Insights into the issue in IPv6 adoption: A view from the Chinese IPv6 Application mix

Chunjing Han* †, Zhenyu Li†, Gaogang Xie†, Steve Uhlig‡, Yulei Wu*,
Liangxiong Li*, Jingguo Ge*, Yunjie Liu†
*Computer Network Information Center, Chinese Academy of Sciences, China
†Institute of Computing Technology, Chinese Academy of Sciences,China
‡Queen Mary University of London, London, UK
Email:chjhan@cstnet.cn, zyli@ict.ac.cn, xie@ict.ac.cn, wuyulei@cstnet.cn, steve.uhlig@gmail.com,,
liliangxiong@cstnet.cn, gejingguo@cstnet.cn, liuyj@chinaunicom.cn

***Abstract*—Although IPv6 has been standardized more than 15 years ago, its deployment is still very limited. China has been strongly pushing IPv6, especially due to its limited IPv4 address space. In this paper, we describe measurements from a large Chinese academic network, serving a significant population of IPv6 hosts. We show that despite its expected strength, China is struggling as much as the western world to increase the share of IPv6 traffic. To understand the reasons behind this, we examine the IPv6 applicative ecosystem. We observe a significant IPv6 traffic growth over the past 3 years, with P2P file transfers responsible for more than 80% of the IPv6 traffic, compared to only 15% for IPv4 traffic. Checking the top websites for IPv6 explains the dominance of P2P, with popular P2P trackers appearing systematically among the top visited sites, followed by Chinese popular services (e.g., Tencent), as well as surprisingly popular third-party analytics including Google. Finally, we compare the throughput of IPv6 and IPv4 flows. We find that a larger share of IPv4 flows get a high-throughput compared to IPv6 flows, despite IPv6 traffic not being rate limited. We explain this through the limited amount of HTTP traffic in IPv6 and the presence of Web caches in IPv4. Our findings highlight the main issue in IPv6 adoption, i.e., the lack of commercial content, which biases the geographic pattern and flow throughput of IPv6 traffic.**

I. Introduction

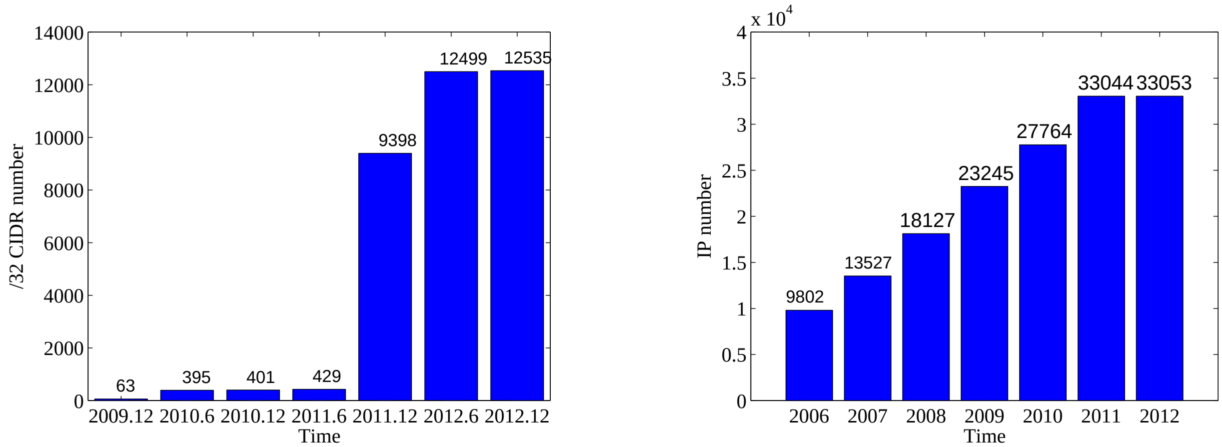
IPv6 has been standardized more than 15 years ago, to solve the issue of the IPv4 address space exhaustion. The adoption of IPv6 has been painstakingly slow [6][30], despite the readiness of most large ISPs. In 2011, the pool of available IPv4 blocks finally ran dry. The RIRs (Regional Internet Registries) are predicted to run out of unallocated IPv4 blocks in the next few years [1]. Despite the inevitability of IPv6 and efforts such as the IPv6 World Day, very little IPv6 traffic is carried across the Internet [2].

Referecce [22] said that by the end of December 2012, China has had 564 million Internet users, with a total of 50.9 million new ones. However, China has fewer IPv4 addresses allocated than a single US institution such as Stanford University[1]. The amount of IPv4 address space allocated to China is way too small to meet the growing popularity of the Internet in such a large country, as illustrated in Fig. 1b [22]. Since 2003, the Chinese government has launched a number of large projects on IPv6 almost every year [32], to build the network and promote its adoption. We can observe on Fig. 1a [22] the sharp increase in IPv6 addresses allocated to China after the exhaustion of IPv4 in 2012. China is now the third largest country in number of IPv6 addresses allocated [22]. Currently, while major ISPs in China support IPv6, most of them are still at the trial stage. Only two Chinese operators provide commercial-grade IPv6 support: CERNET (China Education and Research Network) and CSTNET (China Science and Technology Network). CERNET mainly provides IPv6 services for campuses. CSTNET mainly provides Internet connectivity mostly to research institutes all over China, and some governmental organizations, high-tech companies, and universities. CSTNET has about two million users, and provides network access as well as services such as data centre hosting.

Given the IPv6 experiences to increase its share of the global Internet traffic, we believe that it is important to study how different subsets of the Internet behave. While there are quite a few works on IPv6 topology, evolution, performance [11], and IPv6 DNS configuration [13], IPv6 application has not been extensively studied, partly because of the limited amount of IPv6 traffic [6]. Previous work on Chinese IPv6 [17], [18], [19] were conducted based mainly on campus traffic in CERNET and the datasets used were collected more than two years ago. Over these last two years, IPv6 traffic has increased significantly in China and the traffic of CSTNET has been never studied before. Moreover, as CSTNET serves different types of subscribers than CERNET, it might exhibit unique characteristics. Motivated by these facts, we in this paper perform an extensive study based on the datasets collected from CSTNET in 2011-2013 with an emphasis on IPv6 application mix pattern. The datasets enables us to analyze the possible discrepancies between IPv6 campus and IPv6 academic network. In addition, we also compare the application mix between IPv4 and IPv6 networks. We believe that our Chinese-centric study will shed some light on the current struggle to increase IPv6 adoption.

