Optimizing Internet Transit Routing for Content Delivery Networks

Faraz Ahmed*, Zubair Shafiq†, Amir Khakpour‡, Alex X. Liu*
*Department of Computer Science and Engineering, Michigan State University. {farazah, alexliu}@cse.msu.edu
†Department of Computer Science, The University of Iowa. zubair-shafiq@uiowa.edu
‡Verizon Digital Media Services. amir.khakpour@verizon.com

***Abstract*—Content Distribution Networks (CDNs) maintain multiple transit routes from content distribution servers to eyeball ISP networks which provide Internet connectivity to end users. Due to the dynamics of varying performance and pricing on transit routes, CDNs need to implement a transit route selection strategy to optimize performance and cost tradeoffs. In this paper, we formalize the transit routing problem using a multi-attribute objective function to simultaneously optimize end-to-end performance and cost. Our approach allows CDNs to navigate the cost and performance tradeoff in transit routing through a single control knob. We evaluate our approach using real-world measurements from CDN servers located at 19 geographically distributed IXPs. Using our approach, CDNs can reduce transit costs on average by 57% without sacrificing performance.**

I. Introduction

Content publishers usually rely on third-party Content Distribution Networks (CDNs) for efficiently delivering content to end users. A significant fraction of web content on the Internet is served by CDNs. Cisco estimates that the share of Internet video traffic served by CDNs will increase from 61% in 2015 to 73% by 2020 [5]. Two major considerations for CDNs are cost and performance of delivering content to end users. CDNs maintain copies of content at cache servers that are deployed at carefully selected geographical locations. CDNs also maintain multiple transit routes from cache servers to access ISPs (or "eyeball networks") which provide Internet connectivity to end users. When a user requests an object, the request is redirected to a nearby cache server containing the requested object. The cache server sends the object to the end user via one of multiple transit routes. Since the performance of transit routes changes over time, end-to-end performance achieved by a CDN is dependent on the choice of transit route [32]. Furthermore, the price of Internet transit also varies from one transit provider to another. Thus, a major challenge faced by CDNs is to develop a transit routing strategy to simultaneously optimize cost and performance.

The dynamic nature of transit pricing and performance makes it challenging to optimize the cost and performance tradeoff. There are thousands of eyeball ISPs which are reachable via different transit routes and different geographical locations. Each choice of transit route for a particular eyeball ISP and geographical location has distinct cost and performance characteristics, which makes the problem of developing a transit routing strategy challenging. Therefore, it is important for CDNs to carefully design and adopt a transit route selection strategy by analyzing the dynamically changing cost and performance tradeoffs.

To the best of our knowledge, the problem of optimal Internet transit routing for CDNs considering both cost and performance tradeoffs is not addressed in prior literature. Prior work on route selection has studied multi-homed access networks [11], [13], [37]. Unlike multi-homed access networks, CDNs have multi-homed servers at IXPs (via multiple transit providers) which provide explicit control over content routing. The state of the art for transit routing at IXPs by CDNs does not use a fully automated approach. CDN operators mostly configure routing strategies manually to optimize cost and performance. Due to performance dynamics across thousands of access ISPs and numerous transit routes, it is practically impossible for CDN operators to manually achieve an optimal tradeoff between cost and performance.

In this paper, we present measurement, analysis, design, and evaluation of optimal transit route selection for CDNs. First, we present an approach to measure end-to-end performance across multiple transit routes simultaneously. We measure end-to-end performance as the delay between the CDN server located at an IXP and the end user. Second, we analyze the temporal and spatial dynamics of performance differences between multiple transit routes. Measuring performance differences across multiple transit routes allows us to characterize performance dynamics of Internet transit. Finally, we present our proposed approach for optimal transit routing. We formulate the transit route selection problem as using a multi-attribute objective function that optimizes cost and performance simultaneously. Using linear programming, we obtain tradeoff curves between cost and performance for various routing strategies. Our results show that CDNs can achieve significant cost and performance benefits using our measurement and optimization approach.

We face two key technical challenges in optimizing Internet transit routing. The first challenge is to obtain simultaneous performance measurements across multiple transit routes. This is necessary because performance at each transit route is dependent on multiple external factors. Performance differences between transit routes can vary due to congestion at the ISP-transit interconnections or congestion at intra-ISP links. Furthermore, performance can vary due to traffic engineering policies of customer ISPs. To solve this challenge, we capture the dynamics in performance differences by exploiting the

multi-homing capability of CDN servers. Specifically, we implement a client-side performance measurement JavaScript that is embedded in client-requested web pages by the CDN. The JavaScript downloads multiple copies of a pixel tag simultaneously via multiple transit routes. This measurement methodology allows us to capture user perceived end-to-end performance via multiple transit routes.

The second challenge is to obtain a transit routing strategy that provides an optimal tradeoff between cost and performance. For each IXP location, thousands of destination ISPs are reachable and for each destination ISP there are numerous transit routes. We solve this challenge by formulating the Internet transit routing problem as a multi-objective optimization problem. The objective function computes the utility of a particular routing strategy and optimizes it over all possible strategies. We describe utility as the sum of overall cost incurred to a CDN and performance weighted by a factor γ. By varying values of γ, we obtain various selection strategies. We obtain an optimal strategy by looking at the tradeoffs between cost and performance for different γ values. The ability to analyze and automatically control transit routing has significant cost and performance benefits for CDNs. Our approach allows CDN operators to analyze cost-performance tradeoffs and accordingly choose an optimal routing strategy.

We summarize our key contributions below.

- First, we characterize Internet transit performance from multiple IXP vantage points. Our comparative analysis of two transit providers reveals that one transit route significantly outperforms the other for more than 50% users at a European IXP and more than 30% of users at a North American IXP.
- Second, we propose an optimization approach that allows CDN operators to navigate cost and performance tradeoffs in transit routing through a control knob. Our results show that CDNs can reduce their transit costs on average by 57% without incurring any performance degradation.

II. Background

A. Architecture

CDN caching infrastructure consists of servers located at multiple geographically distributed locations. A content provider pushes copies of its content to the CDN. The CDN maintains copies of the content at geographically distributed cache servers. The client requests for content are typically directed to suitable CDN cache servers based on geographic proximity using DNS redirection or anycast [21].

