Delft University of Technology
Parallel and Distributed Systems Report Series

Correlating Topology and Path Characteristics of Overlay Networks and the Internet

A. Iosup, P. Garbacki, J.A. Pouwelse and D.H.J. Epema
{A.Iosup,P.Garbacki,J.A.Pouwelse,D.H.J.Epema}@ewi.tudelft.nl

report number PDS-2005-002

ISSN 1387-2109

Published and produced by:
Parallel and Distributed Systems Section
Faculty of Information Technology and Systems Department of Technical Mathematics and Informatics
Delft University of Technology
Zuidplantsoen 4
2628 BZ Delft
The Netherlands

Information about Parallel and Distributed Systems Report Series:
`reports@pds.twi.tudelft.nl`

Information about Parallel and Distributed Systems Section:
`http://pds.twi.tudelft.nl/`

Abstract

Real-world IP applications such as Peer-to-Peer file sharing are now able to benefit from network and location awareness. It is therefore crucial to understand the relation between underlay and overlay networks and to characterize the behavior of real users with regard to the Internet. For this purpose, we have designed and implemented MultiProbe, a framework for large-scale P2P file sharing measurements. Using this framework, we have performed measurements of BitTorrent, which is currently the P2P file sharing network with the largest amount of Internet traffic. We analyze and correlate these measurements to provide new insights into the topology, the connectivity, and the path characteristics of the Internet parts underlying P2P networks, as well as to present unique information on the BitTorrent throughput and connectivity.

Keywords: peer-to-peer, overlay networks, Internet measurement, BitTorrent topology.

Contents

1 Introduction **4**

2 Background **4**
2.1 A brief description of BitTorrent 4
2.2 Related work 5

3 The MultiProbe framework **6**
3.1 Pre-processing 6
3.2 Measurement 7
3.3 Post-processing 7
3.4 Implementation 7

4 Experimental Setup **7**
4.1 Pre-setup 7
4.2 Measurements 8

5 Measurement Results **8**
5.1 Measurements summary 8
5.2 Location analysis 9
5.2.1 Continent/country location 9
5.2.2 City location 9
5.2.3 Internet Organization location 10
5.2.4 Autonomous System location 10
5.3 Route and connectivity analysis 10
5.3.1 IP path hop count distribution 11
5.3.2 AS traversals distribution 11
5.3.3 Intra-AS path hop counts 11
5.3.4 Round trip time 12
5.3.5 Application-level connectivity 12
5.4 Traffic break-down 13
5.4.1 Traffic per TCP port distribution 13
5.4.2 Application-level bandwidth 13

6 Conclusions and future work **14**

7 Availability **14**

List of Figures

1 The life cycle of a BitTorrent download . . . 5
2 The MultiProbe framework . . . 6
3 The Internet Organizations location . . . 10
4 The Autonomous Systems location . . . 10
5 The distribution of the RTT . . . 12
6 Application-level connectivity . . . 12
7 Application-level bandwidth . . . 13

List of Tables

1 The number of torrents (files) and users per measured web site . . . 7
2 A summary of the characteristics of the tracked files. . . . 8
3 Statistics of the size of our measurements. . . . 9
4 The percentages of users per continent and country. . . . 9
5 The distribution of the IP path hop count . . . 11
6 The distribution of the AS traversal count . . . 11
7 The distribution of the intra-AS path hop count . . . 11
8 The distribution of the RTT, per classes . . . 12
9 The distribution of users and user weights per TCP port. . . . 13

1 Introduction

The topology and the characteristics of the Internet have a large impact on the operation of Internet applications such as multicast overlays [12], web hosting [7], and Peer-to-Peer (P2P) file-sharing systems [19]. In order to understand and improve the performance of these applications, it is essential to evaluate the way the underlay network supports the application overlays. In this work we present a measurement framework that allows joint large-scale measurements of Internet and the BitTorrent P2P overlay, which is arguably the largest current Internet application.

Large P2P networks continuously have more than 1,000,000 users[1], and P2P file-sharing networks account now for more than one third of the total Internet traffic [6, 9]. Over the past couple of years, BitTorrent has become the largest P2P file-sharing network in terms of generated traffic [9]. We therefore target our study on correlating the the characteristics of BitTorrent and its Internet underlay. For this, we have designed and implemented MultiProbe, a large-scale P2P measurement framework. With MultiProbe, we build on the experience accumulated in our previous work [14], where we have investigated the high level characteristics and the user behavior of BitTorrent. Here, we complement our previous work by focusing on the following research topics:

Measuring underlay/overlay networks How to build a large-scale distributed infrastructure that can measure at the same time P2P and Internet characteristics? How to select a small part of the P2P network such that the measurement results are meaningful?

Characterizing overlay networks and their users Where are the overlay network users located? What is the geographical distribution of traffic? What is the connectivity among the users? What is the application throughput?

Correlating underlay/overlay measurements What is the topology, and what are the characteristics of those parts of the Internet that act as the BitTorrent underlay? What is the size of BitTorrent, for various Internet metrics? What are the ports on which BitTorrent traffic is present?

The main contribution of our work is that we present a correlated view of the dominant overlay network and the Internet, based on measurements taken in May 2005. We show evidence that the majority of BitTorrent users are located in Europe but, at the same time, that BitTorrent traffic is globally spread. We correlate the users' geographical locations with the generated traffic. Our results confirm that BitTorrent users are well connected by the Internet in terms of latency, hop count, and bandwidth. Finally, we find that while standard ports are still regularly used, the highest traffic volume is generated on non-standard ports.

