# Actively Measuring Personal Cloud Storage

Raúl Gracia-Tinedo, Marc Sánchez Artigas, Adrián Moreno-Martínez, Cristian Cotes and Pedro García López
Universitat Rovira i Virgili, Tarragona (Spain)
{raul.gracia|marc.sanchez|adrian.moreno|cristian.cotes|pedro.garcia}@urv.cat

***Abstract*—The Personal Cloud model is a mainstream service that meets the growing demand of millions of users for reliable off-site storage. However, despite their broad adoption, very little is known about the quality of service (QoS) of Personal Clouds.**

**In this paper, we present a *measurement study* of three major Personal Clouds: DropBox, Box and SugarSync. Actively accessing to free accounts through their REST APIs, we analyzed important aspects to characterize their QoS, such as transfer speed, variability and failure rate. Our measurement, conducted during two months, is the first to deeply analyze many facets of these popular services and reveals new insights, such as important performance differences among providers, the existence of transfer speed daily patterns or sudden service breakdowns.**

**We believe that the present analysis of Personal Clouds is of interest to researchers and developers with diverse concerns about Cloud storage, since our observations can help them to understand and characterize the nature of these services.**

***Index Terms*—Cloud Storage; Personal Clouds; Quality of Service (QoS); Measurement**

## I. Introduction

The Personal Cloud model defines a ubiquitous storage facility enabling the unified and location-agnostic access to information flows from any device and application. Commercial providers such as Dropbox, SugarSync or Google Drive are offering very popular Personal Cloud solutions that maintain in sync the information from different devices. The popularity of these killer applications lies behind their easy-to-use SaaS storage facade to ubiquitous IaaS Cloud storage resources like Amazon S3 and others. In a recent report, Forrester research [1] forecasts a market of $12 billion in the US in paid subscriptions to Personal Clouds by 2016.

However, despite their broad adoption, very little is known about the QoS of Personal Clouds. Furthermore, there is no public information about the control policies that Personal Clouds enforce, as well as the factors impacting on their service performance. In our view, exploring these services is specially interesting in the case of *free accounts*. Recent reports show that most Personal Cloud users are *freemium* users. For example, *from the* 50 *million of DropBox users, only the* 4% *pay for storage* [2].

In this paper, we present a measurement study of various Personal Clouds. Concretely, during two months, we have actively measured the REST API service of DropBox, Box and SugarSync *free accounts*. We gathered information from more than 900,000 storage operations, transferring around 70TB of data. We analyzed important aspects to characterize their QoS, such as in/out transfer speed, service variability and failure rate. To our knowledge, this work is the first to deeply explore many facets of these popular services and reveals new insights.

We contribute all of our research observations, including:

- The transfer performance of these services *greatly varies from one provider to another*, which is a valuable piece of information for designers and developers.
- The *geographic location* of a client importantly impacts on the speed of transfers. For instance, North American clients experience transfers several times faster than European ones for the same Personal Cloud.
- In general, transfer speeds of files can be *approximated using well-known statistical distributions*. This opens the door to create Personal Cloud simulation environments.
- The *variability* of transfers depends on several factors, such as the *traffic type (in/out) or the hour of the day*. Actually, we found daily patterns in the DropBox service.
- These services are in general *reliable* and, in some cases, service failures can be modeled as a Poisson process.
- We observed a radical change in the transfer speed of SugarSync in late May 2012. This suggests that Personal Clouds *may change their freemium QoS unexpectedly*, due to internal policy changes or agreements.

After the analysis, we discuss the most important technical findings of this work and their implications. Finally, we contribute the collected data set and we make it publicly available for the research community[1].

The rest of the paper is organized as follows. We discuss the related work in Section II. Our methodology is described in Section III. The measurement data analysis appears in Section IV. Section V summarizes the most important technical findings of this work and we conclude in Section VI.

## II. Related Work

The performance evaluation of Cloud services is a current hot topic with several papers appearing recently [3], [4]. A large corpus of works have focused on measuring distinct aspects of a Cloud service, such as computing performance [5], service variability [3], and the nature of datacenter traffic [4], among others issues.

Unfortunately, only few works have turned attention to measure the performance of Cloud storage services. To wit, the authors in [6] explore the performance of Microsoft Azure, including storage. In this line, the authors in [7] execute an extensive measurement against Amazon S3 to elucidate whether Cloud storage is suitable for scientific Grids or not. Similarly, [8] presents a performance analysis of the Amazon Web Services, with no insights regarding Personal Clouds.

File hosting and file sharing Cloud services have been analyzed in depth by several works [9], [10]. They provide an interesting substrate to understand both the behavior of users and the QoS of major providers (e.g. RapidShare).

The comparison of public Clouds have recently aroused much interest. Concretely, the authors in [11] present CloudCmp, a systematic performance and cost comparator of Cloud providers. The authors validate CloudCmp by measuring the

[1] Available at http://ast-deim.urv.cat/trac/pc_measurement

elastic computing, storage, and networking services offered by several public Clouds based on metrics reflecting the impact on the performance delivered to customer applications.

In a similar fashion, the authors of [12] compare Dropbox, Mozy, Carbonite and CrashPlan backup services. However, their analysis of the performance, reliability and security levels is rather lightweight, and more work is needed to characterize these services with enough rigor.

Recently, the authors in [13] presented an extensive measurement of DropBox in two scenarios: in a university campus and in residential networks. They analyzed and characterized the traffic transmitted by users, as well as the functioning and architecture of the service. Instead of analyzing the behavior of users using a specific Personal Cloud, we focused on characterizing the service that many providers offer.

In contrast to previous works, we analyze in depth many aspects of Personal Clouds that remain unknown (e.g. variability, failures). We believe that this analysis will help researchers and developers to better understand these popular services.

## III. Measurement Methodology

From May 10, 2012, to July 15, 2012, we installed several vantage points in our university network (*Universitat Rovira i Virgili*, Spain) and PlanetLab [14] to measure the performance of three of the major Personal Cloud services in the market: DropBox[2], Box[3] and SugarSync[4]. The measurement methodology was based on the REST interfaces that these three Personal Cloud storage services provide to developers.