There are two common server deployment strategies employed by most commercial CDNs [18]. CDNs either deploy servers inside many access ISPs that are closer to users (*enter-deep* strategy) or deploy servers at a few carefully chosen geographical locations (*bring-home* strategy). For example, Akamai has adopted the enter-deep deployment strategy, and has deployed 100,000+ servers across thousands of ASes [1], [31]. Several major content providers have also adopted the enter-deep deployment strategy. For example, Netflix's Open Connect Appliance (OCA) servers are deployed at ISPs

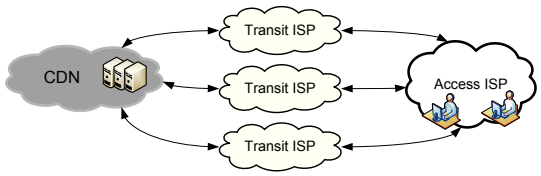


Fig. 1. A CDN interconnects with multiple transit ISPs at IXPs

delivering over 5Gbps in peak daily Netflix traffic [3]. Google Global Cache (GGC) servers are also installed inside large ISP networks [2]. In the bring-home strategy, CDNs deploy large clusters of servers at fewer sites and connect these sites with high-speed connectivity. Instead of deploying these clusters inside large ISPs, these CDNs strategically place their server clusters near Internet Exchange Points (IXPs). At IXPs, CDNs can interconnect with a large number of ISPs using peering or transit [4], [9], [22]. According to a snapshot of the PeeringDB in August 2013, 76% of ASes use Open peering, 21% use Selective, and 3% use Restrictive [22]. For example, Limelight has 18,000+ servers at dozens of Points of Presence (POPs) around the world. CDNs interconnect with major ISPs, including backbone transit ISPs, at IXPs to efficiently deliver content to end users.

Figure 1 provides an architectural overview of the bring-home CDN that we study in this paper. As discussed earlier, we note that the CDN cache servers are located near major IXPs, where they interconnect with backbone *transit providers* (or transit ISPs).[1] CDNs buy transit services from multiple transit providers. CDNs can use one or simultaneously use multiple transit providers in order to minimize their transit costs and maximize performance for end users. Unlike enter-deep strategy, a bring-home CDN has more control over content distribution servers because cache servers are located at a small number of key geographical locations. However, as shown in prior literature [18], the CDN has to deal with larger end-to-end delay and higher transit costs as compared to enter-deep CDNs like Akamai. Therefore, it is crucial for bring-home CDNs, like the one discussed here, to carefully choose transit routes to optimize both performance and cost.

B. Internet Transit Dynamics

Pricing and performance of transit providers vary with respect to time and geographical location. Below, we provide an overview of both pricing and performance dynamics in the Internet transit market.

Pricing Dynamics. Internet transit prices have steadily decreased over the years due to technological advances and increased competition in the Internet transit market. Usage based and tiered pricing models are commonly used in the Internet transit market [34]. In the usage based pricing model, Internet transit is a metered service, i.e., transit providers

[1]Note that the CDN can peer with an access ISP (or "eyeball network") to eliminate transit costs if the access ISP has presence at the IXP. However, small access ISPs typically do not have presence at multiple large IXPs [25]. Moreover, large access ISPs may not directly peer with the CDN due to the intricacies of peering [25]. In this paper, unless stated otherwise, we restrict ourselves to transit routing for CDNs.

charge their customers by measuring the amount of traffic sent or received during the billing period. Some transit providers may charge customers differently based on traffic volume and destination. In the tiered pricing model, transit providers charge customers based on geographical region, traffic commit levels, type of traffic i.e, on-net vs off-net, etc. [34]. The customers who commit higher bandwidths are able to negotiate lower per-Mbps costs as compared to the customers who commit lower bandwidths [24]. The most commonly used pricing scheme in the Internet transit market is called 95th-percentile pricing [30]. In this scheme, usage over a fixed billing period (typically one month) is measured on a megabit per second basis using the 95th percentile value. Unlike capped or fixed billing, where customers pay a fixed amount regardless of usage, 95th percentile charging is flexible. The service providers do not have to implement various charging policies and the customers pay only for what they utilize. Note that customers with bursty traffic are likely to pay higher costs than customers with consistent bandwidth utilization, even though overall traffic volume transferred by bursty customers may be less than the consistent ones. When considering the costs incurred due to effects of bursty traffic on traffic engineering policies, 95th percentile charging method balances the tradeoff between flexibility and the amount charged to customers.

Performance Dynamics. There are several factors that cause performance differences across transit routes. For instance, a transit route may simply be longer (more IP hops) than others, resulting in consistently higher propagation delays. A CDN can easily identify such cases when a transit route is consistently worse than others. Transit performance is also affected by congestion at ISP-transit interconnections or congestion at intra-ISP links resulting in larger queueing delays and packet losses due to buffer overflows. The congestion can be temporary (e.g., during peak hours) or long-lasting indicting link under-provisioning. Such changes in transit performance are not in control of CDNs because contractual agreements between ISPs and changes in inter- and intra-domain routing policies are considered confidential information. From a CDN's perspective, it is important to continually monitor performance across different transit providers and choose transit routes accordingly to optimize end-to-end performance.

Overall, in addition to optimizing performance, CDNs also have to consider financial aspects of Internet transit routing. Different transit providers may charge differently, use different pricing models, and set up contractual agreements with/without performance SLAs. Thus, CDNs have to navigate the cost-performance tradeoff. A CDN can choose the cheapest transit route by sacrificing performance or can pay more for better performance. In this paper, our goal is to understand the cost performance tradeoff from the perspective of CDNs.

III. Performance Measurements & Analysis

A. Measurement Methodology

In this section, we discuss our methodology to measure and analyze performance of transit providers. To measure the performance of multiple transit providers serving a particular ISP (e.g., AS in a specific geographical region), we utilize the multi-homing capabilities of the CDN's cache servers located at IXPs. Specifically, we embed a JavaScript in client-requested web pages to conduct active performance measurements. The client-side JavaScript generates HTTP requests to the multi-homed IXP server and downloads a pixel tag via different transit providers. The same pixel tag is downloaded simultaneously via multiple transit providers; this allows us to capture the performance differences of various transit routes at a given time instance.

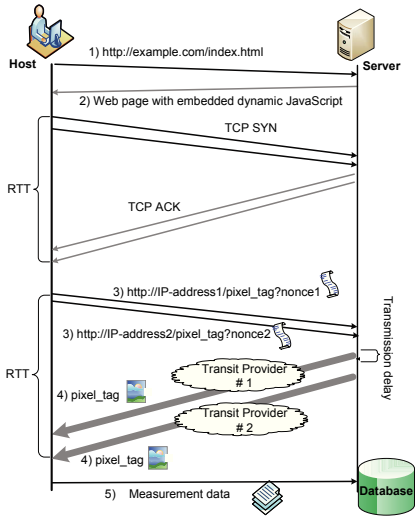
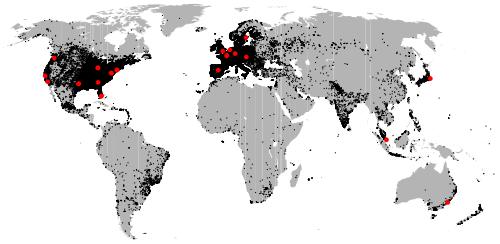
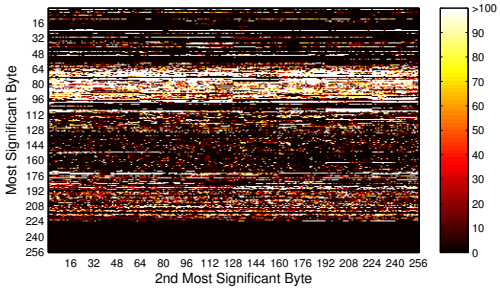


Fig. 2. Transit performance measurements



(a)



(b)

Fig. 3. (a) Red dots represent the locations of measurement servers at IXPs and black dots represent the location of end users. (b) Each dot indicates count of measurements from a /16 IPv4 range block.

