(Big)Data in a Virtualized World: Volume, Velocity, and Variety in Enterprise Datacenters

Robert Birke, Mathias Bjoerkqvist, and Lydia Y. Chen, *IBM Research Zurich Lab;*
Evgenia Smirni, *College of William and Mary;* Ton Engbersen *IBM Research Zurich Lab*

https://www.usenix.org/conference/fast14/technical-sessions/presentation/birke

This paper is included in the Proceedings of the
12th USENIX Conference on File and Storage Technologies (FAST '14).
February 17–20, 2014 • Santa Clara, CA USA
ISBN 978-1-931971-08-9

(Big)Data in a Virtualized World:
Volume, Velocity, and Variety in Cloud Datacenters

Robert Birke$^+$, Mathias Björkqvist$^+$, Lydia Y. Chen$^+$, Evgenia Smirni*, *and* Ton Engbersen$^+$
$^+$*IBM Research Zurich Lab,* **College of William and Mary*
$^+$*{bir, mbj,yic,apj}@zurich.ibm.com,* **esmirni@cs.wm.edu*

Abstract

Virtualization is the ubiquitous way to provide computation and storage services to datacenter end-users. Guaranteeing sufficient data storage and efficient data access is central to all datacenter operations, yet little is known of the effects of virtualization on storage workloads. In this study, we collect and analyze field data from production datacenters that operate within the private cloud paradigm, during a period of three years. The datacenters of our study consist of 8,000 physical boxes, hosting over 90,000 VMs, which in turn use over 22 PB of storage. Storage data is analyzed from the perspectives of volume, velocity, and variety of storage demands on virtual machines and of their dependency on other resources. In addition to the growth rate and churn rate of allocated and used storage volume, the trace data illustrates the impact of virtualization and consolidation on the velocity of IO reads and writes, including IO deduplication ratios and peak load analysis of co-located VMs. We focus on a variety of applications which are roughly classified as app, web, database, file, mail, and print, and correlate their storage and IO demands with CPU, memory, and network usage. This study provides critical storage workload characterization by showing usage trends and how application types create storage traffic in large datacenters.

1 Introduction

Datacenters provide a wide spectrum of data related services. They feature powerful computation, reliable data storage, fast data retrieval, and, more importantly, excellent scalability of resources. Virtualization is the key technology to increase the resource sharing in a seamless and secure way, while reducing operational costs without compromising performance of data related operations.

To optimize data storage and IO access in virtualized datacenters, storage and file system caching techniques have been proposed [13, 18, 28], as well as data duplication and deduplication techniques [22]. The central theme is to move the right data to the right storage tier, especially during periods of peak loads of co-located virtual machines (VMs). Therefore, it is crucial to understand the characteristics of IO workloads of individual VMs, as well as the workload seen by the hosting boxes. There are several storage-centric studies that have shed light on file system volume [14, 20, 31] and IO velocity, i.e., read/write data access speeds [15, 17, 28]. Despite these studies, it is unclear how virtualization impacts storage and IO demands at the scale of datacenters, and what their relationship to CPU, memory, and network demands are.

The objective of this paper is to provide a better understanding of storage workloads in datacenters from the following perspectives: storage volume, read/write velocity, and application variety. Using field data from production datacenters that operate within the private cloud paradigm, we analyze traces that correspond to 90,000 VMs hosted on 8,000 physical boxes, and containing over 22 PB of actively used storage, covering a wide range of applications, over a time span of three years, from January 2011 to December 2013. Due to the scale of the available data, we adopt a black-box approach in the statistical characterization of the various performance metrics. Due to the lack of information about the system topologies and the employed file system architectures, this study falls short in analyzing latency, file contents, and data access patterns at storage devices. Our analysis provides a multifaceted view of representative virtual storage workloads and sheds light on the storage management of highly virtualized datacenters.

The collected traces allow us to look at the volume of allocated, used, and free space in virtual disks per VM, with special focus on the yearly growth rate and weekly churn rate. We measure velocity by statistically characterizing the loads of read and write operations in GB/h as well as IO operations per second (IOPS) in multiple time scales, i.e., hourly, daily, and monthly, focusing on characterization of the time variability and peak load analysis. We deduce the efficiency of storage deduplication in a virtualized environment, by analyzing the IO workload of co-located VMs within boxes. To see how storage and IO workloads are driven by different applications, we perform a per-application analysis that allows us to focus on a few typical applications, such as web, app, mail, file, database, and print applications, highlighting their differences and similarities in IO usage. Finally, we present a detailed multi-resource dependency study that centers on data storage/access and provides insights for the current state-of-the-practice in data management in datacenters.

Our findings can be quickly summarized as follows: VM capacity and used space have annual growth rates of 40% and 95%, respectively. The fullness per VM has a growth rate of 19%, though the distribution of storage fullness remain constant across VMs over the three years of the study. The lower bound of VM storage space churn rate is 17%, which is slightly lower than the churn rate of 21% reported from backup systems [31].

Regarding IO velocity, the IO access rates of boxes scales almost linearly with the number of consolidated VMs, despite the non-negligible overhead from virtualization. Both VMs and boxes are dominated by write workloads, with 11% of boxes experiencing higher virtual IO rates than physical ones. Deduplication ratios grow linearly with the degree of virtualization. Peak loads occur at off-hours and are contributed to a very small number of VMs. VMs with high velocity tend to have higher storage fullness and higher churn rates.

Regarding IO variety, different applications use storage in different ways, with file server applications having the highest volume, fullness, and churn rates. Databases have similar characteristics but low fullness. Overall, we observe that high IO demands strongly and positively correlate with CPU and network activity.

The outline of this work is as follows. Section 2 presents related work. Section 3 provides an overview of the dataset. The volume, velocity and variety analysis are detailed in Sections 4, 5 and 6, respectively. A data-centric multi-resource dependency study is discussed in Section 7, followed by conclusions in Section 8.

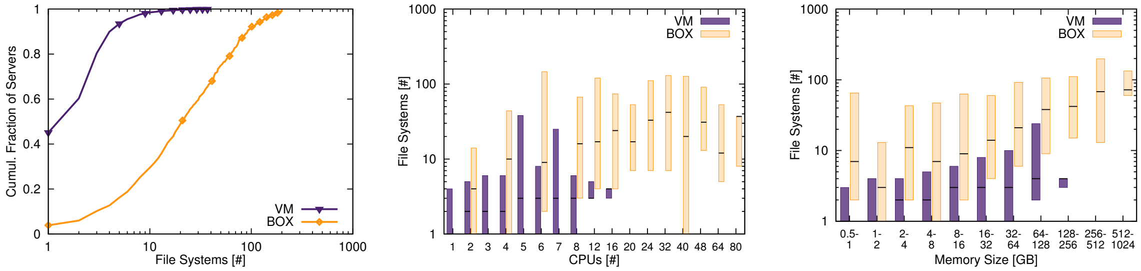
2 Related Work

Managing storage is an expensive business [19]. Coupled with the fact that the cost of storage hardware is several times that of server hardware, efficient use of storage for datacenters becomes critical [29]. Workload characterization studies of storage/IO are pivotal for the development of new techniques to better use systems, but it is difficult to define what is truly a representative system due to the wide variety of workloads. In general, from the various studies on file system workloads, those that stand out are the ones based on academic prototypes and those based on personal computers, in addition to a plethora of lower level storage studies. Virtualization adds additional layers of complexity to any storage media [10, 16]. As virtualization is indeed the standard for datacenter usage, workload studies of virtualized IO are important and relevant. Nonetheless, analyzing all relevant features of all relevant virtualized IO workloads is outside the scope of this work. Here, given the collected trace data, we conduct a statistical analysis with the aim of better understanding how IO occurs in a virtualized environment of a very large scale.