Our analysis in this paper relies on datasets collected from CSTNET between 2011 and 2013. The datasets consist of a combination of SNMP-based traffic statistics, sampled flow aggregate records, raw flow information, and packet-level traces from both native IPv6 as well as dual-stack IPv4-IPv6. We make the following observations about Chinese IPv6

[1] http://www.internetnews.com/infra/article.php/3605501/Were+Running+Out+of+IP+Addresses.htm



(a) IPv6

(b) IPv4

Fig. 1: Address allocations in China: (a) IPv6 and (b) IPv4.

traffic:

- Traffic volume: IPv6 traffic has grown very fast over the last three years in CSTNET. At the end of 2013, the total in/out traffic is as high as 8 Gbps. In comparison, in October 2011, it was barely 2 Gbps (see Fig. 3).
- Transport protocols: TCP dominates over UDP, both for IPv4 and IPv6. Still, the share of UDP is larger than that shown in the refenrence [19] and other IPv4 network [23], with between 25% and 30% of the total traffic.
- Application mix: More than 80% of the IPv6 traffic is generated by P2P applications. IPv4 traffic on the other hand is dominated by HTTP.
- Popular sites: The most popular two IPv6 commercial sites are Google and Tencent, which have only 3% and 15% of the IPv6 traffic. Surprisingly, top popular IPv6 sites include P2P trackers and analytics. Commercial sites are absent from the heavy-hitters.
- Geographic pattern: Most of the IPv6 traffic in China is domestic, i.e., bound to China. This is partly a consequence of the dominance of P2P, as well as the limited number of commercial websites supporting IPv6.
- Flow throughput: IPv6 flow throughput is higher than that for IPv4. On the other hand, IPv6 flow throughput for P2P and HTTP applications are comparable.
- IPv6 adoption: Our findings highlight the main issue in IPv6 adoption, i.e., the lack of commercial content, which biases the geographic pattern of IPv6 traffic and impedes on end-to-end flow performance.

The remainder of this paper is organized as follows. In Section II, we provide details about our three datasets. We investigate the traffic volume, application mix, geographic and temporal distribution of IPv6 traffic in Section III. We examine IPv6 flow throughput and IPv6 headers in Section IV. We present related work in Section V and summarize our results in Section VI.

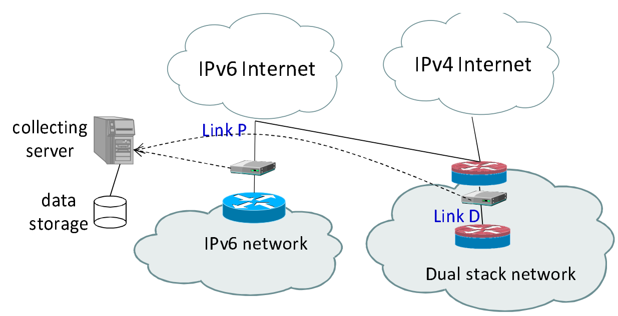


Fig. 2: Data collection setup.

II. Datasets

Our datasets were collected from the IPv6 backbone of CSTNET. CSTNET provides network-wide native IPv6 connectivity. We captured traffic from two backbone links in CSTNET: a 10GE link carrying native IPv6 traffic (called ”Link P”) and another 10GE link carrying IPv4-IPv6 dual-stack traffic (called ”Link D”) to the Internet. The link P mainly servers universities and therefore carries campus traffic, while the traffic carried in link D is mainly from academic community and offices. The traffic probes on both links are equipped with a specialized network card and run nProbe [3]. The probes perform DPI (Deep Packet Inspection) and export flow records using the Netflow format [10]. Monitoring statistics are reported to a collecting server for further aggregation and analysis.

We collected 4 datasets (see Table I) using the traffic monitoring system. The first dataset is a 3-year long daily traffic trace collected from the outbound routers of Link P (see Fig. 2) using SNMP. The IPv6 traffic statistics of Link D are not included because the dual-stack router stores the statistics of IPv4 and IPv6 traffic in the same SNMP MIB entry two years ago.

TABLE I: Summary of the datasets.

Dataset	Duration	Type (and size)
1	3 years (2011-2013)	RRd file
2	1 week (June 2013)	Netflow (1083.6M flows)
3	2 hours (both day and night)	Netflow (66.05M flows)
4	1 hour	Packet headers (87.85 GBytes)

The second dataset is a 1-week long trace in June 2013, consisting of Netflow (v9) IPv4/IPv6 sampled flow information, with a packet sampling rate of $1/1,000$. In total, we observed 1083.6 million flows and 0.28 million unique IPv6 addresses during this week. The sampling allows the relatively long observation period.

The third dataset is made of unsampled flow statistics in Netflow v9 format. Due to the large amount of traffic, we take 2-hour long traces for both link P and D. The dataset contains 6 subsets, 3 of which (link P IPv6 (PD6), link D IPv6 (DD6), link D IPv4 (DD4)) were collected in the afternoon (2PM-4PM) and 3 were collected during the night (9PM-11PM), called "Link P IPv6 (PN6)", "link D IPv6 (DN6)", "link D IPv4 (DN4)" respectively. Note that we separated the IPv6 flows in Link D from the IPv4 flows. In total, the dataset contains 37 million IPv6 flows and 39 million IPv4 flows.

The fourth dataset is a one-hour IPv6 packet level trace, which was collected on both link P and D. We use this dataset for IPv6 packet analysis in Section III.

The four datasets provide complementary views in terms of their time duration. The second and third datasets provide us with per-flow statistics, including flow size, duration and throughput. We define flow size as the total number of bytes transferred in the flow (including headers). The duration of a flow is the time elapsed between the first and last packets seen for this flow. Finally, while being aware of the limitations of IP geolocation, we use Maxmind GeoIP [5] to map IP addresses to countries/regions.

