Correlating Topology and Path Characteristics of Overlay Networks and the Internet

Alexandru Iosup, Paweł Garbacki, Johan Pouwelse, and Dick Epema
Delft University of Technology, Faculty of Electrical Engineering, Mathematics, and Computer Science
P.O. Box 5031, 2600 GA Delft, The Netherlands
Email: p2p@ewi.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.**

I. INTRODUCTION

The topology and the characteristics of the Internet have a large impact on the operation of Internet applications such as multicast overlays [1], web hosting [2], and Peer-to-Peer (P2P) file-sharing systems [3]. 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 the 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 [4], [5]. Over the past couple of years, BitTorrent has become the largest P2P file-sharing network in terms of generated traffic [5]. 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 [6], where we have investigated the high-level characteristics and the user behavior of BitTorrent, without correlating the underlaying Internet. This work complements our previous investigations, adding the network level insights. For this, we focus on the following research topics:

Measuring underlay/overlay networks
How to build a large-scale distributed infrastructure that can actively measure at P2P and Internet characteristics the same time? How to select a representative part of the P2P network for measuring?

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 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. Furthermore, we release to the community the acquired data sets, enabling researchers to validate their models against real data and/or to improve their models with realistic parameters.

From the methodological point of view, we find that while standard ports are still regularly used, the highest traffic volume is generated on non-standard ports; this should be taken into account in future BitTorrent-related studies. From the modeling point of view, our main results are that (i) BitTorrent users are not distributed evenly across the Internet (we find that the majority of BitTorrent users are located in Europe), (ii) the average download bandwidth has reached 500 Kbps (double the value observed in our previous investigation [6]) and should be modeled differently based on the continent location of peers, and (iii) only a relatively small number of Autonomous Systems and Internet Organizations are used by the majority of peers. Finally, we correlate the information to provide strong evidence that BitTorrent users are well connected by the Internet in terms of bandwidth and hop count.

II. BACKGROUND

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

A. 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 [5], thus generating more

[1] Slyck.com lists size information for many large P2P networks.

than 15% of the total Internet traffic[2]. Furthermore, BitTorrent can efficiently handle *flashcrowd* effects [6] (offering popular data to a large number of users, immediately after the data release), and offers a pollution-free sharing environment to its users [6].

There are three levels in the BitTorrent world: the peer level, the tracker level, and the web-site level. Data, be it single files or file archives, exist in BitTorrent in the form of *torrents*. To facilitate the exchange process, torrents are split in smaller parts, called *chunks*. A user (*peer*) has a complete file only after obtaining all the chunks composing that file. BitTorrent peers are using an advanced *tit-for-tat* (bartering) mechanism 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 *bartering-partners selection protocol* aims at filling up all the available bandwidth of a peer; for popular files, the bandwidth filling property is guaranteed [7]. For this reason, and because of the lack of search features at the peer level, BitTorrent is bandwidth-limited. As opposed to other file-sharing protocols like DirectConnect and FastTrack, network latency is not an important issue for BitTorrent. By using the services of a *tracker*, BitTorrent peers can discover other peers that have chunks of the desired file. Since BitTorrent has no searching mechanism, metadata descriptions of the torrents are created in the form of `.torrent` files, which are accessible through *web sites*.

In time, BitTorrent has attracted a very specific, yet very large, audience. As a result of the excellent *flashcrowd* support, of the optimal pollution-removal system, of the existence of centralized services (the trackers), of a lack of search capabilities integrated in the peer functionality, BitTorrent has become a network for sharing fresh information. From the torrent files that can be found through the BitTorrent web sites, only the newest ones can be reliably and efficiently obtained from the network; the other files may be incomplete or even non-existent, or the number of peers that offer them can be too small to fill the desired download bandwidth. For non-fresh files, users typically resort to other networks, e.g., KaZaA or eDonkey, where locating files is easier, and old information can be more easily made available, at the expense of increased pollution [8].

B. Related work

The measurement-based analysis of Internet or large-scale P2P networks has been the subject of numerous investigations for over two decades.