Typically, related work covers aspects of volume [2, 14, 20, 30], velocity [17] and variety, with a focus on file systems. Regarding file system volume, there are several studies that focus on desktop computers [2,14,20]. Using file system metadata during periods of four weeks [20] and five years [2], performance trends and statistics that shed light on fullness, counts of files/directories, file sizes, and file extensions are provided. Recognizing the need to better understand the behavior of backup workloads, Wallace et al. [31] present a broad characterization analysis and point out that the data churn rate is roughly 21% per week. Their study shows that the capacity of physical drives approximately doubles annually while their utilization only drops slightly. The study compares backup storage systems with primary storage ones and finds that their fullness is $60-70\%$ and $30-40\%$, respectively. Characterization of backup systems has been traditionally used to drive the development of deduplication techniques [20, 24].

Most works on IO characterization analysis focus on specific file systems within non-virtualized environments, e.g., NFS [7], CIFS [17], Sprite/VxFs [9], NTFS [25], and the EMC Data Domain backup system [31]. Common characteristics include large and sequential read accesses, increasing read/write ratios, bursty IO, and a small fraction of jobs accounting for a large fraction of file activities. Self-similar behavior [9] is identified and proposed to use to synthesize file system traffic. Backup systems [31] have been observed to have significantly more writes than reads, whereas file systems for primary applications have twice as many reads as writes [17].

Following the advances in virtualization technologies, several recent works focus on optimizing data storage and access performance in virtualized environments with an emphasis on novel shared storage design [11, 13] and data management [15, 18, 28]. To reduce the load on shared storage systems, distributed-like VM storage systems such as Lithium [13] and Capo [28] are proposed. Gulati et al. design and implement the concept of a stor-



(a) CDF (b) Number of CPUs (c) Memory size [GB]

Figure 1: Number of file systems associated with a VM and a box: (a) cumulative distribution, (b) boxplots of file systems as a function of the number of CPUs, and (c) boxplots of file systems as a function of memory size. The boxplots present the 10^{th}, 50^{th}, and 90^{th} percentiles.

age resource pool, shared across a large number of VMs, by considering IO demands of multiple VMs [11]. Systems that aim at improving IO load balancing for virtualized datacenters using performance models have been proposed [10, 23]. Combining intelligent caching, IO deduplication can be achieved by reducing duplicated data across different storage tiers, such as VMs, hosting boxes [18], and disks [15]. Everest [21] addresses the challenges of peak IO loads in datacenters by allowing data written to an overloaded volume to be temporarily off-loaded into a short-term virtual store. Nectar [12] proposes to interchangeably manage computation and data storage in datacenters by automating the process of generating data, thus freeing space of infrequently used data. Workload characterization that focuses on specific server workloads (i.e., application variety) such as web, database, mail, and file server, has been done for the purpose of evaluating energy usage [27]. Till now, only a rather small scale virtual storage workload characterization has been presented [28], pointing out that virtual desktop workloads are defined by their peaks.

The workload study presented here presents a broad overview of virtual machine storage demands at production datacenters, covering IO volume, velocity, and variety, and how these relate to the degree of virtualization as well as usage of other resources. The analysis presented here compliments many existing IO and file system studies by using a very large dataset from production datacenters in highly virtualized environments.

3 Statistics Collection

We surveyed 90,000 VMs, hosted on 8,000 physical servers in different data centers dispersed around the globe, serving over 300 corporate customers from a wide variety of industries, over a three year period and accounting for 22 PB of storage capacity. The servers use several operating systems, including Windows and different UNIX versions. VMware is the prevalent virtualization technology used. For a workload study on current virtualization practices, we direct the interested readers to [5].

The collected trace data is retrieved via `vmstat`, `iostat` and supervisor specific monitoring tools, and is collected for VMs as well as for physical servers, termed *hosting boxes*. Each physical box may host multiple (virtual) file systems, which are the smallest units of storage media considered in this study. To characterize data workloads in virtualized datacenters, we focus on three types of IO-related statistics for VMs.

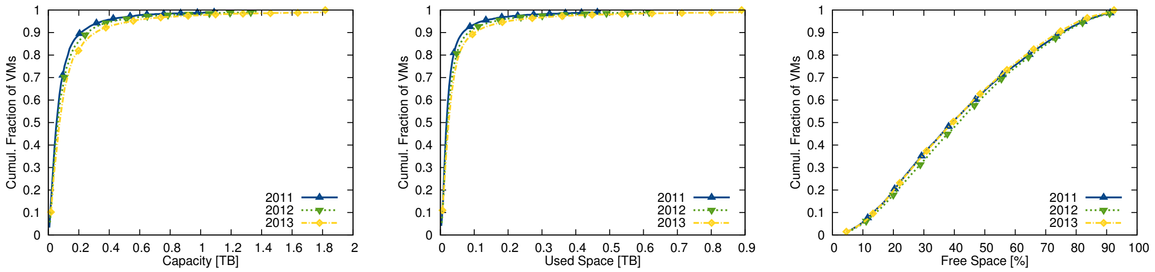
Volume refers to the allocated space, free space, and degree of *fullness*, defined as the ratio between the total used space and the total allocated space, of a VM after aggregating all of its file systems. Here, we focus on long-term trends, i.e., growth rates, and short-term variations, i.e., churn rates.

Velocity refers to read and write speeds measured in number of operations and transfered bytes per time unit, as IOPS and GB/h, respectively. We compare virtual IO velocity, measured at the VMs, with physical IO velocity, measured at the underlying boxes.

Variety refers to volume and velocity of *specific* applications, i.e., app, web, database, file, mail, and print, on specific VMs. To conduct storage-centric multi-resource analysis, we also collect CPU utilization, memory usage, and network traffic for VMs as well as boxes.

The trace data is available in two granularities: (1) in 15-minute/hourly averages from April 2013 and (2) coarse-grain monthly averages from January 2011 to December 2013. When exploring the differences between VMs in a day, we use the detailed traces with 15-minute/hourly granularity from 04/17 and 04/21. Monthly averages are used to derive long-term trends.