2 Background

In this section we provide a brief description of the BitTorrent P2P system, and we review related work.

2.1 A brief description of BitTorrent

BitTorrent is a P2P file-sharing network focusing on high data transfer speed rather than on search capabilities. BitTorrent is currently the largest P2P file-sharing network with over one third of the world's P2P traffic [9], thus generating more than 15% of the total Internet traffic[2].

BitTorrent focuses on high data transfer speed rather than search capabilities. Data, be it single files or file archives, exist in BitTorrent in the form of *torrents*. To facilitate the file exchange process, data is split in smaller parts, called *chunks*. A user (*peer*) has a complete file only after obtaining all the chunks composing that file.

There are three levels in the BitTorrent world: the peer level, the tracker level, and the web-site level. The BitTorrent users are active at the *peer level*. They are using an advanced *tit-for-tat* (bartering) protocol to share file chunks. This mechanism ensures that for a user the amount of incoming data is roughly equal to the amount of outgoing data. The BitTorrent advanced *bartering partners selection protocol* favors users

[1] Slyck.com lists size information for many large P2P networks.
[2] Sprint. Packet Trace Analysis. http://ipmon.sprintlabs.com/

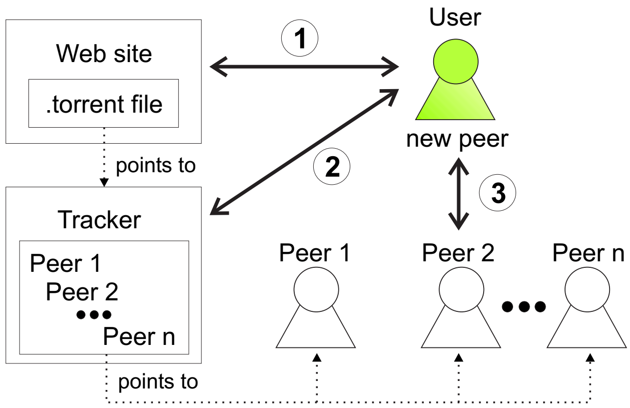

Figure 1: The life cycle of a BitTorrent download (the steps are explained in the text).

with higher bandwidth and users that have the rarest (or unique) file chunks, and guarantees filling up all the available bandwidth of a peer. The *trackers* are semi-centralized components (their activity is centralized but their number makes the whole level work as a decentralized network) used to keep track of the network transfers. Each file in the network is monitored (*tracked*) by a tracker. By using the services of a tracker, BitTorrent peers can discover the IP addresses of other peers that have chunks of the desired file, at a limited rate, typically 25 peers/minute. *Web sites* are used to locate data. For this, the web sites provide pages with BitTorrent content descriptions (movie and music titles, authors, number of user who have the complete file or parts of it, and so on) that a network user may need to retrieve the desired data. Since BitTorrent has no searching mechanism, metadata descriptions of the torrents are created in the form of `.torrent` files, which are stored on web sites. The metadata description contains the location of at least one tracker.

Figure 1 depicts the life cycle of a BitTorrent download. The user would first go to a BitTorrent web site, find the files that she desires (movies, music, or any other content), then download the description of the file tracker in the form of a `.torrent` file (step 1). The `.torrent` file contains the location of at least one tracker. The user then connects to the tracker with her BitTorrent client (peer software). The client software automatically handles the BitTorrent bartering partners selection protocol and tit-for-tat mechanism: whenever needed, it gets a small and random list of network peers from the tracker (step 2), and contacts as many as 50 of them, in order to find 5-10 peers that are willing to exchange chunks with it (step 3). The user would just have to remain online until the complete file has been downloaded. As the BitTorrent file sharing protocol is based on transferring chunks, BitTorrent clients can also resume interrupted downloads, or re-download incorrect chunks.

2.2 Related work

There are many measurement studies of either the Internet or large P2P networks.

An important effort to characterize the global Internet infrastructure is described [1]. The goals of this work are to acquire infrastructure-wide connectivity information, collect round-trip-time and path data, analyze the frequency and patterns of routing changes, and visualize network-wide connectivity. More results of this work were presented in [4]. The Mercator Internet mapping tool [2] attempts to map the Internet from a single, arbitrary location. The Rocketfuel Internet mapping tool [17] uses traceroutes sourced at over 750 vantage points to explore the network topology of 10 ISPs in different continents.

One of the first studies regarding the characterization of P2P file-sharing networks is the work of Saroiu et al. [15]. They characterize the one-point-to-target latency, the bottleneck bandwidth, the user connection/disconnection frequency, and the number of files in the network, and correlate this data. However, their results suffer from the *vantage point* effect [10], and are representative for different types of P2P networks, i.e., Napster and Gnutella. A comprehensive model of the KaZaA P2P network is presented in [3]. Both studies have a P2P-modeling perspective, and do not attempt to also characterize the underlying network. In [5], a 5-month trace of a single file shared using the BitTorrent protocol is presented. The file comes from the operating system domain, thus not being representative for P2P, where users download mostly