Personal Clouds provide REST APIs, along with their client implementations, to make it possible for developers to create novel applications. These APIs incorporate authorization mechanisms (OAuth [15]) to manage the credentials and tokens that grant access to the files stored in user accounts. A developer first registers an application in the Cloud provider website and obtains several tokens. As a result of this process, and once the user has authorized that application to access his storage space, the Personal Cloud storage service gives to the developer an *access token*. Including this *access token* in each API call, the application can operate on the user data.

There are two types of API calls: *meta-info* and *data management* calls. The former type refers to those calls that retrieve information about the state of the account (i.e., storage load, filenames), whereas the latter are those calls targeted at managing the stored files in the account. We will analyze the performance of the most important data management calls: `PUT` and `GET`, which serve to store and retrieve files.

### A. Measurement Platform

We employed two different platforms to execute our tests: University laboratories and PlanetLab. The reason behind this is that our labs contain *homogeneous and dedicated machines* that are under our control, while PlanetLab allows the analysis of each service from *different geographic locations*.

*University laboratories*. We gathered 30 machines belonging to the same laboratory to perform the measurement. These machines were Intel Core2 Duo equipped with 4GB DDR2 RAM. The employed operating system was a Debian Linux distribution. Machines were internally connected to the same switch via a 100Mbps Ethernet links.

| Location | Op. Type | Operations | Transferred Data |
|---|---|---|---|
| University Labs | `GET` | $168,396$ | 13.509 TB |
| | `PUT` | $247,210$ | 15.945 TB |
| PlanetLab | `GET` | $354,909$ | 31.751 TB |
| | `PUT` | $129,716$ | 9.803 TB |

TABLE I
Summary of Measurement Data (May 10 – July 15)

*PlanetLab*: We collected 40 PlanetLab nodes divided into two geographic regions: Western Europe and North America. This platform is constituted by heterogeneous (bandwidth, CPU) machines from several universities and research institutes. Moreover, there were two points to consider when analyzing data coming from PlanetLab nodes: i) Machines might be concurrently used by other processes and users, and ii) The quota system of these machines limited the amount of in/out data transferred daily.

Specifically, we used the PlanetLab infrastructure for a high-level assessment of Personal Clouds depending on the client's geographic location. However, the mechanisms to enforce bandwidth quotas in PlanetLab nodes may induce the appearance of artifacts in bandwidth traces. This made PlanetLab not suitable for a fine-grained analysis in our context.

### B. Workload Model

Usually, Personal Cloud services impose file size limitations to their REST interfaces, for we used only files of four sizes to facilitate comparison: 25MB, 50MB, 100MB and 150MB[5]. This approach provides an appropriate substrate to compare all providers with a large amount of samples of equal-size files. Thanks to this, we could observe performance variations of a single provider managing files of the same size.

We executed the following workloads:

*Up/Down Workload*. The objective of this workload was twofold: Measuring the maximum up/down transfer speed of operations and detecting correlations between the transfer speed and the load of an account. Intuitively, the first objective was achieved by alternating upload and download operations, since the provider only needed to handle one operation per account at a time. We achieved the second point by acquiring information about the load of an account in each API call.

The execution of this workload was continuously performed at each node as follows: First, a node created synthetic files of a size chosen at random from the aforementioned set of sizes. That node uploaded files until the capacity of the account was full. At this point, that node downloaded all the files also in random order. After each download, the file was deleted.

*Service Variability Workload*. This workload maintained in every node a nearly continuous upload and download transfer flow to analyze the performance variability of the service over time. This workload provides an appropriate substrate to elaborate a time-series analysis of these services.

The procedure was as follows: The upload process first created files corresponding to each defined file size which

[2]http://www.dropbox.com
[3]http://www.box.net
[4]http://www.sugarsync.com

[5]Although the official limitation in some cases is fixed to 300MB per file, we empirically proved that uploading files larger than 200MB is highly difficult. In case of Box this limitation is 100MB.

were labeled as “reserved”, since they were not deleted from the account. By doing this we assured that the download process was never interrupted, since at least the reserved files were always ready for being downloaded. Then, the upload process started uploading synthetic random files until the account was full. When the account was full, this process deleted all files with the exception of the reserved ones to continue uploading files. In parallel, the download process was continuously downloading random files stored in the account.

Finally, we executed the experiments in different ways depending on the chosen platform. In the case of PlanetLab, we employed the *same machines in each test*, and therefore, we needed to sequentially execute all the combinations of workloads and providers. This minimized the impact of hardware and network heterogeneity, since all the experiments were executed in the same conditions. On the contrary, in our labs we executed in parallel a certain workload for all providers (i.e. assigning $10$ machines per provider). This provided two main advantages: The measurement process was substantially faster, and fair comparison of the three services was possible for the same period of time.

We depict in Table I the total number of storage operations performed during the measurement period.

### C. Setup, Software and Data Collection

Prior to the start of our experiments, we created around $150$ new user free accounts from the targeted Personal Clouds. That is $120$ new accounts for PlanetLab experiments ($40$ nodes $\times$ $3$ Personal Clouds), and $30$ accounts for the experiments in our labs ($10$ accounts per Personal Cloud deployed in $30$ machines). We also registered as developers 35 applications to access the storage space of user accounts via REST APIs, obtaining the necessary tokens to authenticate requests. We assigned to every node a single new free account with access permission to the corresponding application. The information of these accounts was stored in a database hosted in our research servers. Thus, nodes executing the measurement process were able to access the account information remotely.

Measurement processes were implemented as Unix and Python scripts that ran in every node. These scripts employed third party tools during their execution. For instance, to synchronize tasks, such as logging and starting/finishing experiments, we used the `cron` time-based job scheduler. To gather bandwidth information we used `vnstat`, a tool that keeps a log of network traffic for a selected interface. Nodes performed storage operations against Personal Clouds thanks to the API implementations released in their websites.

The measurement information collected in each storage operation was sent periodically from every node to a database hosted in our research servers. This automatic process facilitated the posterior data processing and exploration. The measurement information that nodes sent to the database describes several aspects of the service performance: operation type, bandwidth trace, file size, start/end time instants, time zone, capacity and load of the account, and failure information.

