

Personal cloud storage: Usage, performance and impact of terminals

Original

Personal cloud storage: Usage, performance and impact of terminals / Bocchi, Enrico; Drago, Idilio; Mellia, Marco. - ELETTRONICO. - (2015), pp. 106-111. (IEEE 4th International Conference on Cloud Networking (CloudNet) Niagara Falls, ON October 2015) [10.1109/CloudNet.2015.7335291].

Availability:

This version is available at: 11583/2625369 since: 2015-12-12T21:52:52Z

Publisher:

IEEE

Published

DOI:10.1109/CloudNet.2015.7335291

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Personal Cloud Storage: Usage, Performance and Impact of Terminals

Enrico Bocchi, Idilio Drago, Marco Mellia
Politecnico di Torino
name.surname@polito.it

Abstract—Personal cloud storage services such as Dropbox and OneDrive are popular among Internet users. They help in sharing content and backing up data by relying on the cloud to store files. The rise of mobile terminals and the presence of new providers question whether the usage of cloud storage is evolving. This knowledge is essential to understand the workload these services need to handle, their performance, and implications.

In this paper we present a comprehensive characterization of personal cloud storage services. Relying on traces collected for one month in an operational network, we show that users of each service present distinct behaviors. Dropbox is now threatened by competitors, with OneDrive and Google Drive reaching large market shares. However, the popularity of the latter services seems to be driven by their integration into Windows and Android. Indeed, around 50% of their users do not produce any workload. Considering performance, providers show distinct trade-offs, with bottlenecks that hardly allow users to fully exploit their access line bandwidth. Finally, usage of cloud services is now ordinary among mobile users, thanks to the automatic backup of pictures and media files.

I. INTRODUCTION

Personal cloud services have become very popular among Internet users. They offer an easy way to synchronize content over multiple devices and to share files with other users. The market of providers is quickly evolving, with well-established players like Dropbox facing the competition with giants like Google and Microsoft, now offering solutions more and more integrated into Windows or Android Operating Systems (OSs). Understanding the typical usage of these services is of primary importance to improve their performance, to point out and address bottlenecks, and to enhance end users' experience.

It is thus not surprising that the research community put a relevant effort in understanding personal cloud storage. Several related works are based on active measurements. Authors of [1] were the first to compare storage providers. By performing a set of active experiments, they measure the backup time of given workloads for four different services. Dropbox bottlenecks are analyzed in [2], while public APIs of three providers are compared in [3]. We introduced in [4] a framework to benchmark cloud storage services. Relying on an automated testing methodology, we compared design choices and benchmarked 11 services in a testbed. Authors of [5] follow a similar methodology, considering also different types of client terminals. Whereas these works have deeply

studied how service design impacts performance, they rely on active experiments in controlled environments. Thus, they miss a characterization of how users interact with each application.

Other works study cloud storage from different perspectives, considering the behavior of clients [6], the quality of experience [7], the type of content people store in the cloud [8], or the pricing practices in the storage market [9].

The only notable work offering a passive characterization of Dropbox usage, focusing on PCs, is presented in [10]. With this paper, we offer an updated picture 30 months after, and extend findings to other terminals used to access the services. We show that the scenario has significantly changed, with competitors challenging the popularity of Dropbox, and with users now accessing cloud storage from mobile terminals and web interfaces. Motivated by these findings, we highlight typical usage patterns and performance of three popular services, namely Dropbox, Microsoft OneDrive, and Google Drive.

We consider (Sect. II) a large passive measurement campaign covering the whole month of October 2014 and performed in a Point-of-Presence (PoP) of an European Internet Service Provider (ISP) where approximately 20,000 residential customers are monitored. One third of them access at least one personal cloud storage service. Contrary to [10], we discover that both Google Drive and especially OneDrive have now acquired a notable market share, contending the throne of the most popular service against Dropbox.

Results show (Sect. III) that Dropbox users produce the highest amount of traffic, while almost half of Google Drive and OneDrive users generate little data. We also observe that a relevant fraction of users now submit their files from mobile devices. This is due to the ever increasing popularity of smartphones and tablets, but also to the deeper integration of storage applications into the OS.