Figure 2 illustrates our measurement methodology. The JavaScript is embedded in the HTML pages served to end users. Once a client downloads the HTML page, the JavaScript executes in the background and sends an HTTP GET request for a pixel tag. To avoid additional delays incurred due to DNS

TABLE I
DATA SET SUMMARY STATISTICS

Transit Provider	#Records x1000	#IPs x1000	#ASes ASes	Median Time(ms)
NTT	5,435.7	2,013.2	15,559	288
TEL	5,217.5	1,923.5	15,616	286
DTA	1,869.9	772.3	7,969	297
PAC	552.9	236.8	1,910	488
PCC	533.7	228.8	1,775	427
AAP	192.3	814.3	189	241

lookups, we hard-code the IP addresses of the measurement server in the JavaScript. To avoid local cache hits, we add a nonce check in the client's HTTP request. This ensures that the pixel tag is served only from the multi-homed measurement server and not from the local browser's cache. We set the size of pixel tag to 10 kilobytes, which is less than the server's initial TCP congestion window. Thus, it takes approximately 2 RTTs to download the pixel tag. The JavaScript records the RTTs at the client side and periodically uploads the measurements to a database server.

Our dataset consists of active measurements conducted from 19 measurement servers which are located at different IXPs. Each record in the dataset consists of time stamps indicating the measurement time, identifiers for transit providers, download time values obtained for each transit provider, client IP address, client AS number, and ISP name. In total, the dataset consists of 6 million entries which were recorded from more than 2 million IP addresses distributed across $16,752$ ASes. All user identifiers in the dataset are anonymized to protect the privacy of users. The measurement time span is distributed over a period of one year. Note that a measurement is recorded only when a client sends a request to download a web page. Therefore, the scope and size of our measurements depends on content popularity and user demographics. However, the CDN discussed in this study has thousands of clients covering a large number of ASes, which include most major residential broadband providers.

As shown in Figure 3(a), 19 measurement servers are located at geographically distributed IXPs and cover all major continents including North America, Europe, Asia, and Australia. We used the publicly available IP geolocation databases to locate host IP addresses. Figure 3(b) visualizes the IP address space of all host IP addresses in our data set. The plot divides the IP space by the first two most significant bytes as axes. The color indicates the number of records from each /16 IP block, where brighter colors represent more records and darker colors represent fewer records. The plot shows that our measurements covers a large chunk of the IP space. Some of the empty portions represent the space for reserved IP addresses. Other large empty portions correspond to unobserved (i.e., not served by the CDN servers), non-allocated, or inactive IP addresses.

Table I provides summary statistics of our measurements. Each row enumerates the total number of measurements, number of unique host IPs, number of unique ASes, and median download time for each transit provider. The three letter abbreviations in the first column denote various transit providers used by the CDN at various IXPs throughout the world. We note that Nippon Telegraph and Telephone (NTT) and Telia (TEL) provide the most coverage (in terms of number of IPs and ASes) to the CDN. Other transit providers are used at IXPs when NTT and TEL do not cover certain IXPs, or when NTT or TEL are temporarily unavailable. The median download times for NTT and TEL are 288 and 286 milliseconds, respectively.

B. Analysis and Discussions

We first analyze the performance characteristics of different transit providers using our measurements. We are particularly interested in understanding *where* and *when* one transit provider outperforms another transit provider. To this end, we explore spatial (where) and temporal (when) variations in transit provider performance difference. Below, we focus our analysis to two major transit providers (NTT and TEL) across four popular POP locations (Paris, Madrid, San Jose, Chicago).

Spatial variations. Figure 4 plots the cumulative distribution of percentage performance difference between transit providers for users across four POP locations. The x-axis represents users and the y-axis represents the percentage performance difference between two transit providers for a given user. The plot indicates the portion of users who receive better performance from one transit provider versus the other transit provider. A positive percentage difference indicates that NTT has better performance, and a negative percentage difference indicates that TEL has better performance. Each curve is aggregated for all ASes that interconnect at the corresponding POP location. We observe that NTT provides significantly better performance than TEL for users on the left of the x-axis, while TEL outperforms NTT for users on the right side of the x-axis. For Madrid POP, we note that NTT outperforms TEL by 10% for approximately 40% of users and TEL outperforms NTT by 10% for approximately 10% of users. In sum, 50% users experience significant performance difference across two transit providers. The remaining 50% users would fare similarly on either transit provider. Using the same 10% performance difference threshold, we note that one of these transit providers outperforms the other for for 55% users at Paris POP, 30% of users at Chicago POP, and 30% of users at San Jose POP. Our findings highlight that there is no outright best transit provider for all users and a careful choice of transit provider is necessary.

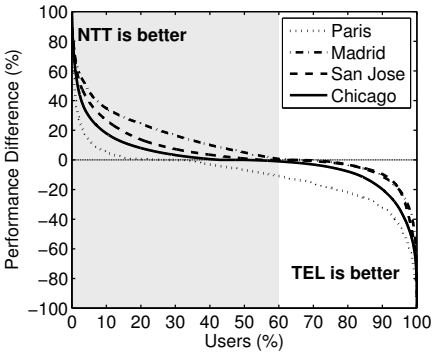
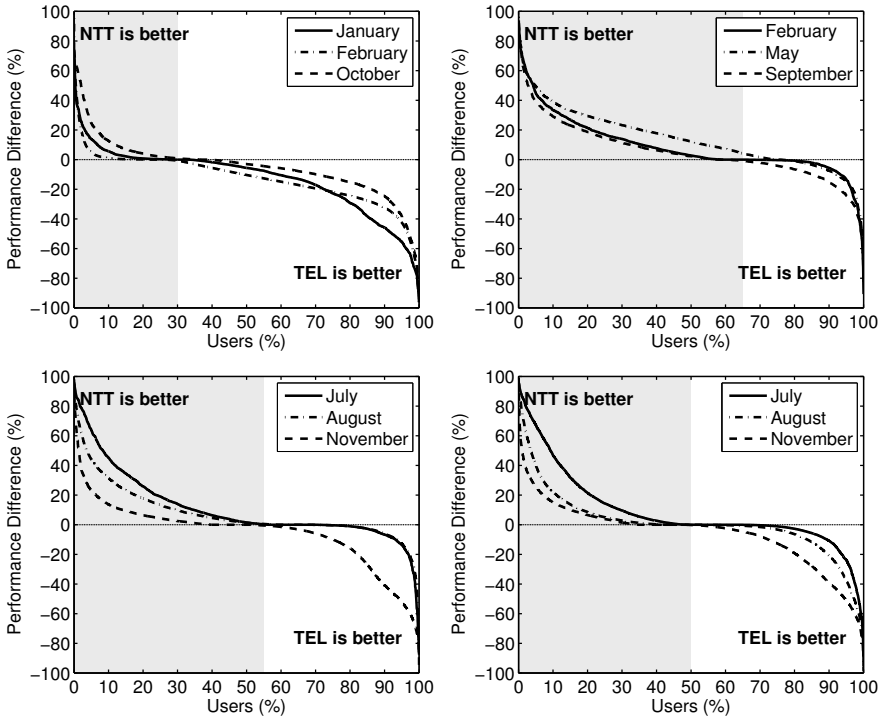