## IV. Measuring Personal Cloud REST APIs

### A. Transfer Capacity of Personal Clouds

In this section, the transfer capacity of Box, DropBox and SugarSync is characterized using the following indicators:

- *File Mean Transfer Speed (MTS)*. This metric is defined as *the ratio of the size of a file, S, to the time, T, that was spent to transfer it:* $MTS = S/T$ *(KBytes/sec)*.
- *Bandwidth Distributions*. We define as a *bandwidth trace* the set of values that reflects the transfer speed of a file at regular intervals of 2 secs. To obtain a single empirical distribution, we aggregated the bandwidth traces of all the transfers separated by uploads and downloads. We refer to the resulting empirical distribution as the *aggregated bandwidth distribution*.

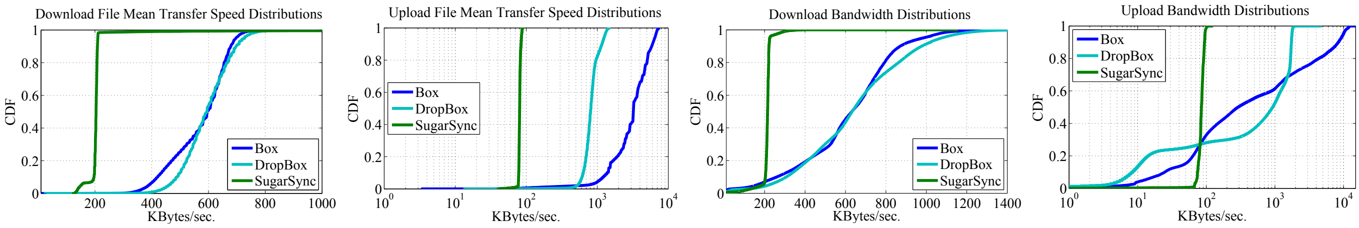
**Transfer speeds**. Fig. 1 reports these metrics for both workloads (up/down and service variability) executed in our university labs during 10 days. First, Fig. 1 evidences an interesting fact: *Personal Clouds are heterogeneous in terms of transfer speed*. For instance, Fig. 1b shows that Box and DropBox present an upload MTS several times faster than SugarSync. The same observation holds for downloads. Moreover, the heterogeneity of these services also depends on the *traffic type* (in/out). This can be appreciated by comparing Fig. 1a with Fig. 1b: *DropBox exhibits the best download MTS while Box presents the fastest uploads.*

This proves that *the transfer performance of these services greatly varies among providers*, and consequently, developers should be aware of this in order to select an adequate provider.

Among the examined Personal Clouds, DropBox and SugarSync are *resellers* of major Cloud storage providers (Amazon S3 and Carpathia Hosting, respectively). On the other hand, Box claims to be owner of several datacenters. In our view, it is interesting to analyze this Cloud ecosystem and the possible implications to the service delivered to end-users.

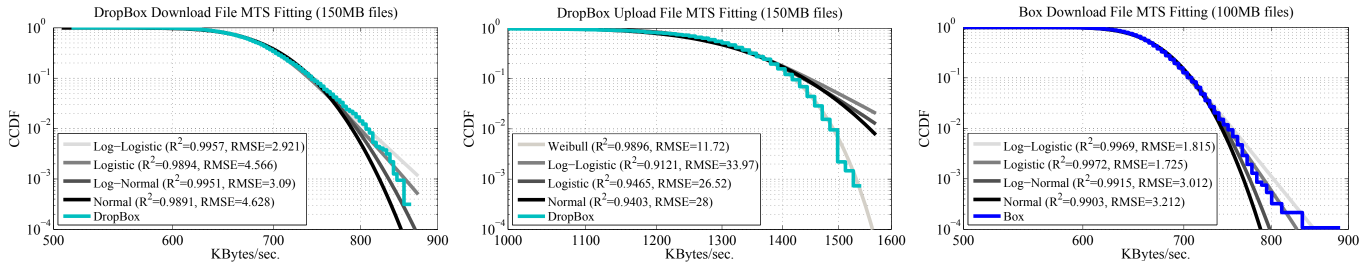
In this sense, in Fig. 1 we observe that Personal Clouds apply *distinct internal control policies* to the inbound/outbound bandwidth provided to users. To wit, both DropBox and Box exhibit an *upload transfer capacity remarkably better than the download capacity*. This means that the datacenter outgoing traffic is more *controlled and restricted* than the incoming traffic. This agrees well with the current pricing policies of major Cloud providers (Amazon S3, Google Storage) which do not charge inbound traffic whereas the outbound traffic is subject to specific rates (see http://aws.amazon.com/en/s3/pricing/).

In SugarSync, both the upload and download transfer speeds are constant and low. Interestingly, SugarSync presents slightly faster downloads than uploads, though only a small fraction of downloads (less than $1\%$) exhibits a much higher transfer speed than the rest. These observations are also supported by Fig. 1c and Fig. 1d: the captured download bandwidth values fall into a small range $[200, 1300]$ KB/sec. Also, the shape of these distributions are not steep, which reflects that there is a strong control in the download bandwidth. On the contrary, upload bandwidth distributions present more irregular shapes and they cover a wider range of values, specially for Box. As a possible explanation to this behavior, the experiments of Fig. 1 were executed from our university labs (Spain) to exclude the impact of geographic heterogeneity. Considering the fact that the majority of Personal Cloud datacenters are located in USA [13], this may have implications in the cost of the traffic sent to Europe. This could motivate the enforcement of more restrictive bandwidth control policies to the outbound traffic.




(a) Download file MTS distributions. (b) Upload file MTS distributions. (c) Download bandwidth distributions. (d) Upload bandwidth distributions.

Fig. 1. Transfer capacity of Box, DropBox and SugarSync free account REST API services. The data represented in these figures corresponds to the aggregation of the up/down and service variability workloads during 10 days (June/July 2012) in our university laboratories.




(a) DropBox down. file MTS fitting (150MB files). (b) DropBox up. file MTS fitting (150MB files). (c) Box down. file MTS fitting (100MB files).

Fig. 2. Distribution fittings of upload/download file mean transfer speeds (MTS) of the examined Personal Clouds (up/down workload, university labs).

**Characterization of transfers**. To characterize the transfer performance of both DropBox and Box (the constant behavior of SugarSync deserved no further analysis), three checks were made to determine the shape of the transfer distributions with sufficient confidence. We used the same methodology of [16].