Next, we assess (Sect. IV) each service performance, unveiling and quantifying the presence of bottlenecks related to (i) the latency to reach data-centers; (ii) shaping policies that severely limit transfer rate (e.g., for OneDrive); and (iii) weaknesses in protocol designs that affect both Dropbox and Google Drive, not allowing users to saturate the high-speed access network today offered by ISPs.

At last, we study (Sect. V) how users access Dropbox taking into account multiple terminal types: From the application running on their PCs, via the web interface, or by apps running on smartphones and tablets. We show surprising differences on both the usage patterns and the performance they obtain.

This research has been funded by the European Union under the FP7 Grant n. 318627 (Integrated Project "mPlane").

TABLE I: Overview of our dataset.

Provider	Households	Upload (GB)	Download (GB)
Dropbox	5,020	681	1,209
OneDrive	4,910	336	132
Google Drive	2,016	164	311
Total	7,349	1,181	1,652

II. DATASETS AND METHODOLOGY

We rely on passive measurements collected in operational networks to characterize how people use different cloud services, and the influence of terminals on both usage and performance. We use the flow-level summaries exported by Tstat [11] in our analysis. A probe running Tstat has been installed in a PoP of an ISP in Europe. The probe monitors the traffic of more than 20,000 households, i.e., people accessing the Internet via ADSL or Fiber To The Home (FTTH) connections, and using PCs, smartphones and tablets connected via WiFi or LAN at home. FTTH offers up to 100 Mb/s and 10 Mb/s in downstream and upstream, respectively. Conversely, ADSL offers up to 24 Mb/s and 1 Mb/s.

For each TCP flow observed in the network, Tstat exports more than 100 metrics, including (i) the anonymized client IP address; (ii) the total number of bytes exchanged with servers; (iii) the timestamp of the first and the last packet with payload in the flow; and (iv) the Fully Qualified Domain Name (FQDN) the client resolved via DNS queries prior to opening the flow [12]. Important to our analysis, anonymized client IP addresses can be considered identifiers of households thanks to the static IP address allocation in place at the ISP. Due to space constraints, we do not provide here all details of the methodology we use. We invite interested readers to refer to our previous works to get a complete description [4], [10], and to contact us in order to access the available dataset.

We focus on three among the most popular personal cloud storage solutions: Dropbox, Google Drive and OneDrive. To isolate traffic generated by these applications, we build a list of FQDNs used by cloud servers. This has been achieved by means of active experiments in a controlled environment. Next, we use this list to filter the records exported by Tstat. For instance, *upload.drive.google.com* isolates the traffic sent from users to Google Drive servers, whereas **.storage.live.com* is used when contacting OneDrive. In some cases, it is possible to distinguish traffic generated by different devices used by customers to access the cloud. For instance, *dl-client*.dropbox.com* identifies traffic related to the Dropbox application running on PCs, while *api-content.dropbox.com* is used by smartphones and tablets applications.¹

In the following, we focus on the traffic observed during the whole October 2014. Table I summarizes our dataset. It shows the number of households running each service, and the total amount of exchanged traffic. Dropbox is the most popular service, with ≈ 1.9 TB transferred in total. OneDrive is as popular as Dropbox, but the amount of traffic exchanged with cloud servers is much smaller, especially considering

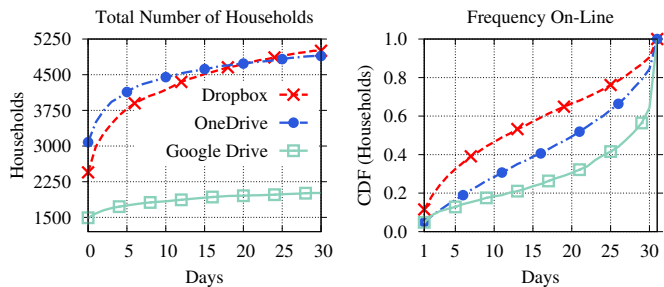


Fig. 1: Number of households using each service and the frequency they connect. Dropbox and OneDrive lead in number of households, but Google Drive users are steadily connected.

the download. Google Drive comes third, with 10% of users having used it at least once in a month. Some households use more than one cloud service. The total number of *unique* households is $\approx 7,300$ (36%). Overall, we observe 1.1 TB of uploads and 1.6 TB of downloads related to cloud storage.

