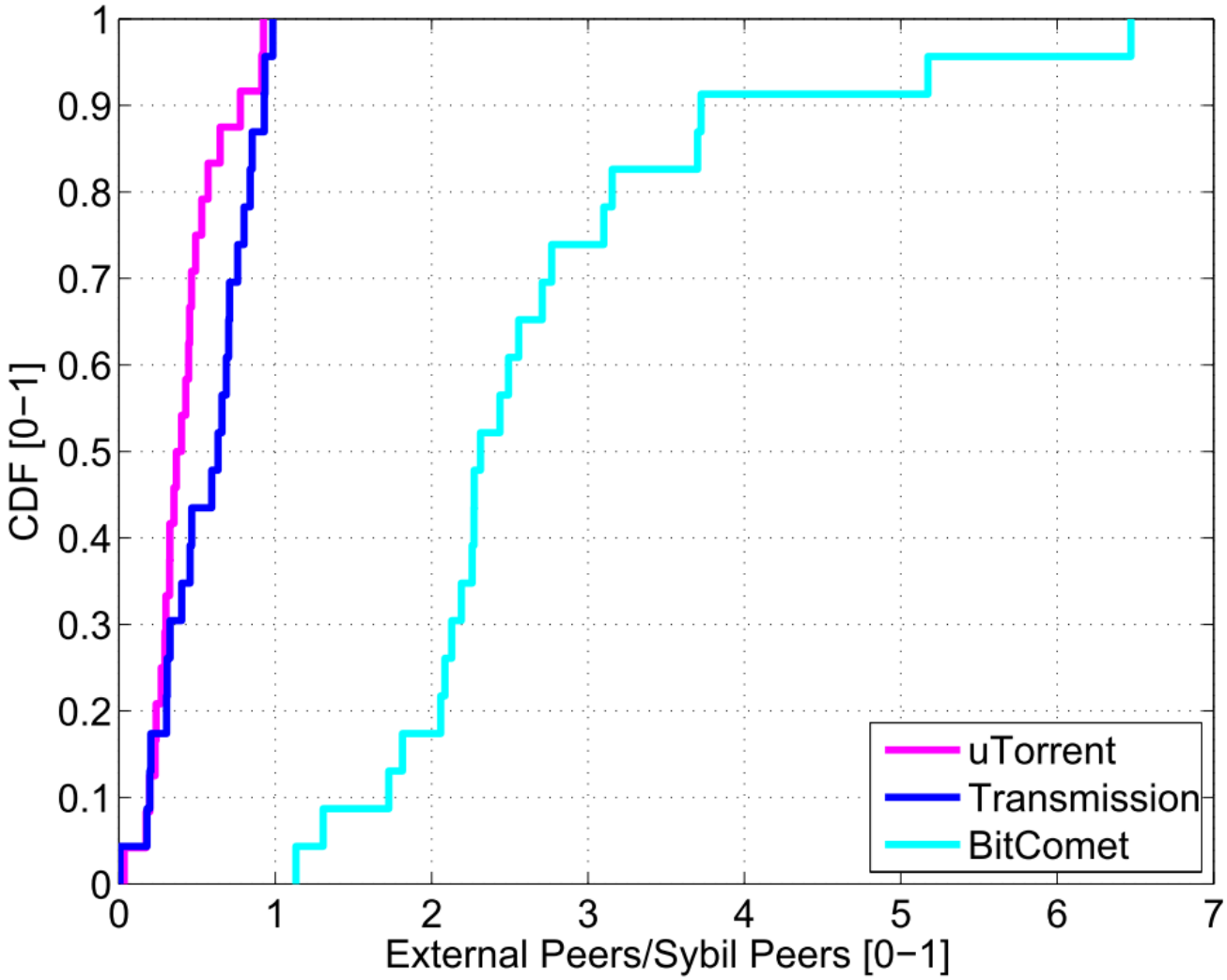
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation – *Impact of the Client*



