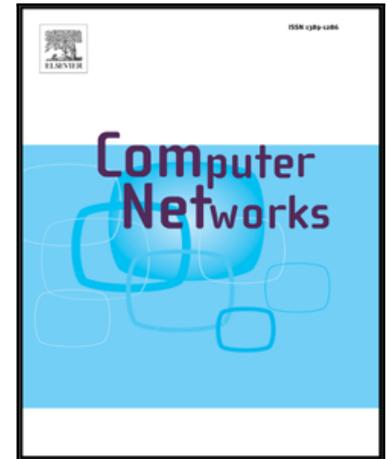


Accepted Manuscript

On the Performance of the Wide-Area Networks Interconnecting
Public-Cloud Datacenters around the Globe

Valerio Persico, Alessio Botta, Pietro Marchetta, Antonio Montieri,
Antonio Pescapé

PII: S1389-1286(16)30353-X
DOI: [10.1016/j.comnet.2016.10.013](https://doi.org/10.1016/j.comnet.2016.10.013)
Reference: COMPNW 6036



To appear in: *Computer Networks*

Received date: 29 February 2016
Revised date: 8 September 2016
Accepted date: 21 October 2016

Please cite this article as: Valerio Persico, Alessio Botta, Pietro Marchetta, Antonio Montieri, Antonio Pescapé, On the Performance of the Wide-Area Networks Interconnecting Public-Cloud Datacenters around the Globe, *Computer Networks* (2016), doi: [10.1016/j.comnet.2016.10.013](https://doi.org/10.1016/j.comnet.2016.10.013)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

On the Performance of the Wide-Area Networks Interconnecting Public-Cloud Datacenters around the Globe

Valerio Persico^a, Alessio Botta^{a,b}, Pietro Marchetta^a, Antonio Montieri^b, Antonio Pescapé^{a,b}

^aUniversity of Napoli "Federico II" (Italy)

^bNM2 s.r.l. (Italy)

Abstract

According to current usage patterns, research trends, and latest reports, the performance of the wide-area networks interconnecting geographically distributed cloud nodes (i.e. *inter-datacenter networks*) is gaining more and more interest. In this paper we leverage only active approaches—thus we do not rely on information restricted to providers—and propose a deep analysis of these infrastructures for the two public-cloud leading providers: *Amazon Web Services* and *Microsoft Azure*. Our study provides an assessment of the performance of these networks as a function of the several configuration factors under the control of the customer and evidences specific cases of particular interest. The analysis of these cases and of their root causes, also related with service fees, provides insights on their impact on both the *Quality of Service* perceived by cloud customers and the outcomes of studies neglecting them.

Our results show that Azure inter-datacenter infrastructure performs better than Amazon's in terms of throughput (+56%, on average). On the other hand, the performance of the two providers is comparable in terms of latency, with the exception of limited specific cases. Moreover, some of the configuration factors cloud customers can leverage (such as larger more expensive VM sizes, advertised to have *better* network performance) may have no effect on the inter-datacenter network performance actually perceived. Counterintuitively, lower performance may even be related to higher costs for the customer. Experimental evidences show that public-cloud providers also rely on external network providers for some geographical regions, which is the cause of lower performance and higher costs. A comparison with previous works show that TCP throughput has not been improved recently, while evidences of higher link capacities have been found.

Keywords: cloud computing, network performance, cloud networks, public clouds, inter-datacenter network

1. Introduction

Enterprise and government organizations increasingly leverage cloud solutions to supply services across the Internet, taking advantage of the ability to scale resources on demand and experimenting unprecedented opportunities in terms of ease of use, reduced costs, and higher reliability [1]. An increasing number of services and applications is now delivered through cloud-based infrastructures, and a large number of companies more and more depend on the cloud for mission critical workloads. For final consumers, this is reflected in ubiquitous access from multiple devices to content and services, delivered to almost anywhere users are located.

On-line service companies such as Amazon, Facebook, Google, Microsoft, and Yahoo! have made huge investments in *networks of datacenters* that host their on-line services and cloud platforms to cope with the increasing demand. While the complexity of these network infrastructures is completely transparent to cloud customers, the performance available to final consumers is deeply affected by it. Unfortunately, cloud

providers often provide no information about the performance a customer should expect from the cloud network, although customers could significantly benefit from details about the Quality of the Service (QoS) guaranteed [2]. In fact, all major providers grant high-performance network connectivity to their customers, but they provide no more than qualitative information about its performance, mainly due to security and commercial reasons [3, 4, 5].

Top players have made huge investments in specific technologies and cutting-edge solutions to connect distributed cloud resources and guarantee proper performance in presence of dramatically dynamic demand. Datacenter operators may purchase transit bandwidth from Telcos (usually paying based on flat or 95th percentile pricing schemes), or own dedicated lines [6]. For instance, the backbone that carries traffic between datacenters is the largest production network at Google and runs on an SDN- and OpenFlow-enabled infrastructure, in order to improve manageability, performance, utilization, and cost-efficiency of such proprietary WAN [7]. More in general, wide-area transit bandwidth costs more than building and maintaining the internal network of a datacenter [8], a topic that has recently received much more attention [9]. Networking costs are estimated to amount to around 15% of a datacenter's total worth, and are more or less equal to its power costs.

Expensive investments in this regard are further justified by

Email addresses: valerio.persico@unina.it (Valerio Persico), a.botta@unina.it (Alessio Botta), pietro.marchetta@unina.it (Pietro Marchetta), montieri@nm-2.com (Antonio Montieri), pescape@unina.it (Antonio Pescapé)

traffic trends recently estimated by sector reports [10]: (i) cloud IP traffic is going to account for a more and more significant part of the overall IP traffic, being estimated to grow at a compound annual growth rate (CAGR) of 23% from 2013 to 2018; (ii) in more details, public-cloud usage is growing faster than private one, and 31% of the cloud workloads will be in public-cloud datacenters, up from 22% in 2013; (iii) finally, traffic between datacenters is growing faster than either traffic to end-users or traffic within the datacenter, and will account for almost 9% of total datacenter traffic by 2018. The rapid growth of this traffic is due to the proliferation of cloud services, the need to shuttle data between clouds, and the growing volume of data that needs to be replicated across datacenters. The effects of this interesting trend can also be spotted in the scientific literature: novel solutions leverage the high network performance offered by public-cloud inter-datacenter WANs to develop high performance applications aimed at transferring contents (e.g., multimedia) among datacenters spread world-wide [11, 12, 13]. This recent literature further extends the range of typical usages of public-cloud inter-datacenter networks, that include bulk-data transfer or on-line content transfer (e.g., from and to storage buckets).

In this situation, however, very little information is available about performance figures offered by public-cloud networks connecting datacenters placed in different geographic regions. Few results are publicly available: public-cloud providers advertise qualitative performance indicators at most or do not disclose them at all; information provided by state-of-the-art, public-cloud monitoring services [14] currently does not include inter-datacenter performance; finally, to the best of our knowledge, the scientific community did not focus on the problem yet, and the poor preliminary results cannot be considered exhaustive. The monetary cost of the experimentations necessary for obtaining this kind of information, not directly unveiled by providers, additionally exacerbates this issue. Hence, it is hard to draw significant conclusion upon the available information. As a consequence, a customer willing to set up a multi-datacenter application in public clouds is not able to cleverly determine the provider or the regions most indicated to host it, based on the inter-datacenter network performance offered for a certain cost.

To fill the existing gap, we have extended the work in [15] and have performed a deep experimental evaluation of the inter-datacenter network of the two leading cloud providers: *Amazon Web Services* (hereafter Amazon) and *Microsoft Azure* (hereafter Azure) [16]. These two providers are a valuable and interesting case of study, as they represent the 40% of the global cloud market [17]. Although our work focused on Amazon and Azure, the proposed methodology is general and therefore it is generally applicable. Indeed, all our experimentations did not rely on providers' support, i.e., have been fulfilled with *non-cooperative* approaches, by adopting the point of view of a generic cloud customer [18]. Note how enforcing these approaches generates a number of non-trivial challenges [19], although it provides a number of advantages not achievable otherwise (e.g., freeing the customers from the will of the provider

to expose detailed information, thus avoiding potential conflicts of interests). The above challenges include the proper identification of the experimental scenarios as well as the accurate prediction of the related experimental costs to fit budget constraints [18]. Above all, non-cooperative approaches have to face the lack of visibility over both the design and the implementation of the cloud network infrastructures that makes the interpretation of the experimental results harder. Addressing these issues, we have collected performance data about network paths interconnecting public-cloud datacenters of the above-cited providers, leveraging active approaches for approximately 800 hours, taking into account a set of geographic regions hosting datacenters for both the providers.

To the best of our knowledge, our work extends the literature in a number of aspects, as reported in the following.

- Experimental results investigate the impact on network performance of several configuration factors under customer control, such as the cloud provider, the region, and the size of the virtual machines.
- Our work is able to depict a clear picture of the inter-datacenter network performance in terms of network throughput and latency, for the two leading public-cloud providers.
- Our analysis provides insights into the communication infrastructure leveraged by cloud providers, also showing the existence of phenomena generated by the management strategies which may impact both the performance experienced by the customers and the results of research investigating these networks.
- Performance results have also been compared to provider-imposed fees, in order to give useful guidelines to customers willing to deploy distributed applications onto the cloud.
- Empirical outcomes confirmed that providers often rely on their own dedicated infrastructure to connect geographically distributed sites, but also show that, in some circumstances, they depend on third-party networks, being forced to provide cloud customers with lower performance at higher costs.
- Our results are compared to those found in previous work, highlighting the changes in terms of performance figures, and analyzing the trend that these infrastructures are subjected to over time.

The paper is organized as follows: Sec. 2 surveys the most-related literature and positions the paper accordingly; Sec. 3 details the methodology adopted for the analysis; Sec. 4 presents the dataset we refer to in our analysis; Sec. 5 shows the results of our work for the two providers taken into account; Sec. 6 ends the paper with the concluding remarks.

2. Related work

Traffic in public-cloud inter-datacenter networks is rapidly growing, as estimated by recent reports [10]. Moreover, novel applications are critically relying on it [11, 12, 13]. Unfortunately limited and poor information is available today about the performance attained and attainable by this traffic.

Most of the works in the literature aimed at providing a broad characterization of the performance of public clouds, i.e., did not directly focus on network performance. Some of these works benchmarked also intra-datacenter network performance—i.e. the performance of the network interconnecting cloud resources deployed within the same site—with different purposes and often providing conflicting results [5, 20, 21, 22, 23]. Two recent works by some of the authors of this paper, tried to shed light on this aspect, considering both Amazon and Azure, and emphasizing the importance of well-defined methodologies [19, 24].

A limited set of works took into account the performance of inter-datacenter networks. Chen *et al.* [25] focused on the interplay of multiple datacenters performing a passive analysis on Yahoo! network flow dataset. Li *et al.* [26] in their broader analysis, benchmarked also cloud inter-datacenter networks, but considering only TCP throughput performance between two US datacenters. They adopted general purpose instances of just one size for their experimentations. They found that the throughput across datacenters is much smaller than the one within the datacenter for all the providers considered, and that both Amazon and Azure show median values of inter-datacenter TCP throughput larger than 200 Mbps.¹ Moreover, they reported a variation of throughput across datacenters larger than the one measured within the same datacenter. Feng *et al.* [12] performed an experimental evaluation of Amazon network paths interconnecting seven different datacenters to support their study on a set of algorithms to minimize operational costs of inter-datacenter video traffic. They considered datacenters in North California, Oregon, Virginia, Sao Paulo, Ireland, Singapore, and Tokyo. They used medium instances and monitored network performance for only 3 minutes. Their results revealed very different throughput values, ranging from 9.6 Mbps (for the path between Sao Paulo and Singapore) up to 545.1 Mbps (for the path between North California and Oregon). Values of end-to-end latency measured were lower than 587.3 ms for the 90% of the paths, while the average was about 349.1 ms. The same authors [11] proposed a protocol to deliver packets in video conferencing, designed for the inter-datacenter network, and tailored to the needs of a cloud-based service. Measurement results supporting this work and obtained adopting small instances placed in the regions reported above ranged from 20.9 Mbps to 130.8 Mbps for throughput and from 11.3 ms to 441.7 ms for latency. Finally, Garcia-Dorado and Rao [27] presented a framework that exploits cloud-pricing schemes to construct overlay distribution networks for bulk-data transfer that proved to be effective

¹Note that the authors used labels instead of names to identify the different providers. We inferred Amazon and Azure looking at the different geographical locations of the datacenters.

from the customer-side perspective. To evaluate their approach they conducted experimentations on both Amazon and Azure, leveraging medium VMs and measuring TCP bandwidth and latency performing two-minute-long measurements during one day. They found low variation of throughput values, especially in paths exposing better performance.