In the area of Internet research, the CAIDA[3] project is actively involved in periodic Internet mapping and characterization, and has arguably participated in the respective research community's major achievements during the past 10 years. We mention two research results outside CAIDA: RouteViews and Rocketfuel. The RouteViews[4] project collects BGP routing perspectives from more than 60 major ISPs worldwide. The combined table typically has nearly 120K globally routable prefixes. The Rocketfuel Internet mapping tool [9] uses traceroutes sourced at over 750 vantage points to explore the network topology of 10 ISPs in different continents. These efforts are not targeting any specific application; our BitTorrent study can offer additional insight into the actual use of the Internet, for instance by uncovering the peers' download bandwidth (a comprehensive survey on the difficulties of accurately measuring the bandwidth available to and consumed by Internet applications is given in [10]), or by showing strong evidence that the P2P traffic of BitTorrent has shifted to non-default ports (answering the 2004 CAIDA question *Is P2P dying or just hiding?* [4] for the BitTorrent network).

Much effort has been put in characterizing protocols that optimize searching in P2P file-sharing networks, like Gnutella, FastTrack, or DirectConnect. Our study complements all these approaches by targeting a file download P2P protocol, BitTorrent. From the measurements targeting BitTorrent, our study is the first to characterize the peers sharing a representative set of files. In [11], the authors use active measurements to determine the bandwidth and latency of the peers of the largest P2P systems at that time, Napster and Gnutella. Their results only refer to P2P file-searching protocols. The passive-measurements of three popular P2P protocols, FastTrack, Gnutella, and DirectConnect, are analyzed in [12]. Again, results target only P2P file-searching protocols. In addition, their work suffers from the methodological aspect of monitoring only the default ports. In [13], a 5-month passive-measurements trace of a single file shared using the BitTorrent protocol is presented. The file comes from the operating system domain, and so is not representative for P2P file-sharing, where users download mostly movies and music.

A large-scale study of the eDonkey network is presented in [14]. The authors' focus is on clustering the peers' preferences; the peers' Internet connectivity is not analyzed, and from the peers' locations only the countries are detailed. Their analysis also displays a clear bias: the US-based users represent only 7% of the total world users, and users from Finland, Sweden, and China — countries strongly represented in terms of number of P2P users[13], [15] — represent each less than 1% of the total world users. In [16], 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 [17]. For example, they show evidence that the eDonkey network occupies more than 85% of the P2P traffic, which is impossible at world level, in the light of other accepted results [5], [11].

An attempt to correlate Internet and P2P network measurements is made in [18], where network connectivity and file-sharing between only 25 broadband users is studied. Therefore, that study lacks breadth, and should be complemented with a broader approach, such as this.

III. THE MULTIPROBE FRAMEWORK

In this section we present the structure of MULTIPROBE, our framework for correlated measurements of large scale overlay

[2]Sprint. Packet Trace Analysis. http://ipmon.sprintlabs.com/

[3]The Cooperative Association for Internet Data Analysis (CAIDA), http://www.caida.org/.

[4]D. Meyer, University of Oregon Route Views project, http://www.routeviews.org/.

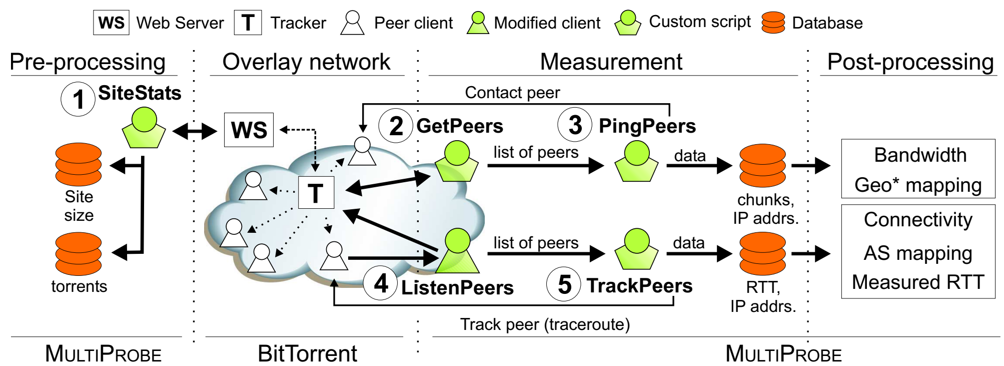

Fig. 1. The MULTIPROBE framework for correlated measurements of large scale overlay networks and the Internet.

networks and the Internet. MULTIPROBE consists of pre-processing, measurement, and post-processing parts, which we discuss below.