We note that the statistics of interest have long tails, therefore we focus on presenting CDFs as well as certain percentiles, i.e., 10^{th}, 50^{th} and 90^{th} percentiles. As the degree of virtualization (i.e., consolidation) on boxes is quite dynamic, we report on daily averages per phys-



(a) Storage Capacity (b) Used Volume (c) % of fullness

Figure 2: CDF of storage volume per VM over three years.

ical box. To facilitate the analysis connecting the per VM storage demands with the per file system storage demands, we present the CDF of the number of file systems across VMs and boxes (see Figure 1 (a)) and also how file system distributions vary across different systems, which we distinguish by the number of CPUs per box and memory (see boxplots in Figure 1 (b) and (c), respectively). Figure 1 (a) shows that boxes typically have a much higher (more than 21) number of virtual file systems than VMs, which have on the average 2 virtual file systems. Such values are very different from desktop studies [2] and underline the uniqueness of our dataset, especially in light of virtualized datacenters. Moreover, looking at the trends of medians in Figure 1 (b) and 1 (c), the number of file systems grows with servers equipped with more CPUs and, particularly, with larger memory.

As our data is obtained by standard utilities at the operating system level, we lack specific information about file systems, such as type, file counts, depth, and extensions. In addition, since the finest-grained granularity of the trace data is for 15-minute/hourly periods, IO peaks within such intervals cannot be observed. For example, the maximum GB/h within a day identified in this study is based on hourly averages, and is much lower than the instantaneous maximum GB/h. The coarseness of the information gathered is contrasted by the huge dataset of this study: 8,000 boxes with high average consolidation levels, i.e., 10 VMs per box, observed over a time-span of three years.

4 Volume

One of the central operations for datacenter management is to dimension storage capacity to handle short term as well as long term fluctuations in storage demand. These fluctuations are further accentuated by data deduplication and backup activities [6, 20]. Surging data demands and data retention periods drive storage decisions; however, existing forecasting studies either adopt a user or a per file system perspective, not necessarily targeting entire datacenters. Here, the aim is to adopt a different perspective and provide an overview on the yearly growth rates and weekly churn rates of storage demand at the VM level. In the following subsections, we focus on the storage demands placed by 90,000 VMs, their used/allocated storage and fullness, followed by statistical analysis of their yearly growth rates and weekly churn rates.

4.1 Data Storage Demands across VMs

Taking yearly averages of the monitored VMs over 2011, 2012, and 2013, we present how storage demands evolve over time and how they are distributed across VMs. Figure 2 (a) and 2 (b) present the CDF of the total sum of allocated and used storage volume per VM over all file systems belonging to each VM. Figure 2 (c) summarizes the resulting fullness. Visual inspection shows that the overall capacity and the used space per VM grow simultaneously, and result in fullness being constant over time. This observation illustrates a similar behavior as the one observed at the file system level [20], and provides information on how to dimension storage systems for datacenters where VMs are the basic compute units.

Via simple statistical fitting, we find that exponential distributions can capture well the VM storage demands in terms of allocated storage capacity and used storage volume. Table 1 summarizes the measured and fitted values, means and 95^{th} percentiles of capacity and used volume are reported. Since there are on average 10 VMs sharing the same physical box [5], a system needs to be equipped with 450 GB of storage space for very aggressive storage multiplexing schemes, i.e., only the used space is taken into account (45×10), or 1120 GB for a more conservative consolidation scheme based on the allocated capacity (112×10). The uniform distribution can approximately model fullness. Since the *relative ratio* of two independently exponential random variables is uniform [26], this further confirms that the exponential distribution is a good fit. Overall, the above analysis gives trends for the entire VM population, which in turn increases over the years, but does not provide any infor-

Table 1: Three year storage volume: measured and fitted data from exponential distribution.

	mean			95^{th}		
Year	2011	2012	2013	2011	2012	2013
Capacity [GB]	122	148	186	365	436	569
Exponential	122	148	186	365	442	556
Used [GB]	47	60	76	128	165	207
Exponential	47	60	76	140	180	228
Fullness [%]	42	44	42	83	83	81

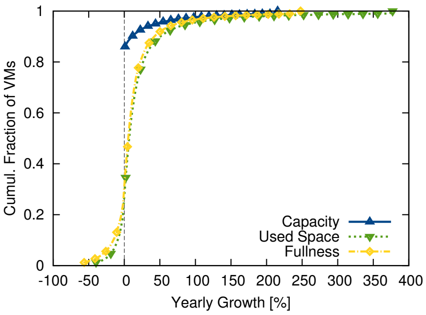


Figure 3: CDF of yearly growth rate of VM storage volume: capacity, used space, and fullness.

mation on how the storage volume of individual VMs changes. In the following subsections, we focus on computing the yearly growth rate and weekly upper bound of the churn rate for each VM by presenting CDFs for the entire VM population.

On average, a VM has 2.55 file systems with a total capacity of 185 GB, of which roughly 42% is utilized, implying that each VM on average stores 77 GB of data. In general, the allocated capacity and free storage space increases over the years, while storage fullness remains constant.

4.2 Yearly Growth Rate

The data growth rate is predicted to double every two years [1]. Yet, it is still not clear how this value translates into growth at the per VM data volume level, or more importantly, whether the existing storage resources can sustain future data demands. Here, we analyze the long-term volume growth rates from two perspectives: supply, i.e., from the perspective of storage capacity, and demand, i.e., from the perspective of used storage volume.

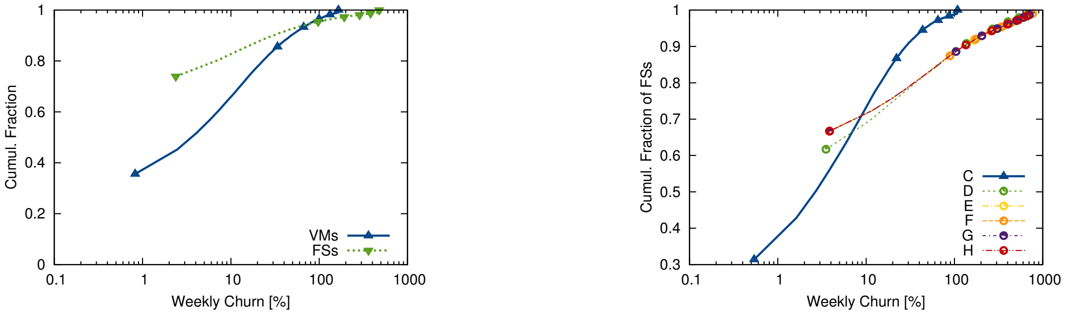
In Figure 3, we show the CDF of the yearly relative growth rate of allocated capacity, used space, and fullness, across all VMs. We compute the relative yearly growth rate as the difference in used capacity between June 2012 and May 2013, and divide it by the start value. A positive (negative) growth indicates an increasing (decreasing) trend. Overall, the CDF of used space is very close to fullness, meaning that the storage space utilization is highly affected by the data demand, rather than by the supply of the capacity.

One can see that most VMs (roughly 86%) do not upgrade their storage, whereas the remaining 14% VMs increase their storage capacity quite significantly, i.e., up to 200%. Due to this long tail, the mean increase is 40.8%. As for the demand of space, almost all VMs increase their used storage. Only a small amount (below 25%) of VMs decrease their used space and have negative growth rates. On the other hand some VMs have a three-fold increase in used space. As a net result, the mean growth of used space is 95.1%. The smallest growth belongs to fullness: the mean rate is 19.1%. Such a value is higher than the fullness trend evaluated across the entire VM population in Figure 2(c). Both storage capacity and used space increase over time for each individual VM with a mean yearly growth rate of 40% and 95%, respectively. The resulting fullness also increases by 19% every year.