III. CHARACTERIZATION OF CLOUD STORAGE USAGE

We study whether users of different cloud providers present different behaviors. We show that users exhibit distinct patterns across providers. Indeed, Dropbox population is more heterogeneous, connecting less, but exchanging more data than Google Drive users. OneDrive users, in contrast, are frequently on-line, but hardly exchanging any data.

A. Connection Frequency

We first investigate if users of Dropbox, OneDrive and Google Drive are equally active. Fig. 1 shows how frequently users access storage services. Left plot shows the cumulative population of unique households for each service. If we consider one day (leftmost point), we see about 1,500 unique households for Google Drive, 2,400 for Dropbox, and 3,000 for OneDrive. Moving to the right, the number of unique households quickly increases in the first 5–10 days, after which the growth rate gets smaller. Significant differences in population sizes are visible – e.g., Dropbox and OneDrive are used in around 25% of the 20,000 monitored households (see Table I). Compared to 2012 [10], Dropbox penetration more than doubled, but competitors are closer now – e.g., OneDrive and Google Drive were far below 1% in 2012. Interestingly, Dropbox and OneDrive curves of unique households intersect in the plot. This suggests that OneDrive users access the service using PC applications that contact the service as soon as they bootstrap the PC. For Dropbox there are instead some occasional users that access the service from other interfaces, e.g., clicking on direct web download links.

To investigate it further, we focus on how many times users access each cloud service. Right plot in Fig. 1 depicts the empirical Cumulative Distribution Function (CDF) of households with respect to the number of days they connect to the cloud provider. Distinct patterns emerge. For Dropbox, the fraction of households occasionally accessing the service is higher. For

¹See the full list at http://www.simpleweb.org/wiki/Cloud_benchmarks

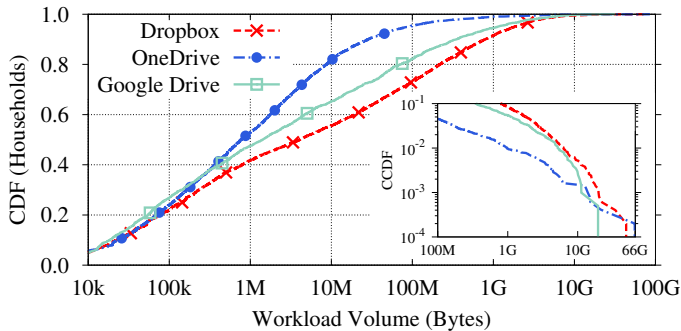


Fig. 2: Total storage workload per household. Dropbox exchanges more data when present in a household. Google Drive comes close. OneDrive seldom synchronizes large volumes.

instance, more than 12% of the households access the service for one day only, while 55% use it for 15 or less days in a month. For OneDrive, these percentages decrease to 4% and 40%, respectively, showing a more frequent access pattern, likely driven by the fact that OneDrive is integrated into recent versions of Microsoft Windows. This results in having a PC or a smartphone contacting OneDrive servers as soon as it is connected to the network. For Google Drive, users connect even more frequently: 60% of the households have accessed servers in 25 out of the 31 days, and 40% access the service on a daily basis. This is explained by the fact that Android devices automatically connect to Google Drive servers for backups.

Results allow us to conclude that users of each cloud service have specific dynamics, with Google Drive ones accessing it more often, even if Dropbox and OneDrive appear to have a higher market presence. Intuitively, these patterns would also reflect in the usage of the services, which we investigate next.

B. Data Volumes

We now compare users in terms of the workload they generate. Fig. 2 provides details of the storage workload each household exchanged in the whole month.

First, about 40–50% of the households generate very little traffic (less than 1 MB). These are *idle* users. Considering households that generate significant workload, we observe that Dropbox users clearly exchange much more data than others. Indeed, 30% of the households that use Dropbox exchange more than 100 MB in a month versus only 4% of OneDrive users. Google Drive is in between. The distribution tail shown in the Complementary CDF (CCDF) in the inset highlights that 1% of Dropbox users exchange 8 GB or more in a month.