Our work significantly differs from the others in recent literature dealing with the performance of the inter-datacenter networks. In spite of the analysis presented in [25], our work does not rely on provider-restricted information. The methodology we propose is completely based on active measurements performed from the point of view of the cloud customer: therefore it is independent of provider's will to disclose information, guarantees the results to be independent from it, and allows to replicate the study at any time. Differently from the analysis in [26], our work explicitly focuses on the performance of inter-datacenter networks. This gives the opportunity of deepening the aspects strictly related to the measurement process. Thanks to this, we investigate also the interesting traffic engineering practices and their impact on both the measurement process and the QoS perceived by cloud customers. Finally, with respect to the measurement data presented in [11], [12], and [27], our study is more systematic, details a repeatable methodology, compares the performance of multiple providers, and takes into account the specific management strategy they enforce.

3. Methodology

Characterizing public-cloud network performance may be very challenging. In this section, we describe the choices we have made to deal with this complexity. We first detail the reference architecture as well as the settings and tools we have adopted (Sec. 3.1). Then we define the factors of interest to unambiguously identify the scenarios considered for the collection of the dataset (Sec. 3.2). Our intent is to ease as much as possible the understanding of the precise conditions in which we perform the analysis and foster its replication in other scenarios, for other cloud providers, or along the time.

3.1. Reference architecture

In this work we aim at measuring the performance of the network paths interconnecting instances (i.e., virtual machines or VMs) deployed onto geographically distributed public-cloud sites. According to the reference architecture reported in Fig. 1, the traffic directed from one side of the communication to the other traverses different and distinct layers. The traffic generated by a sender VM normally traverses (i) the devices composing the intra-datacenter, high-performance network at sender side first. Then, it enters and traverses (ii) the inter-datacenter WAN, and, before being delivered to the receiver VM, it passes through (iii) the intra-datacenter network at receiver side. Note that the internals of both intra- and inter-datacenter networks are out of our knowledge, as we adopt the point of view of the general customer. In fact, our approach is aimed at measuring the performance experimented by customers' traffic. Although

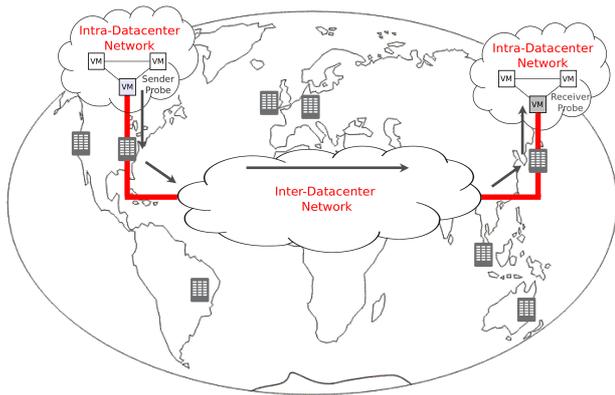


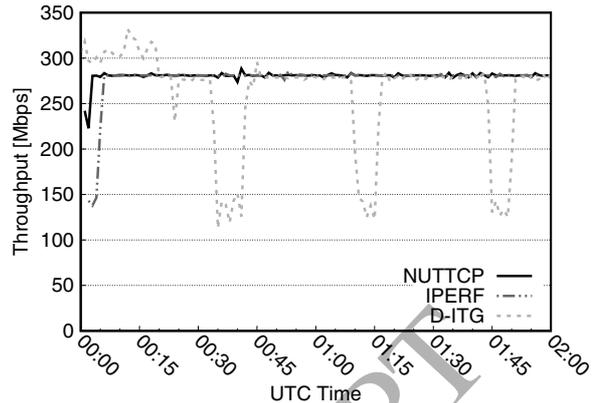
Figure 1: Reference architecture. Before being delivered to the receiver VM, synthetic traffic traverses different layers, i.e. the intra-datcenter network at sender side, the inter-datcenter network, and the intra-datcenter network at receiver side.

different layers exist, the inter-datcenter network performance is assumed to be the bottleneck of the communication due to practical, technological, and physical limitations. Our results have therefore to be intended as related to these networks, if not stated otherwise.

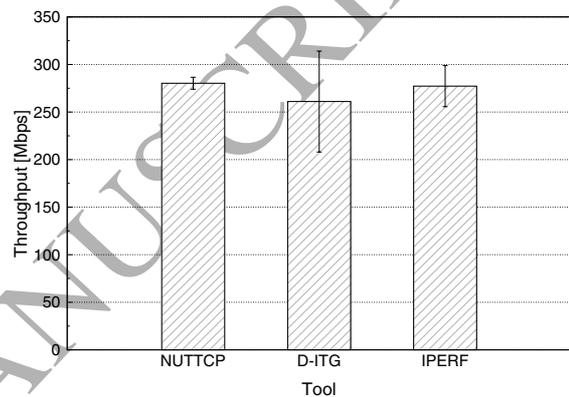
For our experimentation, geographically distributed VMs have been instrumented with Ubuntu 14.04 operating system and all the necessary measurement and diagnostic tools needed for estimating the network performance. As already done in previous works [19, 24], we have used the network measurement tool named *nuttcp* [28] to inject synthetic traffic into the network from a sender to a receiver VM. This tool allows us to measure the raw UDP and TCP application layer throughput and latency, transferring memory buffers from the sender to the receiver VM.

We chose *nuttcp* after an initial experimental campaign comparing the most widely used similar tools. We found that *nuttcp* was able to stably generate synthetic traffic in the virtualized environments taken into account, thus fitting well the requirements of our analyses. An example of this comparison is reported in Fig. 2. Fig. 2a shows how the measured throughput varies over two minutes for the different tools tested in the Amazon virtualized environment. The figure is related to the case of an intra-datcenter measurement within a US datacenter. Fig. 2b provides additional information showing the average throughput achievable over a 5-minute-long experiment and the related variability. We point the reader to [19] for more details about the issues related to the generation of synthetic traffic in cloud environments. While we chose *nuttcp* since it reported results in agreement (on average) with other tools tested, we are currently investigating the root causes that lead to the higher traffic generation instability observed with D-ITG in cloud virtualized environments.

The reference architecture described above has been modified in different ways, according to factors described in the following section, giving birth to a set of scenarios of interest.



(a) Time series.



(b) Maximum achievable throughput measured over a 5-minute-long experiment (mean and standard deviation of 1-second samples). Nuttcp is able to stably generate synthetic traffic, exhibiting a standard deviation markedly lower than iPerf and D-ITG.

Figure 2: Measuring intra-datcenter network throughput with different tools led to similar results (Amazon M-sized VMs, TCP traffic).

Table 1: Summary of factors and considered values.

Factors	Values
Provider	Amazon, Azure
Region	Europe (EU), North Virginia (US), South America (SA), Asia-Pacific (AP)
VM Size	medium (M), extra-large (XL)
Transport Protocol	TCP, UDP

Table 2: Cost for transferring data to another region, as of Sep.'15. Only for Azure, data-transfer costs scale with volume.

Region	Amazon (€/GB)	Azure (€/GB)
EU	0.02	0.0734–0.0422
US	0.02	0.0734–0.0422
SA	0.16	0.1527–0.1350
AP	0.09	0.1164–0.1012

3.2. Factors for identifying scenarios of interest

The inter-datacenter network performance can be measured and monitored in different scenarios when adopting the point of view of the general customer. In fact, a set of factors exists that may heavily influence the perceived performance, as confirmed by the outcome of our experimentations. In this perspective, our work significantly extends the surveyed literature carefully analyzing the impact of these factors. Without claiming to be exhaustive, we believe an important contribution of our work is the identification of the factors to be carefully taken into account when performing similar analyses. These factors—summarized in Tab. 1—will be shortly discussed in the following.

According to latest reports about public IaaS cloud computing [16], the market is dominated by only a few global providers among the huge number of offers. In this work we take into account the IaaS of two **providers**: EC2 for Amazon [29] and *virtual machines* for Azure [30]. The former is the clear market leader (over a million active customers in more than 190 countries), while the latter is the only clear challenger, also due to the continual investments in the latest infrastructure technologies. Our reference architecture reported in Sec. 3.1 is general—thus making this experimental methodology applicable to any IaaS provider—as the comparison with other minor providers would help to broaden the proposed picture of these high performance inter-datacenter infrastructures. We purposely deepen our analysis for the two leading public-cloud providers, to obtain a clear and detailed picture of the two popular cloud offers largely adopted by most of the customers. Both the considered providers are steadily expanding their global infrastructure whose growth is backed by billion investments: infrastructural expansion is claimed to be a priority because of the direct benefits generated for the customers. As of today, Amazon (Azure) has datacenters in 11 (17) regions around the world. For our experimental campaigns we have identified 4 **geographic regions**, where both providers have deployed datacenters: Ireland (hereafter EU), North Virginia (US), Sao Paulo (SA), and Singapore (AP). Being forced to select a subset of all possible regions—mainly due to cost constraints—we have picked a region per continent, in order to ensure geographical diversity to our dataset. We investigated the network performance of all the paths connecting all the regions in this selection. Hereafter, we will adopt the notation $A \rightarrow B$ to refer to the path from region A to region B. $A \leftrightarrow B$ will be used to refer to both the paths $A \rightarrow B$ and $B \rightarrow A$ at once—i.e., when both directions are taken into account. Note that traffic moving outside from a region is subjected to costs that vary with such source region (see Tab. 2). Amazon customers can further choose an **availability zone** once a region has been selected, i.e. a specific, independent, and isolated location inside the chosen region. In our study we have taken into account also the impact of different availability zones inside a region.

Customers can then choose **VM type and size** for both providers. The VM type indicates a family of VMs optimized for a given task (storage, computation, etc.). Once the type is selected, the customer can decide the size of the VM to further specify storage and computation capabilities. Both VM

Table 3: Selected sizes and details (price may vary with regions).

	VM size	Type and Size	CPU cores	RAM (GB)	Network Performance	Min-Max Hourly Cost (€/h)
Amazon	M	m3.medium	1	3.75	Moderate	0.070 - 0.098
	XL	m3.xlarge	4	15	High	0.280 - 0.392
Azure	M	A2	2	3.5	n/a	0.1192 - 0.1460
	XL	A4	8	14	n/a	0.4767 - 0.5839

type and size influence the hourly cost of the VM. In terms of type, we used last-generation, general-purpose VMs for both providers. In terms of size, we considered two different ones named *m3.medium* and *m3.xlarge* for Amazon, and *A2* and *A4* for Azure. Hereafter we simply refer to them as *medium* (M), and *extra-large* (XL), respectively, for both providers. Tab. 3 reports further characteristics and the costs of the VMs adopted in our analyses. Note that both providers provide details only regarding RAM and CPU. Regarding the network characteristics, Amazon only provides a qualitative description of the expected performance, Azure completely hides this information.

Finally, we have considered the two most popular **L4 protocols** in our experiments: UDP and TCP. UDP is useful to analyze the performance of the raw IP traffic, as it adds no closed loop-control and leaves the complete control to the application, no matter of the state of the network. On the other hand, TCP is governed by flow and congestion control, and makes the generated traffic subjected to the status of the network path. Therefore it provides information on the performance the numerous TCP-based applications will experiment.

4. Experimental Dataset

In this section we describe how the factors introduced above have been combined to collect our dataset, whose characteristics are detailed in the following.

According to the non-cooperative approach we have proposed, we did not rely on provider-restricted information. All our analysis is based on the data collected between March and November 2015 adopting the methodology introduced in Sec. 3. The collected process required more than 790 hours of traffic generation (i.e. we have injected synthetic traffic into the inter-datacenter network for more than 790 hours). We considered the 12 combinations of the four regions selected for each provider. Experiments have been run between VMs of the same size (M or XL). Repeated, 5-minute-long experiments have been performed in the same conditions, equally spaced in 24-hour intervals. According to the presence of multiple availability zones in each region for Amazon, we run around 8.6K and 880 5-minute-long experiments for Amazon and Azure, respectively. In line with experimental needs, different availability zones have been tested in parallel. Beside performance measures, path information about each scenario has also been collected to complement the view on the performance.