Fig. 4. Spatial performance variations for different POPs



(c) San Jose (d) Chicago

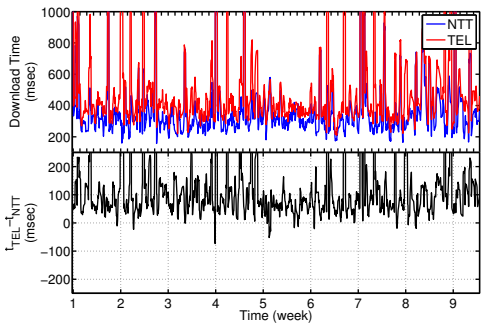
Fig. 5. Performance difference between transit providers for different billing periods

Temporal variations. To investigate temporal variations in performance differences between transit providers, we analyze our performance measurements over time. For each POP location, we again plot the distribution of percentage performance difference between transit providers for users in different one-month billing periods. Figure 5 plots the curves for four popular POP locations and for three different billing periods. We observe that for Paris and Madrid POPs, one transit provider outperforms the other for a vast majority of users across all billing periods. For example, TEL outperforms NTT for up to 70% of users at the Paris POP and NTT outperforms TEL for up to 65% of users at the Madrid POP. While the overall trend remains the same for all POPs, we observe changes across different billing periods. For instance, the performance difference between NTT and TEL increases over time for San Jose POP. More specifically, NTT has equal or better performance as compared to TEL for most users in July. However, TEL's performance improves as compared to NTT's performance over the next billing months. By November, TEL outperforms NTT for more than 25% of users while NTT outperforms TEL for approximately 10% of users. Similar temporal variations in performance difference can be observed for all POPs. During certain consecutive months two transit providers may have similar performance difference characteristics, as a result we may not observe variations over consecutive months. Therefore, we choose different months for different POPs to show variations across different geographical locations at different times.

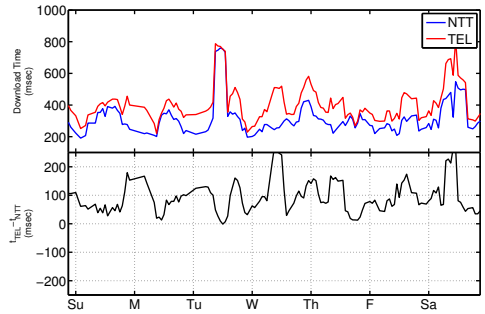
To understand finer-grained temporal variations in transit performance, we analyze the time series of percentage difference between transit providers across AS-POP pairs. We plot the hourly performance time series for the Telefonica-Madrid AS-POP pair in Figure 6. A positive value of $t_{\text{TEL}} - t_{\text{NTT}}$ indicates that NTT's performance is better than TEL's performance. For this particular AS-POP pair, we observe that NTT's performance is consistently better than TEL's performance. We plot the hourly performance time series for the Comcast-San Jose AS-POP pair in Figure 7. We observe that NTT's performance is generally better than TEL's performance during the first three weeks. However, the trend is quickly reversed in the fourth week. The reversal in performance seems to be due to improved performance of TEL (e.g., infrastructure upgrade) rather than degraded performance of NTT (e.g., traffic pattern shift). For rest of the measurement period, TEL consistently outperforms NTT. Figure 6 (b) and 7 (b) show the zoomed time series plots for the fourth week. We note diurnal variations for $t_{\text{TEL}} - t_{\text{NTT}}$, which indicate congestion during peak hours.

IV. Problem Formulation

Due to the rich connectivity at IXPs, a CDN can choose from multiple (often dozens) of transit options to route traffic to end users. In particular, the traffic to an AS can be routed via one of many transit providers. As discussed earlier, the dynamics of cost and performance makes manual solution out of scope. We need to optimize both cost and performance for each destination AS. Our discussions with network engineers revealed that this optimization is often done manually on an hourly basis, and misconfigurations and performance issues are quite common. Due to a large number of ASes, network engineers typically do the manual optimization only for large

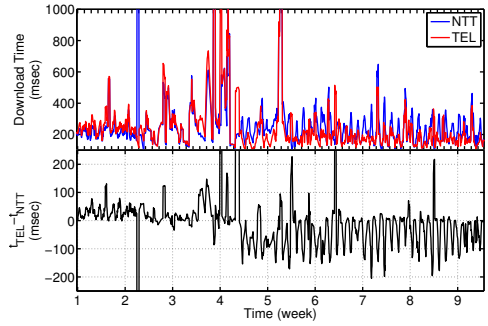


(a) Performance during 9 weeks

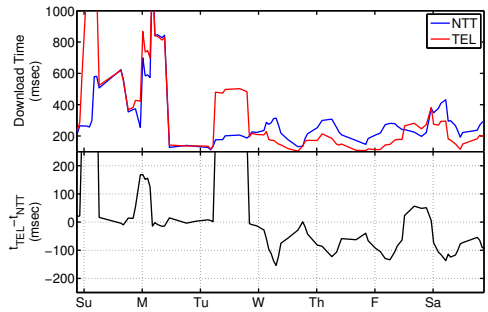


(b) Zoomed in plot for week no. 2

Fig. 6. Performance difference between transit providers for AS-POP pair: `Telefonica-Madrid`



(a) Performance during 9 weeks



(b) Zoomed in plot for week no. 4

Fig. 7. Performance difference between transit providers for AS-POP pair: `Comcast-San Jose`

ASes or the ones with poor performance. Thus, we need automated solutions to the cost-performance optimization problem that can be configured for different cost-performance tradeoffs.

