2006

Algorithms for Optimizing Bandwidth Costs on the Internet

Micah Adler
Ramesh Sitaraman, *University of Massachusetts - Amherst*
Harish Venkataramani

Algorithms for Optimizing Bandwidth Costs on the Internet

Micah Adler, Ramesh Sitaraman, and Harish Venkataramani
Department of Computer Science,
University of Massachusetts Amherst,
Amherst, MA 01003, USA.
Email: {micah,ramesh,harish}@cs.umass.edu

Abstract— **Content Delivery Networks (CDNs) deliver web content to end-users from a large distributed platform of web servers hosted in data centers belonging to thousands of Internet Service Providers (ISPs) around the world. The bandwidth cost incurred by a CDN is the sum of the amounts it pays each ISP for routing traffic from its servers located in that ISP out to end-users. A large enterprise may also contract with multiple ISPs to provide redundant Internet access for its origin infrastructure using technologies such as multihoming and mirroring, thereby incurring a significant bandwidth cost across multiple ISPs. This paper initiates the formal study of bandwidth cost minimization in the context of a large enterprise or a CDN, a problem area that is both algorithmically rich and practically very important. First, we model different types of contracts that are used in practice by ISPs to charge for bandwidth usage, including average, maximum, and 95^{th}-percentile contracts. Then, we devise an optimal offline algorithm that routes traffic to achieve the minimum bandwidth cost, when the network contracts charge either on a maximum or on an average basis. Next, we devise a deterministic (resp., randomized) online algorithm that achieves cost that is within a factor of 2 (resp., $\frac{e}{e-1}$) of the optimal offline cost for maximum and average contracts. In addition, we prove that our online algorithms achieve the best possible competitive ratios in both the deterministic and the randomized cases. An interesting theoretical contribution of this work is that we show intriguing connections between the online bandwidth optimization problem and the seemingly-unrelated but well-studied ski rental problem where similar optimal competitive ratios are known to hold. Finally, we consider extensions for contracts with a committed amount of spend (CIRs) and contracts that charge on a 95^{th}-percentile basis.**

I. Introduction

Over the past years, the Internet has emerged as a business-critical medium for an enterprise to communicate with their vendors and clients. However, the Internet itself was designed as a best-effort delivery network with no guarantees on availability or performance. The Internet is a network of networks, where each network is managed independently by an Internet Service Provider (ISP) who builds and manages the routers, links, and other networking infrastructure. As such, there are tens of thousands of ISPs that constitute the Internet today, ranging from large Tier-1 providers with a global presence (such as Level 3, ATT, Sprint), national providers (such as China Telecom, VSNL in India, SingTel in Singapore), regional providers (such as Earthnet), and local ISPs. An enterprise requiring high-levels of availability for their Internet services faces a fundamental challenge. It is not sufficient for the enterprise to obtain their Internet connectivity from a single ISP, as any single ISP is prone to failure caused by router mishaps, fiber cuts, and configuration snafus. Therefore, many enterprises use strategies such as mutihoming and mirroring that allow them to access the Internet using multiple ISPs and data centers. In addition, many major enterprises also use a Content Delivery Network (CDN) that is a large fault-tolerant distributed platform of web servers hosted in potentially thousands of ISPs. Examples of such CDNs include Akamai [6] and Limelight [5]. More than 15% of the web traffic today use CDNs, including most major media, entertainment, ecommerce and extranet portals.

The model and results of this paper apply in several general technological contexts where cost-efficient traffic management is critical, but we will pick CDN technology as a concrete application to illustrate the value of the results. A CDN works as follows. A CDN negotiates network contracts to buy Internet bandwidth from a large number of ISPs and co-locates its servers in those ISPs. An end-user accessing web content hosted on the CDN is directed[1] by the CDN to an appropriate server at one of the contracted ISPs, so as to optimize availability and performance for the end-user and minimize bandwidth costs for the CDN. The browser of the end-user then downloads the content from that server. Thus, a CDN operates as an "Internet traffic-cop" by controlling which portion of the traffic is served from which ISP. The traffic assignments happen in an online and "real-time" fashion where assignments are changed periodically at the time granularity of minutes (say, every 5 minutes). For more details on CDN architecture, please see an overview paper co-written by one of the authors of this paper [1].