The BitTorrent network uses web sites as content directories. The largest of these sites handle tens of thousands of shared files and serve millions of users. MULTIPROBE can select the largest such web site, with respect to the number of users, at the start of the experiments (the `SiteStats` module, see Figure 1). Besides their directory service, the largest web sites also provide per-file information about the number of users uploading and downloading file chunks (*swarm size*). After sorting the shared files according to their swarm size, the measurements setup module, selects the *top-N* files from the largest web site, and their respective torrents are downloaded. The *number of files* measurement parameter (N) reflects the physical resources in the measurement infrastructure: the more measurement machines, the larger the possible value of N. In practice, the maximum value of N also depends on the number of files with swarms of sizes of a few tens of peers, since only these files can actually be reliably and efficiently obtained from the network [7].

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 1). They subsequently contact the corresponding peers using the BitTorrent protocol, and learn about what chunks are owned by these peers (the `PingPeers` module). In passive-start measurements, probes register themselves to a tracker, and wait to be contacted by other peers (the `ListenPeers` module). 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); we call this process *multi-source traceroute*.

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.

TABLE I
THE NUMBER OF TORRENTS (FILES) AND USERS PER BT 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

IV. EXPERIMENTAL 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 [6], which is now defunct), Novatina, Pirate Bay, and TorrentPortal. At the time of the measurements, these sites were the top four BitTorrent web sites in the `slyck.com`'s rankings[5]. Table I 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 topology and path characteristics of overlay networks and the Internet, we set up MULTIPROBE to measure the BitTorrent/Pirates Bay environment.

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[6], located in The Netherlands. For our passive-start measurements, we used 50 PlanetLab [19] nodes as *peer points*, and 300 PlanetLab nodes for traceroutes towards the peers that contacted the peer points. The 300 PlanetLab nodes were scattered around the world, but were mostly part of the GREN network [20]. `Scriptroute` [21] was of inestimable help for distributed software deployment and operation.

[5] `Slyck.com` ranks the BitTorrent web sites, `http://www.slyck.com/bt.php?page=3`
[6] `http://www.cs.vu.nl/das2/`

TABLE II
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

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 II 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.

In this paragraph we quantify the bias introduced by our experimental infrastructure. We show that, unlike the previous approaches discussed in Section II-B, our bias in negligible. *First*, the selection of the BitTorrent web site could induce a strong bias to the measurements, for instance if the selected web site is small-sized or serves only a special type of community, e.g., operating-systems developers. By selecting the largest BitTorrent web site, which covers a broad range of file types and sizes, we have ensured that we capture the characteristics of a world-wide community of users. However, as Pirates Bay is located in Sweden, we could not preclude the bias of the analysis of users' locations towards Sweden. In Section V-B we quantify this bias, and we show evidence that it is negligible. *Second*, the selection of nodes located within the PlanetLab network could lead to biased results, as the PlanetLab nodes are much better connected than the average users of BitTorrent (see [20]). Because of this, there is a clear bias in round-trip time measurements; as BitTorrent is bandwidth-limited network, this bias is of minor importance. Furthermore, because of the BitTorrent's bartering-partners selection protocol, connections tend to be formed with better-connected nodes. Indeed, a bias is introduced by the use of PlanetLab nodes; we argue that this bias is in fact required to ensure the representativeness of our results: we perform experiments from the points to which any peer in the BitTorrent network would naturally try to connect. This is key to a representative analysis of path structure, Internet Organizations distribution, and Autonomous Systems distribution (see Section V-B). Also, the location of the DAS and the PlanetLab nodes does not influence the download bandwidth measurements (see also Section V-C), nor the application-level connectivity measurements (see also Section V-C). *Third*, the selection of the measurement parameter *number of files* (*N*) could influence our findings, particularly in the case of observed bandwidth. Since we have selected only files with swarms with size over 40 (see Table II), the reliability and efficiency of the data download are increased, and the average bandwidth is high. We argue that this is the real use of BitTorrent, and this is the context in which it should be evaluated.