First, visual inspection of per-file MTS distributions against the most similar standard distributions was performed. Second, we performed a linear regression analysis on the best-fit lines of the quantile-quantile plots from the fitted distributions and empirical data. From this analasys, we obtained the coefficient of determination, $R^2 \in [0, 1]$. A value of $R^2$ close to 1 signals that the candidate distribution fits the data. Finally, we used the Kolmogorov-Smirnov (KS) test to assess the statistical validity of the fittings. Essentially, this test is used to check whether a fitted distribution matches the empirical distribution by finding the maximum differences between both distributions[6].

As seen in Fig. 2a and 2c, both DropBox and Box download file MTS *can be approximated using log-logistic or logistic distributions, respectively*. This argument is supported by the coefficient of determination, $R^2$, which in the case of Box is $R^2 = 0.9972$, and for DropBox is $R^2 = 0.9957$. However, we observe that these fittings differ from the empirical data in the tails of highest transfer speed values. Further, we performed fittings depending on the file size, obtaining closer fittings as the file size grew. The heavier tails found in empirical data but not captured well in the fittings led the KS test to reject the null hypothesis at significance level $\alpha = 0.05$, although in the case of DropBox, this rejection is borderline (KS-test=0.0269, critical value=0.0240, $p$-value=0.197).

Regarding uploads, we find that DropBox file MTS *can be modeled by a Weibull distribution* with shape parameter $\mu = 1339.827$ and scale parameter $\sigma = 14.379$ (Fig. 2b). In addition to the high $R^2 = 0.9896$, the KS test *accepted the null hypothesis* at significance level $\alpha = 0.05$ (KS-test=0.0351, critical value=0.0367, $p$-value=0.0025).

Due to the high variability, we found that Box uploads do not follow any standard distribution. The implications of these observations are relevant. With this knowledge, researchers can model the transfer speed of Personal Cloud services employing specific statistical distributions.

[6]Chi-square test was not used since it works well only when the number of items that falls into any particular bin is approximately the same. However, it is relatively difficult to determine the correct bin widths in advance for different measured data sets, and thus the results of the this test can vary depending on how the data samples are divided [16].

**Transfers & geographic location**. Next, we analyze transfer speeds depending on the geographic location of vantage points.

In Fig. 3, we illustrate the file MTS obtained from executing the up/down workload during 3 weeks in PlanetLab. As can be seen in the figure, *Personal Clouds provide a much greater QoS in North American countries than in European countries*. Intuitively, the location of the datacenter plays a critical role in the performance of the service delivered to users. Observe that this phenomenon is orthogonal to all the examined vendors.

Finally, we quantify the relative download/upload transfer performance delivered by each service as a function of the geographic location of users. To this end, we used a simple metric, what we call the *download/upload ratio* ($D/U$), which is the result of dividing the download and upload transfer speeds of a certain vendor. In Table II, we calculated this ratio over the mean ($\bar{U}, \bar{D}$) and median ($\tilde{U}, \tilde{D}$) values of the file MTS distributions of each provider depending on the geographic location of nodes.

In line with results obtained in our university labs, *European nodes receive a much higher transfer speed when uploading than when downloading* ($D/U < 1$). However, contrary to conventional wisdom, *North American nodes exhibit just the opposite behavior*. This is clearly visible in DropBox and Box. However, this ratio is constant in SugarSync, irrespective of the geographic location.

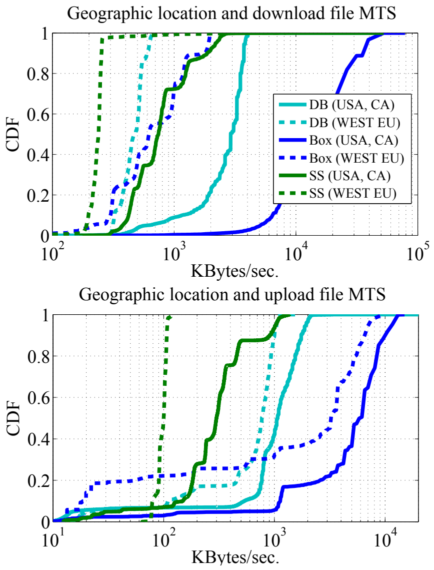



Fig. 3. File MTS distributions of PlanetLab nodes from June 22 to July 15 2012 depending on their geographic location (up/down workload). Clearly, USA and Canada nodes exhibit faster transfers than European nodes.

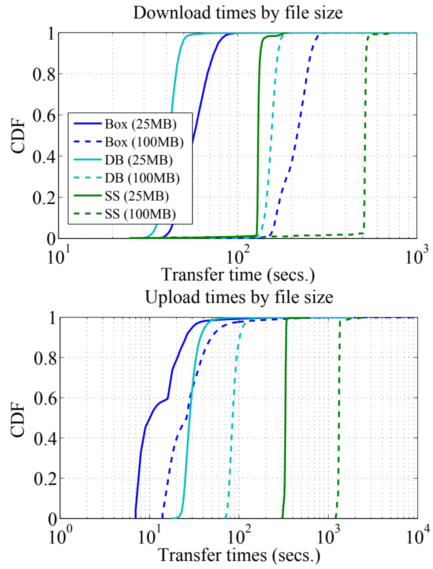



Fig. 4. Transfer times distributions by file size.

| Geo. Location | *Metric* | Box | DropBox | SugarSync |
|---|---|---|---|---|
| USA & CA | $D_{MTS}/U_{MTS}$ | 3.198 | 2.482 | 2.522 |
| | $\tilde{D}_{MTS}/\tilde{U}_{MTS}$ | 2.550 | 2.722 | 2.500 |
| WEST EU | $D_{MTS}/U_{MTS}$ | 0.255 | 0.681 | 2.589 |
| | $\tilde{D}_{MTS}/\tilde{U}_{MTS}$ | 0.190 | 0.682 | 2.387 |

TABLE II
DOWNLOAD/UPLOAD TRANSFER SPEED RATIO OF PERSONAL CLOUDS DEPENDING ON THE CLIENT'S GEOGRAPHIC LOCATION.

### *B. Variability of Transfer Performance*

In this section, we analyze which factors can contribute to the variance in transfer speed observed in Personal Clouds. We study three potential factors, which are *the size of file transfers*; *the load of accounts*; and *time-of-day effects*.

**Variability over file size**. We first investigate the role that file size plays on *transfer times* and *transfer speeds*. Fig. 4 and Table III report the results for both metrics as function of file size, respectively. Unless otherwise stated, results reported in this subsection are based on executing the up/down workload in our university labs during 5 days.