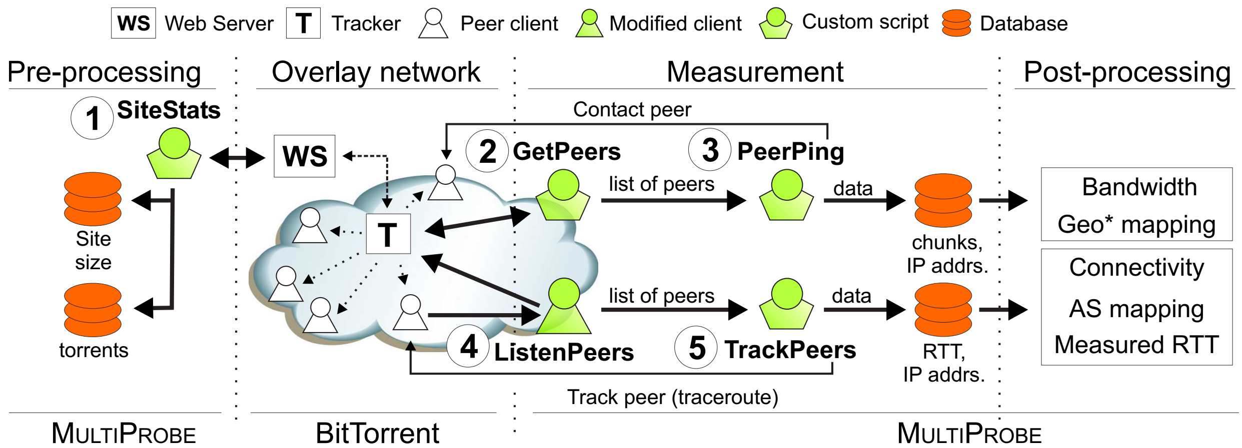

Figure 2: The MultiProbe framework for correlated measurements of large scale overlay networks and the Internet.

movies and music. Their results need confirmation from a boarder study, such as this. In [13], the authors evaluate the traffic of ADSL users from an undisclosed location, and offer detailed statistics for a number of P2P networks. Their results are strongly influenced by the *vantage point* effect. A comprehensive study of BitTorrent's characteristics has been performed by three of the authors of this work in [14]. Their work investigates with the high level characteristics the user behavior of BitTorrent. This paper adds the needed peer and network level insights, being therefore the natural complement of this previous analysis.

An attempt to correlate Internet and P2P network measurements is made in [8], where network connectivity and file-sharing bandwidth between 25 broadband users is studied. Besides the reduced size of the experimental setup, the peers were placed only in US, which reduces the representativeness of their results. In [16], the authors detail traffic characteristics for three popular P2P file-sharing networks– FastTrack, Gnutella, and DirectConnect. The traffic information is collected from the border routers of a single, though large-scale, ISP. Due to router limitations in generating traffic statistics, standard ports are used to identify P2P traffic. In this work we show that such an approach would be ineffective for the case of BitTorrent (see Section 5.4.1).

To summarize, our work is the first to correlate the characteristics of BitTorrent and its Internet underlay, it is based on large-scale measurements of a real P2P network, and has a much broader diversity of analysis results.

3 The MultiProbe framework

In this section we present the structure of MultiProbe, our framework for correlated measurements of large scale overlay networks and the Internet. MultiProbe consists of pre-processing, measurement, and post-processing parts, which we discuss below.

3.1 Pre-processing

The BitTorrent network uses web sites as content directories. The largest of these sites serve millions of users. MultiProbe can select the largest such web site at the start of the experiments (the `SiteStats` module, see Figure 2, bullet 1). The web site selection module is linked with the measurements setup module, in which the top files from the largest web site are selected, and their respective torrents are downloaded. The files are sorted according to the number of users downloading or offering chunks of these files (*swarm size*).

Table 1: The number of torrents (files) and users per measured web site.

Web-site	Number of torrents	Number of users
Pirates Bay	119,026	1,702,429
Mininova	21,790	1,014,332
TorrentPortal	26,580	848,704

3.2 Measurement

For our measurements, we inserted monitoring nodes (*probes*) into the BitTorrent network. We define two types of measurements: *active-start measurements*, in which probes initiate the contact with BitTorrent peers, and *passive-start measurements*, in which probes wait for an externally initiated contact from BitTorrent peers. The need for the second approach is twofold. First, some measurement platforms limit the number of contacts a user can initiate during a fixed period of time. Second, the first approach does not allow to contact peers behind firewalls, while the second can overcome this limitation.

In active-start measurements, probes contact BitTorrent trackers repeatedly, acquiring lists of peer contacts (the `GetPeers` module, see Figure 2, bullet 2). They subsequently contact the corresponding peers using the BitTorrent protocol, and learn about what chunks are owned by these peers (the `PingPeers` module, see Figure 2, bullet 3). In passive-start measurements, probes register themselves to a tracker, and wait to be contacted by other peers (the `ListenPeers` module, see Figure 2, bullet 4). Discovery packets from multiple sources within the measurement infrastructure are sent to the peers discovered in this way, to characterize the Internet parts that are used by BitTorrent (the `TrackPeers` module, see Figure 2, bullet 5). We call this process *multi-source traceroute*.

3.3 Post-processing

Finally, MultiProbe has a wide-range of post-processing modules, which can process and correlate detailed information about both the overlay network and the Internet characteristics. Currently, MultiProbe can analyze for all data and per file: geographical and geopolitical location, RTT, IP path hop count, AS and ISP mapping, port traffic break-down, application-level bandwidth, application connectivity, and various time patterns.

3.4 Implementation

We have implemented the MultiProbe framework in `Python`[3], `Ruby`[4], and `C`. We used `Python` for the measurement scripts, except for the multi-source traceroute, which was implemented in `Ruby`, due to deployment restrictions. For the `ListenPeers` module, we adapted an open-source `Python` implementation of a BitTorrent client. We used `Python` and `C` for pre- and post-processing.

4 Experimental Setup

This section presents the experimental setup for our measurements.