III. IPv6 TRAFFIC CHARACTERISTICS

We begin this section by examining the growth of IPv6 traffic, and then proceed to analysis of IPv6 traffic patterns from the perspective of transport protocols, application mix, popular sites and geographic distribution.

A. *Growth of IPv6 traffic*

We use the first dataset to illustrate the growth of IPv6 traffic as seen from link P. Figure 3 provides the evolution of traffic volume in both directions of link P. At the end of October 2013, the total in/out traffic was as high as 10 Gbps. The three periods when the traffic sharply decreases (as low as 100 Mbps) correspond to the Chinese Spring Festival vacations, where network usage of research and academic networks is minimal. We see a sharp increase of traffic in September 2011, due to an upgrade of link P from 1 to 10 Gbps. While the overall in and out traffic is relatively symmetric, we observe asymmetric traffic during two periods: Sep.-Oct. 2011 and July-Sep. 2012. Investigations with the

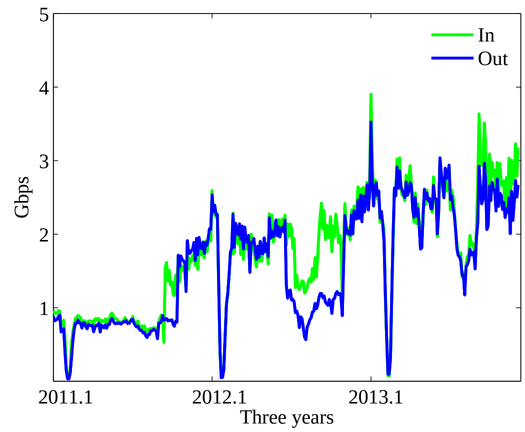


Fig. 3: Growth of IPv6 traffic on Link P.

TABLE II: Transport protocol breakdown.

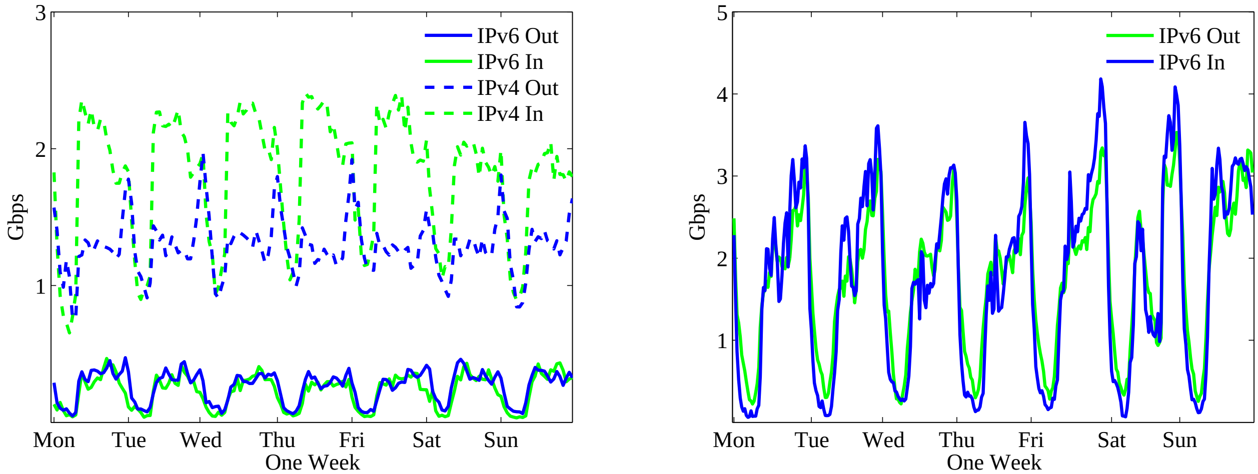
	IPv6 Link P			IPv6 Link D			IPv4 Link D		
Protocol	TCP	UDP	Other	TCP	UDP	Other	TCP	UDP	Other
Traffic(%)	64.28	30.23	4.88	70.80	27.68	1.53	64.41	34.60	0.97

network operator concluded that these were caused by BGP (IPv6) configuration changes between CSTNET and CERNET.

Figure 4 provides a comparison of the traffic pattern over a one week period for IPv4 (Figure 4a) and IPv6 traffic (Figure 4a and 4b) in June 2013. The IPv4 traffic in Figure 4a displays the traffic of a typical research and office network [21], which is different from campus traffic in Figure 4b and the result in [19]. Indeed, we observe that peak hours are the work office hours, with very little traffic during the evening. Interestingly, the incoming IPv4 traffic is larger than the outgoing, suggesting that requests from Chinese academic institutions towards the rest of the world is much higher than what is consumed by users of outside networks. Contrary to IPv4, we observe that IPv6 is symmetric. Such a symmetry is curious, and is rooted in the peculiar application mix for IPv6 traffic. We are now heading towards studying the IPv6 application mix. We determine the application of a flow through a DPI-based application traffic identification engine located on the probes [4]. We use the second dataset (1-week long flow-level traces) for this analysis.

B. *Transport protocols*

We first examine the IPv4/IPv6 traffic carried by different transport protocols in Table II. For both IPv4 and IPv6, the majority of the traffic is carried by TCP. 4.88% of IPv6 campus traffic is carried by transport protocols other than TCP and UDP, compared to only 0.97% for IPv4 traffic and the 1.53% for IPv6 research traffic. Further investigations reveal that most of this IPv6 traffic is fragmented (3.45%) or corresponds to IPv6 address announcement over ICMPv6 (1.42%). Compared to previously reported in [19] (with 1.63%), we observe a lot of fragmentation, indicative of a poor application software transition to IPv6.



(a) Weekly IPv4 and IPv6 traffic of Link D in June 2013.

(b) Weekly IPv6 traffic of Link P in June 2013.

Fig. 4: Weekly traffic in June 2013: (a) IPv4 Link D and (b) IPv6 Link P.

TABLE III: Breakdown of main applications for IPv6 traffic.