As clearly shown, the vast majority of OneDrive users seldom exchanges any data, reinforcing the intuition that a relevant amount of OneDrive users are actually Windows users accessing the service, but not using it. Note the existence of *loyal* OneDrive users in the CCDF, where few households exchanged more than 60 GB in the whole month. Google Drive users usually exchange less traffic than Dropbox, even though they access the service almost every day.

C. Usage Scenarios

Both Dropbox and Google Drive users generally download more from the cloud than what they upload (respective ratio 1.8 and 1.9 – see Table I). OneDrive users show the opposite pattern: 2.5 times more upload volume than download. We investigate this difference more in details.

Fig. 3 provides a finer view into the storage workload by characterizing the total data stored and retrieved per household in the complete month. Each dot in the figure represents a household. By setting a threshold at 1 MB, we coarsely identify four behaviors: (i) *idle*, i.e., households with little upload and download traffic; (ii) *upload only*; (iii) *download only*; or (iv) *loyal*, i.e., household with large amounts of data for both download and upload. Households with less than 10 kB are placed on the axes. Numbers embedded in Fig. 3 report the percentage of households and bytes in each group, respectively.

Focus first on Dropbox results. We see that a large percentage (37%) of households is idle and exchanges a negligible amount of data in one month. 91% of the workload is produced by loyal users (30% in population). Interestingly, the amount of downloaders (25%) is much higher than uploaders (8%).² Thus, Dropbox users tend to download more than what they upload, possibly to synchronize content among several PCs or to retrieve content shared by other users.

Focus now on OneDrive. More than half (56%) of the households are idle, while only around 17% are loyal. The latter generates 97% of the overall volume, with few households uploading close to 100 GB, hence biasing usage toward more upload than download. We conjecture these households are using the service to perform large backups.

Finally, observe Google Drive: 45% of the households are idle. The percentage is higher than Dropbox likely because of Android users, which make little usage of Google Drive, but still connect to it on background. The fraction of loyal users (15%) is similar to OneDrive, but they exchange less data. 31% of the households fall in the download-only class. They tend to download a lot, generating 28% of total volume. While the results do not allow us to speculate about the precise reasons behind this latter usage scenario, they decisively illustrate how users’ service usage varies in practice.

IV. PERFORMANCE

We focus on performance, taking throughput as a main index, and contrasting it across both services and terminals. We observe several bottlenecks that impair performance: (i) TCP slow-start dominates short transfers; (ii) services (e.g., OneDrive) artificially limit the throughput; and (iii) protocol designs (per client terminal) have a great impact.

A. Performance of Storage Clients

Given a TCP flow, we consider the size of the application payload, and the timestamp of the first and the last data packet to calculate the flow duration. We then define the throughput as the ratio between payload size and flow duration.

²Results for Dropbox are in-line with what has been reported in [10].

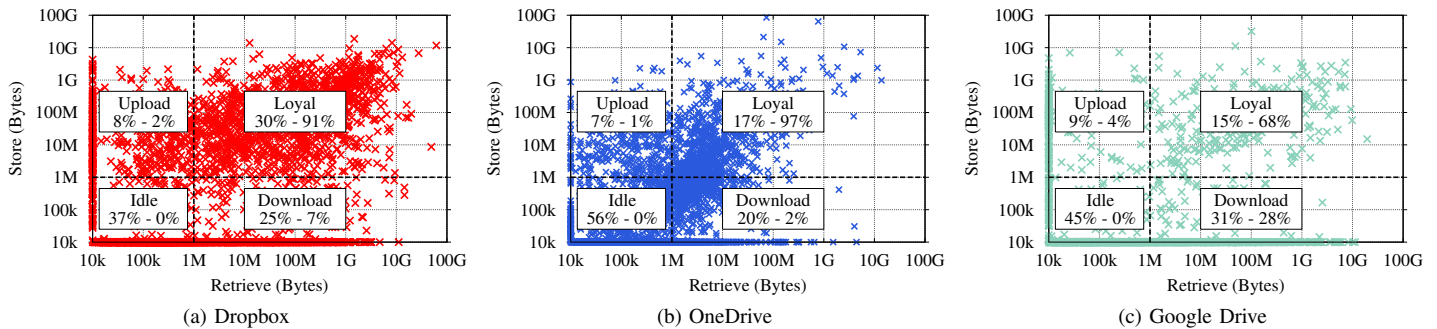


Fig. 3: Volume stored and retrieved by each household (in log axes). Households are divided in groups with limits at 1 MB. Labels report the percentage of households and bytes per group. Many idle households are seen in OneDrive and Google Drive.