A CDN can be viewed as a reseller of bandwidth, where it pays each ISP for the traffic served from that ISP to end-users. A CDN in turn gets paid by the enterprises for the traffic the CDN delivered on their behalf. A significant portion of the variable costs of operating a CDN is the total

[1]The mechanism for directing an end-user to a particular server at a particular ISP is typically through the Domain Name Service (DNS), where a domain name such as "www.yahoo.com" is translated to the IP address of the selected server by the CDN's DNS servers.

1-4244-0595-5 /06/$20.00 ©2006 IEEE.

bandwidth costs that it pays the ISPs, and minimizing this cost is the primary focus of this paper. In particular, CDNs buy bandwidth from ISPs using network contracts that fall into three categories depending on how billable traffic (in Mb/s) is computed over the billing period (typically, a month). Given a monthly profile of the traffic sent from an ISP, the billable traffic is either computed on an average basis (AVG), a maximum basis (MAX), or as the 95^{th} percentile (95^{th}) of the traffic profile. While most real-world contracts are either AVG or 95^{th}, MAX is highly important from a practical system design perspective, since traffic cannot be controlled in a precise enough fashion[2]to take advantage of the 5% window for free traffic in a 95^{th} percentile contract. Therefore, real-life traffic optimizers view 95^{th} percentile contracts as MAX contracts for purposes of the optimization, and hence studying the MAX contract model is very important.

The overall bandwidth cost to be minimized is the sum of the costs incurred at each individual ISP, where each ISP charges for the billable traffic at a specific contracted rate. Because of these contract type differences, a given traffic profile for a billing period can have a significantly different cost depending on how it is assigned to the ISPs. The primary focus of this paper is the problem of how a CDN should assign traffic to multiple ISPs to minimize the overall bandwidth costs incurred.

While the model and results presented in this paper use CDNs as a motivating example, the results are also applicable to other important technologies such as multihoming [2], where an enterprise contracts with multiple ISPs to provide redundant Internet access for its origin infrastructure. The enterprise would then route traffic to and from its origin via the multiple uplinks, so as to minimize bandwidth costs and maximize availability and performance. Recently, commercial offerings from RouteScience [3] and Internap [4] offer products that help the enterprise optimize traffic across the different ISPs that provide network connectivity.

The astute reader will recognize that while this paper considers optimizing cost in isolation, real-world technologies such as CDNs and multihoming aim to first optimize a notion of performance (such as minimizing web download time by reducing latency and loss) while striving to optimize cost. However, pure cost optimization is an important first step for the following reasons.

- From an algorithmic standpoint understanding pure cost optimization is a major stepping stone for the more general bi-criteria cost-performance optimization that we plan to do in future work. We believe that the algorithmic ideas generated in this study will shed light on the more complex bi-criteria optimization framework.
- Different types of traffic have different sensitivities to performance and cost. Delivering a real-time application is extremely performance sensitive but also less cost sensitive as customers are willing to pay more for higher performance. However, other types of traffic such as (non-realtime) background downloads of large files is less performance sensitive but also more cost sensitive as customers expect to pay much less. The latter situation is more closely alligned with the pure cost optimization regime presented in this paper.
- The pure cost optimization studied in this paper provides a lower bound on the bandwidth cost achievable by any real-world system that simultaneously optimizes performance and cost. Comparing the actual incurred cost with this lower bound delineates the portion of the actual cost that is intrinsic to the contracts and traffic from the remaining additional cost premium attributable to providing performance and other considerations. Understanding this cost premium and how it varies with different types of traffic is critical to understanding the cost structure of service.

[2]The imprecision comes from several sources. For instance, some browsers don't comply with TTLs in a precise fashion, and traffic moved away from an ISP by the optimizer will decay slowly over time instead of falling sharply.