V. MEASUREMENT RESULTS

In this section we present the measurement results; all the results are available on the MULTIPROBE web site (see Section VIII). For space reasons, we focus here on three main aspects. First, we correlate the users' geographical locations with the generated traffic. 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 [16], there is a wide amount of hidden traffic on non-standard ports. These results confirm the initial findings of [4], and corroborate them with quantitative data. We complete our traffic analysis with a detailed break-down of the application-level bandwidth.

A. 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 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 III 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.

TABLE III

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

B. 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 publicly-available libraries and databases provided MaxMind's GeoIP[7] and WebLogExpert's[8] to geo-locate the BitTorrent users. In our previous work [15] we have established the superior performance in both speed and IP address mapping accuracy of the combined use of these tools, when compared with two other similar approaches, CAIDA's NetGeo[9] and IP:Country; the reason is the more up-to-date information present in the GeoIP databases. From the original data, 3-4% of the IP addresses could not be mapped to continents and/or countries. As many as 21% of the IP addresses could not be mapped to a city, a known problem of the GeoIP (and NetGeo) database building approach [22]. 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.

Table IV 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. This result is not the result of the measurements bias; we have shown evidence of a similar phenomenon, but for a very limited number of files, in our previous work [15]. 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). In order to evaluate the bias, we resort to our previous results [15]: there, Sweden was occupying the 8th place in the world, with the percentage of users and weights per country of 2.6 and 3.3, respectively. Assuming these results as correct, the bias induced by the selection of Pirates Bay as the measurements target is +6.0% and +10.4%. Eliminating this bias, Europe remains (by far) the dominant continent, and Sweden remains in the world Top-10. We conclude that the bias was negligible.

Major cities like Madrid, Amsterdam, and Toronto are strongly represented in the top 50 cities both by number of users and by their weight. However, we found that *buffer cities*–small cities that host important network junctions, such as Oldenburg (DE), and Herndon (USA)–are also highly ranked, and cover 10% of the total users.

More than 6500 distinct Internet Organizations (*IO*) carry BitTorrent traffic. Figure 2 shows the CDF of the number of users and users' weights per IO. As expected, there is no dominant IO; 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.

Similarly to the IOs, 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. For space considerations, data regarding AS distribution is only available on the web site (see Section VIII).

C. Route and connectivity analysis

Path IP hop count is a natural connectivity metric in the Internet [23]. Here, we also consider the number of AS traversals, as high latencies can be incurred at the interconnections between different ASs. Table V shows the distribution of IP path hop count, AS traversals, and intra-AS IP path hop count for more than 900,000 paths explored from PlanetLab peers to overlay network peers that have contacted them. The row *Connection* defines six categories of count ranges, with values given for IP path hop count, AS traversals, and intra-AS IP path hop count respectively in rows *No. IP path hops*, *No. AS traversals*, and *No. intra-AS IP path hops*. More than 66% of the users have a good connection to other peers or better, similar to the best hop count results reported in [23]. We also observe that a surprisingly high number of ASs have long routing paths inside: over 1% of the intra-AS paths have over 10 IP hops, due to faulty router configurations.

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 3 shows the detailed distribution of RTT. More than 75% of the measured RTTs fall in the 100ms-1s category. Less than 10% of the measured RTTs are outside the 2.2 seconds margin. Since BitTorrent is not latency-limited, we conclude that most BitTorrent peers are close to each other, if the distance metric is RTT.

[7] MaxMind, http://www.maxmind.com

[8] WebLog Expert, http://www.weblogexpert.com

[9] `http://www.caida.org/tools/utilities/netgeo/`

TABLE IV
THE PERCENTAGES OF USERS PER CONTINENT AND COUNTRY.

Continent, 2-letter	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 (ISO3166)	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+

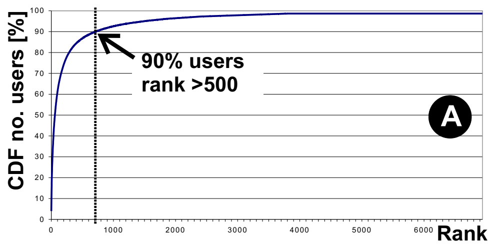

Fig. 2. The users' Internet Organizations location: (a) CDF of number of users; (b) CDF of users' weights. The horizontal axis shows the rank of Internet organizations, with respect to their percentage.