We formally define the dynamic cost-performance optimization problem for Internet transit selection using constraint programming. Constraint programming allows us to formulate the optimization problem in a form that can be used to obtain a tradeoff curve between cost and performance. More specifically, we design an objective function that minimizes a utility cost function. The function is a weighted sum of cost and performance of a transit provider selection strategy. Based on the tradeoff curve, CDNs can adapt transit provider selection to obtain the desired tradeoff.

For a POP P_j, consider a transit provider i that charges r_i units of cost to a CDN. Each POP serves multiple ASes and each AS can be served through multiple transit providers. For a AS-POP pair $< A_k, P_j >$, let $b^i_{k,j}$ and $d^i_{k,j}$ respectively be the estimated bandwidth usage and performance through transit provider i. For usage based pricing r_i is a fixed amount that transit provider i charges on hourly basis. For $95th$ percentile pricing, we consider a billing period of one week and compute the $95th$ percentile of hourly bandwidth usage. We use this bandwidth usage to compute the transit costs. Specifically, for transit provider i the transit cost is $\sum_{k,j} r_i \times b^i_{k,j}$, where $b^i_{k,j}$ represents the $95th$ percentile value. We use number of records observed in each hour as an estimate of hourly bandwidth usage. The performance $d^i_{k,j}$ is measured in terms of download time in milliseconds. The CDN needs to select a transit provider for each AS-POP pair $< A_k, P_j >$. Let $x^k_{i,j}$ be the optimization variable that assigns transit provider i to AS-POP pair $< k, j >$, thus $x^k_{i,j} \in \{0, 1\}$ We formulate the minimization objective function as $\sum_{k,j} \sum_i r^i x^i_{k,j} b^i_{k,j} + \gamma x^i_{k,j} d^i_{k,j}$. The objective function minimizes the utility cost of assigning transit providers to POP-AS pairs. The utility cost function consists of two terms. The first term computes the cost in dollars charged by the selected transit providers weighted by the bandwidth usage. The second term is the median download time observed at the transit provider multiplied by a tunable parameter γ. We choose to model performance as part of the objective function because modeling performance as a constraint may render the problem unsolvable as performance of all transit providers may not satisfy the constraint. It is also possible that multiple transit providers may satisfy the constraint resulting in selecting the transit provider with poorer performance. The CDN operator can vary γ to obtain a desired tradeoff between cost and performance. Essentially, overall utility is the weighted sum of cost and performance. Smaller values of γ push the optimizer towards a minimum cost solution whereas larger values of γ push towards a better performance solution. The objective function is subject to the following constraints: $\sum_i x^i_{k,j} = 1 \forall k, j$, $x^i_{k,j} \in \{0, 1\}$, and $\sum_{k,j} x^i_{k,j} b^i_{k,j} \leq C^i$, where C^i is the capacity of transit provider i. The first constraint ensures that only one transit provider is assigned to a POP-AS pair. The second constraint ensures that assignments are integral. The third constraint sets the limits on maximum bandwidth utilization by transit providers.

Note that the second constraint for ensuring integral assignments results in a combinatorial optimization problem, where the optimal solution can be obtained by evaluating the

objective function over all possible assignments. We can utilize openly available Mixed-Integer Programming (MIP) solvers to obtain the optimal assignment. However, in practice each POP connects to thousands of ASes through multiple transit providers. Since MIP is NP-hard, it is not feasible to solve it for large input sizes. Thus, we relax the constraint to allow fractional assignments i.e., $x_{i,j}^k \in [0,1]$. This relaxation allows us to obtain optimal solution with fractional assignments in linear time using standard Linear Programming (LP). Following lemmas state the relationships between solution obtained through LP relaxation and Integer LP (ILP) [28], [35].

Lemma 4.1: Let S_{LP} be set of feasible solutions in the LP and let S_{ILP} be set of feasible solutions in ILP. All feasible solutions in ILP are also feasible solutions in LP; i.e., $S_{ILP} \subset S_{LP}$. For the minimization problem, solution obtained through relaxed LP has smaller value than the solution obtained through original ILP i.e., min(LP)$\leq$ min(ILP).

An optimal solution in LP may be found at the boundaries of the convex hull formed by the linear constraints. However, the feasible solutions in ILP are given by a set of points inside the convex hull formed by the linear constraints. These set of points do not form a convex set. Hence, value of the optimal solution in a minimization problem in LP is less than or equal to the value of an ILP optimal solution. Following the inequality in Lemma 4.1, we next state the condition under which both optimal solutions for LP and ILP are equal.

Lemma 4.2: Let X^o denote the optimal solution of the original ILP problem. Let X^l denote the optimal solution obtained from LP. If the optimization variable $x_{i,j}^k$ of the LP takes integral values in X^l, then it is also an optimal solution for ILP and hence the optimal solution for the original problem; i.e, $X^l = X^o$.

Optimal solution of a LP as stated in the lemma generally does not exist. Several heuristics have been proposed in the literature to obtain an integral solution from the relaxed LP [10]. In this paper, we use the relaxed version of the minimization problem. For our empirical evaluation, we obtain a small percentage of fractional values in the solution to the relaxed version of the original problem. The fractional assignments can be dealt heuristically through randomized rounding or greedy assignments [27]. Another approach to deal with fractional assignments is to employ fractional routing, which can be realized by hash-based splitting or through multi-homing agents [6], [23].