4.3 Weekly Churn Rate: Lower Bound

Here, we study short-term fluctuations of storage volume utilization, defined by the percentage of bytes that have been deleted during a time period of a week with respect to the used space. Note that this value represents a lower bound on the churn rate, since what is available in the trace is total volume in 15-minute intervals, i.e., if a VM writes and deletes the same amount of data within the 15-minute interval, there is no way to know how much is *truly* deleted during that period. We therefore report here a lower bound on the churn rate; the true value may be larger than the one reported here. The inverse of the lower bound of the churn rate reveals the upper bound of the data retention period. For example, a 20% weekly churn rate here means that the data is kept up to 5 weeks before being deleted. We base our computation of the weekly churn rate of VMs on the 15-minute data collected between 04/22/2013 to 04/28/2013. The churn rate is computed as the sum of all relative drops in used space, i.e., all negative differences between two adjacent 15-minute samples. We note that as data is also added over this one week time frame and we consider the sum of all deleted data, this value can go over 100%.

We present CDFs of two types of weekly churn rates in Figure 4 (a): by VMs and by file systems (FSs). The former gives the data volume deleted by VMs and the latter focuses on data volume deleted from an individual file system. Seen from the starting point and long tail of file system's CDF, a high fraction of file systems have a churn rate of zero, while a small fraction of file systems have very high churn rates. Thus a higher variability of churn rates is observed at file systems than at VMs. To



(a) VM and file system

(b) Specific file systems

Figure 4: CDF of the weekly churn rate computed based on single VM and a single file system: the x-axis is the percentage of storage space deleted in a week; the y-axis are the cumulative fraction of VMs, file systems, and specific file systems.

further validate this observation, we compute the churn CDF of the most commonly seen volume labels of file systems, i.e., C, D, E, F, G, and H, from Windows systems, that account for roughly 87% of the entire VM population. Shown in Figure 4 (b), one can clearly see that volume label C has very low churn rate, compared to the other labels. Such an observation matches with the common practice that C drives on Windows systems store program files that are rarely updated and other drives are used to store user data.

Overall, the churn rates of VMs have a mean around 17.9%, whereas churn rates of file systems have a mean around 20.8%. This value, being a lower bound, is on par with previous results in the literature, where a true churn rate, computed from detailed file system traces, is 21% [31]. Most VMs have rather low churn rate lower bounds; from Figure 4 (a) one can see that 75% of VMs have churn rates below 15%. However, 10% of VMs have a churn rate higher than 50%. VMs with high churn rates pose challenges for the storage system, because a large amount of space needs to be reclaimed and written.

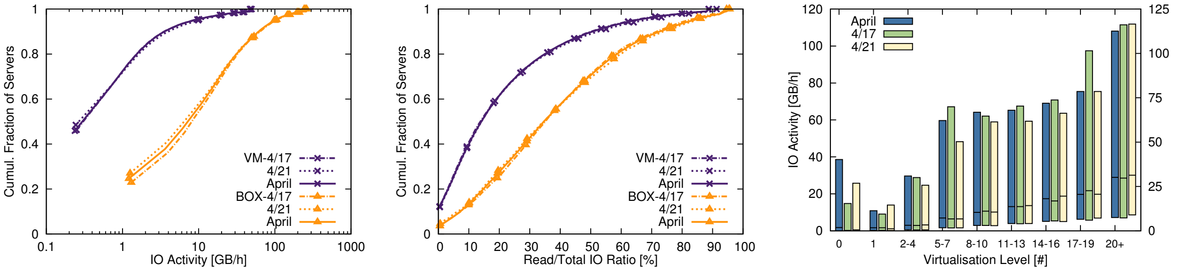
5 Velocity

The most straight-forward performance measure for storage systems is the IO speed, which we term velocity within the context of VMs accessing big data in data centers. The performance at peak loads [21] has long been a target focus for optimization. To expedite IO operations, caching [28] and IO deduplication [15] algorithms are critical. This is especially true within the context of virtualized data centers where the system stack, e.g., the additional hypervisor layer, for IO activities becomes deeper and more complex. The evaluation of caching and IO deduplication schemes in virtualized datacenters is usually done at small scale or lab-like environments [15, 28]. We quantify VM velocity via the speed by which data is placed in and retrieved from datacenter storage, and further pinpoint “hot” or “cold” VMs from the IO perspective. The statistics presented in the following subsections are based on hourly averages from 04/17/2013, which is shown representative for IO velocity in Section 5.1. The focus is on understanding their variability over time and their dependency on the virtualization level (i.e., on the number of simultaneous executing VMs), as well as on peak IO load analysis.

5.1 Overview

We start this section by presenting an overview of the daily velocity of VMs (and their corresponding boxes) in terms of (1) transferred data per hour (GB/h) including both read and write operations; and (2) the percentage of transferred data associated with read operations. Figure 5 depicts the aforementioned information in three types of statistics: the hourly average based on 04/17/2013 (weekday), 04/21/2013 (weekend), and daily average computed over the entire month of April 2013. The aim is to see if the IO velocity of a randomly selected date is sufficiently representative. Overall, the statistics of the daily velocity on 04/17/2013 are very close to those of a weekend day and to the statistics aggregated from the daily average over the entire April, see the almost overlapping lines in all three subfigures of Figure 5. Hence, in the rest of this paper we focus on a specific day 04/17/2013, which we consider as representative.

Shown by a lower CDF in Figure 5 (a), boxes have higher IO velocity than VMs. The average IO velocity for boxes and VMs are 26.7 GB/h and 2.9 GB/h, respectively, i.e., the velocity for boxes is larger roughly by a factor of 9. This factor is in line with the average consolidation level [5], i.e., 10 VMs per box and hints to a linear scaling of IO activity. Regarding the percentage of read operations, boxes have heavier read workloads than VMs do, as shown by the CDF curve in Figure 5 (b)



(a) IO in GB/h (b) Percentage of reads (c) IO by virtualization level

Figure 5: Daily velocity: IO read and write activities per VM and box on 4/17, 4/21, and the entire April.

that corresponds to boxes. There is roughly 12% of VMs having only write workloads, as indicated by the leftmost point of the VM CDF. Meanwhile, less than 1% of VMs have read workloads only. Indeed, the mean read ratio of boxes and VMs are 38% and 21%, respectively. Overall, the velocity of VMs and boxes is dominated by write workloads.

To verify how the virtualization level affects the box IO activity, we group the box IO activity by virtualization level and present the 10^{th}, 50^{th}, and 90^{th} percentiles, see the boxplots in Figure 5 (c). The box IO activity increases almost linearly with the virtualization level, this can be seen by the 50^{th} percentile. When further normalizing the IO velocity of a box by the number of consolidated VMs, the average values per box drop slightly with the virtualization level. This implies that there is a non-negligible fixed overhead associated with virtualization. We omit this graphical presentation due to lack of space.