TABLE V
THE DISTRIBUTION OF IP PATH HOP COUNT, AS TRAVERSALS, AND INTRA-AS PATH HOP COUNT 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
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
No. intra-AS IP path hops	1	2	3	4-5	6-9	10+
No. paths [%]	58.1	12.9	8.0	12.2	7.6	1.2

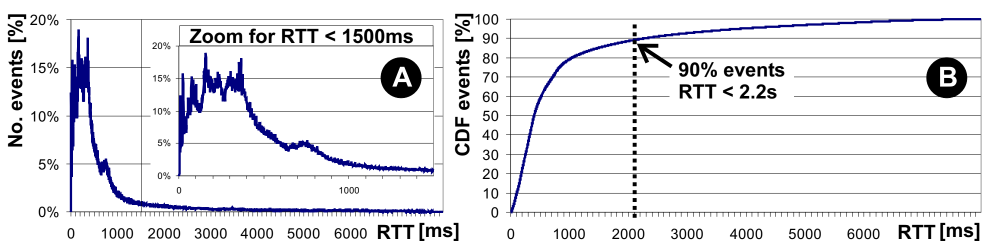

Fig. 3. The distribution of measured RTT: (a) detailed distribution and zoom to RTTs below 1.5 seconds; (b) CDF of RTT. The vertical axis shows the respective percentages of numbers of RTTs.

We define the *application-level connectivity* of a P2P network peer as the number of incoming and 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 and 10 minutes, for all observed swarm sizes. For this, we first used the *peer points* from PlanetLab. According to the BitTorrent protocol, peers first contact a tracker, obtain a list of peers that are sharing the same file, then periodically initiate contact with a randomly selected peer from the list. The *peer points* are waiting for incoming connections and record them. We have then split the recorded data in 5 and 10 minutes intervals. Finally, we have correlated this information with the observed swarm size for the same intervals. We found that, regardless of the swarm size, for time intervals of 5 and 10 minutes the average number of incoming connections is around 10 and 20, respectively.

D. *Traffic break-down*

In order to monitor and control effectively P2P traffic, it is important to have a realistic traffic break-down per TCP port. Table VI shows the distribution of users and user weights per TCP port. These results show that, contrary to reported distribution of users results from 2004 [16], 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

TABLE VI
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

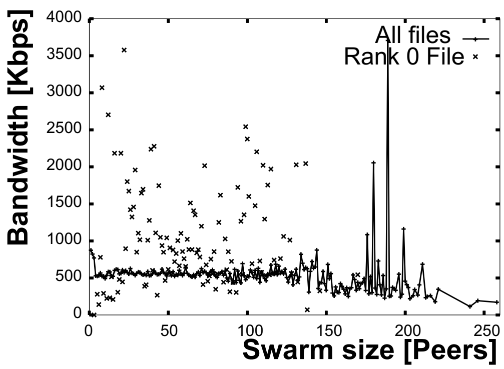

Fig. 4. Average bandwidth per swarm size in BitTorrent.

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.

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 II-A), we argue that this measure is equivalent with the bottleneck bandwidth.

Figure 4 presents the influence of the swarm size on users' bandwidth. The connected dots represent the average bandwidth computed for all files. The dots represent the bandwidth computed for the file with the largest swarm size. We find that the average bandwidth of swarm members does not depend on the swarm size. The bursty behavior observed for larger swarm sizes is due to the smaller number of observations of those swarm sizes. The aggregate average bandwidth fluctuated around 500 Kbps, double the bandwidth reported in [6]. We explain this difference by the bias of the previous measurements: except for 10% of the files, the measurements in [6] targeted files with too small swarm sizes.

Figure 5 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 or uploading the same file, and the *swarm size* as the number of users from a swarm, during a fixed time interval.

VI. DISCUSSION

In this section we discuss possible ways of using the data presented in this paper for modeling realistic P2P workloads, improving the way the Internet serves BitTorrent, and devising good measurement methods.