Application mix (bytes, %)	Link P IPv6	Link D IPv6	Link D IPv4
P2P	85.84	63.36	14.92
HTTP	0.90	6.51	44.35
Streaming & VoIP	4.02	4.10	6.39
File transfer	1.70	0.75	1.69
Game	0.67	0.40	1.40
Chat	0.05	0.97	1.75
Network operation	0.05	0.02	3.34
Mail	0.02	0.00	0.53
Security	0.01	0.00	1.09
Other	6.78	23.90	24.54

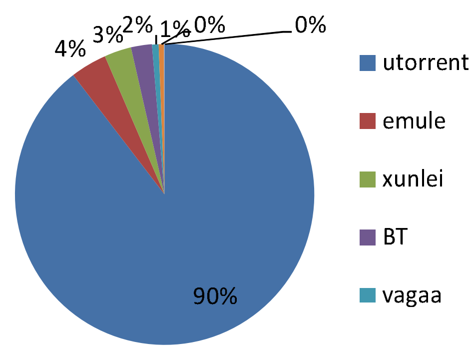


Fig. 5: Flavors of P2P protocol of Link P.

C. Application mix

Table III provides the distribution of traffic for the major applications. We observe that the application mix is less diverse that the one reported in [2]. P2P (Peer-to-Peer) file sharing is the dominant application for IPv6, confirming observations made in China [19] and in US [8]. Moreover, P2P applications account for as much as 86% of the total traffic for Link P, which is much higher than in the US (61%) [8] and by previous work in China (78%) [19]. Link P carries more P2P traffic than link D because users of Link P are mostly university students, while the population using Link D are mostly researchers and engineers. As IPv6 is not charged and not rate-limited, it is used preferentially for file downloads, especially high-definition video.

Fig. 5 provides a breakdown of the P2P flavors in Link P, which is as similar as Link D. We observe that a limited number of P2P applications are responsible for most of the IPv6 traffic, with 90% of the P2P traffic generated by utorrent. This suggests a loss of diversity in the P2P ecosystem [2], [7], [19]. For the "Other" application of Link D, by analyzing the organizations on both sides of communications, we find that there are some scientific data transmission tasks. Because the IPv6 transition of file transferring software is not as difficult as HTTP application [28], and the demand for traffic volume is very huge, both of which improve the IPv6 adoption.

We observe widely different application mixes for IPv4 and IPv6 in Figure 6. IPv4 is dominated by HTTP while IPv6 by P2P. HTTP applications account for as much as 40% of the IPv4 traffic, while P2P applications only account for 15% of IPv4 traffic. Similar figures have been reported in the west, e.g. US [8]. Two factors contribute to the difference in the application mix between IPv4 and IPv6 network. First, there is no negative incentive against P2P in the IPv6 network. One reason could be the lack of IPv6-capable traffic management tools [8], or a willingness to stimulate IPv6 traffic at all costs. Second, P2P applications in IPv6 provide content (high-definition video content, software and game) that is actually not freely available in IPv4 commercially provided Internet content provider. This is in contrast to the availability of HTTP content, with very limited HTTP content available on IPv6. Indeed, as we will show next, P2P trackers in IPv6 network are very popular websites, accounting for more than 70% of HTTP visits. Further, commercial HTTP traffic is extremely low in IPv6.

TABLE IV: Example of popular websites in IPv6 and their hosting organization.

Domain	Application type	Hosting organization
ipv6.neubt.com	P2P tracker	Northeastern university
tracker.cgbt.cn	P2P tracker	Beijing Jiaotong University
mat1.gtimg.com	HTTP (pictures)	Tencent
www.google-analytics.com	Web analytics	Google
googleads.g.doubleclick.net	Web advertising	Google
zhenghongjuan12.appspot.com	Google app	Google
img1.gtimg.com	HTTP (images)	Tencent
pagead2.googlesyndication.com	Advertisement	Google
www.52v6.com	P2P tracker	Beijing University of Science and Technology
bt.byr.cn	P2P tracker	Beijing University of Posts and Telecommunications

TABLE V: Example of popular websites in IPv4 and their hosting organization.

Domain	Application type	Hosting organization
baidu.com	Search engine	Baidu
qq.com	Social-networking	Tencent
webterren.com	Web analytics	Webterren
sina.com	News portal	Sina
qiyi.com	online video	Baidu
360.cn	Network security	Qihu360
taobao.com	E-commerce	Alibaba
sogou.com	Search engine	Tencent
weibo.com	microblogging sites	Sina
ac.cn	academy	CAS

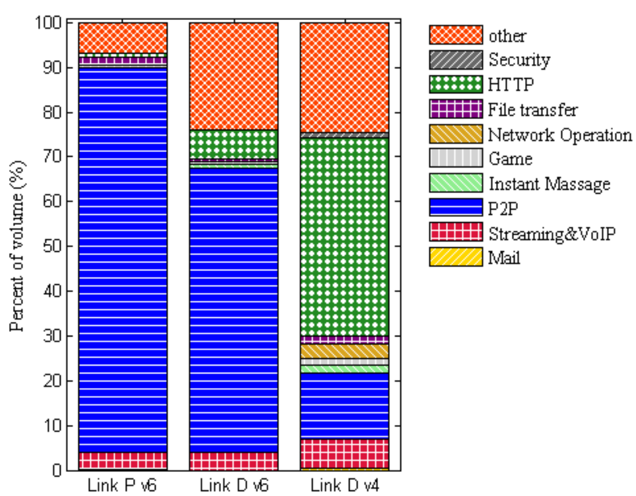


Fig. 6: Application mix.

D. Popular sites

We now dig into the most visited websites. Table IV lists the top 10 most popular IPv6 websites of (aggregating the packets of the Link D and Link P together) based on a one hour trace using the fourth database. While different times of the day see a different ordering of the most popular websites, the type of sites in the top 10 are stable over time across the traces. Moreover, we observe that P2P trackers are well represented, consistently with the dominance of P2P traffic in IPv6. HTTP is also represented, through Tencent. Finally, and maybe surprisingly, analytics, advertising, and apps hosted by Google are also present. Despite the mixed relationship between Google and China, its good IPv6 support brings it among the popular sites. Despite the strong presence of P2P among the top sites, the presence of third-party analytics gives some hope that more and more web traffic will be seen in the future once commercial popular websites migrate to IPv6.