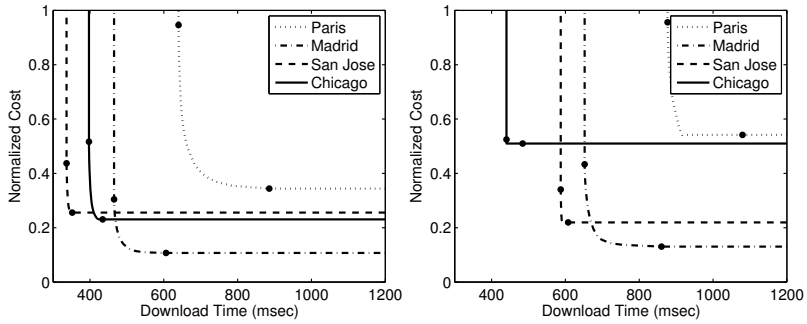
V. Results

In this section, we discuss the cost and performance benefits of our proposed optimization approach for transit selection by CDNs. For pricing, we use usage-based and $95th$ percentile pricing models to compute costs incurred by the CDN in using various transit providers. On an hourly basis, we compute the median bandwidth usage for each AS-POP pair. For usage-based pricing, we study transit provider selection for two different scenarios; first, when all transit providers charge equal pricing rates (per Mbps), and second when all transit providers charge different pricing rates. The optimization framework works on hourly basis, and the optimization works for each PoP independently over all ASes and transit providers. Therefore, differences in pricing for different POP locations (which is very common) does not impact transit providers across POPs. For $95th$ percentile charging, we use one week as the billing period in our experiments. 95th percentile charging is based on 95th percentile of all measurements over the billing period. Therefore, the optimization framework works on weekly basis under this charging model. For performance, we use download time of active measurements over a year as input to the optimization problem. For bandwidth, we use number of records observed in each hour as an estimate of bandwidth usage in that hour. As number of records are dependent on clients requesting the pixel tags, larger number of requests indicate greater bandwidth usage, therefore, we use number of records as an estimate for bandwidth usage. Specifically, we compute the median download time in every hour for all transit providers for each POP-AS pair. For a given transit provider, our system records several measurements for each AS-POP pair during a particular hour. We use the median value of these measurements in our experiments. We use an open source implementation of Embedded Conic Solver (ECOS) for solving the optimization problem. Given the solver output, we can compute the performance and cost of traffic via transit providers. While our optimization framework is scalable for more than two transit providers, we limit our analysis to two transit providers (namely `NTT` and `TEL`) for simplicity.

A. Tradeoff analysis

We first study the tradeoffs between cost and performance in the Internet transit selection. By varying γ, we obtain a tradeoff curve between cost and performance for each POP. For each hour, we compute the cost and download time using the first and second terms of our objective function, respectively. We repeat this process for different γ values and obtain a tradeoff curve between total cost and average download time per AS. Each point on the tradeoff curve represents a transit provider selection strategy for all ASes served by the POP. We solve the optimization problem every hour and use average to obtain a tradeoff curve.

Figure 8 plots the cost-performance tradeoff curves for selection between two transit providers at four POP locations. We plot the tradeoff curves for usage-based pricing models and $95th$ percentile pricing models in Figures 8(a) and (b), respectively. The x-axis is the average download time per AS (in milliseconds) and the y-axis represents the incurred transit costs (in dollars). Each point on the curve represents the average cost and performance over a period of one month for a given γ value. To plot the tradeoff curves, we vary the values of γ from 0 to 5 with an increment of 0.1. For smaller values of γ, we obtain solutions with lower costs but worse performance. For $\gamma = 0$, we get the lowest cost solution that selects transit providers with minimum costs irrespective of their performance. As we increase the value of γ, we give higher weight to performance and we obtain higher cost solutions with better performance. For $\gamma = 5$, we get the best



(a) Usage-based pricing (b) 95th percentile pricing

Fig. 8. Cost-performance tradeoff curves for different POPs and pricing models, with γ ranging from 0 to 5.

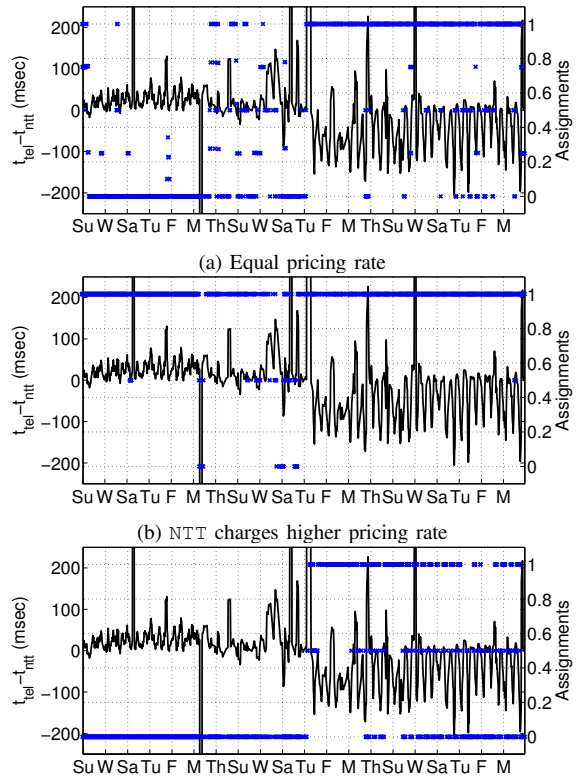
performance solution that always selects the transit providers with the lowest download time. The pairs of dots on each curve represent the minimum cost solution for $\gamma = 0$ and best performance solution for $\gamma = 5$. The lines are extrapolated to show solutions that can be obtained for values of γ less than 0 and greater than 5.

Comparing Figures 8(a) and (b), we note that the CDN ends up paying more under the $95th$ percentile pricing model. For Paris and Madrid POPs, we observe smooth tradeoff curves, indicating that the CDN can select among a number of strategies by varying γ around the knee of the tradeoff curves. We also observe that regardless of the pricing model, the CDN can save more than 50% on cost with a performance degradation of less than 40 milliseconds. For Chicago and San Jose POPs, we observe sharp knees of the tradeoff curves which indicate that slightly changing γ values has a substantial impact on cost and performance. For these POPs, the CDN can save on average 30% on cost with a performance degradation of less than 20 milliseconds. For large values of γ, the CDN may end up paying a lot more even though there is much room for savings without significantly degrading the performance. Thus, there is very limited flexibility for CDNs in terms of optimal transit provider selection. In summary, the tradeoff curves under the usage based pricing model and the $95th$ percentile charging model show that the CDN can reduce transit cost by 57% and 35% respectively, without significant degradation in performance.

B. Transit selection analysis

Recall that transit provider selections are outputs of our optimization framework, i.e., the values of the target variable $x^k_{i,j}$ provides selection results. We analyze transit provider selections for three different usage-based pricing scenarios: (1) Both transit providers charge equal rate, (2) NTT charges higher rate than TEL, and (3) TEL charges higher rate than NTT. For equal pricing scenario, the optimization framework provides the performance-optimal solution. For unequal pricing scenarios, the optimization framework selects the best transit provider in terms of both cost and performance.

Figure 9 plots the timeseries of performance difference between TEL and NTT for an example POP-AS pair. The primary y-axis (left side) shows the difference in performance measurements between the two transit providers. The secondary y-axis (right side) shows the output selection values for transit provider TEL, where 1 indicates TEL is selected and 0 indicates NTT is selected. Figure 9(a) shows selections when both transit providers charge equal rates. We observe that during the first four weeks TEL performs worse than NTT, as a result TEL is not selected. For the later half, the performance difference becomes negative indicating that TEL is performing better than NTT, thus TEL is selected during the later half. Figures 9(b) and (c) show selections when NTT and TEL charge different rates. In both cases, we selected a γ value that provides the best tradeoff between performance



(a) Equal pricing rate

(b) NTT charges higher pricing rate

(c) TEL charges higher pricing rate

Fig. 9. Transit provider selection for different pricing scenarios. Blue dots represent transit provider selections (right y-axis); assignments of 1 indicate TEL selection and assignments of 0 indicate NTT selection. (a) NTT and TEL charge equal pricing rate and the solution is performance-optimal. (b,c) NTT and TEL charge difference pricing rates and the solution is a tradeoff between cost and performance.