Fig. 4 plots the transfer time distribution for all the evaluated Personal Clouds. As shown in the figure, for the same provider, all the distributions present a similar shape, which suggests that *the size of file transfers is not a source of variability*. As expected, the only difference is that the distributions for large file sizes are shifted to the right towards longer time values. Significant or abnormal differences were not observed when transferring large files compared to small data files. This observation is applicable to all evaluated Personal Clouds. This leads us to the conclusion that these Personal Clouds *do not perform aggressive bandwidth throttling policies to large files*.

An interesting fact appreciable in Table III is that *managing larger files report better transfer speeds than in case of small files*. Usually, these improvements are slight or moderate (0.5% to 25% higher MTS); however, uploading 100MB files to Box exhibits a MTS 48% higher than uploading 25MB files to this service. In our view, this phenomena is due to the variability in the incoming bandwidth supplied by Box, and the TCP slow start mechanism, which makes difficult for small file transfers to attain high performance [17].

Further, we found that all the measured Cloud vendors tend to perform a more restrictive bandwidth control to outgoing traffic. This can be easily confirmed by inspecting the obtained standard deviations $\sigma$ of file MTS listed in Table III. Clearly, *the inbound traffic in Dropbox and Box is much more variable than the outbound traffic*. On the contrary, despite its limited capacity, the source of highest transfer variability in SugarSync is in the outbound traffic, which a clear proof of the existing heterogeneity in Personal Clouds.

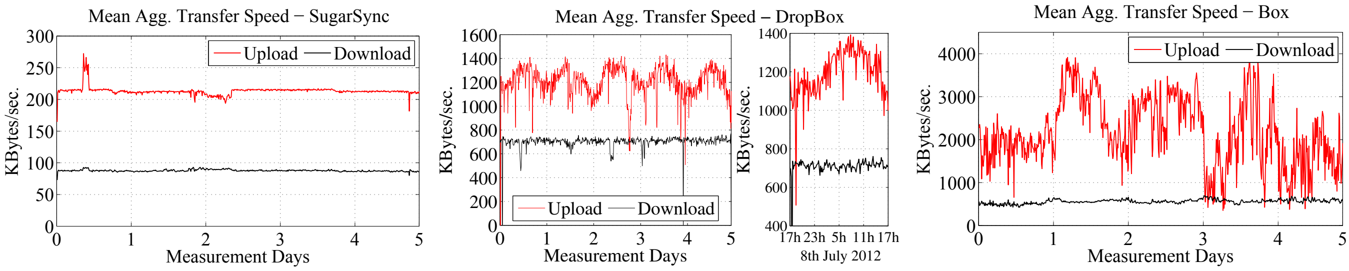
**Variability over load account**. Next we explore if Personal Clouds perform any control to the transfer speed supplied to users based on the amount of data that users have in their accounts. To reveal any existing correlation, dispersion graphs were utilized to plot the relationship between the MTS and the load of an account at the instant of the storage operation.

As shown in Fig. 6, *we were unable to find any correlation between the file MTS and the load of an account* in any of the measured Personal Clouds. This suggests that the transfer speed delivered to users remains the same irrespective of the current amount of data stored in an account. This conclusion is important to characterize which types of control mechanisms are actually applied to these storage services.

**Variability over time**. We now analyze how the transfer speed varies over time. To better capture these variations, we used the data from the *service variability workload*, which was aimed to maintain a constant transfer flow and was executed at our university labs. The results are shown in Fig. 5 where the *mean aggregated bandwidth* of all nodes as a whole is plotted in time intervals of 600 seconds. As expected, we found that the transfer speed of these services behave differently depending on the provider. To wit, while SugarSync exhibits a *stable service for both uploads and downloads*, at the price of a modest transfer capacity (Fig. 5a), the upload transfer speed varies significantly over time for Dropbox and Box.

| | | Upload File MTS Distribution (KBps) | | | | | | | | Download File MTS Distribution (KBps) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Size | Provider | Min. | Q1 | Median | Q3 | Max. | Mean ($\mu$) | Std. Dev. ($\sigma$) | CV ($\sigma/\mu$) | Min. | Q1 | Median | Q3 | Max. | Mean ($\mu$) | Std. Dev. ($\sigma$) | CV ($\sigma/\mu$) |
| 25MB | DropBox | 13.54 | 819.20 | 903.94 | 1008.24 | 1456.36 | 896.28 | 151.56 | 0.1691 | 24.89 | 582.54 | 624.152 | 672.16 | 970.90 | 626.94 | 71.23 | 0.1136 |
| | Box | 14.70 | 1379.71 | 2383.13 | 3276.80 | 3744.91 | 2271, 29 | 973.06 | 0.3963 | 163.84 | 397.19 | 459.90 | 534.99 | 794.38 | 463.72 | 87.76 | 0.0837 |
| | SugarSync | 41.87 | 78.25 | 78.96 | 80.17 | 86.23 | 79.26 | 2.82 | 0.0356 | 136.53 | 198.59 | 200.11 | 201.65 | 1048.57 | 201.35 | 37.89 | 0.1882 |
| 50MB | DropBox | 213.99 | 970.90 | 1092.27 | 1191.56 | 1497.97 | 1069.12 | 152.23 | 0.1424 | 210.56 | 624.15 | 663.66 | 699.05 | 888.62 | 661.55 | 58.02 | 0.0877 |
| | Box | 5.26 | 2496.61 | 4369.07 | 4766.25 | 5825.42 | 3721.12 | 1357.18 | 0.3647 | 14.15 | 623.16 | 647.26 | 672.16 | 887.42 | 646.22 | 44.33 | 0.0686 |
| | SugarSync | 40.27 | 78.72 | 79.44 | 80.41 | 86.95 | 79.59 | 3.08 | 0.0387 | 144.43 | 200.88 | 202.43 | 204.00 | 2496.61 | 216.57 | 149.28 | 0.6893 |
| 100MB | DropBox | 250.26 | 1127.50 | 1219.27 | 1310.72 | 1519.66 | 1205.69 | 143.05 | 0.1186 | 25.09 | 647.27 | 676.50 | 708.49 | 1497.97 | 680.32 | 50.94 | 0.0749 |
| | Box | 4.71 | 2912.71 | 3883.61 | 6168.09 | 7489.83 | 4350.37 | 1797.32 | 0.3252 | 14.43 | 436.91 | 487.71 | 579.32 | 1233.62 | 507.82 | 89.36 | 0.0539 |
| | SugarSync | 42.23 | 78.96 | 79.62 | 80.66 | 87.31 | 79.64 | 3.74 | 0.0470 | 145.64 | 202.03 | 204.00 | 205.20 | 3744.91 | 223.49 | 219.50 | 0.9822 |

TABLE III
SUMMARY OF FILE MTS DISTRIBUTIONS BY FILE SIZE.