Fig. 4 shows the results. For each flow, we plot throughput (y -axes) versus flow size (x -axes). Recall that the three considered services either download or upload data in a single TCP flow, without mixing operations. Red and green markers are used in the figure for download and upload, respectively. Dropbox, OneDrive, and Google Drive are shown from left to right in top plots. Bottom plots split Dropbox results per client terminal, which we will discuss in the next section. Plots are in log scale.

First, notice the linear dependence between the maximum achieved throughput and flow size for small transfers. This is due to TCP slow-start, which requires several Round Trip Times (RTT) to end. The transfer of short workloads ends before TCP can grow the congestion window, thus without saturating the path capacity. In such cases, latency dominates the throughput. Interestingly, both Dropbox and OneDrive have data-centers in the U.S. [4], and the RTT to our vantage point is higher than 100 ms. Google Drive data-center is instead much closer (RTT=15 ms). Notice how the maximum throughput for Google Drive is higher than others when committing short workloads. Yet, short transfers face the TCP bottleneck.

Second, focus on large transfers (≥ 1 MB for uploads and ≥ 10 MB for downloads). Here TCP reaches steady state and can saturate the available capacity. The upload rate limits of 1 Mb/s (ADSL customers) or 10 Mb/s (FTTH customers) are clearly visible. For downloads, both Dropbox and Google Drive allow customers to saturate their downlink capacity at ≈ 20 Mb/s (100 Mb/s) for ADSL (FTTH) customers in some occasions. On the contrary, OneDrive artificially forces the download throughput to less than 10 Mb/s.

At last, Dropbox and Google Drive performance is affected by complex client policies. For example, large files are exchanged in chunks not exceeding 16 MB each by Google Drive at PCs [4]. Chunks are sent in separate flows. Note the drop in the number of flows after this threshold.³ This policy results in a strongly limited download rate for files larger than the chunking threshold.

³Flows larger than the threshold are related to web and mobile clients.

B. Performance per Terminal Type

Bottom plots in Fig. 4 depict the performance of Dropbox according to the type of device used to access the service. We split the data shown in Fig. 4a into three categories: PC, mobile and web. The latter includes both the Dropbox main web interface and direct links sent among users to share files.

Two major considerations hold. First, focusing on Dropbox at PCs (Fig. 4d), no single flow is able to saturate the FTTH line capacity, especially for downloads. We notice a clear correlation between flow size and the achieved performance. Recall that Dropbox splits large files into chunks of 4 MB. When files larger than 4 MB are to be transferred, the PC client slices them into chunks and may send chunks over independent flows [4]. When the flow size exceeds this threshold, the larger the flow is, the lower is the maximum achieved throughput. This can be explained by a combination of factors, including the complex operations performed by the client to optimize network usage (e.g., delta encoding), the delay introduced by commit operations when dealing with the several chunks that compose large file sets, etc. At last, flows are limited to around 400 MB by design, which penalizes the transfer of workloads larger than that.

Second, notice in Figs. 4e and 4f how the bottlenecks identified at the PC client are no longer present in mobile and web accesses. We see in particular that the web interface is the only one able to take full advantage of the high-speed lines at households' disposal. This is explained by the mechanisms adopted to download large files from the web interface. Clients perform a simple HTTPS download that reaches up to 100 Mb/s on FTTH connections. Note also the difference in the amount of downloads and uploads among the client terminals, better detailed later.

In conclusion, although personal cloud storage presents high network demands, artificial shaping policies, client design choices and terminal limits can affect performance.