Tab. 4 reports all the details about the dataset collected, showing both the cumulative and the detailed information

Table 4: Experimental dataset details. For both providers, the actual experimental duration, the number of 5-minute-long experimental samples, the start date of the experimental campaign, and its overall duration are detailed for every combination of the factors taken into account (i.e. VM size, cloud regions, and L4 protocol). The *Duration* and *Experimental Duration* fields show the overall time span of the experimental campaign and the actual generation time in hours. Note that the *Experimental Duration* may be greater than the overall *Duration*, since the experiments related to different AZs have been run in parallel.

(a) Amazon

VM size	Regions	TCP				UDP			
		Exp. Duration	#Samples	Start Date	Duration	Exp. Duration	#Samples	Start Date	Duration
M	EU-US	48	576	3 Apr	25.76	48	576	19 Apr	27.55
	US-SA	48	576	4 Mar	25.78	48	576	11 Apr	27.40
	EU-SA	48	576	7 Mar	25.78	0.53	4	15 Mar	25.78
	EU-AP	48	576	31 Mar	25.82	48	576	11 Apr	27.21
	US-AP	48	576	7 Apr	25.77	0.53	4	15 Mar	25.36
	SA-AP	48	576	5 Mar	25.90	0.53	4	14 Mar	25.53
	<i>Total</i>		288	3456		154.81	145.59	1740	158.83
XL	EU-US	48	576	15 Apr	25.77	0.53	4	14 Mar	25.61
	US-SA	48	576	15 Apr	25.77	0.53	4	15 Mar	25.33
	EU-SA	48	576	16 Apr	25.83	0.53	4	15 Mar	25.41
	EU-AP	48	576	14 Apr	25.84	0.53	4	14 Mar	25.77
	US-AP	48	576	17 Apr	25.90	0.53	4	15 Mar	25.52
	SA-AP	48	576	18 Apr	25.93	0.53	4	14 Mar	25.35
	<i>Total</i>		288	3456		155.04	3.18	36	152.99

(b) Azure

VM size	Regions	TCP				UDP			
		Exp. Duration	#Samples	Start	Duration	Exp. Duration	#Samples	Start	Duration
M	EU-US	16	192	22 May	25.76	5.8	70	25 May	23.73
	US-SA	4	48	29 May	23.73	0.5	6	3 Nov	2.17
	EU-SA	4	48	30 May	23.64	0.5	6	3 Nov	2.17
	EU-AP	4	48	29 May	23.72	0.5	6	3 Nov	2.17
	US-AP	4	48	30 May	23.70	0.5	6	3 Nov	2.17
	SA-AP	4	48	23 May	23.73	3.2	38	26 May	18.61
	<i>Total</i>		36	432		144.28	11	132	51.02
XL	EU-US	3.3	40	24 May	19.60	0.5	6	27 May	4.50
	US-SA	4	48	6 Jun	23.71	0.5	6	2 Nov	2.16
	EU-SA	4	48	7 Jun	23.67	0.5	6	2 Nov	2.16
	EU-AP	4	48	1 Jun	23.72	0.5	6	2 Nov	2.16
	US-AP	4	48	2 Jun	23.71	0.5	6	2 Nov	2.16
	SA-AP	4	48	24 May	23.74	0.5	6	27 May	4.50
	<i>Total</i>		23.3	280		138.15	3	36	17.64

(number of samples, start date, actual and overall duration) about each of the scenarios taken into account. Our experiments have been subjected to providers' fees, and according to their terms of service, inter-datacenter traffic is subjected to volume-based charging (see Tab. 2). We limited the number of UDP runs—especially for larger VM sizes—because of budget constraints and the high volumes transferred during experiments.

It is worth noting that we publicly release the entire dataset, to foster further analyses and replication, and to support longi-

tudinal studies.²

5. Inter-datacenter network performance

We discuss in this section the most interesting results stemming out from the analysis of data we have collected. Firstly, we provide an assessment of the performance of the network paths interconnecting geographically distributed cloud sites for the two providers taken into account, comparing their performance, and also showing how it is influenced by a set of fac-

²<http://traffic.comics.unina.it/cloud>.

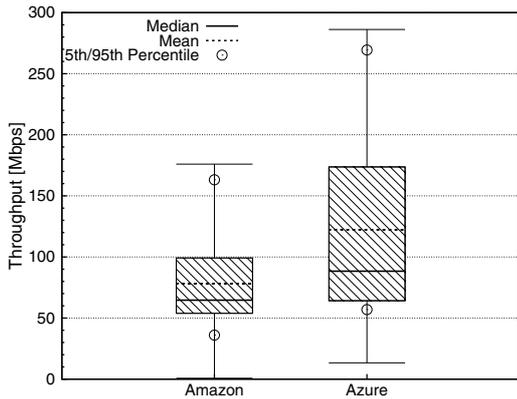


Figure 3: TCP throughput distribution across different regions. Each sample represents the mean of a 5-minute-long experiment. Azure performs better on average (+56%).

tors. This broad analysis provides useful information for customers willing to deploy distributed applications onto the cloud, and helps them with their setup choices, according to application needs. Moreover, this analysis provides insights useful for researchers and practitioners willing to perform similar analyses. Secondly, we further deepen the analysis for the interesting cases, providing insights into the infrastructures leveraged for inter-datacenter communications.

In more details, TCP and UDP throughput performance is discussed in Sec. 5.1 and Sec. 5.2, respectively. Sec. 5.3 provides more details about the variability of the throughput observed. In Sec. 5.4 we discuss how fees and performance are related. The performance in terms of latency is analyzed in Sec. 5.5. Finally the impact of choosing different availability zones is deepened in Sec. 5.6.

5.1. TCP throughput

Fig. 3 reports an overall picture of the throughput for both providers in all the experiments (see Tab. 4). Each sample considered in the plot is the mean value of a 5-minute-long experiment. It is worth noting that even the values placed far from the global average (i.e. the dashed line in Fig. 3) well represent the instantaneous values measured during that particular 5-minute measurement. In fact, the *coefficient of variation*³ (*CoV*) within each experiment is very low: the 95th percentile of its distribution along all the experiments is about 0.2. Fig. 3 shows that Azure inter-datacenter network performs better than Amazon's one in terms of network throughput, achieving TCP throughput values 56% higher (78.2 Mbps vs. 122.2 Mbps, on average). Almost the same proportion is kept when considering maximum throughput achieved (286.2 Mbps vs 176.0 Mbps). Interestingly, only 25% of samples collected on Amazon's infrastructure have values higher than 99 Mbps, while 95% of samples collected on Azure's one have values higher than 57 Mbps.

³ $CoV(X) = \frac{\sigma}{\mu}$, where X is a set of experimental samples, σ is its standard deviation, and μ is its mean value.

Finally, TCP throughput values for Amazon can be as small as 1 Mbps, while for Azure they are never smaller than 13 Mbps.

Fig. 4 provides a breakdown of the performance of TCP. The bar chart reports mean and standard deviation values across different regions and different VM sizes, for the two providers. A few important observations can be made from this figure. We can immediately observe the large performance differences across different regions, up to about 80% in the worst case. Interestingly, ordering the regions according to the achievable throughput, we obtain the same ranking for the two providers, with the only exception of US↔AP pair, which performs better than EU↔AP pair for Azure, on average. The achievable TCP throughput is not clearly affected by the size of the VM, for both providers, despite the different fees imposed. Note that small performance differences between differing sizes are observed. But, they are not always biased towards the larger, and they are always associated to higher variability. The standard deviation inside a region pair is normally very low, although some pairs show a higher throughput variability only for Azure XL VMs (e.g., SA→US, US→SA, and SA→EU).

Recall that M and XL VMs are advertised to have *Moderate* and *High* network performance respectively, according to Amazon's documentation. Our results show that differing performance figures across different VM sizes, highlighted in a previous work [19], are achievable only by communications that involve VMs both deployed inside the same region, while they do not hold for inter-datacenter TCP performance. This is likely due to TCP dynamics.

Finally, performance figures appear to be roughly symmetric in the majority of the cases examined, although during our experimentations we also encountered severe degradations involving only one direction of the communication. Fig. 5 reports some of these interesting cases. As shown, intermittent but heavy performance degradation—exposing throughput settling down to less than 10 Mbps—has been observed between AP and EU for Amazon and between SA and EU for Azure. Interestingly, although referring to different days and providers, these results are both related to only one direction of the communication, i.e. the downlink of VM inside EU.

We believe this broad assessment can be very useful to cloud customers willing to draw upon public clouds to deploy their distributed architectures. Thanks to these results, customers can wisely select among regions, when possible. Otherwise, this analysis provides them with a quantification of the significant network performance differences among regions. Moreover, relying on larger sizes to increase TCP inter-datacenter performances has not effect. Also, comparing the two providers we found that Azure performs better on average, while asking higher costs for the VMs and for the data transfer. Due to this trade-off, we believe that the choice of the provider should be tailored according to regions of interest and should be driven by the specific characteristics of the application. Finally, considerations about performance symmetry may also be considered to properly place nodes in the different regions, according to the specific application the inter-datacenter network is leveraged for, and to the different roles of the counterparts involved in a communication.

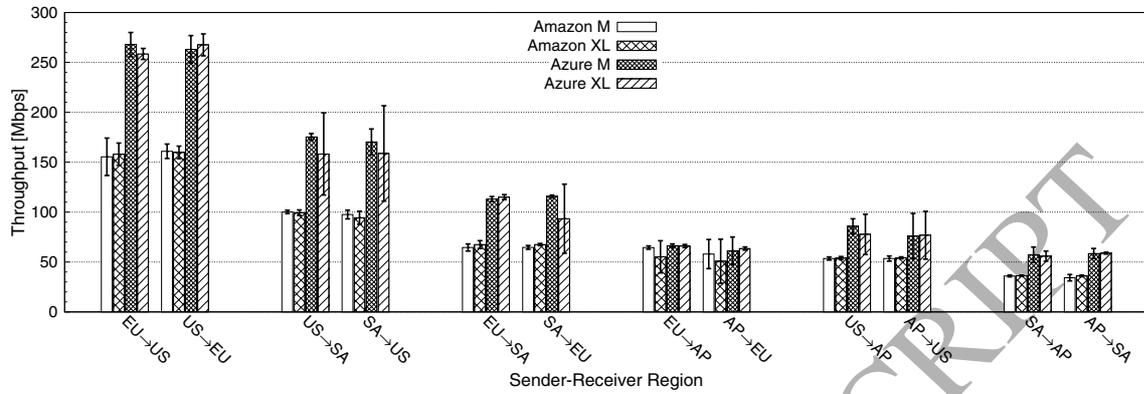


Figure 4: TCP throughput breakdown on different region pairs (mean and standard deviation), for different providers and VM sizes. For both providers, region pair is the factor with the highest impact, while VM size appears non influential.

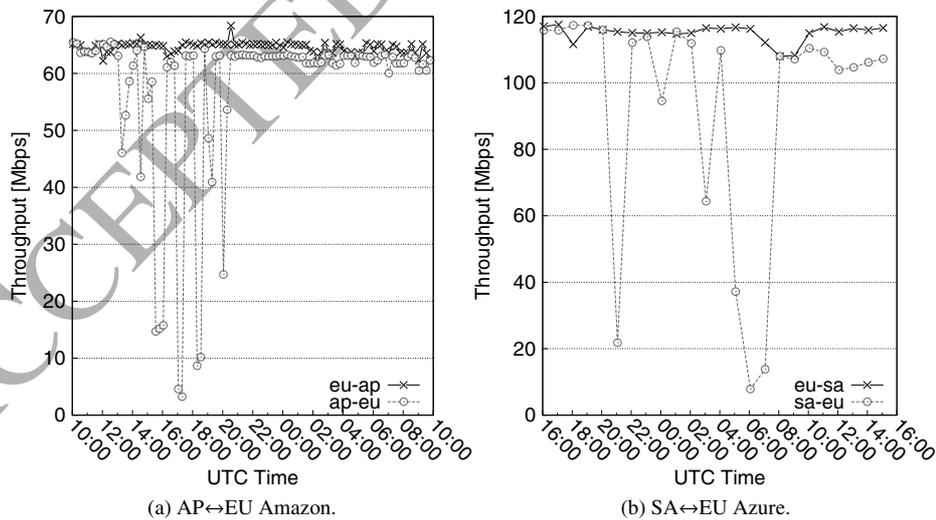


Figure 5: Relevant examples of performance asymmetry for different directions.