5.2 Deduplication of Virtual IO

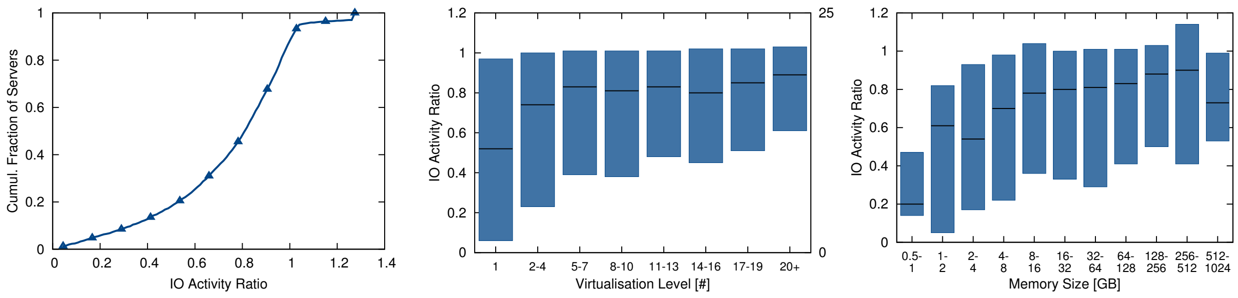
IO deduplication techniques [15] are widely employed to reduce the amount of IO. The discussion in this section is limited to virtual IO since, from the traces, there is no way to distinguish how and where the data is deduplicated and/or cached. We compare the sum of all virtual IO activity aggregated over all consolidated VMs within a box, termed virtual IO, divided by the IO activity measured at the underlying physical box, termed box IO, and call this ratio the *virtual deduplication ratio*. In contrast to the rest of the paper, we here use IOPS as the measurement of velocity, instead of GB/h. When the deduplication ratio is greater (or smaller) than one, the virtual IO is higher (or lower) than the physical box IO, respectively. A deduplication ratio of one is used as the threshold between deduplication and amplification.

We summarize the CDF of the deduplication ratio in Figure 6 (a). Roughly 50% of boxes have a deduplication ratio ranging from 0.8 to 1.2, i.e., close to one, indicating similar IO activities at the physical and virtual levels. Another observation is that most boxes experience amplification, as indicated by deduplication ratios less than one (including close to one), i.e., virtual IO loads are lower than physical IOs. This can be explained by the fact that hypervisors induce IO activities due to VM management, e.g., VM migration.

There is a very small number of boxes (roughly 11%) experiencing IO deduplication, as indicated by the boxes having deduplication ratios greater than one. To understand the cause of such deduplication, we compute the separate deduplication ratio for read and write activities. We see that the observed deduplication stems more from read than write operations, as indicated by a higher fraction of boxes (roughly 18%) having deduplication read ratios greater than one. One can relate this observation to the fact that read caching techniques are more straightforward and effective than write caching techniques.

To see how virtualization affects deduplication ratios, we group the deduplication ratios by their virtualization level and present them using boxplots, see Figure 6 (b). Looking at the lower and middle bars of each boxplot, i.e., the 10^{th} and 50^{th} percentiles, we see that the deduplication ratios increase with the virtualization level. Such an observation can be explained by the fact that IO activities of co-located VMs have certain dependencies that further present opportunities for reducing IO operations for hypervisors. Higher virtualization levels can lead to better IO deduplication. We note that similar observations and conclusions can be deduced by using IO in GB/h, with the deduplication ratios roughly ranging between 0 to 3.

In addition to virtualization, the effectiveness of IO deduplication can be highly dependent on the cache size. Unfortunately, our data set does not contain information about cache sizes, only memory sizes, which in turn are often positively correlated to the cache sizes. Therefore, to infer the dependency between cache size and IO deduplication ratio, we resort to memory size and categorize deduplication ratios by box memory sizes, see Figure 6 (c). The trend is that the IO deduplication ratio increases with increasing memory size, though with a drop for systems having memory greater than 512 GB.

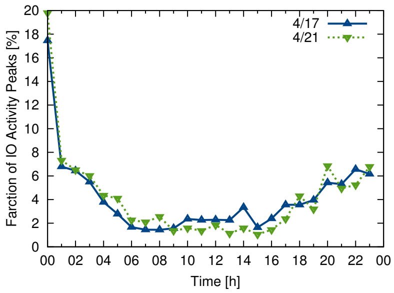


(a) CDF

(b) By virtualization level

(c) By memory size

Figure 6: Virtual IO deduplication/amplification per box: $\frac{Virtual\_IOPS}{Physical\_IOPS}$



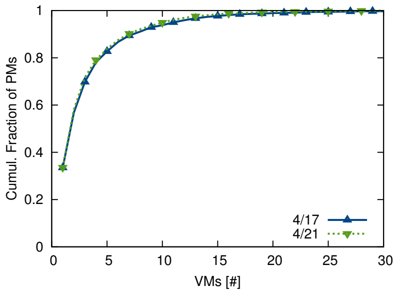


Figure 7: PDF of virtual loads peak times in a day over all consolidated VMs.

Figure 8: Number of VMs to reach 80% of peak load over all consolidated VMs.

5.3 Peak Velocity of Virtual IO

Virtualization increases the randomness of access patterns due to the general lack of synchronized activity between the VMs and the larger data volume accessed, which in turn imposes several challenges to IO management [8]. The first question is how IO workloads fluctuate over time. To such an end, for each VM and box, we compute their coefficient of variation (CV) of the IO activity in GB/h during a day using the hourly data. The higher the CV value, the higher the variability of the IO workload during the day. Our results show that boxes have rather stable IO velocity with an average CV of around 0.8, while VMs have an average CV of around 1.3.

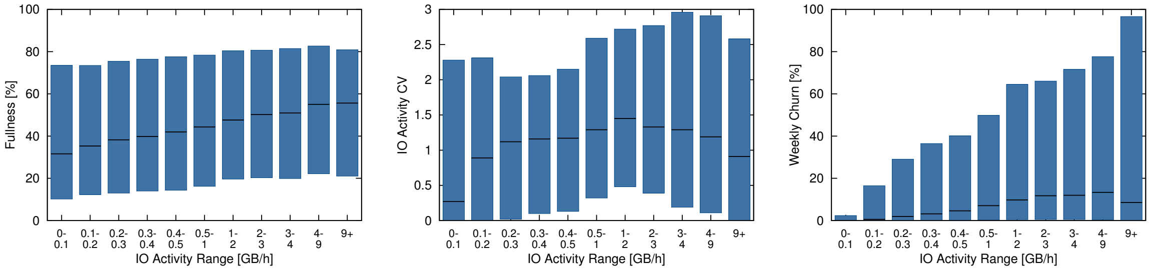
The confirmation of higher time variability of VMs lead us to focus on the characteristics of virtual IO aggregated over all VMs hosted on the same box, in particular their peak loads. We try to capture when the peaks of aggregated velocity happen, and how each VM contributes to the peak. We do this both for a Wednesday (04/17/2013) and a Sunday (04/21/2013) based on the hourly IO activity data.