With the full release of our data (see Section VIII), we hope that the research community dealing with Internet-aware P2P technologies [24], [25] will be able to test and/or improve their models with more realistic parameters. Our measurements address only the peer-level of a bandwidth-limited file-sharing network, but give insight in both the characteristics of the application (file types and sizes, bandwidth), and the way the application interacts with the underlay network (traffic size, paths, ports, even RTTs). The relatively small number of ASs and IOs needed to cover the majority of users enables full simulations even on small experimental setups. Other communities can use the information presented in this work: since BitTorrent has been shown to have better performance than the traditional file-transferring methods — in [26], BitTorrent and a vanilla flavor of it are shown to achieve much better scalability and transfer speed then (Grid) FTP, when used in an Internet-deployed (Desktop) Grid, and for files that exceed 20MB — we argue in favor of using our results for more realistic workload and environment modeling.

While the BitTorrent traffic is spread across the world, we find that only a small number of ASs and IOs are used by the majority of users. In addition, over 25% of the intra-AS paths have 3-9 hops. Thus, making AS-level nodes BitTorrent-aware may lead to a reduction of intra-AS resource contention (routing), to an increase in the BitTorrent download performance (caching), and to advanced capabilities of traffic shaping (filtering).

BitTorrent is a bandwidth-dominated and bandwidth-aware application. Furthermore, BitTorrent connections are made to diverse port values. The implications of these properties on the design and conduction of experiments are twofold: the measurement infrastructure must be set such that it is well-connected to the Internet, and all the TCP flows must be observed, regardless of the source/destination port.

VII. 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.

Our main 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 path characteristics; (3) the average BitTorrent download bandwidth is 500 kbps; (4) the average BitTorrent application-level connectivity does not depend on the swarm size; (5) BitTorrrent has shifted in the last year from static to random TCP port selection.

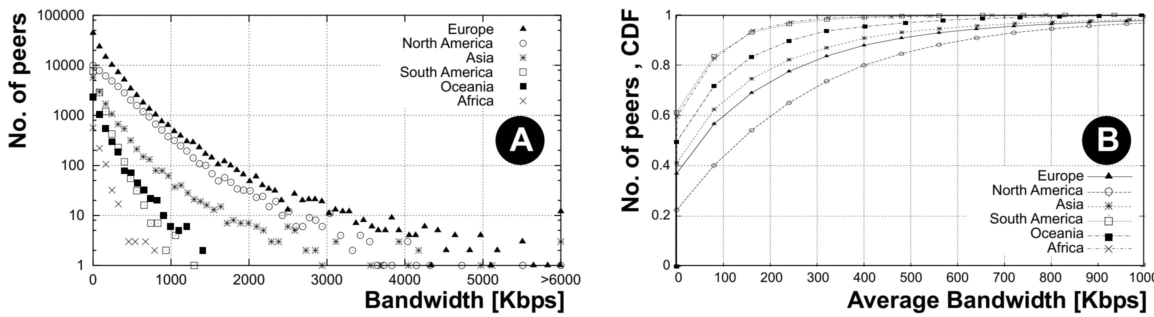

Fig. 5. Application-level bandwidth: (a) number of users per bandwidth, value; (b) number of users per bandwidth, CDF.

For the future, we plan to enable our framework to measure several other large-scale P2P networks, and to use MULTIPROBE for measurements in the new environments.

VIII. AVAILABILITY

For space reasons, a number of graphs, tables and other results could not be presented 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. The first author would also like to help Ana Lucia Varbanescu, who has been of utmost help.

REFERENCES

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

[2] P. Krishnan, D. Raz, and Y. Shavitt, "The Cache Location Problem," *IEEE/ACM Transactions of Networking*, vol. 8, no. 5, pp. 568–582, October 2000.

[3] B. Zhao, A. Joseph, and J. Kubiatowicz, "Locality aware mechanisms for large-scale networks," in *Proceedings of FuDiCo*, June 2002, pp. 229–238.

[4] T. Karagiannis, A. Broido, N. Brownlee, k. claffy, and M. Faloutsos, "Is P2P dying or just hiding?" in *Globecom*, Dallas, Texas, USA, Dec 2004.

[5] A. Parker, "The True Picture of Peer-To-Peer File-Sharing," Sep. 2005, panel Presentation, IEEE 10th Int'l. Workshop on Web Content Caching and Distribution, Sophia Antipolis, France.

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

[7] D. Qiu and R. Srikant, "Modeling and performance analysis of bittorrent-like peer-to-peer networks," in *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*. New York, NY, USA: ACM Press, 2004, pp. 367–378.