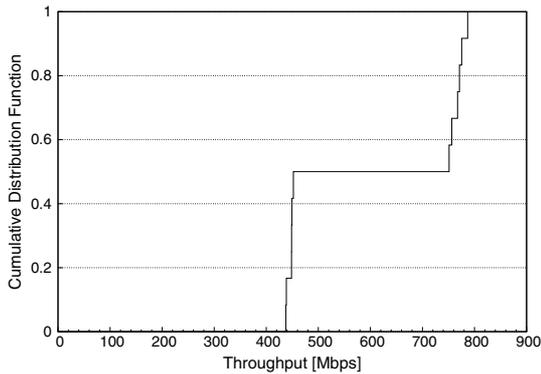
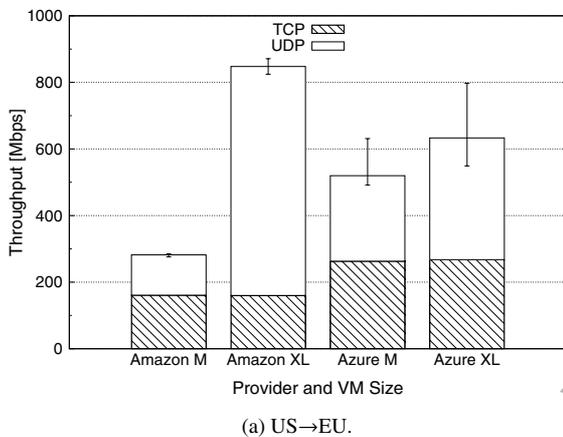
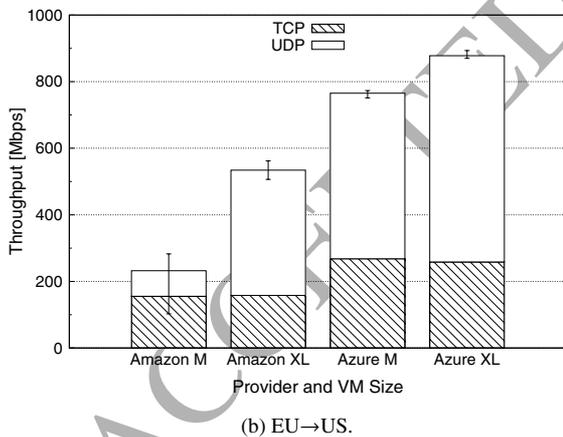


Figure 6: Empirical cumulative distribution for UDP throughput (Azure, XL VMs, EU→US). Each sample is the average of a 5-minute-long experiment. VMs have been terminated and recreated at each experiment.



(a) US→EU.



(b) EU→US.

Figure 7: TCP and UDP inter-datacenter average throughput for US↔EU with M and XL VMs (whiskers report maximum and minimum). While VM size does not affect the achievable TCP throughput, UDP is able to reach higher end-to-end performance, giving evidence of path capacities as large as more than 800 Mbps.

In general, our analysis revealed TCP throughput values smaller than the ones reported in previous works. Indeed, authors of [26] observed TCP (median) throughput higher than 200 Mbps for both Amazon and Azure. Results are hard to

compare because the authors disclosed no information about the VM size adopted for the experimentations and restricted inter-datacenter throughput analyses to only two regions placed in the same continent (United States). Several interpretations are available for this discrepancy, therefore. The different performance figures may be explained by the fact that experimentations in [26] involved datacenters placed at geographical sites closer to each other and hence backed by infrastructures implementing technologies guaranteeing different performance levels. Another potential cause is related to the presence of less competing traffic across the inter-datacenter networks at the time when experimentations conducted in [26] were performed, i.e. around 5 years before ours [10]. Interestingly, the performance reduction observed after these 5 years is substantially larger for Amazon than for Azure. This could be further justified by different catchment areas for the two providers, considering also the impact a higher number of users on TCP congestion-control dynamics. Finally, the authors of [26] also found throughput variability markedly higher than ours. This observation holds although the analysis in [26] refers to data collected over a more limited observation period (one single day) and is related to measurements between two datacenters both placed in the US. This reduced variability is in line with the increase of the competing traffic already hypothesized above.

5.2. UDP throughput and end-to-end path capacity

Before digging into the details of this analysis and of the related results, we show how UDP measurement accuracy may be biased by specific phenomena related to the resource management enforced by providers. Properly understanding the impact of these phenomena allows also to improve the quality of the analysis. In more details, we found how UDP throughput values measured for Azure inter-datacenter paths can be impacted by intra-datacenter limitations already documented [24]. TCP performance analysis is immune from these limitations instead, since the end-to-end bottleneck located along the paths proved to be tighter than them. Fig. 6 shows the distribution of the UDP throughput measured in different 5-minute-long experiments between two XL VMs (EU→US). The value of the end-to-end throughput measured along the path between the same pair of regions settles to two well-defined values when adopting different VMs pairs: 450 Mbps and 800 Mbps. The former appears to be not related to inter-datacenter performance as it reflects intra-datacenter limitations. Note that the measured value does not change if the VMs involved in the measurement process are not terminated and then launched again from scratch [24]. In the specific case shown, we encountered the lower value 50% of the times: this may induce to heavily under-estimate the maximum end-to-end UDP throughput if not properly taken into account. Although the influence of intra-datacenter limitation cannot be avoided *a priori*, restarting the VMs involved in the measurement process helps refine the estimation done. Note that the phenomenon described, although not representing a peculiarity of the inter-datacenter network, can impact the performance transparently perceived by the cloud customer. Since this work is aimed at characterizing the inter-datacenter network, we did not focus on the intra-

datacenter limitations already investigated in [24]. With this aim, we purposely iterated the VM termination and recreation process to bypass intra-datacenter limitations and unveil the peculiarities of the inter-datacenter network. We believe that the one adopted represents a good practice to adopt when performing this kind of analyses for Azure infrastructure. We can now analyze the obtained results. In summary, they show that the UDP throughput proved to be significantly higher than TCP's for all the source-destination pairs considered. Although this result was expected (as the UDP protocol is not subjected to congestion control dynamics typical of TCP), interestingly we have identified cases in which UDP inter-datacenter throughput durably reaches the intra-datacenter performance figures reported in [19] and [24]. In these cases, the UDP throughput appears mainly limited by the bottlenecks imposed by providers at source.

In detail, Fig. 7 compares UDP and TCP average throughput obtained between the pair of regions with the best performance for both providers: US and EU. While TCP performance does not vary with the size of the VMs, UDP throughput reaches much larger values. Note also how UDP maximum values are compatible with limitations imposed by providers at source and based on VM size. On the one hand, these results suggest that worse TCP performance is determined by network congestion across datacenters. On the other hand, the better performance of UDP gives evidence of the network capacity of the inter-datacenter paths. We can find further justifications for this empirical result considering the impact of the higher number of customers Amazon has on TCP congestion control dynamics. Although VMs are allowed to inject traffic into the inter-datacenter network at a rate as high as the throughput measured in [19] and [24], and the inter-datacenter network is able to deliver traffic at this speed, network congestion represents the main bottleneck when relying on TCP. This result is generalizable across different regions, even if the actual UDP throughput values change from case to case.

We have also estimated the capacity of each path as described in the following to better understand how network resources are deployed across regions. We calculated the average throughput over 5-second-long non-overlapping windows considering the timeseries of each UDP experiment involving XL VMs in the dataset. This approach allowed us to mitigate the effect of throughput spikes—potentially caused by queuing and buffering dynamics generated by the coexistence of multiple layers of virtualization—that could exceed end-to-end capacity [31]. We then estimated the end-to-end capacity of each path extracting the maximum of this time series. Note that the values obtained with the approach proposed, represent a lower bound of the capacity, i.e. the end-to-end path that a user can leverage is able to deliver *at least* traffic at this throughput. The results are shown in Fig. 8: Fig. 8a and Fig. 8b show the lower bounds for Amazon and Azure inter-datacenter paths, respectively. Estimated capacities are always larger than 800 Mbps for most cases. More in details, lower bounds for Amazon capacities ranged from around 560 Mbps to 986 Mbps, with only two cases exposing values smaller than 800 Mbps. For what concerns Azure, estimated capacities were lower on average: they

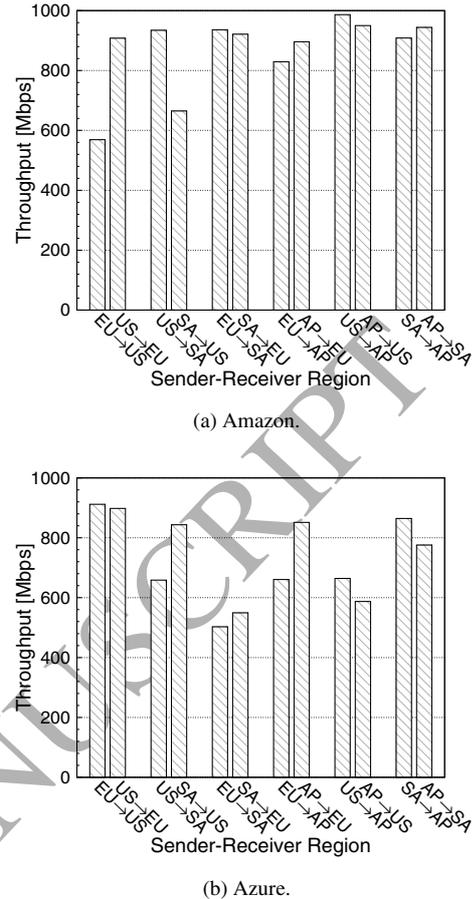


Figure 8: Lower bounds of path capacity for inter-datacenter paths, measured as the maximum UDP throughput achievable between two regions.

ranged from 502 Mbps to 912 Mbps.

Interestingly, Amazon capacity values reported in Fig. 8a result significantly larger than the homologous reported in [12], that are always smaller than 300 Mbps. Fig. 9 compares these values. As shown in the figure, EU→US surprisingly exposed the largest value according to the analysis performed in [12] but the lowest in ours. The results of our analysis show larger values than previous ones also limiting our dataset to data related to M-sized VMs (whose UDP throughput values reach 280 Mbps in all the circumstances considered).

These discrepancies could be justified by substantial infrastructure enhancements. However, we believe that they could also be impacted by the measurement methodology adopted in [12] (single 3-minute-long experiments leveraging M-sized VMs have been run). This further highlights the need to detail the adopted methodology, to compare and validate results through further analyses.

5.3. Throughput variability

Our further analyses have shown that inter-datacenter performance for both providers may not be simply regulated by limitations imposed at source: restrictions can be imposed by

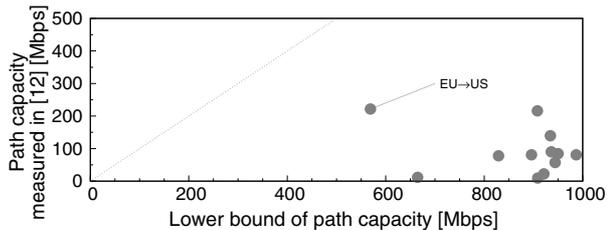


Figure 9: Path-capacity comparison. Lower bounds estimated in this work are higher than values reported in one of the previous [12].

providers also along the path, also because of the several layers traversed, none of which is under the direct control of the cloud customer. Our dataset revealed some interesting cases in which UDP throughput measured over time is not stable. They are mainly related to Amazon M for EU→US and Azure M and XL for US→EU. This inflates the variance of the throughput, as reported by the larger error bars in Fig. 7. We discuss these representative examples in the following.