4.1 Pre-setup

In the pre-setup phase of the experiments, we first select the largest BitTorrent web site. We have developed size evaluators for four well-known BitTorrent web sites: Mininova (the replacement of the previously largest web site, SuprNova [14], which is now defunct), Novatina, Pirate Bay, and TorrentPortal. Table 1 shows the number of torrents (files) and users of each of these web sites; as Novatina was reporting exaggerated numbers of users and contained many broken links, we have removed it from further consideration. At the time of our measurements, the dominant BitTorrent web site was Pirates Bay. Therefore, for measuring

[3] Python official web site. `http://www.python.org`.
[4] Ruby official web site. `http://www.ruby-lang.org`

Table 2: A summary of the characteristics of the tracked files.

Measurement type	Active-start			Passive-start		
	Avg	Min	Max	Avg	Min	Max
BitTorrent swarm size	132	46	4,187	307	129	4,187
Torrent size	2.6GB	0.1MB	41GB	3.0GB	24MB	41GB
Number of torrent chunks	2,463	1	34,591	2,623	92	34,410
Torrent chunk size	1.1MB	32KB	4MB	1.2MB	128KB	4MB
Number of files in torrent	47	1	5,265	53	1	1,565

topology and path characteristics of overlay networks and the Internet, we set up MultiProbe to measure the BitTorrent/Pirates Bay environment.

4.2 Measurements

We have used two platforms for our measurements. For our active-start measurements, we have deployed MultiProbe on 100 of the 400 processors available on our Distributed ASCI Supercomputer[5]. For our passive-start measurements, we used 50 PlanetLab [11] nodes as *peer points*, and 300 PlanetLab nodes for traceroutes towards the peers that contacted the peer points. `Scriptroute` [18] was of inestimable help for distributed software deployment and operation.

We successfully tracked the top 2,000 files in the active-start measurements. Due to machine failures on PlanetLab, we successfully tracked only 695 of the originally intended top 750 files in the passive-start measurements (files ranked 155, 331-340, 405-410, and 716-750 were not tracked). The two sets of tracked files overlap. Table 2 shows the statistics of the tracked files. We define a *shared file* as a file shared using BitTorrent. The set of tracked shared files presents a wide range of values for shared file size, number of chunks, chunk size, and files in the torrent. The average shared file size is around 2-3GB, which shows that BitTorrent is mostly used for large-sized transfers. We also find that most torrents are archives, and the average number of files per archive is around 50.

5 Measurement Results

In this section we present the measurement results; all the results are available on the MultiProbe web site (see Section 7). We focus here on three main aspects. First, we correlate the users' geographical locations with the generated traffic; while the users' geographical location has been previously detailed in several published results, for example in [5], our results are the first published statistics of a correlation of these locations and the generated traffic, for a large number of files. The second category of results refers to route analysis and application-level connectivity; we report path characteristics of the overlay network flows, including path hops and AS mapping, and analyze application-level connectivity. The third and final category are related to traffic analysis. The break-down of traffic per TCP port shows that, contrary to reported traffic analysis results from 2004 [13], there is a wide amount of hidden traffic on non-standard ports. These results confirm the initial findings of [6], and corroborate them with quantitative data. We complete our traffic analysis with a detailed break-down of the application-level bandwidth.

5.1 Measurements summary

We record two types of events in active-start measurements: `ContactOK` events, observed when the probe successfully contacted a target peer, and `ContactErr` events, observed when a probe did not manage to contact a target peer. For `ContactOK` events, we record the target peer's IP address, contact port number, and number of owned chunks. For `ContactErr` events, we record the target peer's IP address, contact port number, and reason of the connection failure (e.g., *connection timed out*).

For passive-start measurement, there are two main types of events: `Contacted` events, observed when the measurement peer is contacted by another peer, and multi-source traceroute events, observed when the

[5] `http://www.cs.vu.nl/das2/`

Table 3: Statistics of the size of our measurements.

Measurement type	Active-start	Passive-start	Measurement type
Measure	Value	Value	Measure
Period	05-11 May, 2005		Period
No. of files	2000	695	No. of files
No. of unique users	229,410	226,441	No. of unique users
Observed traffic	193.78 TB	909,961	No. of successful traceroutes
No. of `ContactOK` events	28,781,347	19,893,684	No. of visited IP addresses
No. of `ContactErr` events	7,240,601	3,729,570	No. of `Contacted` events
Data size, compressed	3GB	8GB	Data size, compressed

Table 4: The percentages of users per continent and country.

Continent	EU								NA		AS		SA		OC	AF
Number	59.4								22.3		8.0		6.9		2.7	0.5
Weight	58.8								22.1		7.4		4.5		2.6	0.3
Country	SE	UK	DE	ES	FR	FI	NL	PL	US	CA	TW	CN	BR	AR	AU	ZA
Number	8.6	7.7	5.4	7.8	5.8	2.5	4.2	4.1	14.4	6.5	1.9	1.0	2.6	2.0	2.3	0.1
Weight	13.7	8.1	4.6	4.5	4.4	4.0	3.9	3.0	14.7	7.0	1.9	0.8	1.8	1.2	2.4	0.1
World rank	2	3	5	6	7	8	9	10	1	4	14	21	15	19	12	50+

reverse path tree towards a peer was completed. We store detailed data relevant for these two types of events in sets of compressed files, one set per traced file.

Table 3 shows statistics of the size of our measurements. The two types of measurements yielded approximately the same number of identified unique users, over 225,000. Overall, we have gathered over 40,000,000 events, and we observed over 193 TB of data traffic, not including control messages or other overhead.