To understand the difference of popular websites in IPv6 and IPv4 network, we further list the top 10 most popular IPv4 websites of Link D in Table V. As expected, search engine and social networking sites dominate the list. The difference in top websites in IPv6 and IPv4 network implies the slow adoption of IPv6 for popular websites. Interestingly, although Tencent hosts two popular websites in both IPv4 and IPv6, they provide different services. The two sites in IPv6 provide image services, while the sites in IPv4 provide social networking service and search portal.

E. Geographic and temporal distribution

We now examine the distribution of IPv6 traffic on a per-country basis. We consider the IPv6 traffic on both Link P and Link D, and compare its geographic distribution with the IPv4 traffic on Link D. A region's traffic is computed as the aggregated traffic of the flows having the country either as source or destination.

We find that 36.7% of the IPv4 traffic is international, i.e., either the source or destination lies outside China. In contrast, only 1% of IPv6 traffic is international. As shown in Table VII, 46.97% of the IPv4 international traffic is exchanged with the US. This number for the US might be inflated due to inaccurate geolocation, e.g., of Google that is geo-located in California. However, it represents well the dominance of US players in IPv4, especially commercial players.

For IPv6 international traffic, we do a similar analysis to [19]. From Table VI, we see that the geographic distribution of IPv6 international traffic differs widely from the one of IPv4. The countries that exchange most IPv6 traffic with China are in Europe, e.g., Switzerland (24.54%). Europe captures 79.16%

of the international IPv6 traffic. Application-level analysis reveals that there is very little P2P traffic exchanged with Germany, Switzerland and the Czech republic, but rather it is HTTP and Other applications. Further digging into the responsible networks shows that this traffic is mostly file transfers of scientific data between European countries and China. For example, TCP and browser-based applications are used to transfer data between the Institute of High Energy Physics of Chinese Academy of Sciences and CERN.

TABLE VI: Top countries for IPv6 international traffic.

Region	IPv6 international traffic (bytes,%)
DE	27.00
CH	24.54
CZ	11.88
US	10.14
FR	7.58
IT	5.93
AU	4.79
TW	3.18
NL	1.26
SE	0.97
Other	2.73

TABLE VII: Top countries for IPv4 international traffic.

Region	IPv4 international traffic (bytes,%)
US	46.94
GB	6.83
AP	5.61
DE	5.53
HK	4.31
KR	3.82
JP	3.76
TW	3.50
FR	3.10
RU	1.83
Other	14.77

Figure 7 shows the application mix for Chinese IPv6 traffic for the top 10 international countries, considering a one week period. We observe that both HTTP and P2P represent a significant share of the traffic. Regarding P2P, US (see Figure 7) and France dominate, because of the popularity of their trackers. As already mentioned, Germany and Switzerland both have a significant amount of traffic using specific TCP-based applications to transfer scientific data. We analyze that the heavy-hitters for HTTP in DE, CH, CZ, and US are just the websites of these IPv6-capable research and technology institutes, which indicates these institutes make full use of their IPv6 network in their website, FTP and customized application tools.

Figure 8 shows the evolution of the application mix for all (domestic and international) IPv6 traffic, throughout a day. We observe a relatively stable application mix, due to the small amount of international traffic and the dominance of P2P (see Table IV), with only minor changes such as a bit more P2P in the early morning. We also observe a slightly higher fraction of streaming traffic in the late night around 9 PM, when users watch online videos.

IV. FLOW PROPERTIES

Flow throughput is an important performance indicator. We measure flow throughput by dividing the flow size by the duration of the flow. The duration of a flow is the time elapsed between the first and last packet. As in [21], we excluded the flows whose duration is less than 100 ms (millisecond), as their small duration biases the throughput. We rely on the third dataset, which contains unsampled flow information.

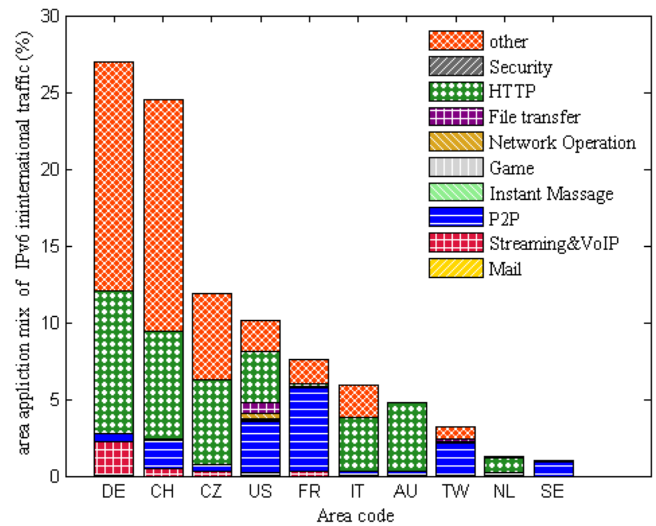


Fig. 7: Traffic mix for international IPv6 traffic.

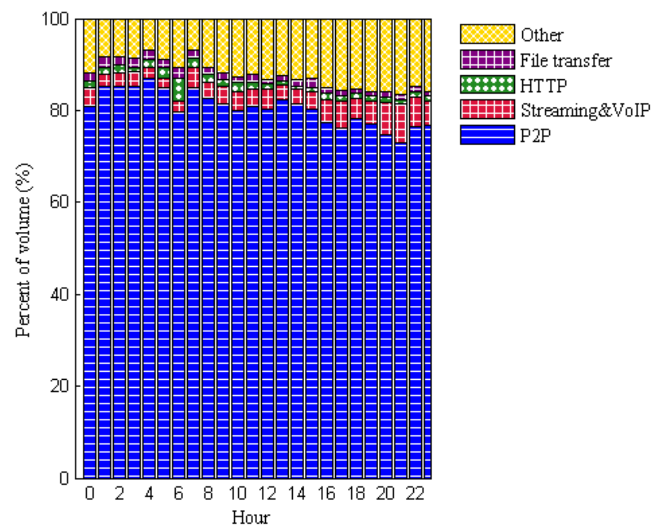


Fig. 8: Evolution over time of all IPv6 traffic (domestic and international).