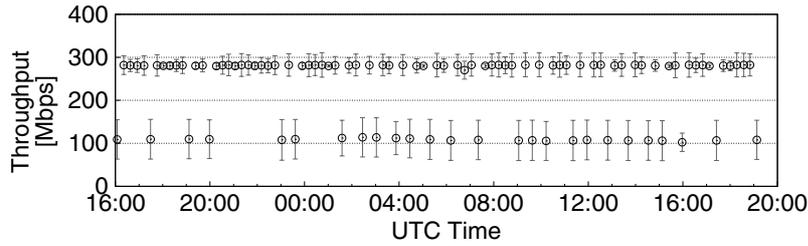
For what concerns Amazon VMs, we found a number of interesting cases in which UDP throughput is not stable, showing substantial changes in the perceived path capacity over the time. An example is related to the path connecting M-sized VMs between EU and US regions, whose minimum, average, and maximum values are reported in Fig. 7b. Deeper details for the spotted phenomenon are shown in Fig. 10. UDP throughput dramatically changes its value during the time—but not during the same experiment (Fig. 10a). It switches between two well-defined values around 280 Mbps and 110 Mbps (Fig. 10b). While the former is in line with maximum performance achievable by M-sized Amazon VMs, the latter reflects a clear and systematic path capacity decrease. This result is in line with management practices commonly adopted [7] (current network technologies, such as SDN, allow to change network configuration on the fly, based on system state and needs). On the other hand, it can also be explained considering the multiple paths that we have identified between these two regions. We uncovered them using state-of-the-art technologies such as *Paris Traceroute Multipath Detection Algorithm (MDA)* [32]. The characteristics of the path inferred by adopting MDA for the case examined are reported in Fig. 10c. Each node in the graph represents a unique IP address discovered by the tool along the path from the source VM (S) to the destination VM (D). Stars represent anonymous hops, i.e. associated to devices whose ICMP error messages did not reach S . According also to common datacenter topologies [9], they can be mapped to IP nodes placed inside the intra-datacenter network, and the root cause of their appearance is likely related to the filtering of incoming ICMP messages enforced at the border of the US datacenter. Edges in the graph connect nodes discovered by leveraging the same traffic flow. As shown in the figure the sole node appearing at the 5th hop acts as a loadbalancer at IP level, systematically distributing incoming traffic among different interfaces (hops 6, 7, and 8). The observed characteristics are confirmed also by the router-level graph [33, 34] obtained through alias resolution—using state-of-the-art techniques [35, 36]—applied

to the output of MDA. Note how in this case performance variation occurs only across different experiments—even though no termination and recreation have been enforced—while within each of them, UDP throughput settles to a well-defined value for all the duration of the experiment.

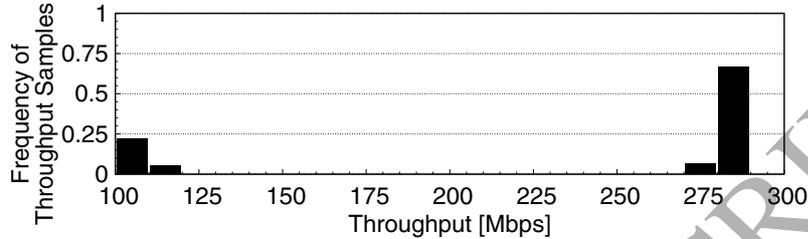
On the other hand, the performance variability observed for the paths interconnecting Azure M or XL VMs between US and EU regions exposes a significantly different nature. Indeed, the variability spotted is due to performance variation *within* each of the 5-minute-long experiments. Fig. 11 shows some experimental evidences for this phenomenon. Considering three experiments related to these two regions, Fig. 11a shows how the throughput typically varies within each experiment: interestingly, two well-defined throughput values can be easily identified. All the experiments between this pair of regions shows the same pattern: the throughput dramatically switches from a high value (between around 700 and 850 Mbps, depending also on the VM size) to a low one (around 400 Mbps). Also the stability of the throughput samples significantly changes after the transition: before the transition the throughput appears very unstable, while after the transition, it stably settles down to about 400 Mbps. The CoV is almost halved after the transition (see the smaller values for the standard deviation in Fig. 10a).

The dramatic throughput variation described above happens *on the fly*, i.e., when a communication is active. Moreover, the high-to-low transition may happen at differing points in time for different experiments. As a consequence, different mean values have been observed, which also generate the larger variability range in Fig. 7a. The distribution of the transition time is shown in Fig. 11b. Interestingly, in more than 75% of the cases, the transition happens at around 100 s, thus exposing a certain deterministic behavior. Differently from the case of Amazon reported above, this empirical result could be mapped to mechanisms that restrict the maximum capacity available to a customer based on the traffic volume previously generated. Fig. 11c shows the distribution of the volume of the traffic transferred before the high-to-low throughput transition happens. For about 75% of the cases, volume transferred before the transition ranges from 8,000 to 10,000 MB. Note that, small differences in the transition time reflect in larger discrepancy in the transferred volume because of the high throughput. Besides Azure, similar cases have been identified also for the paths connecting Amazon XL VMs.

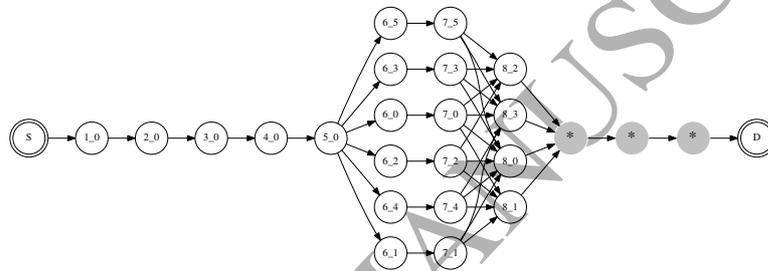
We verified that no significant variation of the actual traffic injected in the network by the sender VMs has been identified for the cases discussed above. Lower throughput values are reflected by a proportionally higher packet loss. Therefore, the differing performance levels identified are not caused by the traffic generation capabilities of the VMs. This further suggests that the observed phenomenon is the consequence of traffic management policies enforced by providers along the paths that interconnect one region to the other. It is worth noting that this phenomenon may impact both the results of the measurement process and the user experience. Firstly, measurements shorter than 100 s are not able to spot the throughput transition. Secondly, longer measurements could lead to misinterpret performance variability if not associated to a deeper analysis.



(a) UDP throughput time series (error bars report the standard deviation).



(b) UDP throughput empirical distribution.



(c) Load-balanced paths inferred by adopting MDA.

Figure 10: Amazon, EU→US, M-sized VMs. UDP throughput switches between two different values over time, suggesting the dynamic variation of the capacity of the path. Note how the distribution of the throughput samples follows a binomial shape. A likely cause is the existence of multiple paths between the source and the destination.

Finally, dramatic throughput drop (around $\sim 50\%$ in the example proposed) can cause non-negligible troubles to customers, heavily impacting the perceived Quality of Service.

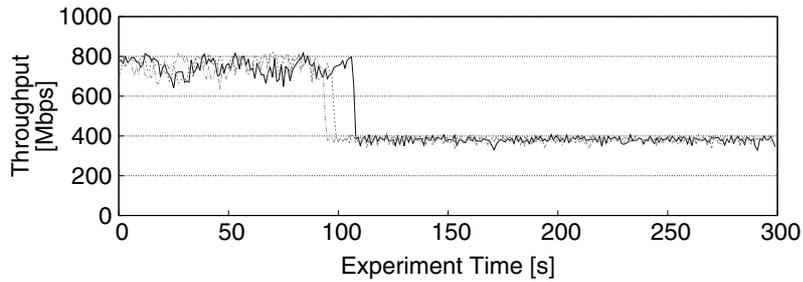
5.4. Performance vs. fees

Our empirical results show that worse performance typically involves two specific regions: AP and SA. Interestingly, these two regions are also the ones with the highest data-transfer costs for the customers (see Tab. 2), thus representing unfavourable choices for them. Indeed, data transfer from AP and SA is subjected to higher costs with respect to EU and US regions, which amounts to $8\times$ and $4.5\times$ for Amazon, and up to $3.2\times$ and $2.3\times$ for Azure, respectively. To better understand these aspects, additional experimentations have been performed. We have traced IP paths between regions by adopting *traceroute*. Note that this analysis has been performed only for Amazon, due to ICMP filtering policies likely enforced at the borders of the Azure datacenters, which prevent the adoption of *traceroute*. The experimental campaigns we set up were designed to trace the region-to-region path while running the performance test. Then we have mapped each IP address identified along the paths to the autonomous system it belongs to. With this aim, we

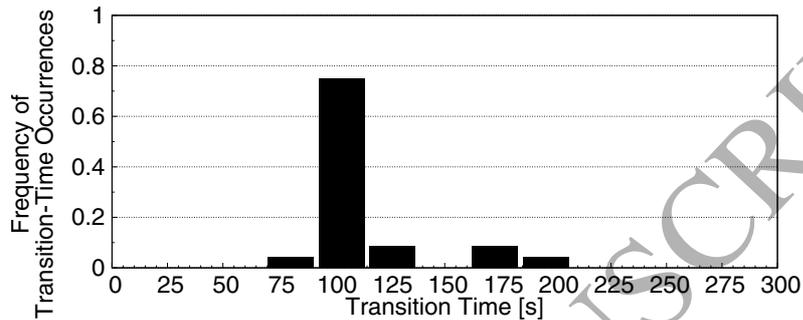
have applied IP to autonomous system number mapping (*IP-to-ASN mapping*) to each IP address collected, by relying on free external services [37] and on the IP-address ranges publicly advertised by the provider itself [38].

Fig. 12 reports the results of this analysis for both M and XL VMs. The bar charts show for each path the number of hops it is composed of, also classifying the hops along the path from the source (hop 0) to the destination (whose position depends on the length of the path). More precisely, they also show the results of the IP-to-ASN mapping, by splitting the hops composing each path in three sets: (i) the ones mapped to AS owned by Amazon itself (#16509/#38895), reported in white; (ii) the ones impossible to map being anonymous or associated to private IP addresses, reported in black; (iii) the ones mapped to ASes other than Amazon, reported in grey. Results appeared stable over multiple experimental repetitions.

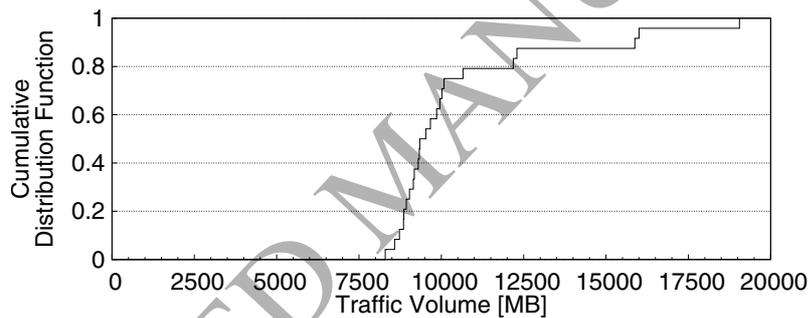
As shown in the figure, we can infer that the length of the path is symmetric in most of the cases—i.e., it does not change when switching source and destination regions. A few cases showing asymmetry have been spotted. For instance: we counted 9 or 10 hops for EU↔SA (M VMs); we counted 15 or 14 and 13 or 20 hops for US↔SA and SA↔AP, respectively



(a) Instantaneous throughput evolution of three 5-minute-long experiments.



(b) Distribution of the high-to-low transition time.



(c) Cumulative Distribution of the volume of traffic transferred before the transition happens.

Figure 11: Azure, US→EU. UDP throughput within 5-minute-long experiments typically switches from a high to a low stable value of 400 Mbps (a). The transition typically happens around 100 s (b). The transferred traffic volume ranges from 8,000 to 10,000 MB for 75% of the cases (c).

(XL VMs).

The black part of the bar charts shows that anonymous hops normally appear close to the edges of the path, and may therefore be intuitively mapped to Amazon (i.e. intra-datacenter hops). Hence, we can consider as part of Amazon infrastructure the hops reported in white and black in Fig. 12a and Fig. 12b. In this hypothesis, our results show that four out of the six paths allow to deliver traffic among geographically distributed datacenters without going out from the infrastructure owned and managed by Amazon. The remaining two paths instead imply the transit through external ASes, from which Amazon probably buys transit bandwidth. All these ASes are tier-1 (namely, *Level 3*, *TATA Communications*, *NTT America*, and *Telefonica*). A schematic view over AS-level-path graph is reported in Fig. 12c. As shown in the figure, in some cases (i.e., where explicitly stated) the chosen VM size may impact the AS path traversed. Results are summarized in Tab. 5. While the num-

ber of hops traversed ranges from 9 to 20 hops, the number of domains traversed varies from 1—when only Amazon AS is traversed—to 3.