V. USAGE PER CLIENT TERMINAL

We focus on Dropbox since we identify a significant number of households using the service from different terminals. Our

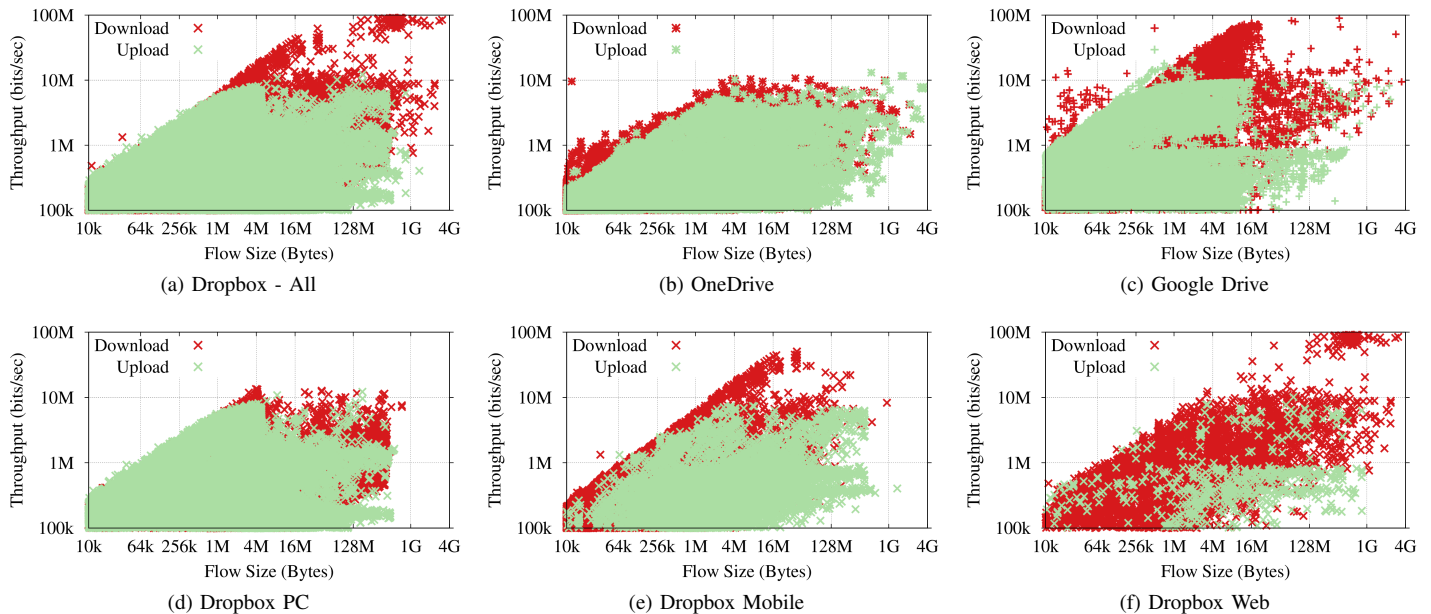


Fig. 4: Throughput for different services and client terminals. Dropbox can saturate the line capacity only with mobile and web clients. OneDrive shapes throughput. Complex client policies limit Google Drive throughput for large workloads.

findings suggest that usage patterns change considerably according to the type of terminal. Uploads from mobile terminals are common, and usually larger than those faced by PC clients. On the other hand, downloads are rare and small on mobile devices. At last, the web interface is hardly used to upload content, being instead preferred for downloads.

A. Connection Frequency and Volume

We evaluate how often households using Dropbox connect from each type of client, and how much data they exchange. We identify 2,196, 1,628 and 3,832 households using PC, mobile and web, respectively. For web, 3,581 households use it to consume direct links people send each other to share files.

The left plot of Fig. 5 depicts the CDF of the number of days in which households interacting with each type of client transfer data. Focusing on the leftmost point in the figure, note how 40% of the households using the web interface connect only once in a full month. This is expected given the direct link usage scenario. Some of these households do not even have Dropbox installed on PCs. Still, they generate workload when receiving web links from others. Similarly, 30% of the households with mobile clients connect only once in a month. Compare it to PC clients, in which 90% (60%) of the households exchange data in more than 1 day (10 days).