5.2 Location analysis

In order to understand the links between overlay networks and the Internet, it is mandatory to geo-locate the users and to compare different locations according to the number of users and their traffic. We define a user's *weight* as the amount of data transferred by that user.

We used the libraries and databases provided MaxMind's GeoIP[6] and WebLogExpert's[7] to geo-locate the BitTorrent users. From the original data, 3-4% of the IP addresses could not be mapped to continents and/or countries, and as many as 21% of the IP addresses could not be mapped to a city. We define an *Internet organization* as a corporate network or an ISP for home users. Less than 1.5% of the IP addresses could not be mapped to an Internet organization.

5.2.1 Continent/country location

Table 4 shows the number of users and their weights for different continents and countries. The *World rank* row shows the relative country rankings. Europe is the dominant continent in both the number of users and the users' weights. We expected Sweden to be placed in the top 10 countries, but its surprising ranking as 2nd in the world can only be explained as a bias of our data source (Pirates Bay is located in Sweden).

5.2.2 City location

Major cities like Madrid, Paris, Amsterdam, and Toronto are strongly represented in the top 50 cities both by number of users and by their weight. However, the fact that over 20% of the users and over 25% of the generated traffic could not be mapped to a city prevents us from generalizing our conclusions.

[6] MaxMind, http://www.maxmind.com
[7] WebLog Expert, http://www.weblogexpert.com

Another aspect prevents an accurate users-to-cities mapping: the existence of *buffer cities*–small cities that host important network junctions–. Examples of buffer cities are Oldenburg and Eschborn in Germany), and Herndon and Mount Laurel in USA. Over 10% of the IP addresses were mapped to buffer cities.

5.2.3 Internet Organization location

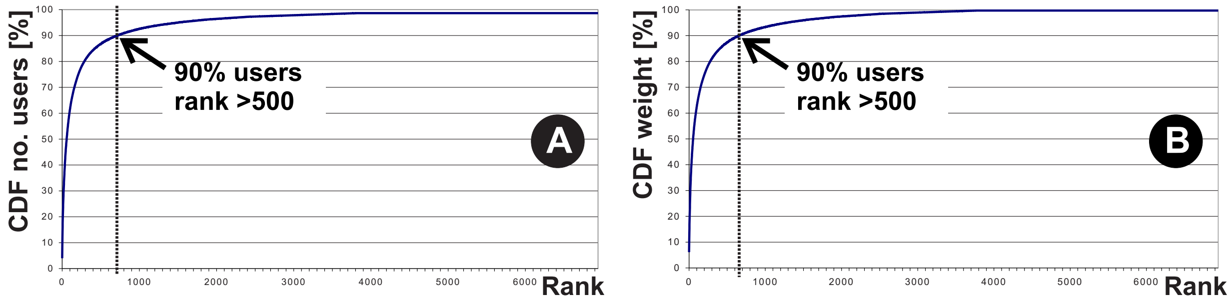

Figure 3: The Internet Organizations location: (a) CDF of number of users; (b) CDF of users' weights. The horizontal axis shows the rank of Internet Organizations, with respect to their percentage.

More than 6500 distinct Internet Organizations carry BitTorrent traffic. Figure 4 shows the CDF of the number of users and users' weights per Internet Organization. As expected, there is no dominant Internet Organization; notably, the top 500 Internet organizations cover less than 90% in both categories. This shows that BitTorrrent has grown to a large-scale, completely distributed state, in which both users and their traffic are scattered around the world.

5.2.4 Autonomous System location

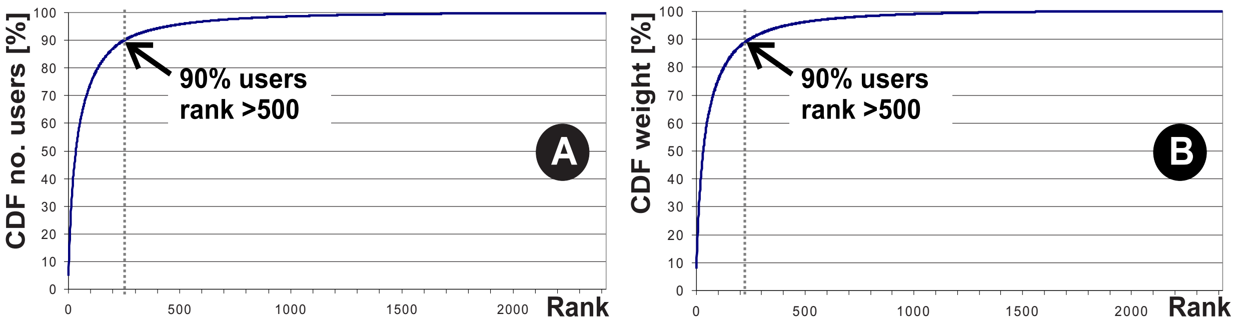

Figure 4: The Autonomous Systems location: (a) CDF of number of users; (b) CDF of users' weights. The horizontal axis shows the rank of Autonomous Systems, with respect to their percentage.

Similarly to the Intern Organizations, more than 2400 different Autonomous Systems (AS) were identified as part of the BitTorrent's underlay infrastructure; the top 250 ASs cover roughly 90% of the users.

5.3 Route and connectivity analysis

The analysis of the Internet routes between the peers of an overlay network could show whether these peers are well-connected or not, and whether the ISPs have optimized their internal traffic and their peering relations with respect to P2P traffic performance. In this section we show the distribution of path IP hop counts, the distribution of AS traversals, the distribution of intra-AS path hop counts, and the distribution of round-trip times. Statistical results are given for more than 900,000 paths explored from PlanetLab peers to overlay network peers that have contacted them.