(a) SugarSync does not present important changes in both in/out traffic speed over time.

(b) We observe daily patterns in the DropBox upload transfer speed. Download transfer speed remains stable.

(c) Box upload transfers are highly variable and probably affected by daily patterns.

Fig. 5. Evolution of Personal Clouds upload/download transfer speed during 5 days. We plotted in a time-series fashion the mean aggregated bandwidth of all nodes (600 secs. time-slots) executing the service variability workload in our university laboratories (3rd−8th July 2012).

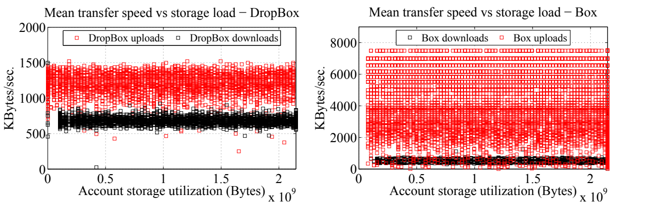



Fig. 6. Relationship between file MTS and the storage load of an account.

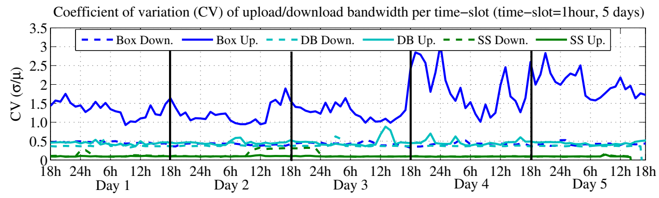



Fig. 7. Evolution of transfer speed variability over time (service variability workload, university labs).

Appreciably, *DropBox exhibits appreciable daily upload speed patterns* (Fig. 5b). Data represented in Fig. 5 was gathered between July 3, 6:00p.m. and July 8, 3:00p.m. Clearly, during night hours (1 a.m.−10 a.m.), transfer speed was between 15% to 35% higher than during diurnal hours. This phenomenon has been also detected in the experiments performed in PlanetLab, thereby discarding any artificial usage pattern induced by our university network. Moreover, considering that DropBox uses Amazon S3 as storage backend, our results are consistent with other recent works [3] that observed similar patterns in other Amazon services.

Further, we found that Box upload service may be subjected to high variability over time. Indeed, we observed *differences in upload transfer speed by a factor of* 5 *along the same day*. This observation is consistent with the analysis of the file MTS distribution where significant heterogeneity was present. More interestingly, Box uploads appear to be also affected by *daily patterns*. Concretely, the periods of highest upload speed occurred during the nights, whereas the lowest upload speeds were observed during the afternoons (3 p.m. −10 p.m.). Due to the huge variability of this service, a long-term measurement is needed to provide a solid proof of this phenomenon, though.

With respect to downloads, we observed no important speed changes over time in any system. This suggests that *downloads are more reliable and predictable, probably due to a more intense control of this type of traffic by the datacenter*.

To specifically compare the variability among services over time, we made use of the *Coefficient of Variation* (CV), which is a dimensionless and normalized measure of dispersion of a probability distribution, specifically designed to compare data sets with different scales or different means. The CV is defined as:

$$CV = \frac{1}{\bar{x}}\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2},$$

where $N$ is the number of measurements; $x_1$, .., $x_N$ are the measured results; and $\bar{x}$ is the mean of those measurements.

Fig. 7 depicts the CV in 1-hour time slots of the *aggregated bandwidth* provided by each Cloud vendor. Clearly, it can be observed important differences across the vendors. Concretely, SugarSync experiences low variability with a CV of only 10%. DropBox with a CV around 50%, however, exhibits a much higher variability than SugarSync, including isolated spikes in the upload bandwidth that reach a CV of 90%. In this sense, the Box download bandwidth capacity exhibits a similar trend. Finally, the highest observed variability was for Box uploads. In the first 3 days of the experiment, Box exhibited a mean CV of 125% approx. However, in the last part of the experiment some spikes reached a CV of 300%, suggesting that it is really hard to predict the behavior of this service.

### C. Service Failures and Breakdowns

Another important aspect of any Cloud storage service is *at what rate* users experience failures, and *whether the pattern of failures can be characterized by a simple failure process like a Poisson process*, which allows researchers to develop tractable analytical models for Cloud storage.

| *Downloads* | DropBox | Box | SugarSync |
|---|---|---|---|
| 25MB | $0.047\%(\frac{5}{10,503})$ | $0.572\%(\frac{68}{11,878})$ | $0.115\%(\frac{2}{1,740})$ |
| 50MB | $0.082\%(\frac{8}{9,745})$ | $0.698\%(\frac{80}{11,445})$ | $0.057\%(\frac{1}{1,727})$ |
| 100MB | $0.044\%(\frac{4}{9,026})$ | $0.716\%(\frac{80}{11,169})$ | $0.059\%(\frac{1}{1,691})$ |
| 150MB | $0.042\%(\frac{3}{7,136})$ | – | $0.076\%(\frac{1}{1,359})$ |
| *Uploads* | DropBox | Box | SugarSync |
| 25MB | $0.384\%(\frac{41}{10,689})$ | $0.566\%(\frac{227}{40,043})$ | $0.889\%(\frac{8}{899})$ |
| 50MB | $0.450\%(\frac{48}{10,663})$ | $1.019\%(\frac{405}{39,719})$ | $1.079\%(\frac{10}{926})$ |
| 100MB | $0.502\%(\frac{54}{10,740})$ | $2.097\%(\frac{836}{39,875})$ | $1.988\%(\frac{18}{905})$ |
| 150MB | $1.459\%(\frac{58}{3,974})$ | – | $3.712\%(\frac{33}{889})$ |

TABLE IV
SERVER-SIDE FAILURES OF API OPERATIONS ($3_{rd} - 8_{th}$ JULY 2012).