To explore this further, the right plot in Fig. 5 depicts the CDF of households with respect to the volume they exchange for different terminals. We see that the web interface creates the lowest amount of traffic. Only 20% of the households transfers more than 10 MB in a month. On the other extreme, PC clients generate more traffic than others – e.g., 70%

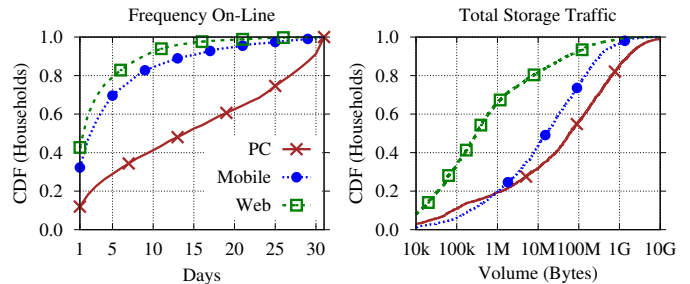


Fig. 5: CDF of households active with each type of client, and their total traffic. Web and mobile clients are rarely active. However, mobile clients produce lots of traffic.

exchanges more than 10 MB. Much more interesting is the behavior for mobile terminals. Although households rarely use Dropbox from mobile terminals (see left-hand plot), the amount of traffic from such clients is noteworthy: 20% (50%) of households exchanges more than 100 MB (10 MB). We conclude that usage in PC, mobile, and web varies greatly, and we provide in the following additional details for each scenario.

B. Workloads

Overall, users uploaded 462 GB, 180 GB and 38 GB from PC, mobile and web, respectively. Here we highlight a major change in cloud storage usage habits. Whereas upload in mobile was marginal in 2012 [10], it is now equivalent to 37% of PC upload volume. In terms of downloads, PCs

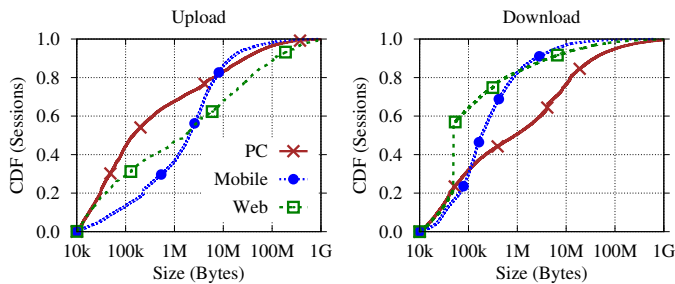


Fig. 6: Bytes per download or upload session. Most PC uploads are short, while downloads are the largest ones. Mobile uploads instead are large, while mobile downloads are rare and short.

still dominate, peaking at 980 GB, while only 51 GB were downloaded on mobile terminals, and 177 GB from the web during the entire month.

We dissect these figures by evaluating the size of workloads people create *at once*. We again split flows into upload or download. Flows on PCs are however biased toward 4 MB due to the chunking mechanism Dropbox adopts. We overcome the bias by evaluating storage *sessions*: Whenever upload or download flows in the same household overlap in time, we merge them into a single storage session.⁴ We then analyze the size of storage sessions per device type.

Fig. 6 depicts the size of upload (left-hand side) and download sessions (right-hand side). Note how most of them on PCs are very small, 70% carrying less than 1 MB. This behavior has been associated [10] to the dominance of small files as well as to the capabilities Dropbox implements to optimize network usage, such as compression and delta encoding. Compare now with the line depicting mobile sessions. Mobile uploads are concentrated between 1 MB and 10 MB. We conjecture this behavior is associated to the usage of cloud storage to store pictures and videos captured with mobile devices. Web upload does not implement smart policies and thus workloads grow smoothly with no clear bias toward specific sizes.

Comparing left and right plots, we see that downloads are larger than uploads on PCs. Again, this behavior has been associated to the large workloads that are spread to many devices in a household. In particular, at boot time PCs often have to download full batches of files, modified or added elsewhere. In contrast, uploads are sent as soon as changes occur. Mobile downloads are much rarer and smaller than uploads. This is explained by the fact that mobile clients do not synchronize content unless users explicitly request it. The CDF for web is biased toward 50 kB, which is the typical size of picture previews on the Dropbox website.