A. Flow rate from transport and application perspectives

Figure 9 plots the Cumulative Distribution Function (CDF) of flow throughput for TCP flows and UDP flows during daytime. The distribution of flow throughput for flows at night is similar (not shown). Overall, TCP flows have much higher rates than UDP, for both IPv4 and IPv6. UDP flow throughput for IPv4 and IPv6 are comparable. While most IPv6 TCP flows get a higher throughput than the IPv4 ones, there are more IPv4

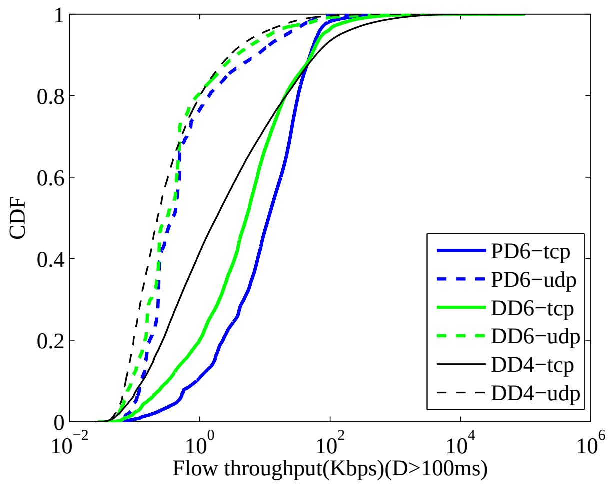


Fig. 9: CDF of flow throughput: TCP vs. UDP.

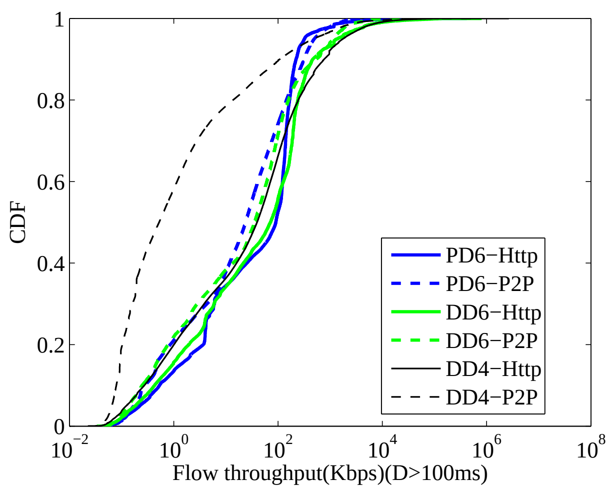


Fig. 10: CDF of flow throughput: HTTP vs. P2P.

flows that get a throughput larger than 100 Kbps compared to IPv6 flows. This is surprising, as IPv6 flows are not rate limited in China.

As IPv4 and IPv6 have quite different application mixes, we now examine the flow throughput for the two most popular applications (see Figure 10): HTTP and P2P. We observe that some HTTP flows in IPv4 obtain a higher throughput than other types of flows. Digging into the flows that obtain a higher throughput, we identified the reason to be web caches [29] inside the network. This explains what we found from Figure 9. Moreover, while the throughput of HTTP flows and P2P flows transmitted through link P (PD6) is comparable, the flow throughput for P2P flows through link D is much lower than the rate for HTTP flows, especially for IPv4 flows. The main cause is traffic shaping/control equipment enabled in the dual-stack network.

We further examine the how the flow throughout relates to flow size in Figure 11 for link D. Note that the x-axis is in

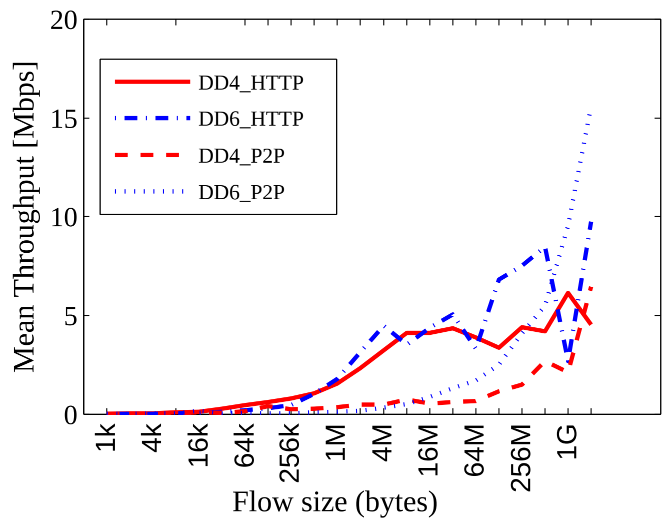


Fig. 11: Flow throughput versus flow size

logarithm scale. The throughput for HTTP flows larger than 256KB in both IPv4 and IPv6 network ramp up quickly and then remains steady around 4Mbps. This is explained by the widely deployed web cache. The large variations for HTTP flows larger than 64MB in IPv6 network are due to fewer samples for large-size flows. In contrary, the throughput for P2P flows grows slowly with flow size, especially for IPv4 flows. We can also observe that P2P flows in IPv6 network have a higher average throughput than P2P flows in IPv4 network, possibly due to the traffic shaping/control effect on IPv4 flows.

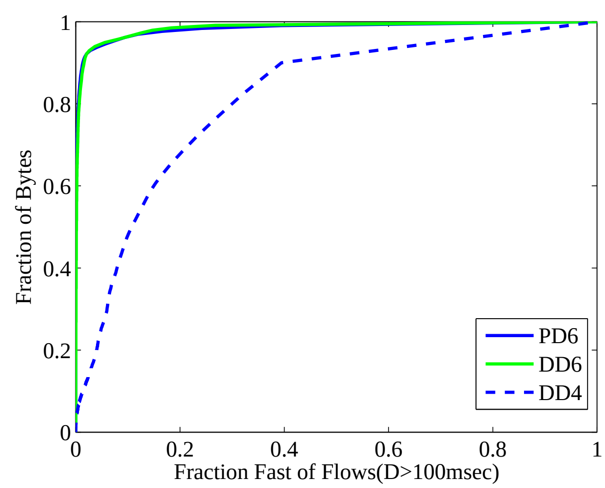


Fig. 12: CDF of traffic in flows (ordered by decreasing throughput).