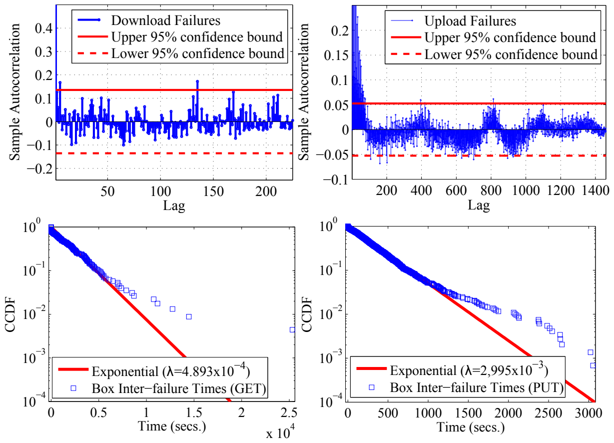



Fig. 8. Failure interarrival times autocorrelation (upper graphics) and exponential fitting of failure interarrival times (lower graphics) for Box.

For this analysis, any event *server-side* notification signaling that a storage operation did not finish *successfully* was counted as a *failure*, thereby excluding any failure, where abnormal or degraded service was observed[7]. Table IV summarizes the server-side failures observed during a 5-day measurement based on the variability workload run at our labs.

**Failure rates**. Table IV illustrates a clear trend: *in general, uploads are less reliable than downloads*. This phenomenon is present in all the Personal Clouds measured and becomes *more important for larger files*. As can be observed, downloads are up to 20X more reliable than uploads (DropBox, SugarSync), which is an important characteristic of the service delivered to users. In this sense, although failures among uploads and downloads are not so high, Box seems to provide the least reliable service. Anyway, failure rates are generally below 1%, which suggests that these *free storage services are reliable*.

**Poissonity of failures**. Now we study whether service failures appear Poisson or not, because Poisson failures allow for easy mathematical tractability. Poisson failures are characterized by interarrival times which are independent of one another and are distributed exponentially [18], and for which the failure rate is constant. In this case, we focused only on Box, since it was the only service for which enough observations were available for the statistical analysis to be significant.

To verify whether failures are independent, we calculated the autocorrelation function (ACF) for consecutive failures in the time series and depicted it in Fig. 8[8]. When the failures are completely uncorrelated, the sample ACF is approximately normally distributed with mean 0 and variance $1/N$, where $N$ is the number of samples. The 95% confidence limits for ACF can then be approximated to $0 \pm \frac{2}{\sqrt{N}}$. As shown in Fig. 8, in the case of download failures, autocorrelation coefficients for most lags lie within 95% confidence interval, which demonstrates that failure interarrival times are independent of one another. However, uploads failures are not independent, since the first lags exhibit high ACF values, which indicates short-term correlation, with alternating positive and negative ACF trends.

To conclude Poissonity for failures, failure interarrival times must be exponentially distributed, for we report the *coefficient of determination*, $R^2$, after performing linear regression on the distribution $log_{10}(1 - \{Pr\{X < x\})$, where $Pr\{X < x\}$ is the empirical failure interarrival time distribution obtained for Box. In the case of downloads $R^2 = 0.9788$ whereas for uploads $R^2 = 0.9735$. This means that failure interarrival times *approximately follow an exponential distribution*, which is evidenced in Fig. 8, where most of the samples match the exponential fitting, with the exception of those at the end of the tail. Hence, Box download failures can be *safely considered as being Poisson*. Although upload interarrival times can be well fitted by the exponential distribution, they are not independent and further analysis is needed to their characterization.

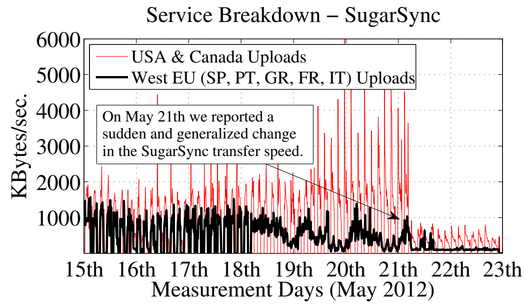



Fig. 9. We observe a radical change in the upload transfer speed of SugarSync from May 21 onwards. After May 21 all the tests performed against SugarSync reported very low transfer speeds. This reflects a change in the QoS provisioned to the REST APIs of free accounts.

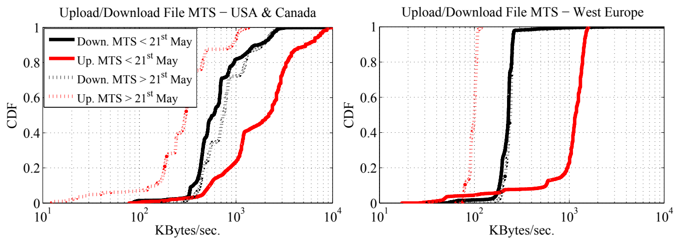



Fig. 10. PlantLab experiments against SugarSync before and after the service breakdown reported in May 21. We observe an important service degradation for uploads, whereas the download service remains unaltered.

**Service breakdowns**. Apart from the “hard” failures, there are other types of “soft” failures related with the deterioration of the QoS. And indeed, we captured a strong evidence of this in late May 2012 (Fig. 9). In Fig. 9 we present a time-series plot of the aggregated upload MTS of PlanetLab nodes against

[7]We filtered the logged error messages depending on their causes as detailed in the API specifications. We considered as errors most of the responses with $5XX$ HTTP status codes as well as other specific errors related with timed out or closed connections in the server side.

[8]Due to lack of space, we refer the reader to [18] for a technical description in depth of this methodology to assess Poissonity.

SugarSync. This information is divided for those nodes located in West Europe and USA & Canada[9].

Clearly, the behavior of the upload speed of SugarSync changed radically from May 21 onwards (Fig. 9). Before that date, SugarSync provided high transfer upload speed, comparable to current performance of Box. However, in May 21 SugarSync bandwidth provisioning policies changed dramatically; the upload MTS was reduced from $1,200$KBps to 80KBps in Western Europe — a similar trend can be observed in USA and Canada. Note that we accessed to the SugarSync service from a variety of nodes and accounts, discarding thus the possibility of IP filtering and account banning.