All in all, we conclude mobile terminals are mostly content producers, web interfaces are often used to fetch content via direct links, whereas PCs dominate personal cloud storage in terms of data volumes for file synchronization.

⁴Flows from different devices in a household could be merged. Manually evaluating Dropbox control traffic in the PoP, which is unencrypted and contains a device identifier, we estimate that our methodology merges flows of different devices in less than 5% of the storage sessions.

VI. CONCLUSIONS

This paper evaluated personal cloud storage using passive measurements collected in an operational network. Whereas Dropbox in PCs used to dominate the market, we showed that the landscape is changing. Usage of cloud storage is increasing both in competitors, such as OneDrive and Google Drive, as well as in mobile terminals.

Our study of cloud storage usage and performance yielded new insights. For example, we observed that usage across providers is distinct, with high numbers of OneDrive and Google Drive users that are active without creating much workload. This is a consequence of the close integration of cloud storage into modern OSs. Performance bottlenecks were highlighted, including explicit traffic shaping policies and the impact of well-known TCP bottlenecks on small workloads. Finally, we noticed a large increase in traffic from mobile terminals, which seems to be related to multimedia content produced by tablets and smartphones.

While measurements in this work are tied to the vantage point we use, we believe they provide valuable information into overall trends and are of interest to understand and track the evolution of personal cloud storage systems. In our ongoing efforts, we plan to extend this work by providing to the research community synthetic models able to produce similar profiles and usage patterns.

REFERENCES

- [1] W. Hu, T. Yang, and J. N. Matthews, "The Good, the Bad and the Ugly of Consumer Cloud Storage," *SIGOPS Oper. Syst. Rev.*, vol. 44, no. 3, pp. 110–115, 2010.
- [2] H. Wang, R. Shea, F. Wang, and J. Liu, "On the Impact of Virtualization on Dropbox-Like Cloud File Storage/Synchronization Services," in *Proceedings of the IWQoS*, 2012, pp. 11:1–11:9.
- [3] R. Gracia-Tinedo, M. S. Artigas, A. Moreno-Martinez, C. Cotes, and P. G. Lopez, "Actively Measuring Personal Cloud Storage," in *Proceedings of the CLOUD*, 2013, pp. 301–308.
- [4] E. Bocchi, I. Drago, and M. Mellia, "Personal Cloud Storage Benchmarks and Comparison," *IEEE Trans. Cloud Comput.*, vol. PP, no. 99, pp. 1–14, 2015.
- [5] Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai, and Z.-L. Zhang, "Towards Network-Level Efficiency for Cloud Storage Services," in *Proceedings of the IMC*, 2014, pp. 115–128.
- [6] G. Gonçalves, I. Drago, A. P. C. da Silva, A. B. Vieira, and J. M. de Almeida, "Modeling the Dropbox Client Behavior," in *Proceedings of the ICC*, 2014, pp. 1332–1337.
- [7] P. Casas and R. Schatz, "Quality of Experience in Cloud Services: Survey and Measurements," *Comput. Netw.*, vol. 68, pp. 149–165, 2014.
- [8] S. Liu, X. Huang, H. Fu, and G. Yang, "Understanding Data Characteristics and Access Patterns in a Cloud Storage System," in *Proceedings of the CCGrid*, 2013, pp. 327–334.
- [9] M. Naldi and L. Mastroeni, "Cloud Storage Pricing: A Comparison of Current Practices," in *Proceedings of the HotTopiCS*, 2013, pp. 27–34.
- [10] I. Drago, M. Mellia, M. M. Munafò, A. Sperotto, R. Sadre, and A. Pras, "Inside Dropbox: Understanding Personal Cloud Storage Services," in *Proceedings of the IMC*, 2012, pp. 481–494.
- [11] A. Finamore, M. Mellia, M. Meo, M. M. Munafò, and D. Rossi, "Experiences of Internet Traffic Monitoring with Tstat," *IEEE Netw.*, vol. 25, no. 3, pp. 8–14, 2011.
- [12] I. Bermudez, M. Mellia, M. M. Munafò, R. Keralapura, and A. Nucci, "DNS to the Rescue: Discerning Content and Services in a Tangled Web," in *Proceedings of the IMC*, 2012, pp. 413–426.