Now, we examine how important, in terms of total traffic, high-throughput flows are, by computing the cumulative fraction of traffic generated by the highest-throughput flows. Figure 12 plots the results for the three different links. First, we observe that IPv6 flows have a much more skewed distribution than IPv4. The top 10% IPv6 flows in throughput account for about 95% of the total IPv6 traffic. In comparison, the top 10% IPv4 flows in throughput only account for 48% of the IPv4 traffic. It seems that the main reason behind the difference between IPv4 and IPv6 is that in IPv6 network, majority of the IPv6 fast flows are P2P flows, while in IPv4 network, HTTP flows constitute most of the fast flows.

B. IPv6 headers

One of the major differences between IPv4 and IPv6 is the packet header. However, as shown in [9], IPv6 header options are not used yet. While their usage is very small, the large amount of traffic we see still observes packets having non-zero FlowLabel, AH, ESP and mobility fields. We report in Table VIII the number of packets and corresponding bytes.

TABLE VIII: IPv6 header options usage

Filled value>0	Packets (percent)	Payload bytes (percent)
Traffic Class	1,158 (0.0036%)	1,189,928 (0.0082%)
Flow-label	82,408 (0.259%)	45,126,042 (0.311%)
Authentication Header	0	0
Encapsulating Security Payload	0	0
Mobility (currently without upper-layer header)>0	0	0
Total statistics	31,709,820	13.54GB

The Flow Label [27] is seen in less than 0.3% of the packets and the Traffic Class is also barely used, no more than 1%. This is not surprising, as the situation is similar to the lack of use of the ToS field in IPv4. Even though RFC 6437 describes the usage of the IPv6 Flow Label Specification, no best practice has been defined for its usage, and therefore its limited usage is not surprising. The security and mobility extension headers are not used at all in the studied network. We believe that the situation is likely to change in the future for the security and mobility extension headers, and that their absence are merely an artifact of the studied network.

From the use of the headers, we can conclude that security and mobility are not the motivation behind the increase in IPv6 traffic. As most of the traffic is IPv6, one should not have expected a much different outcome. However, as more commercial sites and applications migrate to IPv6, we expect the situation of the header use to change drastically.

V. RELATED WORK

Several studies [17][18][19] examined the IPv6 traffic of CERNET2. The data they used for analysis were collected two years ago and they mainly focus on campused IPv6 traffic. While geographical distribution of IPv6 traffic is examined in [19], it neither compared IPv4 and IPv6 traffic under the same user group, nor compares campus IPv6 and research IPv6 traffic, as we do in this paper. In addition, we in this paper use a DPI-based application identification engine, which shows to be more accurate than port-based engine used in [18]. Note that currently, only two Chinese network operators, CSTNET and CERNET, provide public-grade native IPv6. Others such as China Mobile, China Telecom, China Unicom, have deployed some IPv6 metropolitan area networks in small scale only for experiment and demonstration.

Google has supported IPv6 in their web services [24]. Users are redirected to Google's IPv6 servers if they use dedicated DNS servers, which return IPv6 addresses, rather than IPv4 addresses, to users. CAIDA's analysis based on datasets collected from two OC-192 commercial backbone links showed that IPv6 traffic accounts for 0.03% of total traffic [6]. At the same time, Colitti et al. [30] also measured IPv6 adoption and found that actual IPv6 deployments are very few, but growing steadily. Akamai [28] examined the challenges web sites will face in trying to provide high quality experiences to their end users during IPv6 transition. They pointed out that website servers should support dual stack in long times.

Karir et al. [15] studied the population using IPv6 by counting active IPv6 addresses. Their results showed that the majority of IPv6 hosts rely on transition mechanisms, but in CSTNET and CERNET[19], most hosts use native IPv6. Dhamdhere et al. [11] analyzed the evolution and development of IPv6 ASs, the corresponding AS-path, and some IPv6-enabled web sites. They reached a similar conclusion to [12], that the performance over IPv6 paths is comparable to the one of IPv4 paths if the AS-level paths are the same. In China, CERNET has a separate AS for IPv6 traffic, and does not use their IPv4 AS to carry IPv6 traffic, while CSTNET uses the same AS to transfer the IPv4 and IPv6 traffic.

Labovitz [8] in 2011 reported that P2P applications account for 61% of the IPv6 traffic in the west. Sandvine et al. [26] collaborated with several Tier-1 North American Cable operators in assessing World IPv6 Day's influence on their own networks. The IPv6 application mix has changed [2] after the "World IPv6 Day" (June 8th, 2011), during which several large content and service providers coordinated a large-scale IPv6 test-run. After this event, about 60% of the studied IPv6 traffic was HTTP. We find a different application mix in the Chinese IPv6 traffic. P2P accounts for more than 80% of the traffic in 2013, with HTTP generating less than 10%.

Alzoubi et al. [16] quantified the performance implications of switching to IPv6 for a website. They found no evidence of performance penalty when doing this. Claffy [14] provided an extensive survey of available data facilitating the tracking of IPv6 deployments and called for more publicly available data. To help in this direction, we are planning to contribute some traces from the Chinese IPv6 network in the near future. Sasanus et al. [31] studied the bandwidth requirements when various applications migrate to IPv6 and the impact of IPv6 protocol for network application.

Cho et al. [20] and Maier et al. [7] investigated the characteristic of residential customer traffic by comparing DSL and fiber users, heavy-hitters and normal users, and geographic traffic matrices as seen from backbone links. Zhang et al. [21] examined Internet flow throughput and the relationship between the throughput and other flow characteristics such as size and duration.

VI. SUMMARY

This paper provides an updated view on IPv6 traffic from the perspective of China. We study IPv6 in China through multiple traces gathered at a large academic network. We reported a sharp growth of IPv6 traffic in China in the past two years. We observe a dominance of P2P applications in IPv6 traffic, much stronger than two years ago [19]. We find that most IPv6 traffic is domestic, i.e., stays within China, contrary to IPv4 traffic. Within the limited amount of non-domestic IPv6 traffic, we find large fraction of this traffic not being carried by P2P nor HTTP, but by transfers of scientific data. Finally, we find that most IPv6 flows have a higher throughput than IPv4, though the highest throughput flows in IPv6 get a lower throughput than those from IPv4. Overall, our findings highlight the main issue in IPv6 adoption, i.e., the lack of commercial content, which biases the geographic pattern of IPv6 traffic and impedes on end-to-end flow performance.