Table 5: The distribution of IP path hops for paths between PlanetLab peers and overlay network peers which have contacted them.

Connection	Direct	Strong	Good	Average	Loose	Very Loose
No. IP path hops	0-1	2-4	5-9	10-14	15-19	20+
No. paths [%]	5.9	20.4	40.2	25.3	6.4	1.1

Table 6: The distribution of AS traversals for paths between PlanetLab peers and overlay network peers which have contacted them.

Connection	Direct	Strong	Good	Average	Loose	Very Loose
No. AS traversals	0-1	2	3	4-5	6-9	10+
No. paths [%]	12.5	44.7	32.0	10.2	0.5	0.1

Table 7: The distribution of intra-AS path hop counts for paths between PlanetLab peers and overlay network peers which have contacted them.

Connection	Direct	Strong	Good	Average	Loose	Very Loose
No. intra-AS hops	1	2	3	4-5	6-9	10+
No. paths [%]	58.1	12.9	8.0	12.2	7.6	1.2

5.3.1 IP path hop count distribution

Path IP hop count is a natural connectivity metric in the Internet [4]. The higher the number of hops, the higher the latency, and the lower the performance of the P2P application. This measure especially affects P2P networks which exchange large numbers of small messages; even if BitTorrent is essentially bandwidth dominated, high latencies can also influence it. Table 5 shows the distribution of IP path. We define six categories of IP path hop count ranges, with values given in row *No. IP path hops*. More than 66% of the users have a good connection to other peers or better, similar to the best IP hop count results reported in [4].

5.3.2 AS traversals distribution

In this section we consider the number of AS traversals, as high latencies can be incurred at the interconnections between different ASs, depending on the peering relations between the AS owners. Table 6 shows the distribution of AS traversals. We define the same six categories of count ranges as in the case of IP path hop counts (see Section 5.3.1), with values given in the *No. AS traversals* row. Again, more than 66% of the users have good connections to other peers or better.

During this analysis we also discovered a number of *traversal loops* – AS traversals in which the same traversal occurs at least twice. This happens when boundary routers from two ASs are misconfigured. At least 0.01% of the analyzed traceroutes have this problem, leading to as many as 56 AS traversals in one IP packet path.

5.3.3 Intra-AS path hop counts

The number of intra-AS path hops gives hints on how well do AS owners optimize the TCP/IP traffic within their AS. This number is essentially similar to the IP path hop count, but can also uncover the AS location of possibly misconfigured routers. Table 7 shows the distribution of intra-AS path hop counts for paths between PlanetLab peers and overlay network peers which have contacted them. We define the same six categories of count ranges as in the case of IP path hop counts (see Section 5.3.1), with values given in the *No. intra-AS hops* row. We find that over 50% of the paths have 1 hop, and that more than 70% of the paths have at most 2 hops, which means that users experience a good connection to other users within the same AS.

During this analysis we also discovered a number of *routing loops* – routes in which the same IP address is visited at least twice. This happens when intra-AS routers are misconfigured. At least 0.001% of the analyzed

Table 8: The distribution of the measured Round-Trip Times, per classes.

	Local	LAN	Campus	Metropolitan				Inter-city		Intercontinental
RTT Min [ms]		15	50	100	200	300	400	500	750	1000
RTT Max [ms]	15	50	100	200	300	400	500	750	1000	
Paths [%]	0.6	2.9	5.8	13.3	13.7	13.8	8.6	12.8	7.5	20.7

traceroutes have this problem, which gives evidence that the misconfiguration problem occurs intra-AS with a probability an order of magnitude lower than in the inter-AS case.

5.3.4 Round trip time

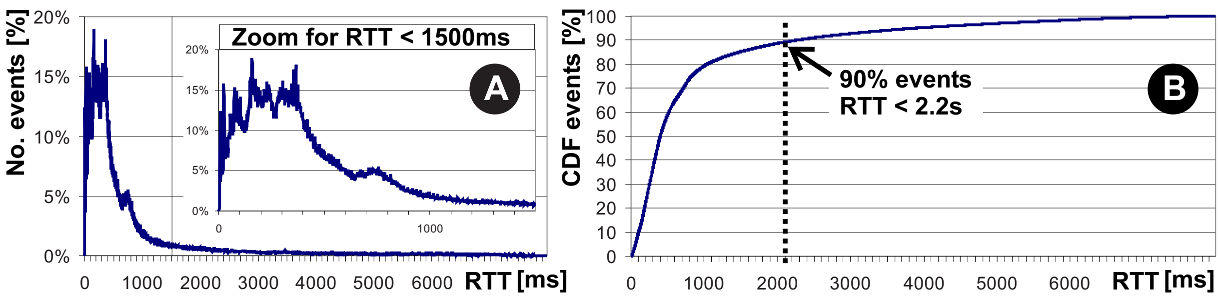

Figure 5: The distribution of measured RTT: (a) detailed distribution and zoom to RTTs below 1.5 seconds; (b) CDF of RTT. The vertical axis shows the respective percentages of numbers of RTTs.

There is much interest in studying the round-trip time (RTT) characteristics of P2P networks. We measured the RTTs using traceroutes between our nodes and contacting peers. We report here only the results for successful traceroutes. Since measurements were conducted from well-connected PlanetLab nodes, there is a *positive bias* in the reported results: the average RTTs are smaller than normal. However, the fact that we used globally spread sources reduces this effect. Figure 5 shows the detailed distribution of RTT. Less than 10% of the measured RTTs are outside the 2.2 seconds margin. We define six classes of RTTs: *local* (0-15ms), *LAN* (15-50), *campus* (50-100), *metropolitan* (100-500), *inter-city* (500-1000), and *intercontinental* (1000+). Table 8 shows the detailed break-down of RTTs, according to their class. *Metropolitan* and *inter-city* connections are dominant, with more than 75% of the measured RTTs falling in these categories. We conclude that most BitTorrent peers are close to each other, if the distance metric is RTT.