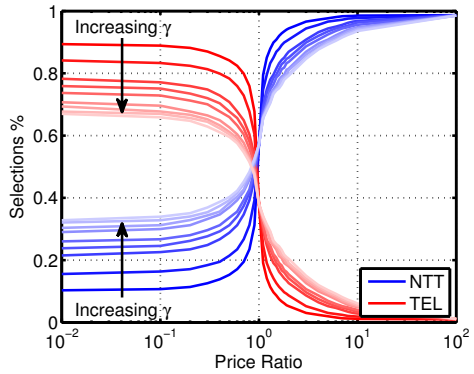


Fig. 10. Transit provider selection for varying price ratio: $\frac{\texttt{TEL}\ \ price}{\texttt{NTT}\ \ price}$.

and cost, i.e., the knee of the tradeoff curve. When NTT charges more, TEL is mostly selected even when TEL has worse performance. Whereas when TEL charges more, we see selections for both NTT and TEL. This shows the impact of performance difference and γ values on transit provider selection. In the later half, performance difference is larger, and as a result the optimal selection is to select the best performing transit provider. Therefore, even though TEL is charging more than NTT, TEL is selected because performance improvement trumps additional cost.

In order to systematically quantify the impact of different pricing on transit route selection, we compare selections for varying pricing rates. Figure 10 plots the percentage of selections obtained from the optimization framework for varying price ratio between NTT and TEL. The y-axis shows the fraction of total number of hours in the input data. Each pair of symmetric red/blue lines correspond to a given γ value. The fraction of NTT selections are shown in blue and the fraction of TEL selections are shown in red and are symmetric to the corresponding blue line. Price ratio is calculated as $\frac{\texttt{TEL}\ \ price}{\texttt{NTT}\ \ price}$. The plot shows that when TEL is 10 times cheaper than NTT, TEL has larger percentage of selections. Furthermore, when TEL is 10 times expensive than NTT, selections for NTT reach more than 90% for smaller γ values and more than 85% for larger γ values. Overall, we observe that as γ increases the fraction of NTT selections also increase regardless of performance. This finding indicates that overall NTT is favorable in terms of performance.

VI. Related Work

There is prior work on route selection from the perspective of stub ISPs, edge networks, and enterprise networks. Several studies have measured and analyzed performance and pricing aspects of such multi-homed networks. In contrast, this paper studies transit route selection at IXPs by CDNs.

Multi-homing: Akella *et al.* presented a large scale measurement based performance analysis of multihoming [6]. The authors evaluated multihoming performance from the perspectives of enterprises and content providers, and found that multihoming improves performance. In [11], Dhamdhere *et al.* proposed a two step methodology for optimizing cost and performance in multi-homing. Tao *et al.* also explored the performance benefits of path switching in multi-homing scenarios and overlay network scenarios [33]. In [14], Yang *et al.* proposed smart routing algorithms to optimize cost and performance for multihomed enterprises and stub ISPs. The customer ISPs are assumed to be multihomed to a set of ISPs and they share ISP links for multihoming purposes. The proposed algorithms dynamically assign traffic to achieve optimal cost and performance. In [13] the authors showed that route control systems in multihomed networks can cause traffic oscillations between available paths which have adverse affects on performance. The authors used bandwidth information and randomization in their algorithms to avoid oscillations. In [36], Wang *et al.* studied monetary aspects of multihoming. The authors proposed a dynamic programming algorithm to obtain a subset of ISPs for multihoming purposes so that the cost is minimized for the customer. They also study dynamics of ISP pricing strategies in response to customer subscriptions. More recently, in [37] and [23], the authors proposed measurement and optimization frameworks for service provider selection for content delivery from content providers to its users.

As compared to prior work on multi-homing, our work is different in two major dimensions: First, we focus on multi-homed CDN servers at IXPs, which offer a large number of possible interconnections between CDNs and ISPs. Prior work on multi-homing assumes the multi-homed network to be an edge network, whereas in our case we deal with multi-homed CDN servers at IXPs. Second, our light-weight measurement methodology is unique as it allows us to measure performance across multiple transit routes simultaneously at a large scale.

Transit pricing: Pricing in the internet transit has been studied from two different perspectives. First, from the transit ISP perspective, where transit ISPs aim to maximize their profit while satisfying service quality and customer traffic demands. Prior work has studied the impact of adopting various pricing strategies on ISP profits such as time-dependent pricing, tiered pricing, and pricing differentiation [16], [19], [20]. In particular, Valancius *et al.* studied the affect of destination based tiered pricing on ISP profit [34]. The authors developed demand and cost models using real-world traffic data and showed the impact of various traffic bundling strategies on ISP profits. In contrast, we study pricing from the transit ISP customer's perspective, where CDNs aim to minimize transit costs while achieving best performance. For transit cost minimization, prior work has studied various transit link sharing strategies to reduce transit costs [7], [8], [15], [29]. Our work focuses on cost reduction by selecting best available transit route which also provides best performance.

Server deployment: Another category of prior work is focused on optimizing server deployments by CDNs. In [26], the authors study the problem of online placement of servers to minimize cost in terms of latency, hop count and economic cost. In [12], the authors propose NetPaaS a system that enables CDN-ISP collaboration leading to informed server placement and end-user to server assignment. In [17], the authors formulate and solve the cache deployment optimization problem to minimize CDN deployment costs while maintaining end-user performance.

VII. Conclusion

In this paper, we studied the problem of optimal Internet transit selection from the perspective of CDNs. To the best of our knowledge, transit route selection at IXPs by CDNs is not studied in prior literature. We make the following key contributions in this paper. First, we propose a method to conduct simultaneous performance measurements across multiple transit routes which are maintained by CDNs for delivering content to access ISPs. Second, we provide spatio-temporal characterization of performance differences observed across multiple transit routes. Our findings highlight that there is not an outright best transit provider. Third, we formulate the optimal transit route selection problem as an optimization problem and analyze tradeoff curves and resulting selection strategies for different pricing of transit routes. We observe sharp knees in tradeoff curves for some geographical regions, indicating that there is very limited flexibility for CDNs in terms of choosing a suitable transit provider. Without an optimal selection strategy, CDNs may end up sacrificing performance and/or cost. Using our proposed approach, CDNs can reduce transit costs on average by 57% without sacrificing performance.

Acknowledgements

We thank our shepherd, Niklas Carlsson, and the anonymous reviewers for their useful feedback on this paper. This work is supported in part by the National Science Foundation under Grant Numbers CNS-1318563, CNS-1524698, CNS-1421407 and 1464110, and the National Natural Science Foundation of China under Grant Numbers 61472184 and 61321491, and the Jiangsu High-level Innovation and Entrepreneurship (Shuangchuang) Program.