It is worth noting that AP and SA are the worse-connected regions in terms of external ASes to be traversed. It is known that the growth of cloud infrastructures is driven by multiple factors such as appealing environmental, financial, and political climates. The high concentrations of Internet exchanges and high-performance network infrastructures are not necessarily the most relevant factors to be taken into account to determine the location of a new datacenter to be deployed. Indeed, the existence of friendly governments, tech sectors, or highly educated populations are also key factors commonly considered. Therefore, the evolution of the communication infrastructure could be also a result of the deployment of the datacenters. This is in accordance with the fact that AP and SA are the regions made available most recently among the ones

Table 5: IP-level hops and domains traversed for each pair of regions in Amazon wide-area inter-datacenter network. IP-level path length and AS-level are symmetric if not explicitly stated otherwise.

Inter-datacenter path	Number of hops traversed	Number of domains traversed	Domains traversed
EU↔US	12	1	Amazon.com, Inc.;
EU↔SA	9/10	1	Amazon.com, Inc.;
US↔SA	14	1	Amazon.com, Inc.;
US↔AP	15	1	Amazon.com, Inc.;
EU↔AP	16	2	Amazon.com, Inc.; NTT America, Inc.;
SA→AP (M)	13	3	Amazon.com, Inc.; Level 3 Communications, Inc.; TATA Communications (AMERICA) Inc.;
AP→SA (M/XL)	13	3	Amazon.com, Inc.; Level 3 Communications, Inc.; NTT America, Inc.;
SA→AP (XL)	20	3	Amazon.com, Inc.; TATA Communications (AMERICA) Inc.; Telefonica International Wholesale Services, SL;

considered—being launched in 2010 and 2011, respectively.

Experimental results obtained with different VM sizes (AP→SA) suggest the adoption of different routing policies for VMs of different sizes. The adoption of external network providers in case of AP or SA region gives also a possible explanation for higher data-transfer costs. They are probably aimed at discouraging intensive network usage by cloud customers, and at guaranteeing proper revenues to the cloud provider according to its business plans also in regions where proprietary network infrastructures have not been deployed yet.

5.5. Latency

Our experimental data shows that latency and throughput are in general not highly correlated. In many cases, high throughput implies lower latency, but low throughput does not necessarily imply high latency and vice versa. In general, latency could not be considered the main cause of the throughput degradations.

Fig. 13 provides an overall view over performance in terms of latency (RTT) for both providers. As expected, latency values appear to be symmetric, although non-negligible differences exist across different regions. US region is the one with the lowest average latency towards the considered destinations, while AP exposes way higher values.

Fig. 14 compares the mean latencies experimented between homologous regions for Amazon and Azure, and their variability. Average latency is equal to 193.61 ms and 214.67 ms for Amazon and Azure, respectively. On average, RTT values appeared to be smaller than the ones reported by previous works [11, 12]. Indeed, these values resulted to be also compatible with the delay constraints hypothesized for the application proposed in [11]. In general, as shown by the limited standard deviation, latency proved to be very stable over time, for both providers. Fig. 15 investigates the variability of the latency from a different angle, showing the cumulative distribution of the CoV of the RTT for different experiments. For both providers we observed CoV values smaller than 0.05 for more than 80% of 5-minute-long experiments, and smaller than 0.1 in all but two of around 70 experiments for Amazon. Interestingly, the worst-performing region pair in terms of latency (i.e. SA↔AP) exposed the stablest performance over time.

In more details, while experimented RTT is similar across providers for five out of six region pairs, EU↔AP shows a

markedly higher latency for Azure (197 ms vs. 315 ms). The latency measured for EU↔AP for Azure is almost equal to the sum of the latencies measured for EU↔US and US↔AP. This suggests that Azure inter-datacenter infrastructure is possibly configured to route traffic from EU to AP through US, thus inflating the length of the path and the perceived latency.

In the following, we propose the results of analyses aimed at further investigating the nature of the delay perceived by the cloud customer. Fig. 16 shows how the average RTT varies with the geographic distance between datacenters. As expected, distance proved to have large impact on RTT performance. However, propagation delay amounts for just a limited portion of the overall delay.⁴ Indeed, Fig. 16 shows how the delay without the propagation quota (filled squares and circles), can be coarsely clustered into 4 ranges: (i) 0–50 ms; (ii) 50–100 ms; (iii) 100–150 ms; (iv) 200–250 ms. This delay—considering negligible the computation done at the destination—is given by transmission, elaboration, and queuing quotas. All these quotas basically depend on the characteristics of the end-to-end connections (link speeds, computational capabilities of the devices along the path, congestion of the queues, etc.)—i.e. on the *performance of the technologies adopted*. They are therefore influenced by the economic investments done in the communication infrastructure. Interestingly, Fig. 16 shows how the geographic distance indirectly influences also the delay without the propagation quota: the larger the distance, the larger the delay, even without the propagation quota. This can be explained by the impact of the distance on the technological choices done by the providers: due to technological or budget constraints, longest paths are provided with less performing technologies. We cannot exclude the traffic across that specific paths is subjected to more complex elaborations (e.g., due to the presence of middleboxes along the path performing specific tasks). We left the analysis of these specific aspects as a future work.

⁴In more detail, the propagation delay has been evaluated as $D_p = \frac{d}{c}$, where d is the distance (calculated by adopting *Vincenty's formulae*) between the source and destination datacenter (whose location has been obtained from [39]), and c is the speed of light. Since the one considered is the *minimum distance* and the actual propagation speed is slightly smaller than the speed of light, the one evaluated represents a lower bound of the actual propagation delay.

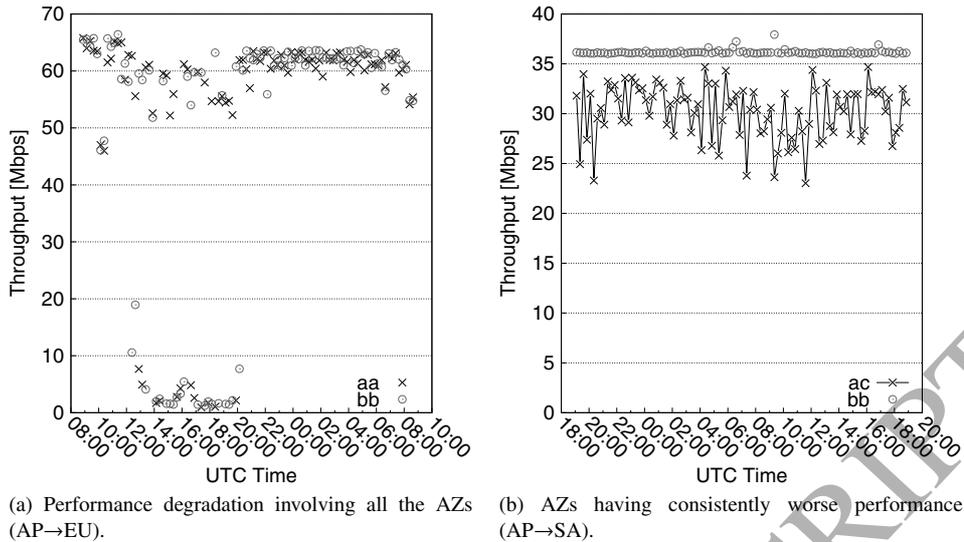


Figure 18: Examples for the interesting cases.

5.6. Impact of the availability zone

In our study we have also taken into account the impact of selecting different availability zones (AZs) inside a region, i.e. we investigated how performance may vary when choosing one of the isolated locations made available inside a region by Amazon.

We set up the experimentations such that 5-minute-long experiments have been launched over multiple distinct AZ pairs at the same time. Therefore we have information about the performance evolution over the time for multiple AZs in parallel. The main outcome of this analysis is that changing AZ gives no clear advantage in terms of achievable throughput. Fig. 17 reports the distribution of the *coefficient of variation of the root mean square error*⁵ (CV_{RMSE}) as an indication for the difference of throughput performance perceived along paths between different AZs. The figure shows this result for both M and XL VMs: CV_{RMSE} results lower than 0.2 for 90% of the samples, underlining how throughput performance along paths connecting differing AZs in the same region is the same, on average. In the following, cases showing higher values for CV_{RMSE} are deepened (Fig. 18).

In a limited number of cases, severe performance degradation lasting for several hours has been identified, showing throughput dropping down to values smaller than 5 Mbps. An example for AP→EU is reported in Fig. 18a. Pairs of homologous samples report different throughput values for different AZs (thus justifying the high CV_{RMSE}). However, in the period between 2:00 p.m. and 8:00 p.m., all the tested logical paths connecting disjoint AZ sets (namely aa and bb) show a severe degradation of performance. This example shows how AZs although guaranteeing site isolation, revealed to be not completely independent from the network point of view. We have observed

different cases similar to the one described. On the other hand, we have also observed cases in which the degradation involves only one AZ pair and not the others. This happened for a single pair of regions in our dataset (AP→SA). Fig. 18b shows how the performance monitored for the AZ pair identified by ac is consistently lower than the homologous identified by bb during the entire 24-hour-long observation period.

Finally, extending latency considerations to Amazon AZs, led us to point out some interesting patterns, for which an example is reported in Fig. 19. The figure shows that different AZ pairs present consistently but slightly distinct latency values. Considering that AZs are independently mapped to identifiers for each customer account [29], and that latency values proved to consistently depend from AZs, we believe that latency information proves to be useful to cloud customers to identify the actual AZ assigned inside a region. This information can be also leveraged to set up resources into the most convenient AZ, according to potentially existing performance discrepancies found. Even more interestingly, intermittent latency deterioration has been spotted in several circumstances, that equally affected the different AZs, i.e., causing a fixed latency shift. Fig. 19 shows an instance of this phenomenon. This behavior suggests how the root cause of this performance deterioration is placed in the portion of the communication path shared by all the traffic between the two regions, thus not depending from the specific AZ chosen.

6. Conclusion

The performance of the network between geographically-distributed cloud datacenters is gaining more and more interest, according to last cloud traffic forecasts, to cutting-edge solutions implemented and technologies developed by cloud providers, and to recent research trends. In this paper we have first provided a performance assessment in terms of throughput and latency for the two leading public cloud providers: Amazon

⁵ $CV_{RMSE}(X, Y) = \frac{\sqrt{E[(X-Y)^2]}}{E[E[X], E[Y]]}$ where X and Y are the empirical distributions of the throughput values collected considering two distinct pairs of availability zones.

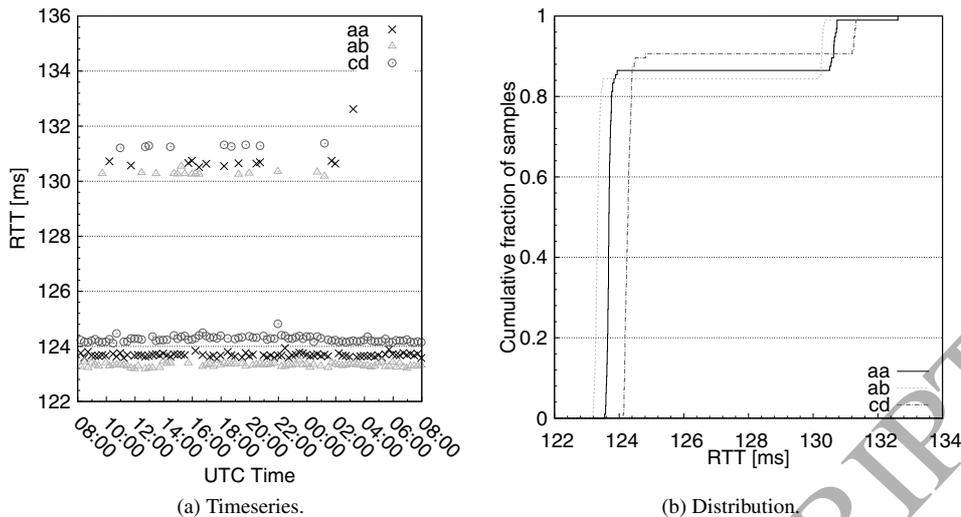


Figure 19: Latency between different AZs (SA→US). A constant additive latency offset is observed in case of performance degradation.