5.3.5 Application-level connectivity

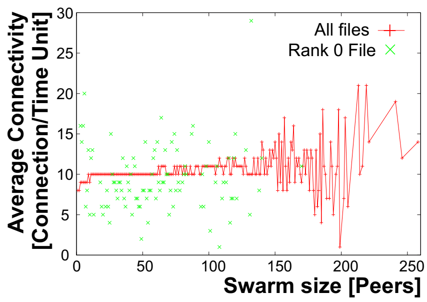

Figure 6: Number of incoming connections for a 5 minutes time unit

We define the *application-level connectivity* of a P2P network peer as the number of incoming and

Table 9: The distribution of users and user weights per TCP port.

Start port		1000	6881	7000	10000	20000	30000	40000	50000	60000
End port	999	6880	6999	9999	19999	29999	39999	49999	59999	
No. users [%]	0.5	3.4	42.1	5.1	19.1	9.7	5.4	7.0	6.8	0.9
Weight [%]	0.6	4.2	11.2	9.7	31.3	16.6	7.0	9.0	9.1	1.3

outgoing connections between that peer and other peers in the network. The maximum number of outgoing connections is a constant in most BitTorrent clients. We analyzed the number of incoming connections per 5 minutes, for all observed swarm sizes. Figure 6 shows evidence that the average number of incoming connections is around 10, regardless of the swarm size. We attribute the deviation from the average for a swarm size of 150 and more to the low number of observations in the respective categories.

5.4 Traffic break-down

The analysis of current BitTorrent traffic characteristics like the distribution of traffic per TCP port and the application-level bandwidth can provide useful insight in how to monitor, provision for, and shape the future BitTorrent traffic. In this section we present detailed statistics for the aforementioned traffic characteristics.

5.4.1 Traffic per TCP port distribution

In order to monitor and control effectively P2P traffic, it is important to have a realistic traffic break-down per TCP port. Table 9 shows the distribution of users and user weights per TCP port. These results show that, contrary to reported distribution of users results from 2004 [13], there is a wide amount of hidden traffic on non-standard ports. The high percentage of observed users of the standard BitTorrent ports (6881 through 6999) is misleading; the standard ports traffic accounts for a mere 11.2% of the total traffic. A number of non-standard ports (data not shown here) are used by high traffic users: ports 16881, 49152, 10000 and 10001 account alone for almost 2% of the total traffic. Given the overall distribution of traffic per TCP ports, we can conclude that BitTorrrent clients have shifted in the last year to random port selection.

5.4.2 Application-level bandwidth

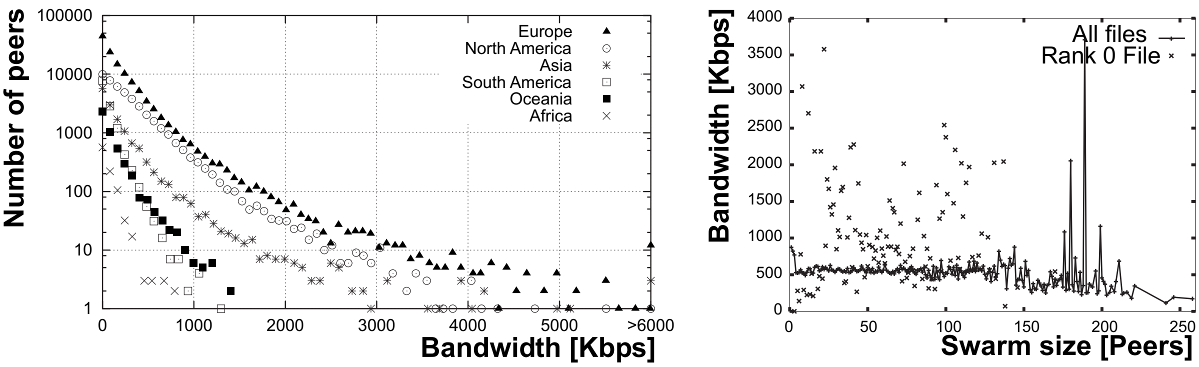

Figure 7: Application-level bandwidth: (a) number of users per bandwidth value; (b) average bandwidth per swarm size

We define the *application-level bandwidth* as the total amount of data downloaded by a peer over a fixed time interval. Based on the BitTorrent's bartering partners selection protocol (see Section 2.1), we argue that this measure is equivalent with the bottleneck bandwidth. Figure 7 (a) shows the distribution of number of users per bandwidth, for different continents. We find two groups with similar bandwidth characteristics: (1) Europe, North America, and Asia, and (2) South America, Oceania, and Africa.

We define a *swarm* as the set of users downloading the same file, and the *swarm size* as the number of users from a swarm, during a fixed time interval. Figure 7 (b) presents the influence of the swarm size on users' bandwidth. We find that the average bandwidth of swarm members does not depend on the swarm size. The aggregate average bandwidth fluctuated around 500 Kbps, for a time interval set to 5 minutes.