5.3.1 Peak Timings

Figure 7 presents the empirical frequencies showing which hour of the day the aggregated virtual peak IO loads happen. Clearly, most VMs have peaks during after-hours, i.e., between 6pm to 6am, for both days. This observation matches very well with timings for peak CPU [4] and peak network [3] activities but does not match the belief that IO workloads are driven by the working hours schedule [18]. Indeed, in prior work [5] we have observed that most VM migrations occur during midnight/early morning hours, which is consistent with the activity seen in Figure 7. Clearly, the intensity of virtual IO workloads is affected by background activities such as backup and update operations that are typically run during after-hours.

5.3.2 Top VM Contributors

Another interesting question is how consolidated VMs contribute to peak loads. Information on top VM contributors to peak loads is critical for improving peak load performance via caching [21,28]. We define as top contributors the co-located VMs having the highest contributions to the peak load in order to reach a certain threshold, i.e., 80% of the peak load in this study. We summarize the distribution of the number of top VM contrib-



(a) Fullness (b) Time variability (c) Weekly churn rate

Figure 9: Cold vs Hot VMs: volume, time variability in a day, and weekly storage space churn. The x-axis is IO in GB/h and y-axes are fullness [%], coefficient of variation (CV), and weekly churn rate.

utors for both days in Figure 8. Interestingly, one can see a clear trend indicating that it is very common that a small number of VMs dominates peak loads for both days. Such a finding is similar to the one reported in [28], where only independent (i.e., not co-located) VMs are considered. These results further show that making a priority the optimization of the IO of a few top VMs may have a large impact on overall performance.

5.4 Characteristics of Cold/Hot VMs

Motivated by the fact that a few number of VMs contribute to peak loads, we try to capture the characteristics of VMs based on their IO activity in GB/h, aiming to classify the VMs as cold/hot. The hotness of the data is very useful to dimension and tier storage systems; e.g., cold data in slow storage media and hot data in flash drives. To this end, we compare the used volume, time variability, and churn rate of VMs grouped by different levels of IO activity, see Figure 9 (a), (b), and (c), respectively. Each box represents a group of VMs having an average activity falling into the IO activity range shown on the x-axis.

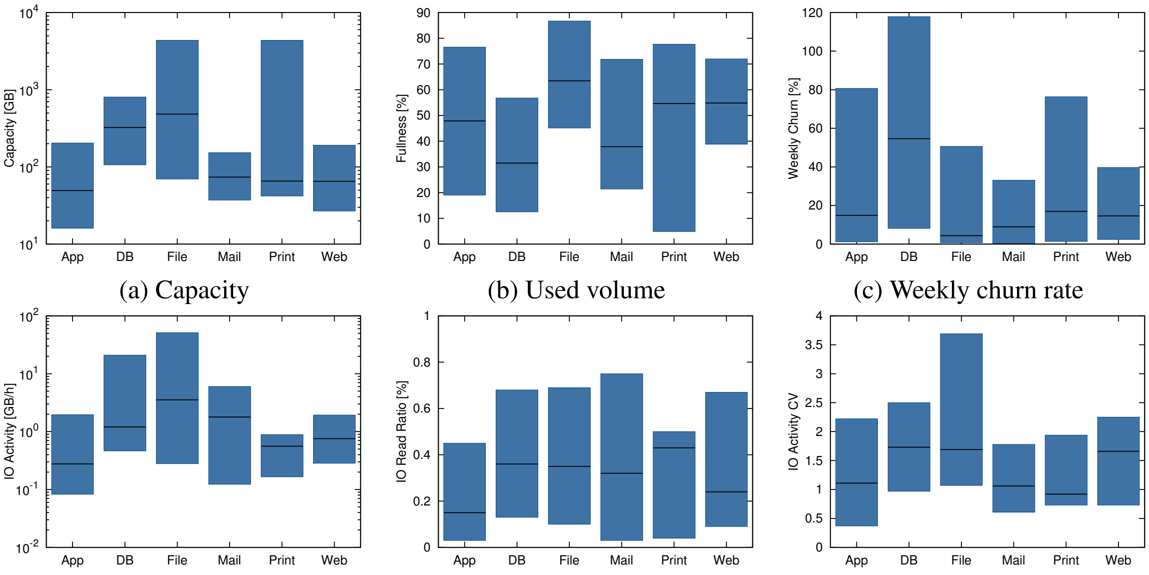
The 50^{th} percentile, i.e., the middle bar in each boxplot, increases with the IO activity level for both the fullness and churn rate. Overall, VMs with high IO activities are also fuller and have higher churn rates, compared to VMs with low IO activities. For fullness, not only the 50^{th} percentile, but also the entire boxes shift with the IO activity level. To see if the reverse is also true, we classify the IO activity level by different levels of used space both in GB and percentage. The data shows that high space usage indeed results in high IO activity, especially when measured in GB. However, VMs with very full storage systems, i.e., 90-100% occupancy, have slightly lower IO activity than VMs with 80-90%. This stems from the fact that most storage systems have optimal performance when they are not completely full. A common rule of thumb is that the best performance is achieved when the used space is up to 80%. Hence, only cold data is placed on disks with a higher percentage of used space. Due to space constraints, we omit the presentation of this set of results.

The time variability shows a different trend, i.e., the CV first increases as IO velocity increases but later decreases, see Figure 9 (b). The hottest VMs, i.e., the ones with IO greater than 9 GB/h, have the second lowest CV, as can be seen from the 50^{th} percentile. We thus conclude that hot VMs have relatively constant, high IO loads across time.

Regarding churn rates, both the 50^{th} and 90^{th} percentiles clearly grow with IO activity levels, indicating strong correlation between IO activity and churn. Such an observation matches very well with common understanding that hot VMs have frequent reads/writes, resulting in frequent data deletion and short data retention periods. This is confirmed by our data showing quantitatively that 50% of hot VMs, i.e., VMs having an IO activity level of 9 GB/h or more, have data retention periods ranging between 11.11 $(1/0.09)$ and 1.02 $(1/0.98)$ weeks. In summary, hot VMs have higher volume consumption (55%) and churn rates (9%).

6 Variety

The trace data allows to distinguish application types for a subset of VMs. Here, we select the following applications: *app*, *web*, *database (DB)*, *file*, *mail*, and *print*, and characterize their volume and velocity. Our aim here is to provide quantitative as well as qualitative analysis that could be used in application-driven optimization studies for storage systems. The app servers host key applications for clients, such as business analytics. DB servers run different database technologies, such as DB2, Oracle, and MySQL. File servers are used to remotely store files. Due to business confidentiality, it is not possible to provide detailed information about these applications. We summarize the storage capacity, used space, weekly churn rate, IO velocity, percentage of read operations, and time variability using boxplots for each application



(a) Capacity

(b) Used volume

(c) Weekly churn rate

(d) IO in GB/h

(e) Read ratio

(f) Time variability (CV)

Figure 10: Application's storage volume and IO velocity.

type, see Figure 10. We mark the 10^{th}, 50^{th}, and 90^{th} percentile of VMs belonging to each application. Most statistics are based on the data collected on 04/17/2013, except for the weekly churn rate that is based on data between 04/22/2013 to 04/28/2013.