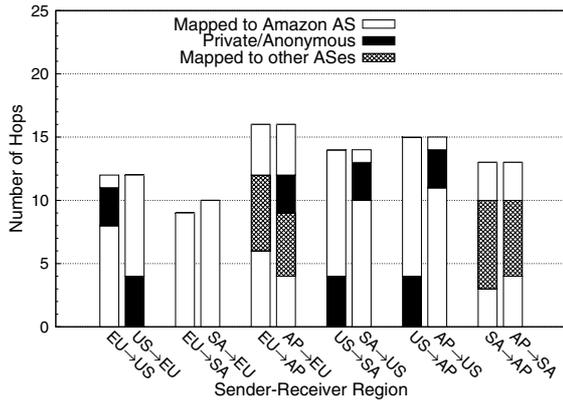
and Azure. Then we have deepened the most interesting outcomes. Experimental results have shown how Azure performs better in terms of maximum achievable throughput (+56%, on average), for slightly higher costs. Moreover, evidences revealing the deployment of high performance infrastructures among cloud datacenters have been found for both providers: both are able to deliver (UDP) traffic between two geographically distributed sites at up to 800 Mbps. The discrepancy in maximum TCP throughput perceived is mainly due to the impact of the presumably high traffic concurrency for Amazon. Interesting cases have been reported to highlight the traffic engineering solutions adopted. Indeed, experimentations have allowed us to infer that providers leverage proprietary networks for inter-datacenter communications in most of the cases. Interestingly, paths with lower performance are characterized by higher data-transfer costs for the customers. Results confirmed that Amazon uses external (i.e. non proprietary) network providers for these connections, thus justifying higher fees imposed to the customers. Performance in terms of latency is similar across different providers, with the exception of one pair of regions for which Amazon outperforms Azure. Finally, different Amazon availability zones (AZs) proved to have the same network performance on average (with the exception of one isolated case) and network performance degradations, when encountered, involve all AZs inside a region. Our results show how the observed performance of the cloud inter-datacenter networks in terms of achievable TCP throughput—thought stable—differs from the performance figure presented in previous work. Indeed, it appears to have decreased over time, suggesting how either the network management policies adopted have become more restrictive, or the growing catchment area has impacted performance. Conversely, results about latency show how RTT values appear to be smaller than the ones reported in previous work.

Acknowledgments

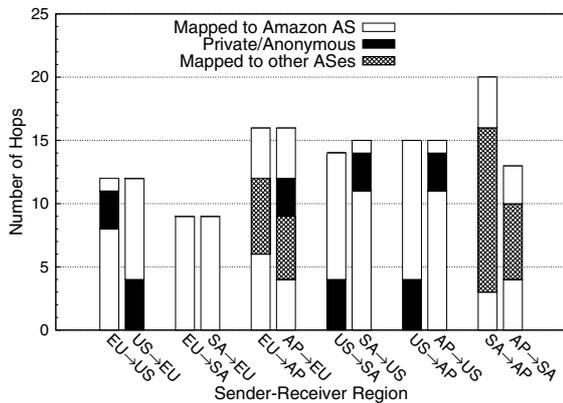
This work is partially funded by art. 11 DM 593/2000 for NM2 srl (Italy).

References

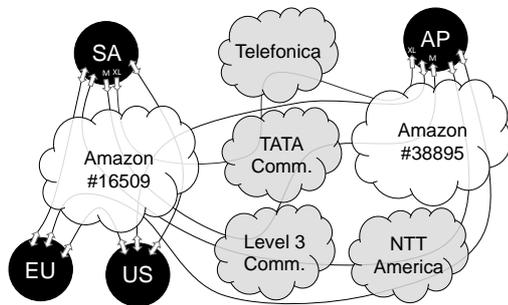
- [1] A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, Above the clouds: A Berkeley view of cloud computing, Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/Eecs 28 (2009) 13.
- [2] G. Aceto, A. Botta, W. de Donato, A. Pescape, Cloud monitoring: Definitions, issues and future directions, in: Cloud Networking (CLOUDNET), 2012 IEEE 1st International Conference on, 2012, pp. 63–67. doi:10.1109/CloudNet.2012.6483656.
- [3] H. Ballani, P. Costa, T. Karagiannis, A. Rowstron, Towards predictable datacenter networks, in: Proceedings of the ACM SIGCOMM 2011 Conference, SIGCOMM '11, ACM, New York, NY, USA, 2011, pp. 242–253. doi:10.1145/2018436.2018465. URL <http://doi.acm.org/10.1145/2018436.2018465>
- [4] J. C. Mogul, L. Popa, What we talk about when we talk about cloud network performance, SIGCOMM Comput. Commun. Rev. 42 (5) (2012) 44–48. doi:10.1145/2378956.2378964. URL <http://doi.acm.org/10.1145/2378956.2378964>
- [5] C. Raiciu, M. Ionescu, D. Niculescu, Opening up black box networks with cloudtalk, in: Proc. 4th USENIX Conference on Hot Topics in Cloud Computing, 2012, p. 6.
- [6] N. Laoutaris, M. Sirivianos, X. Yang, P. Rodriguez, Inter-datacenter bulk transfers with netstitcher, ACM SIGCOMM Computer Communication Review 41 (4) (2011) 74. doi:10.1145/2043164.2018446.
- [7] G. Inc., Inter-datacenter wan with centralized te using sdn and openflow, <https://www.opennetworking.org/images/stories/downloads/sdn-resources/customer-case-studies/cs-googlesdn.pdf> (2012).
- [8] A. Greenberg, J. Hamilton, D. A. Maltz, P. Patel, The cost of a cloud: research problems in data center networks, SIGCOMM Comput. Commun. Rev. 39 (1) (2008) 68–73.
- [9] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, S. Sengupta, VI2: A scalable and flexible data center network, SIGCOMM Comput. Commun. Rev. 39 (4) (2009) 51–62. doi:10.1145/1594977.1592576. URL <http://doi.acm.org/10.1145/1594977.1592576>
- [10] Cisco, Cisco global cloud index: Forecast and methodology, 2013–2018, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html (2013).



(a) Length and hop classification for paths between M VMs.



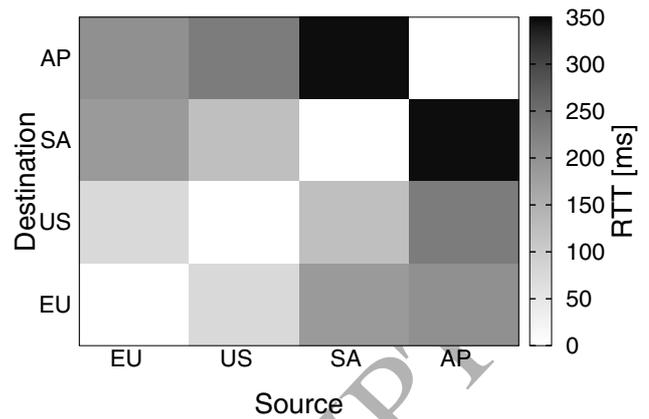
(b) Length and hop classification for paths between XL VMs.



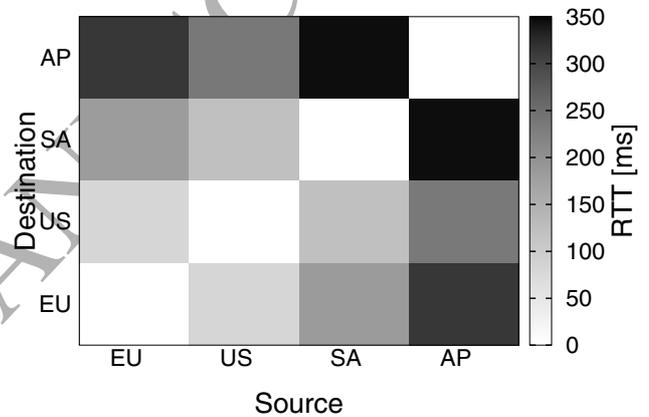
(c) AS level graph. VM size impacts the traversed path only where explicitly stated. Incoming and outgoing paths can be discerned on the basis of the direction the arrows reported.

Figure 12: IP-to-ASN mapping and length for Amazon inter-datacenter paths. Paths lay on provider-owned infrastructure for four out of six cases.

- [11] Y. Feng, B. Li, B. Li, Airlift: Video conferencing as a cloud service using inter-datacenter networks, in: Network Protocols (ICNP), 2012 20th IEEE International Conference on, IEEE, 2012, pp. 1–11.
- [12] Y. Feng, B. Li, B. Li, Jetway: minimizing costs on inter-datacenter video traffic, in: Proceedings of the 20th ACM international conference on Multimedia, ACM, 2012, pp. 259–268.
- [13] F. Wang, J. Liu, M. Chen, Calms: Cloud-assisted live media streaming for globalized demands with time/region diversities, in: INFOCOM, 2012 Proceedings IEEE, IEEE, 2012, pp. 199–207.
- [14] CloudHarmony, <https://cloudfarmy.com/>.
- [15] V. Persico, A. Botta, A. Montieri, A. Pescapé, A first look at public-cloud inter-datacenter network performance, in: 2016 IEEE Global Commu-



(a) Amazon. Average values in ms (towards destination): 154 (EU), 143 (US), 218 (SA), 258 (AP).



(b) Azure. Average values in ms (towards destination): 193 (EU), 147 (US), 217 (SA), 298 (AP).

Figure 13: Latency between different regions. Azure exposes slightly higher values than Amazon. For both providers, US region is the one with the lowest average latency towards the considered destinations, while AP expose way higher values.

nications Conference: Communication QoS, Reliability and Modeling (Globecom2016 CQRM), Washington, USA, 2016.

- [16] L. Leong, D. Toombs, B. Gill, Magic quadrant for cloud infrastructure as a service, worldwide, <http://www.gartner.com/technology/reprints.do?id=1-2G205FC&ct=150519&st=sb> (2015).
- [17] Synergy research, press release, Aws remains dominant despite microsoft and google growth surges, <https://www.srgresearch.com/articles/aws-remains-dominant-despite-microsoft-and-google-growth-surges> (2016).
- [18] V. Persico, A. Montieri, A. Pescapé, CloudSurf: a platform for monitoring public-cloud networks, in: 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI) (IEEE RTSI 2016), Bologna, Italy, 2016.
- [19] V. Persico, P. Marchetta, A. Botta, A. Pescapé, Measuring network throughput in the cloud: the case of amazon ec2, Elsevier, Computer Networks, Special Issue on Cloud Networking and Communications II.
- [20] K. LaCurts, S. Deng, A. Goyal, H. Balakrishnan, Choreo: Network-aware task placement for cloud applications, in: Proceedings of the 2013 Conference on Internet Measurement Conference, IMC '13, ACM, New York, NY, USA, 2013, pp. 191–204. doi:10.1145/2504730.2504744. URL <http://doi.acm.org/10.1145/2504730.2504744>

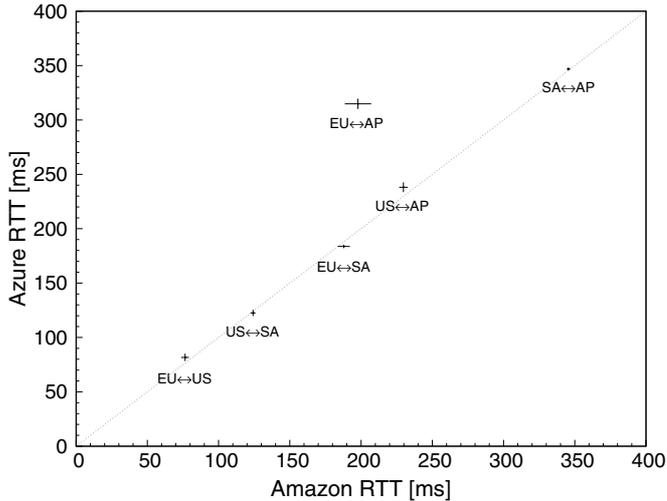


Figure 14: Comparison of inter-datacenter latencies experienced when relying on different providers. Crossing points identify mean values, while bars report standard deviation for both Amazon (x axis) and Azure (y axis). Latency across homologous pairs resulted to be comparable except than for EU↔AP. Little variability has been observed.