6 Conclusions and future work

This paper presents a correlated view of overlay networks and the Internet. For that purpose, we have designed, implemented, and deployed MultiProbe, a large-scale P2P measurement framework. Large-scale joint measurements of BitTorrent and the Internet were conducted in May 2005, and correlated into comprehensive statistical data, in four categories: location, route, connectivity, and traffic.

The main new results are: (1) the majority of BitTorrent users are located in Europe but, at the same time, BitTorrent traffic is globally spread; (2) BitTorrent peers are well connected by the Internet in terms of latency, hop count, and bandwidth; (3) BitTorrent peer connectivity does not depend on the swarm size; (4) BitTorrrent has shifted in the last year from static to random TCP port selection.

For the future, we plan to enable our framework to measure several other large-scale P2P networks, and to repeat our experiments in the new context.

7 Availability

For brevity reasons, a number of graphs, tables and other results were not included in this paper. All this material as well as the data presented in this study can be found at the MultiProbe web site: `http://multiprobe.ewi.tudelft.nl/`.

Acknowledgments

Many thanks go to Neil Spring. We cannot stress enough how his excellent work on Scriptroute, and his ready help, allowed us to finish this project. Other people have contributed as well, and should not be forgotten. We thank John van Baarle, Ivaylo Haratcherev, Jacco Taal, and Ana Lucia Varbanescu.

References

[1] K. C. Claffy, T. E. Monk, and D. McRobb. Internet Tomography. Nature Web Matters, January 1999. (also CAIDA TR at http://www.caida.org/outreach/papers/1999/webmatters99/).

[2] R. Govindan and H. Tangmunarunkit. Heuristics for Internet Map Discovery. In *IEEE INFOCOM 2000*, pages 1371–1380, Tel Aviv, Israel, March 2000. IEEE.

[3] K. Gummadi, R. Dunn, S. Saroiu, S. Gribble, H. Levy, and J. Zahorjan. Measurement, modeling, and analysis of a peer-to-peer file-sharing workload, 2003.

[4] B. Huffaker, M. Fomenkov, D. Moore, and K. Claffy. Macroscopic analyses of the infrastructure: Measurement and visualization of Internet connectivity and performance. In *Proceesings of the 2nd International Workshop on Passive and Active Network Measurement (PAM'01)*, Amsterdam, The Netherlands, April 2001.

[5] M. Izal, G. Urvoy-Keller, E. W. Biersack, P. Felber, A. A. Hamra, and L. Garcés-Erice. Dissecting BitTorrent: Five Months in a Torrent's Lifetime. In *Proceesings of the 5th International Workshop on Passive and Active Network Measurement (PAM'04)*, pages 1–11, April 2004.

[6] T. Karagiannis, A. Broido, N. Brownlee, k. claffy, and M. Faloutsos. Is P2P dying or just hiding? In *Global Internet and Next Generation Networks (Globecom 2004)*, Dallas, Texas, USA, Dec 2004.

[7] P. Krishnan, D. Raz, and Y. Shavitt. The Cache Location Problem. *IEEE/ACM Transactions of Networking*, 8(5):568–582, October 2000.

[8] K. Lakshminarayanan and V. N. Padmanabhan. Some Findings on the Network Performance of Broadband Hosts. In *ACM/USENIX Internet Measurement Conference*, Miami, Florida, USA, October 2003.

[9] A. Parker. The True Picture of Peer-To-Peer File-Sharing, Sep. 2005. Panel Presentation, IEEE 10th International Workshop on Web Content Caching and Distribution, Sophia Antipolis, France.

[10] V. Paxson. Strategies for Sound Internet Measurement. In *ACM/USENIX Internet Measurement Conference*, pages 263–271, Taormina, Italy, October 2004.

[11] L. Peterson, T. Anderson, D. Culler, and T. Roscoe. A Blueprint for Introducing Disruptive Technology into the Internet. In *Proceedings of HotNets–I*, Princeton, New Jersey, October 2002.

[12] G. Phillips, S. Shenker, and H. Tangmunarunkit. Scaling of multicast trees: comments on the Chuang-Sirbu scaling law. In *Proceedings of the SIGCOMM*, pages 41–51, New York, NY, USA, 1999. ACM Press.

[13] L. Plissonneau, J.-L. Costeux, and P. Brown. Analysis of Peer-to-Peer Traffic on ADSL. In *Proceesings of the 6th International Workshop on Passive and Active Network Measurement (PAM'05)*, pages 69–82, March 2005.

[14] J. Pouwelse, P. Garbacki, D. Epema, and H. Sips. The BitTorrent P2P File-sharing System: Measurements and Analysis. In *Proceedings of the 4th International Workshop on Peer-To-Peer Systems (IPTPS'05)*, Ithaca, New York, USA, February 2005.

[15] S. Saroiu, P. Gummadi, and S. Gribble. A Measurement Study of Peer-to-Peer File Sharing Systems. In *Proceedings of Multimedia Computing and Networking*, 2002.

[16] S. Sen and J. Wang. Analyzing peer-to-peer traffic across large networks. *IEEE/ACM Transactions on Networking*, 12(2):219–232, 2004.

[17] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson. Measuring ISP topologies with Rocketfuel. *IEEE/ACM Transactions of Networking*, 12(1):2–16, 2004.

[18] N. Spring, D. Wetherall, and T. Anderson. Scriptroute: A public Internet measurement facility. In *the 4th USENIX Symposium on Internet Technologies and Systems (USITS '03)*, March 2003.

[19] B. Zhao, A. Joseph, and J. Kubiatowicz. Locality aware mechanisms for large-scale networks. In *Proceedings of Workshop on Future Directions in Distributed Computing (FuDiCo)*, pages 229–238, June 2002.