Storage Capacity: File VMs have the highest capacities, followed by DB VMs – see the relative values of their respective 50^{th} percentiles. Mail, print, web, and app have similar storage capacities, but print VMs have the highest variance – see the height of the boxplot.

Volume: Fullness shows a slightly different trend from the allocated storage capacity. File VMs are also the fullest, hence they store the largest data volume. Database VMs that have the second highest allocated capacities are now the least full, hinting to large amounts of free space. In terms of variability of fullness across VMs in the same application type, print VMs still have very different storage fullness.

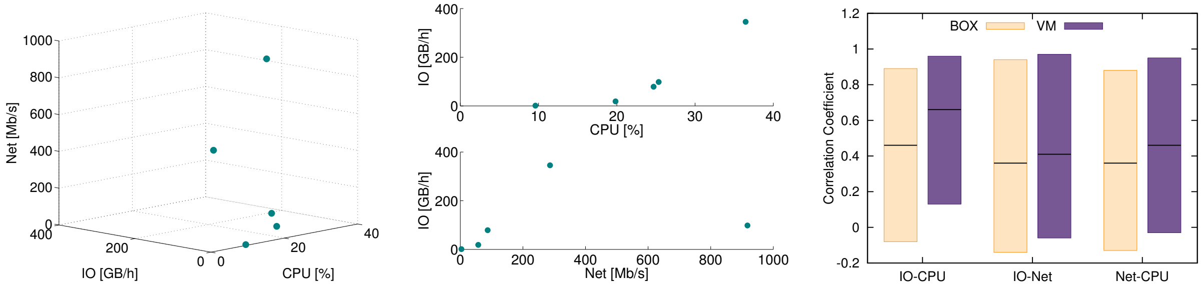
Weekly Churn Rate: DB VMs have the highest weekly churn rate, with some VMs having churn rates greater than 120%, hinting to frequent updates where a lot of storage volume is deleted and reclaimed. Unfortunately, due to the coarseness of the trace data, we cannot confirm whether this is due to the `tmp` space used for large queries, although this is a possible explanation. Such an observation goes hand-in-hand with low fullness of DB. Based on the value of 50^{th} percentile, print VMs have the second highest churn rate, as print VMs store many temporary files, which are deleted after the print jobs are completed. Due to dynamic contents, app and web VMs have high churn rates as well, i.e., similar to the mean churn rate of 17.9% shown in Section 4.3.

IO Velocity: Applying characteristics of hot/cold VMs summarized in Section 5, it is no surprise that file VMs have the highest IO velocity, measured in GB/h. According to the 50^{th} percentile, mail and DB VMs have the second and third highest IO velocity. Print, web, and app VMs experience similar access speeds.

Read/Write Ratio: All application VMs have their 50^{th} percentile of read ratio less than 50%, i.e., all application types have more write intensive operations than read operations. Indeed, as discussed in Section 5, VMs are more write intensive. Among all, app VMs have the lowest read ratio, i.e., lower than 20%. In contrast, print VMs have the highest read ratio close to 50%, which is reasonable as print VMs have rather symmetric read/write operations, i.e., write files to storage and read them for sending to the printers.

Time Variability: To see the IO time variability per application, we use their CV across a day, computed from 24 hourly averages. DB and file show high time variability by their 50^{th} percentile being around 1.8. As web VMs frequently interact with users who have strong time of day patterns, web VMs exhibit time variability as high as file and DB VMs. Mail, print, and app VMs have their CV slightly higher than 1, i.e., IO activities are spread out across the day.

In summary, file VMs have the highest volume, velocity and IO load variability, but with a rather low weekly churn rate around 10%. DB VMs have high volume, velocity, IO load variability and churn rate, but with very low fullness. Mail VMs have moderate volume, and high velocity evenly across the day. All application VMs are write intensive.



(a) VM workload centroids (b) IO-CPU and IO-Net projections (c) Correlation coefficients

Figure 11: Dependency among IO [GB/h], CPU[%], Network [Mb/s].

7 Interdependency of CPU and Network

Since the statistical analysis presented here is based on the perspective of VMs and boxes, it is possible to correlate the storage workloads with those of other resources, in particular CPU and network. Using hourly averages from 04/17/2013, we capture the dependency of VM IO activities on CPU utilization and network traffic measured in megabits per second (Mb/s). We focus on the following two questions: (1) what are the most representative patterns of IO, CPU, and network usage; and (2) what is the degree of dependency among these three resources. For the first question, we use K-means clustering to find the representative VM workloads. For the second question, we use the correlation coefficients for each VM for any pair of IO, CPU, and network, and summarize their distributions.

7.1 Representative VM Workloads

When presenting the VMs' daily average IO, CPU, and network by means of a three dimensional scatter plot, there are roughly 90,000 VM points. Due to the unavoidable over-plotting, there is no obvious pattern that can be identified via visual inspection. To identify representative VM workloads, we resort to K-means clustering. Due to the lack of a priori knowledge on the number of VM clusters, we first vary the target number of clusters from 3 to 20 to observe clustering trends over an increasing number of clusters. Our results show that the overall trajectories of cluster centroids are consistent across different number of clusters. In Figure 11 (a), we present the centroids of 5 clusters. When the cluster number further increases beyond 5, more centroids appear on the line between the first two lowest centroids.

To take an IO-centric perspective, we analyze the representative VM workloads by looking at projections of VM centroids on the IO-CPU and IO-network planes, see Figure 11 (b). When looking at the IO-CPU plane, we see that IO workloads increase with CPU utilization in an exponential manner. The VM centroid with the highest IO (around 342GB/h), i.e., the rightmost point, has the highest CPU utilization (around 36%). In the IO-network plane the trend is less clear. One can observe that the first four VM centroids roughly lie on a line having their network traffic increasing at the same rate as their IO velocity. However, the last VM centroid with the highest network traffic (around 917Mb/s) has a relatively low IO activity (around 97GB/h). Overall, the majority of representative VMs have IO workloads that increase commensurately with CPU loads and network traffic, while very IO intensive VMs tend to heavily utilize the CPU but not the network.

7.1.1 Correlation Coefficients

In Figure 11 (c), we present the 10^{th}, 50^{th}, and 90^{th} percentiles of the correlation coefficients of IO-CPU, IO-network, and CPU-network. To compute correlation coefficients of the aforementioned three pairs, for each VM/box, we use three time series of 24 hourly averages: IO GB/h, CPU Utilization, and network traffic.

Among all three pairs, IO-CPU shows the highest correlation coefficients, especially for VMs. The 50^{th} percentile of the IO-CPU correlation coefficient for VMs and boxes is around 0.65 and 0.45, respectively. This indicates that IO activities closely follow CPU activities. Such an observation is consistent with the clustering results. The correlation coefficients for boxes are slightly lower than for those of VMs. Indeed, there is a certain fraction of boxes and VMs that exhibit negative dependency, and this is observed more prominently between IO and network. As for the network-CPU pair, VMs and boxes demand both resources roughly in a similar manner, supported by that fact that the correlation coefficient values are mostly above zero.