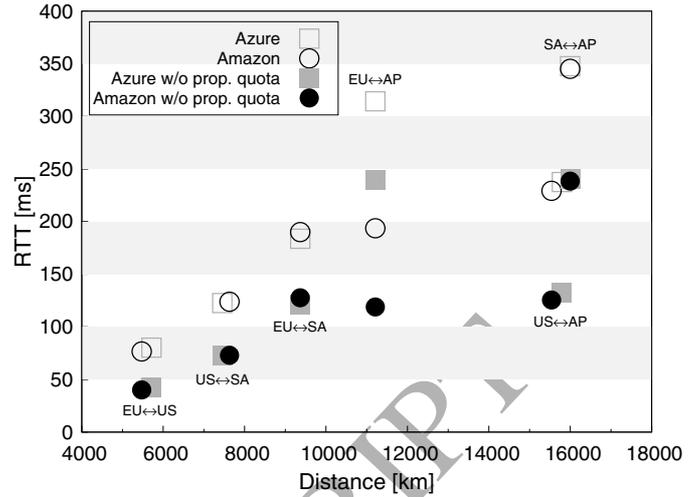


Figure 16: Latency vs. distance. As expected, latency is impacted by geographical distance between datacenters (empty squares and circles). However, propagation delay amounts for just a limited portion of the overall delay. Also excluding the propagation quota—known to be the quota mainly impacted by geographic distance—the delay grows with the distance (filled squares and circles). Results suggest that technologies with different performance levels are deployed.

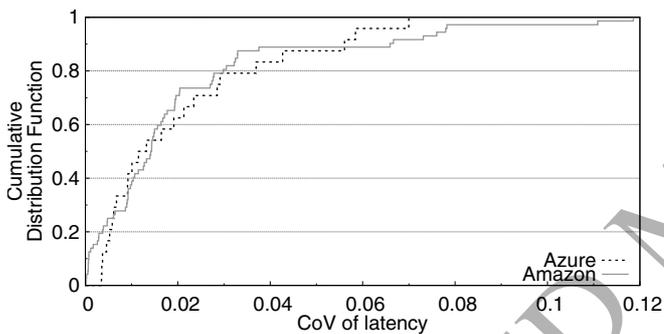


Figure 15: CoV distribution of latency (RTT) across different 5-minute-long experiments. Both providers expose little variation for latency.

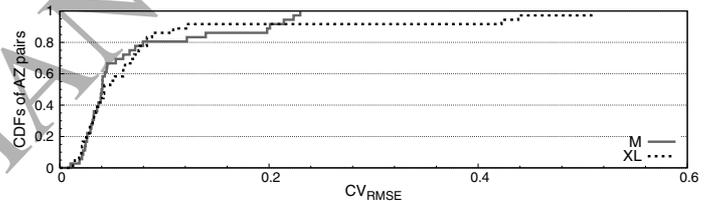


Figure 17: CV_{RMSE} distribution of the throughput across different Amazon AZs. 90% of samples has values lower than 0.2, showing how the choice of AZ is non-influential for throughput performance in most of the cases.

- [21] Z. Hill, J. Li, M. Mao, A. Ruiz-Alvarez, M. Humphrey, Early observations on the performance of windows azure, in: Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, ACM, 2010, pp. 367–376.
- [22] J. Schad, J. Dittrich, J.-A. Quiané-Ruiz, Runtime measurements in the cloud: Observing, analyzing, and reducing variance, Proc. VLDB Endow. 3 (1-2) (2010) 460–471. doi:10.14778/1920841.1920902. URL <http://dx.doi.org/10.14778/1920841.1920902>
- [23] R. Tudoran, A. Costan, G. Antoniu, L. Bougé, A performance evaluation of azure and nimbus clouds for scientific applications, in: Proceedings of the 2nd International Workshop on Cloud Computing Platforms, ACM, 2012, p. 4.
- [24] V. Persico, P. Marchetta, A. Botta, A. Pescapé, On network throughput variability in microsoft azure cloud, in: 2015 IEEE Global Communications Conference (GLOBECOM), 2015, pp. 1–6. doi:10.1109/GLOBECOM.2015.7416997.
- [25] Y. Chen, S. Jain, V. Adhikari, Z.-L. Zhang, K. Xu, A first look at inter-data center traffic characteristics via yahoo! datasets, in: INFOCOM, 2011 Proceedings IEEE, 2011, pp. 1620–1628. doi:10.1109/INFCOM.2011.5934955.
- [26] A. Li, X. Yang, S. Kandula, M. Zhang, Cloudcmp: Comparing public cloud providers, in: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, IMC '10, ACM, New York, NY, USA, 2010, pp. 1–14. doi:10.1145/1879141.1879143. URL <http://doi.acm.org/10.1145/1879141.1879143>
- [27] J. Garcia-Dorado, S. Rao, Cost-aware Multi Data-Center Bulk Transfers in the Cloud from a Customer-Side Perspective, IEEE Transactions on Cloud Computing 7161 (c) (2015) 1–1. doi:10.1109/TCC.2015.2469666. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7208810>
- [28] B. Fink, R. Scott, nuttcp, v6.1.2. URL <http://www.nuttcp.net/>
- [29] Amazon Web Services website, <http://aws.amazon.com/> (Sep. 2015).
- [30] Microsoft Azure website, <http://azure.microsoft.com/> (Sep. 2015).
- [31] G. Wang, T. Ng, The impact of virtualization on network performance of amazon ec2 data center, in: INFOCOM, 2010 Proceedings IEEE, 2010, pp. 1–9. doi:10.1109/INFCOM.2010.5461931.
- [32] B. Augustin, T. Friedman, R. Teixeira, Measuring multipath routing in the Internet, IEEE/ACM Transactions on Networking 19 (3) (2011) 830–840. doi:10.1109/TNET.2010.2096232. URL <http://dx.doi.org/10.1109/TNET.2010.2096232>
- [33] P. Marchetta, P. Merindol, B. Donnet, A. Pescapé, J. Pansiot, Topology discovery at the router level: A new hybrid tool targeting isp networks, Selected Areas in Communications, IEEE Journal on 29 (9) (2011) 1776–1787. doi:10.1109/JSAC.2011.1110003.
- [34] P. Marchetta, A. Montieri, V. Persico, A. Pescapé, I. Cunha, E. Katz-Bassett, How and how much traceroute confuses our understanding of network paths, in: 2016 IEEE International Symposium on Lo-

- cal and Metropolitan Area Networks (LANMAN), 2016, pp. 1–7. doi:10.1109/LANMAN.2016.7548847.
- [35] P. Marchetta, V. Persico, A. Pescapé, Pythia: yet another active probing technique for alias resolution, in: Proceedings of the ninth ACM conference on Emerging networking experiments and technologies, ACM, 2013, pp. 229–234.
- [36] K. Keys, Y. Hyun, M. Luckie, K. Claffy, Internet-scale ipv4 alias resolution with midar, *IEEE/ACM Transactions on Networking (TON)* 21 (2) (2013) 383–399.
- [37] Team Cymru. IP to ASN mapping service, <http://www.team-cymru.org/IP-ASN-mapping.html>.
- [38] Amazon Web Services. AWS IP address ranges, <http://docs.aws.amazon.com/general/latest/gr/aws-ip-ranges.html>.
- [39] G. Cook, D. Pomerantz, Clicking clean: A guide to building the green internet, <http://www.greenpeace.org/usa/wp-content/uploads/legacy/Global/usa/planet3/PDFs/2015ClickingClean.pdf> (2015).

ACCEPTED MANUSCRIPT



Valerio Persico is a PhD student at the Department of Electrical Engineering and Information Technology (DIETI) of the University of Napoli Federico II (Italy), where he received his MS degree in 2013, defending a thesis about a novel technique for topology discovery of IP networks. He is a member of the research group called Traffic, working in the area of computer networks and multimedia. His research interests fall in the area of networking and of IP measurements; in particular his past and present work focuses on cloud monitoring, traffic classification, IP topology discovery, IP path tracing, and IP alias resolution. During his PhD he coauthored several journal and conference publications, being awarded with the Best Student Paper Award for his paper "Don't Trust Traceroute (Completely)" at CoNext 2013. He also served and serves as peer reviewer for several international conferences (IFIP Networking, IEEE Globecom, IEEE ICC, IEEE ISCC, IEEE CCNC, IEEE IC2E, etc.) and journals (Elsevier SIMPAT, Elsevier JNCA, etc.) in the area of networking.



Alessio Botta is a postdoc at the Department of Computer Engineering and Systems of the University of Napoli Federico II (Italy). He graduated in Telecommunications Engineering (M.S.) and obtained the Ph.D. in Computer Engineering and Systems, both at University of Napoli Federico II. His research interests are in the area of networking and, in particular, in the area of network performance measurement and improvement, with a specific focus on wireless and heterogeneous systems. Alessio Botta has coauthored more than 40 international journal (IEEE Communications Magazine, IEEE Transactions on Parallel and Distributed Systems, Elsevier Computer Networks, etc.) and conference (IEEE Globecom, IEEE ICC, IEEE ISCC, etc.) publications. He has served and serves several technical program committees of several international conferences (IEEE Globecom, IEEE ICC, etc.) and he acts as reviewer for different international conferences (IEEE Infocom, etc.) and journals (IEEE Transactions on Mobile Computing, IEEE Network,

IEEE Transactions on Vehicular Technology, etc.) in the area of networking. In 2010 he was awarded with the best local paper award at IEEE ISCC 2010.



Pietro Marchetta received his PhD in Computer Engineering at University of Napoli in 2014. Currently, he holds a post-doctoral position at the Department of Electrical Engineering and Information Technology of the University of Napoli Federico II (Italy). His main research activities focus on methodologies, techniques and large-scale distributed platforms for Internet Measurements with a specific focus on Internet routing and performance. These activities have been conducted in the context of international collaborations with leading global research groups. In 2010, he visited the IP Networking Lab at the University of Louvain-la-neuve (UCL), where he contributed to design, develop and evaluate a novel Internet topology discovery approach based on IGMP. In 2013, he joined as visiting researcher the Networked Systems Laboratory of the University of Southern California (USC) working on path prediction techniques and performance evaluation approaches. He served and serves a reviewer for a dozen of conferences and journals (e.g. Elsevier's Computer Networks and Future Generation Computer Systems). For his research, he received some awards including the first place at the ACM Student Research Competition at SIGCOMM 2012 and the Best Student Workshop Paper Award in CoNEXT 2013.



Antonio Montieri received his M.S. Laurea Degree in Computer Engineering in 2015 at the University of Napoli Federico II (Italy). Currently, he is a research associate at NM2 srl (Italy). His research interests are in the networking field with focus on network measurements, traffic analysis, and cloud network assessment.



Antonio Pescapé is an Associate Professor at the Department of Electrical Engineering and Information Technology of the

the University of Napoli Federico II (Italy) where he teaches courses in Computer Networks, Computer Architectures, Programming, and Multimedia and he has also supervised and graduated more than 170 among BS, MS, and PhD students. He received the M.S. Laurea Degree in Computer Engineering and the Ph.D. in Computer Engineering and Systems at University of Napoli Federico II, Napoli (Italy). His research interests are in the networking field with focus on Internet Monitoring, Measurements and Management and on Network Security. Antonio Pescapé has co-authored over 165 journal (IEEE ACM Transaction on Networking, Communications of the ACM, IEEE Communications Magazine, JSAC, IEEE Wireless Communications Magazine, IEEE Networks, etc.) and conference (SIGCOMM, Infocom, NSDI, Conext, IMC, PAM, Globecom, ICC, etc.) publications and he is co-author of a patent. He has served and serves as workshops and conferences Chair (including IEEE ICC (NGN symposium)) and on more than 170 technical program committees of IEEE and ACM conferences. He serves

as Editorial Board Member of Journal of Network and Computer Applications and has served as Editorial Board Member of IEEE Survey and Tutorials (2008-2011) and was guest editor for the special issue of Computer Networks on Traffic classification and its applications to modern networks and guest editor for the special issue of the Journal of Future Generation Computer Systems on "Internet of Things and Cloud Services". For his research activities he has received several awards, comprising a Google Faculty Award, several best paper awards and two IRTF (Internet Research Task Force) ANRP (Applied Networking Research Prize). Antonio Pescapé has served and serves as independent reviewer/evaluator of research and implementation projects and project proposals co-funded by the EU Commission, Sweden government, several Italian local governments, Italian Ministry for University and Research (MIUR) and Italian Ministry of Economic Development (MISE). Antonio Pescapé is a Senior Member of the IEEE.