References

[1] Facts & Figures - Akamai Technologies, Inc. http://www.akamai.com/html/about/facts_figures.html.
[2] Google Global Cache (GGC). https://peering.google.com/about/ggc.html.
[3] Netflix Open Connect Content Delivery for ISPs. https://openconnect.netflix.com.
[4] PeeringDB. http://www.peeringdb.com.
[5] White paper: Cisco visual networking index forecast and methodology 2015-2020. Technical report, Cisco.
[6] A. Akella, B. Maggs, S. Seshan, A. Shaikh, and R. Sitaraman. A measurement-based analysis of multihoming. In *Proc. of the SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pages 353–364, 2003.
[7] I. Amigo, P. Belzarena, and S. Vaton. On the problem of revenue sharing in multi-domain federations. In *Proc. of the IFIP Networking*, pages 252–264, 2012.
[8] I. Castro and S. Gorinsky. T4P: Hybrid interconnection for cost reduction. In *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 178–183, 2012.
[9] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger. There is more to IXPs than meets the eye. *ACM SIGCOMM Computer Communication Review*, 43(5):19–28, 2013.
[10] Y. Crama and P. Hammer. *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*. 2010.
[11] A. Dhamdhere and C. Dovrolis. ISP and egress path selection for multihomed networks. In *Proc. of the IEEE INFOCOM*, pages 1–12, 2006.
[12] B. Frank, I. Poese, Y. Lin, G. Smaragdakis, A. Feldmann, B. Maggs, J. Rake, S. Uhlig, and R. Weber. Pushing CDN-ISP Collaboration to the Limit. *ACM SIGCOMM Computer Communication Review*, 43(3):34–44, 2013.
[13] R. Gao, C. Dovrolis, and E. W. Zegura. Avoiding oscillations due to intelligent route control systems. In *Proc. of the IEEE INFOCOM*, pages 1–12, 2006.
[14] D. K. Goldenberg, L. Qiuy, H. Xie, Y. R. Yang, and Y. Zhang. Optimizing cost and performance for multihoming. *ACM SIGCOMM Computer Communication Review*, 34(4):79–92, 2004.
[15] L. Gyarmati, R. Stanojevic, M. Sirivianos, and N. Laoutaris. Sharing the cost of backbone networks: cui bono? In *Proc. of the ACM IMC*, pages 509–522, 2012.
[16] P. Hande, M. Chiang, R. Calderbank, and J. Zhang. Pricing under constraints in access networks: Revenue maximization and congestion management. In *Proc. of the IEEE INFOCOM*, pages 1–9, 2010.
[17] S. Hasan, S. Gorinsky, C. Dovrolis, and R. K. Sitaraman. Trade-offs in optimizing the cache deployments of CDNs. In *Proc. of the IEEE INFOCOM*, pages 460–468, 2014.
[18] C. Huang, A. Wang, J. Li, and K. W. Ross. Measuring and evaluating large-scale CDNs. In *Proc. of the ACM IMC*, pages 15–29, 2008.
[19] L. Jiang, S. Parekh, and J. Walrand. Time-dependent network pricing and bandwidth trading. In *Proc. of the IEEE Network Operations and Management Symposium Workshops, NOMS*, pages 193–200, 2008.
[20] G. Kesidis, A. Das, and G. de Veciana. On flat-rate and usage-based pricing for tiered commodity internet services. In *Proc. of the IEEE CISS*, pages 304–308, 2008.
[21] R. Krishnan, H. V. Madhyastha, S. Srinivasan, S. Jain, A. Krishnamurthy, T. Anderson, and J. Gao. Moving beyond end-to-end path information to optimize CDN performance. In *Proc. of the ACM IMC*, pages 190–201, 2009.
[22] A. Lodhi, N. Larson, A. Dhamdhere, C. Dovrolis, and kc claffy. Using PeeringDB to understand the peering ecosystem. *ACM SIGCOMM Computer Communication Review*, 44(2):20–27, 2014.
[23] S. Narayana, W. Jiang, J. Rexford, and M. Chiang. Joint server selection and routing for geo-replicated services. In *Proc. of the IEEE/ACM Conf. on Utility and Cloud Computing*, pages 423–428, 2013.
[24] W. B. Norton. *The Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press, 2011.
[25] W. B. Norton. *The 2014 Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press, 2014.
[26] L. Qiu, V. N. Padmanabhan, and G. M. Voelker. On the placement of Web server replicas. In *Proc. of the IEEE INFOCOM*, pages 1587–1596, 2001.
[27] P. Raghavan and C. D. Tompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.
[28] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.
[29] R. Stanojevic, I. Castro, and S. Gorinsky. CIPT: Using tuangou to reduce IP transit costs. In *Proc. of the ACM CoNEXT*, page 17, 2011.
[30] R. Stanojevic, N. Laoutaris, and P. Rodriguez. On economic heavy hitters: Shapley value analysis of 95th-percentile pricing. In *Proc. of the ACM IMC*, pages 75–80, 2010.
[31] A.-J. Su, D. R. Choffnes, A. Kuzmanovic, and F. E. Bustamante. Drafting behind Akamai: Inferring network conditions based on CDN redirections. *IEEE/ACM Transactions on Networking (TON)*, 17(6):1752–1756, 2009.
[32] P. Sun, M. Yu, M. J. Freedman, and J. Rexford. Identifying performance bottlenecks in CDNs through TCP-level monitoring. In *Proc. of the ACM SIGCOMM workshop on Measurements up the stack*, pages 49–54, 2011.
[33] S. Tao, K. Xu, Y. Xu, T. Fei, L. Gao, R. Guerin, J. Kurose, D. Towsley, and Z. L. Zhang. Exploring the performance benefits of end-to-end path switching. In *Proc. of the IEEE ICNP*, pages 304–315, 2004.
[34] V. Valancius, C. Lumezanu, N. Feamster, R. Johari, and V. V. Vazirani. How many tiers?: Pricing in the Internet transit market. In *Proc. of the ACM SIGCOMM*, pages 194–205, 2011.
[35] R. Vanderbei. *Linear Programming: Foundations and Extensions*. International Series in Operations Research & Management Science. 2007.
[36] H. Wang, H. Xie, L. Qiu, A. Silberschatz, and Y. R. Yang. Optimal ISP subscription for Internet multihoming: Algorithm design and implication analysis. In *Proc. of the IEEE INFOCOM*, pages 2360–2371, 2005.
[37] Z. Zhang, M. Zhang, A. G. Greenberg, Y. C. Hu, R. Mahajan, and B. Christian. Optimizing cost and performance in online service provider networks. In *Proc. of the USENIX NSDI*, pages 33–48, 2010.