[8] N. Christin, A. Weigend, and J. Chuang, "Content availability, pollution and poisoning in file sharing peer-to-peer networks," in *ACM Conference on Electronic Commerce, Vancouver, Canada*, June 2005. [Online]. Available: http://www.weigend.com/ChristinWeigendChuang2005.pdf

[9] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with Rocketfuel," *IEEE/ACM Transactions of Networking*, vol. 12, no. 1, pp. 2–16, 2004.

[10] R. Prasad, C. Dovrolis, M. Murray, and K. Claffy, "Bandwidth estimation: metrics, measurement techniques, and tools," *Network, IEEE*, vol. 17, no. 6, pp. 27–35, 2003.

[11] S. Saroiu, P. K. Gummadi, and S. Gribble, "A measurement study of peer-to-peer file sharing systems," in *Multimedia Computing and Networking (MMCN '02)*, January 2002. [Online]. Available: http://www.cs.washington.edu/homes/gribble/papers/mmcn.pdf

[12] S. Sen and J. Wang, "Analyzing peer-to-peer traffic across large networks," in *IMW '02: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*. New York, NY, USA: ACM Press, 2002, pp. 137–150.

[13] M. Izal et al., "Dissecting BitTorrent: Five Months in a Torrent's Lifetime." in *Proceedings of PAM*, Antibes Juan-les-Pins, France, April 2004, pp. 1–11.

[14] F. Le Fessant, S. Handurukande, A.-M. Kermarrec, and L. Massouli, "Clustering in peer-to-peer file sharing workloads," in *3rd International Workshop on Peer-to-peer systems (IPTPS 04)*, San Diego, CA, February 2004. [Online]. Available: http://www.irisa.fr/paris/Biblio/Papers/Kermarrec/LefHanKerMas04IPTPS.pdf

[15] A. Iosup, P. Garbacki, J. Pouwelse, and D. Epema, "A scalable method for taking detailed and accurate geo* snapshots of large P2P networks," TU Delft, Tech. Rep. PDS-2005-002, ISSN 1387-2109, 2005.

[16] L. Plissonneau, J.-L. Costeux, and P. Brown, "Analysis of Peer-to-Peer Traffic on ADSL," in *PAM*, 2005, pp. 69–82.

[17] V. Paxson, "Strategies for Sound Internet Measurement," in *ACM/USENIX Internet Measurement Conference*, Taormina, Italy, October 2004, pp. 263–271.

[18] 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.

[19] L. Peterson, T. Anderson, D. Culler, and T. Roscoe, "A Blueprint for Introducing Disruptive Technology into the Internet," in *HotNets–I*, October 2002.

[20] S. Banerjee, T. G. Griffin, and M. Pias, "The Interdomain Connectivity of PlanetLab Nodes," in *Proceedings of the Passive and Active Measurement Workshop (PAM2004)*, Antibes Juan-les-Pins, France, April 2004.

[21] N. Spring, D. Wetherall, and T. Anderson, "Scriptroute: A public Internet measurement facility," in *the 4th USENIX USITS*, March 2003.

[22] V. N. Padmanabhan and L. Subramanian, "An Investigation of Geographic Mapping Techniques for Internet Hosts," *Proceedings of SIGCOMM*, p. 13, 2001. [Online]. Available: citeseer.ist.psu.edu/padmanabhan01investigation.html

[23] B. Huffaker, M. Fomenkov, D. Moore, and K. Claffy, "Macroscopic analyses of the infrastructure: Measurement and visualization of Internet connectivity and performance," in *PAM*, Amsterdam, The Netherlands, April 2001.

[24] M. Castro, P. Druschel, Y. C. Hu, and A. I. T. Rowstron, "Topology-aware routing in structured peer-to-peer overlay networks." in *Future Directions in Distributed Computing*, ser. Lecture Notes in Computer Science, A. Schiper, A. A. Shvartsman, H. Weatherspoon, and B. Y. Zhao, Eds., vol. 2584. Springer, 2003, pp. 103–107.

[25] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in *IEEE INFOCOM'02*, 2002.

[26] B. Wei, G. Fedak, and F. Cappello, "Scheduling independent tasks sharing large data distributed with BitTorrent," in *IEEE/ACM Grid'2005 Workshop*, 2005.