In this sense, Fig. 10 shows the upload/download MTS distributions for measurements performed before and after the service breakdown —executing the same workload (up/down workload) over the same nodes. Clearly, the change in the transfer speed of SugarSync was focused on uploads, that previously exhibited a good performance. On the other hand, we see that the download service was almost unaltered after May 21. These observations apply to both geographic regions. This means that Personal Clouds may change their *freemium* QoS unexpectedly, due to internal policy changes.

## V. Lessons Learned

Here we summarize the most relevant technical observations obtained from this measurement:

**Characterization of transfers**. In some cases, we observed that transfer time distributions can be characterized by known statistical distributions like the *log-logistic* and the *logistic* for downloads in Dropbox and Box, respectively. We also found that upload transfer times are Weibull distributed in Dropbox. In SugarSync, we observed a constant and very limited transfer performance. This characterization opens the door to create Personal Cloud modeling and simulation environments.

**High service variability**. The variability of Personal Cloud services is significant and induced by many factors. To wit, we discovered that uploading to DropBox is substantially faster at nights ($15\% - 35\%$), which proves the presence of daily usage patterns. We also found that the magnitude of the variation is not constant over time. An example of this is Box. While Box uploads exhibited a mean variability of $125\%$ at the beginning of our experiment, the CoV reached $300\%$ at the end. Further, we found that *uploads are more variable than downloads*.

**Reliability and Poissonity of failures**. In general, we found that Personal Clouds are *reliable*, exhibiting failure rates below $1\%$. We also found that for Box, failure interarrival times approximately follow an exponential distribution. Moreover, Box download failures can be modeled as a Poisson process, which is analytically simple.

**QoS changes and data lock-in**. We found that SugarSync changed its *freemium* QoS unexpectedly. Concretely, the mean upload speed delivered by SugarSync suddenly dropped from $1,200$ KBps to 80 KBps in EU. This emphasizes the relevance of the *data lock-in* problem, where a customer gets trapped in a provider whose service is unsatisfactory but cannot move to a new one because of the amount of data stored in it.

[9]Spikes present in Fig. 9 are due to the PlanetLab quota system, which limits the amount of data that users can transfer daily.

## VI. Conclusions

In this work, we have examined central aspects of Personal Cloud storage services to characterize their performance, with emphasis put on the data transfers. We have found interesting insights such as the high variability in transfer performance depending on the geographic location, the type of traffic, namely inbound or outbound, the file size, and the hour of the day. We have examined their failure patterns and found that these services are reliable, but susceptible to unpredictable changes in their quality of service as that witnessed in SugarSync.

This work is the first step to the characterization of Personal Clouds and will help researchers and developers to understand the behavior of these popular storage services.

## Acknowledgment

This work has been partly funded by the Spanish Ministry of Science and Innovation through projects DELFIN (TIN-2010-20140-C03-03) and RealCloud (IPT-2011-1232-430000) and by the European Union through project FP7 CloudSpaces ($\mathrm{FP7} - 317555$).

## References

[1] F. Research, "The personal cloud: Transforming personal computing, mobile, and web markets," 2011. [Online]. Available: http://www.forrester.com/rb/Research/personal_cloud_transforming_personal_computing\%2C_mobile\%2C_and/q/id/57403/t/2
[2] [Online]. Available: http://en.wikipedia.org/wiki/Dropbox_(service)
[3] A. Iosup, N. Yigitbasi, and D. Epema, "On the performance variability of production cloud services," in *CCGRID'11*, 2011, pp. 104–113.
[4] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken, "The nature of data center traffic: measurements & analysis," in *SIGCOMM IMC'09*, 2009, pp. 202–208.
[5] E. Walker, "Benchmarking amazon ec2 for high-performance scientific computing," *Usenix Login*, vol. 33, no. 5, pp. 18–23, 2008.
[6] Z. Hill, J. Li, M. Mao, A. Ruiz-Alvarez, and M. Humphrey, "Early observations on the performance of windows azure," in *HPDC '10*, 2010, pp. 367–376.
[7] M. R. Palankar, A. Iamnitchi, M. Ripeanu, and S. Garfinkel, "Amazon s3 for science grids: a viable solution?" in *DADC'08*, 2008, pp. 55–64.
[8] A. Bergen, Y. Coady, and R. McGeer, "Client bandwidth: The forgotten metric of online storage providers," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, 2011, pp. 543–548.
[9] D. Antoniades, E. P. Markatos, and C. Dovrolis, "One-click hosting services: a file-sharing hideout," in *IMC'09*, 2009, pp. 223–234.
[10] J. Sanjuàs-Cuxart, P. Barlet-Ros, and J. Solé-Pareta, "Measurement based analysis of one-click file hosting services," *Journal of Network and Systems Management*, vol. 20, pp. 276–301, 2012.
[11] A. Li, X. Yang, S. Kandula, and M. Zhang, "Cloudcmp: comparing public cloud providers," in *ACM IMC'10*, 2010, pp. 1–14.
[12] W. Hu, T. Yang, and J. Matthews, "The good, the bad and the ugly of consumer cloud storage," *ACM SIGOPS Operating Systems Review*, vol. 44, no. 3, pp. 110–115, 2010.
[13] I. Drago, M. Mellia, M. M. Munafo, A. Sperotto, R. Sadre, and A. Pras, "Inside dropbox: understanding personal cloud storage services," in *ACM IMC'12*, 2012, pp. 481–494.
[14] B. Chun, D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. Bowman, "Planetlab: an overlay testbed for broad-coverage services," *SIGCOMM Comput. Commun. Rev.*, vol. 33, no. 3, pp. 3–12, 2003.
[15] E. Hammer-Lahav, "The OAuth 1.0 Protocol," http://tools.ietf.org/html/rfc5849, 2010.
[16] H. Balakrishnan, M. Stemm, S. Seshan, and R. H. Katz, "Analyzing stability in wide-area network performance," *SIGMETRICS Perform. Eval. Rev.*, vol. 25, no. 1, pp. 2–12, 1997.
[17] M. Allman, V. Paxson, W. Stevens *et al.*, "Tcp congestion control," 1999.
[18] M. S. Artigas and E. Fernández-Casado, "Evaluation of p2p systems under different churn models: Why we should bother," in *Euro-Par*, 2011, pp. 541–553.