REFERENCES

[1] G. Huston: IPv4 Address Report, 2013. http://www.potaroo.net/tools/ipv4/index.html.

[2] N. Sarrar, G. Maier, B. Ager, R. Sommer and S. Uhlig. Investigating IPv6 Traffic: What Happened at the World IPv6 Day? Proc. of PAM, 2012.

[3] nProbe v6. http://www.ntop.org/products/nprobe/

[4] G. Xie, G. Zhang, J. Yang, Y. Min, V. Issarny, and A. Conte. The survey on traffic of metro area network with measurement on-line. Proc. of the 20th International Teletraffic Congress, 2007.

[5] Maxmind GeoIP, http://www.maxmind.com/zh/ipv6.

[6] P. Hicks, Fraction of IPv6 traffic (in packets and bytes) for monthly passive traces, May 2011. http://www.caida.org/data/passive/trace_stats/ipv6_traffic.xml.

[7] G. Maier, A. Feldmann, V. Paxson, M. Allman. On Dominant Characteristics of Residential Broadband Internet Traffic. Proc of ACM IMC, 2009.

[8] G Labovitz.: Six month, six providers and IPv6. Tech. rep. (March 2011), http://www.monkey.org/~labovit/papers/v6sixmonths.pdf.

[9] S. Deering and R. Hinden, ”RFC 2460. Internet Protocol, Version 6 (IPv6) Specification,” 1998. http://tools.ietf.org/html/rfc2460.

[10] P. Hoffman, C. Bonatti, A. Eggen. ”RFC 3854. Netflow v9. Cisco Systems NetFlow Services Export Version 9,” 2004. http://www.ietf.org/html/rfc3954.txt.

[11] A. Dhamdhere and M. Luckie. Measuring the deployment of IPv6 topology, routing and performance. Proc. of ACM SIGCOMM IMC, 2012.

[12] M. Nikkhah, R. Guérin, Y. Lee, R. Woundy. Assessing IPv6 Through Web Access: A Measurement Study and Its Findings. Proc. of ACM CoNEXT, 2011.

[13] S. Park, J. Jeong, C. Hong. DNS Configuration in IPv6: Approaches, Analysis, and Deployment Scenarios. IEEE Internet Computing, vol. 17, no. 4, pp. 48-56, July-Aug. 2013.

[14] Kc Claffy. Tracking IPv6 evolution: data we have and data we need. SIGCOMM Comput. Commun. Rev., 2011.

[15] M. Karir, G. Huston, G. Michaelson, and M. Bailey. Understanding IPv6 Populations in the Wild. Proc of PAM, 2013.

[16] H. A. Alzoubi, M. Rabinovich, and O. Spatscheck. Performance Implications of Unilateral Enabling of IPv6. Proc of PAM, 2013

[17] W. Shen, Y. Chen, Q. Zhang, Y. Chen, B. Deng, X. Li, G. Lv, Observations of IPv6 Traffic. Proc. Computing, Communication, Control, and Management, (CCCM), 2009, pp. 278-282.

[18] L. Gao, J. Yang, H. Zhang, D. Qin, B. Zhang. What's Going On in Chinese IPv6 World. Proc. of Network Operations and Management Symposium (NOMS), 2012, pp. 534-537.

[19] F. Li, C. An, J. Yang, J. Wu, H. Zhang. A study of traffic from the perspective of a large pure IPv6 ISP. Computer Communications 37(2014) 40-52

[20] K. Cho, K. Fukuda, H. Esaki, A. Kato. The Impact and Implications of the Growth in Residential User-to-User Traffic. Proc. of ACM SIGCOMM, September 11-15, 2006.

[21] Y. Zhang, B. Lee, V. Paxson and S. Shenker. On the Characteristics and Origins of Internet Flow Rates. Proc. of ACM SIGCOMM, 2002.

[22] CNNIC, The 31th Statistical Report on Internet Development in China, 7 Jan, 2013 http://www1.cnnic.cn/IDR/ReportDownloads/201302/P020130312536825920279.pdf.

[23] M. Zhang, M. Dusi, W. John, C. Chen, Analysis of UDP traffic usage on internet backbone links. Proc. Applications and the Internet (SAINT), 2009

[24] Google, ”Google ipv6 implementors conference,” 2012, https://sites.google.com/site/ipv6implementors/.

[25] J. Wu, J. H. Wang, J. Yang, CNGI-CERNET2: an IPv6 deployment in China. Proc. J, ACM Computer Communication Review 41(2):48-52 (2011)

[26] Sandvine, Global Internet Phenomena Spotlight-World IPv6 Day. http://www.sandvine.com/downloads/documents/Sandvine.

[27] S. Amante,B. Carpenterr,S. Jiang,J. Rajahalme,”RFC 6437. IPv6 Flow Label Specification,” 2011. http://tools.ietf.org/html/rfc6437

[28] Akamai. IPv6: What the Transition Means for Content and Application Delivery. Technical report, 2012

[29] M. Dong, L. Zheng, K. Ota, S. Guo, M. Guo, L. Li: Improved Resource Allocation Algorithms for Practical Image Encoding in a Ubiquitous Computing Environment. JCP 4(9): 873-880 (2009)

[30] L. Colitti, S.H. Gunderson, E. Kline, T. Refice, Evaluating IPv6 Adoption in the Internet, Proc. PAM, 2010, pp. 141-150.

[31] S. Sasanus, R. Mortier. Differences in Bandwidth Requirements of Various Applications due to IPv6 Migration, Proc. Information Networking (ICOIN), 2012

[32] J. Wu, J. Wang, J. Yang, CNGI-CERNET2: an IPv6 Deployment in China, ACM SIGCOMM Computer Communication Review, Vol. 41, No. 2, April 2011: 48-52