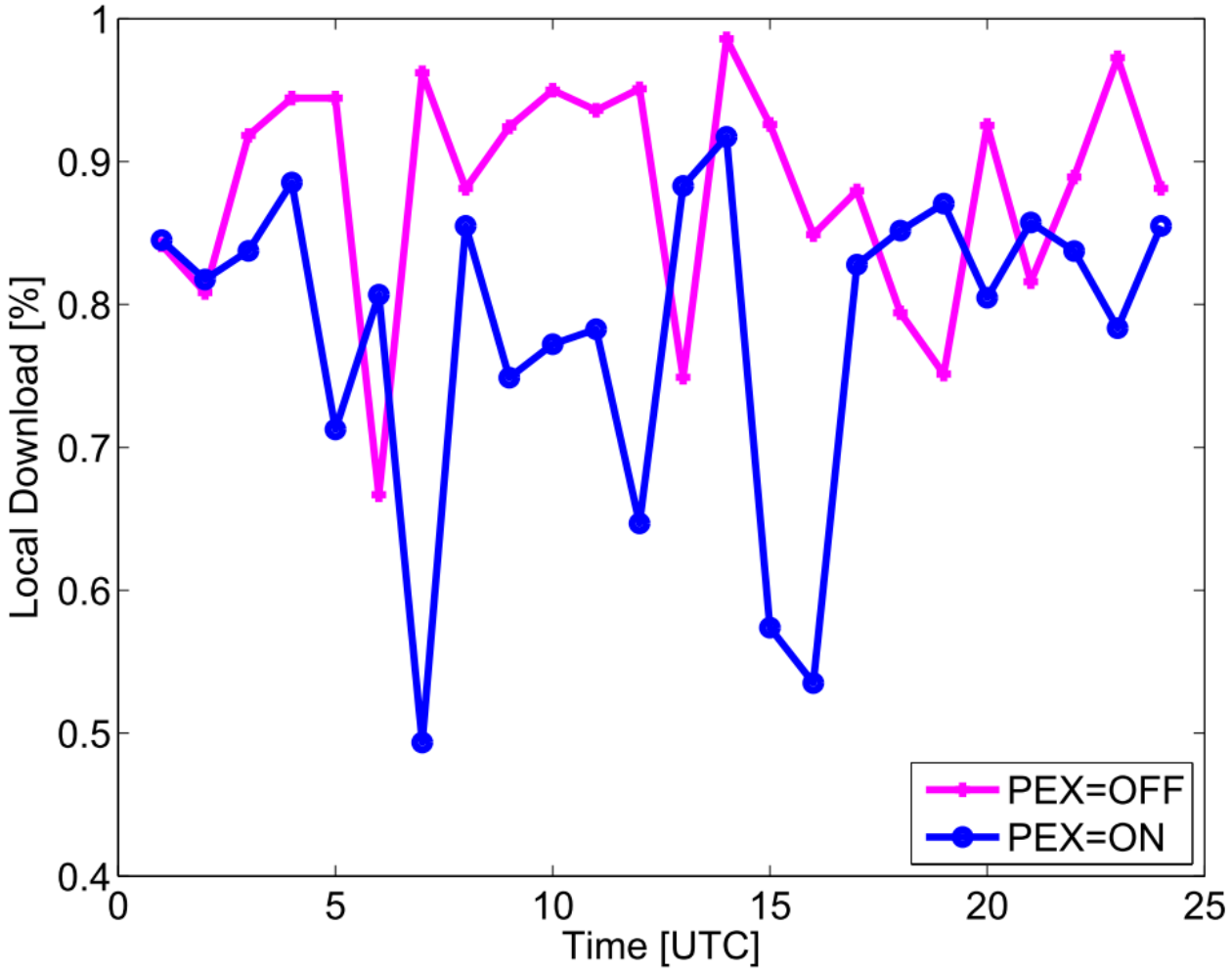
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation – *Competition Sybils vs. Other Peers*



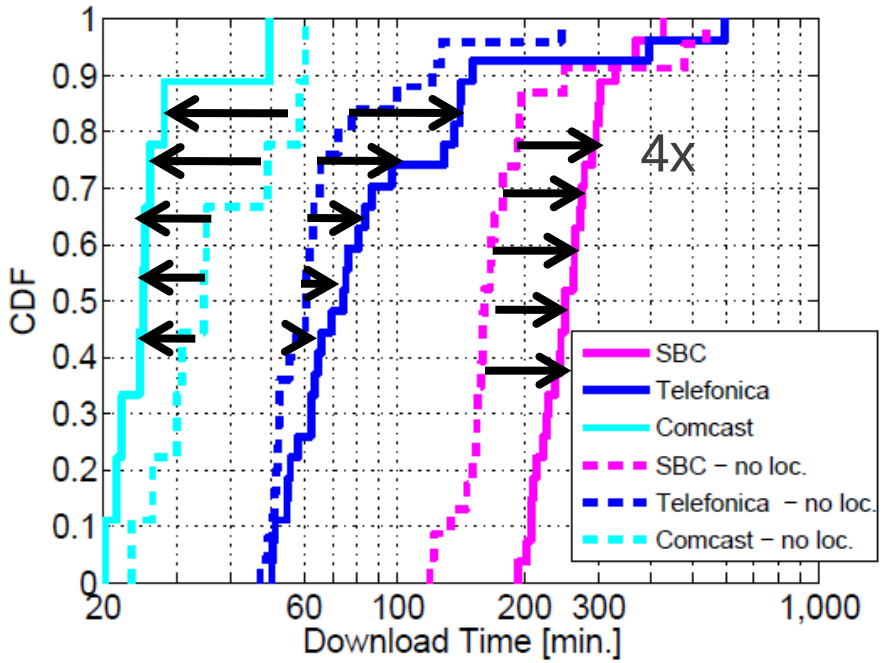
COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation – *Impact of PEX*



COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Evaluation – *Download Speed Analysis*



- CAUSES
  - Different local popularity of F
  - Peers located at Telefonica have low upload rate and cannot contribute much to the swarm

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Conclusions

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Conclusion

- Today, the role of the trackers is distributed among peers via a Distributed Hash Table (DHT)
  - Existing solutions for BitTorrent traffic localization leveraging central trackers will become soon ineffective
- We design, prototype, and preliminary evaluate the first traffic localization mechanism for DHT-based BitTorrent networks
  - Resource utilization scales with the # of peers per torrents.
  - Can run on any machine with connection to the Internet.
  - Issues with flawed protocol implementations.
  - Works very well for popular torrents in fast networks.

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Questions?

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

What is a sybil attack?

- Sybils - A single entity can join a P2P network many times with many distinct logical identities (nodeIDs)
- Eclipse attack - Eclipse a file with info-hash **F** from a P2P network
- How? - Insert in a DHT several sybils with nodeIDs close to **F**
  - Crawl the surrounding zone of **F** ,**Z**, to learn about the peers in **Z**
  - Send HELLO messages to each peer in **Z** from the sybils
    - Each peer in **Z** believes the sybils are the closest peers to **F**
    - Announce and search requests for **F** falling into **Z** are routed to one of the sybils

COPYRIGHT © 2011 ALCATEL-LUCENT. ALL RIGHTS RESERVED.