8 Conclusions

We conducted a very large scale study in virtualized, production datacenters that operate under the private cloud paradigm. We analyze traces that correspond to the activity across three years of 90,000 VMs, hosted on 8,000 physical boxes, and containing more than 22 PB of actively used storage. IO and storage activity is reported from three viewpoints: volume, velocity, and variety, i.e., we take a holistic view of the entire system but also look at individual applications. This workload characterization study differs from others from its sheer size both from observation length and number of traced systems. Yet while some of our findings confirm those reported on smaller studies, some others provide a different perspective. Overall, the degree of virtualization is identified as an important factor in perceived performance, ditto for the per application storage requirements and demand, pointing to directions to focus on for better resource management of virtualized datacenters.

Acknowledgements

We thank the anonymous referees and our shepherd, Garth Gibson, for their feedback that has greatly improved the content of this paper. This work has been partly funded by the EU Commission under the FP7 GENiC project (Grant Agreement No 608826). Evgenia Smirni has been partially supported by NSF grants CCF-0937925 and CCF-1218758, and by a William and Mary Plumeri Award.

References

[1] *Big Data Drives Big Demand for Storage* `http://www.idc.com/getdoc.jsp?containerId=prUS24069113`. 2013.

[2] Agrawal, N., Bolosky, W., Douceur, J., and Lorch, J. A five-year study of file-system metadata. In *FAST* (2007), pp. 3–3.

[3] Birke, R., Chen, L. Y., and Minkenberg, C. A datacenter network tale from a server's perspective. In *IEEE IWQoS* (2012), pp. 1–10.

[4] Birke, R., Chen, L. Y., and Smirni, E. Data centers in the cloud: A large scale performance study. In *IEEE CLOUD* (2012), pp. 336–343.

[5] Birke, R., Podzimek, A., Chen, L. Y., and Smirni, E. State-of-the-practice in data center virtualization: Toward a better understanding of vm usage. In *IEEE/IFIP DSN* (2013), pp. 1–12.

[6] Dubnicki, C., Gryz, L., Heldt, L., Kaczmarczyk, M., Kilian, W., Strzelczak, P., Szczepkowski, J., Ungureanu, C., and Welnicki, M. Hydrastor: A scalable secondary storage. In *FAST* (2009), pp. 197–210.

[7] Ellard, D., Ledlie, J., Malkani, P., and Seltzer, M. Passive NFS tracing of email and research workloads. In *FAST* (2003).

[8] Evans, C. *Dedicated VM storage emerges to meet virtualisation demands.* `http://www.computerweekly.com/feature/Dedicated-VM-storage-emerges-to-meet-virtualisation-demands`. 2013.

[9] Gribble, S. D., Manku, G. S., Roselli, D. S., Brewer, E. A., Gibson, T. J., and Miller, E. L. Self-similarity in file systems. In *SIGMETRICS* (1998), pp. 141–150.

[10] Gulati, A., Shanmuganathan, G., Ahmad, I., Waldspurger, C. A., and Uysal, M. Pesto: online storage performance management in virtualized datacenters. In *SoCC* (2011), p. 19.

[11] Gulatiand, A., Shanmuganathan, G., Zhang, X., and Varman, P. Demand based hierarchical qos using storage resource pools. In *USENIX ATC* (2012), pp. 1–14.

[12] Gunda, P. K., Ravindranath, L., Thekkath, C. A., Yu, Y., and Zhuang, L. Nectar: Automatic management of data and computation in datacenters. In *OSDI* (2010), pp. 75–88.

[13] HANSEN, J. G., AND JUL, E. Lithium: virtual machine storage for the cloud. In *SoCC* (2010), pp. 15–26.

[14] J.DOUCEUR, AND BOLOSKY, W. A large-scale study of file-system contents. In *SIGMETRICS* (1999), pp. 59–70.

[15] KOLLER, R., AND RANGASWAMI, R. I/O deduplication: utilizing content similarity to improve I/O performance. In *FAST 2010*, pp. 16–16.

[16] LE, D., HUANG, H., AND WANG, H. Characterizing datasets for data deduplication in backup applications. In *FAST* (2012), pp. 1–10.

[17] LEUNG, A. W., PASUPATHY, S., GOODSON, G. R., AND MILLER, E. L. Measurement and analysis of large-scale network file system workloads. In *USENIX ATC* (2008), pp. 213–226.

[18] LI, M., GAONKAR, S., BUTT, A. R., KENCHAMMANA, D., AND VORUGANTI, K. Cooperative storage-level de-duplication for I/O reduction in virtualized data centers. In *MASCOTS* (2012), pp. 209–218.

[19] MERRILL, D. R. *Storage economics: Four principles for reducing total cost of ownership. http://www.hds.com/assets/pdf/four_principles_for_reducing_total_cost_of_ownership.pdf.* 2009.

[20] MEYER, D., AND BOLOSKY, W. A study of practical deduplication. In *FAST* (2011), pp. 1–1.

[21] NARAYANAN, D., DONNELLY, A., THERESKA, E., ELNIKETY, S., AND ROWSTRON, A. I. T. Everest: Scaling down peak loads through i/o offloading. In *OSDI* (2008), pp. 15–28.

[22] NG, C.-H., MA, M., WONG, T.-Y., LEE, P. P. C., AND LUI, J. C. S. Live deduplication storage of virtual machine images in an open-source cloud. In *Middleware* (2011), pp. 81–100.

[23] PARK, N., AHMAD, I., AND LILJA, D. J. Romano: autonomous storage management using performance prediction in multi-tenant datacenters. In *SoCC* (2012), p. 21.

[24] PARK, N., AND LILJA, D. J. Characterizing datasets for data deduplication in backup applications. In *IISWC* (2010), pp. 1–10.

[25] ROSELLI, D., LORCH, J., AND ANDERSON, T. A comparison of file system workloads. In *USENIX ATC* (2000), pp. 41–54.

[26] ROSS, S. *A First Course in Probability*. 2004.

[27] SEHGAL, P., TARASOV, V., AND ZADOK, E. Evaluating performance and energy in file system server workloads. In *FAST* (2010), pp. 253–266.

[28] SHAMMA, M., MEYER, D., WIRES, J., IVANOVA, M., HUTCHINSON, N., AND WARFIELD, A. Capo: Recapitulating storage for virtual desktops. In *FAST* (2011), pp. 31–45.

[29] SIMPSON, N. *Building a data center cost model. http://www.burtongroup.com/Research/DocumentList.aspx?cid=49.* 2009.

[30] VOGELS, W. File system usage in windows nt 4.0. *SIGOPS Oper. Syst. Rev. 33*, 5 (Dec. 1999), 93–109.

[31] WALLACE, G., DOUGLIS, F., QIAN, H., SHILANE, P., SMALDONE, S., MARK, M. C., AND HSU, W. Characteristics of backup workloads in production systems. In *FAST* (2012), pp. 4–4.