A. *Prior Work*

Considering the practical importance of the problem in recent years, heuristic implementations exist. However, this is the first formal study of algorithms for bandwidth cost minimization across multiple ISPs. Recently, there has been some interesting work on cost minimization from a multihoming perspective [15] where AVG and 95^{th} percentile contracts are considered and empirically evaluated. However, our work is unique in considering the typical CDN situation where the optimizer *simultaneously* routes traffic to ISPs with *bounded* capacities and a mix of contract types, and formal bounds for optimality are shown in the competitive ratio framework for online algorithms. There is extensive literature on online algorithms [13], [14]. Prior research on online algorithms for ski-rental and related problems [12], [11] is particularly relevant as we show interesting connections between our problem and this class of problems.

B. *Our Contributions*

The first contribution of the paper is the modeling and formulation of an area of great practical importance with a rich potential for future algorithmic investigation. The model and algorithms presented here are immediately relevant to commercial technologies of today, advancing the current state-of-the-art. In Section II, we derive an optimal offline algorithm that routes traffic to a set of ISPs with AVG and MAX contracts such that the total cost is minimized. The offline algorithm assumes that the traffic that needs to be routed for the entire billing period is known in advance. While this is not an assumption that holds in practice, note that the offline optimal algorithm produces a lower bound on the cost against which any online algorithm can be compared at the end of each billing period.

In Section III-A, we turn to online algorithms that know only the current and the past traffic levels, and are unaware of any events in the future. Specifically, we devise a deterministic online algorithm that is at most a factor of 2 in cost from the optimal offline solution. Further, in Section III-B, we devise a randomized online algorithm that has an expected cost that is a factor of at most $\frac{e}{e-1}$ from optimal. In both cases, we show that the competitive ratios are the best possible. An interesting theoretical contribution of this work is that we show intriguing connections between the online bandwidth optimization problem and the seemingly-unrelated but well-studied ski rental problem. Specifically, our work shows that the online decision to route through a MAX versus an AVG contract is a generalized form of the buy-versus-rent decision in the ski-rental problem. This furthers our understanding of the class of online problems where competitive ratios of 2 and $e/(e-1)$ are optimal for deterministic and randomized online algorithms respectively. Other problems in this class include previously known generalizations of ski rental, such as the Bahncard problem [9] and the TCP Acknowledgment problem [8], [11] where the same competitive ratios apply. Next, in Section IV, we extend the contract framework to include the notion of a committed information rate (CIR), where the CDN must send a certain committed amount of traffic through an ISP. We extend our results of Section II to provide an optimal offline algorithm for MAX and AVG contracts with CIR. Finally, we show that optimizing costs for 95^{th} percentile contracts is NP-hard, differentiating it from the MAX and AVG contracts.

C. Problem Description

1) Network Contracts: A first important step in our study is accurately modeling a network contract with an ISP. While a network contract is a complex legal document, there are three important parameters that provide a simple yet realistic model for designing applicable optimization algorithms.

Type. The contract type dictates how the ISP will bill for the traffic that is sent over its link. The billing period (typically a month) is divided into M 5-minute time buckets (typically, $M = 8640$), and the total traffic sent on the link is averaged within each 5-minute interval. The three types of contracts we study are AVG, MAX, and 95th contracts where the billable traffic is computed as the average, maximum or the 95th percentile respectively of the 5-minute-bucket-averages in the billing period. The AVG and 95^{th} contracts account for most network contracts in existence today. As noted earlier, routing traffic demand on the Internet is imprecise, since the offered traffic load is often hard to estimate and the controls are imprecise (for instance, when web traffic is moved from one ISP to another ISP, it may take several minutes for the move to take effect depending on DNS TTLs and browser behavior). Due to the imprecision in both traffic estimation and control, 95^{th} contracts are handled as though they were MAX contracts in practice. Hence, the great importance of studying MAX contracts. Further, as we will see MAX contracts are more tractable and provide good insights into the underlying optimization.

Unit Cost. *Unit cost* is the cost per Mbps of billable traffic. Let $x_1 \geq x_2 \geq \cdots \geq x_M$ be the average traffic within each of the M 5-minute buckets during the billing period, placed in descending order. For an AVG contract, the bill for the month is $C_{AVG} * (\sum_i x_i/M)$, where C_{AVG} is the unit cost. For a MAX contract, the bill for the month is $C_{MAX} * (x_1)$, where C_{MAX} is the unit cost. Likewise, for a 95^{th} contract, the bill for the month is $C_{95th} * (x_{\frac{M}{20}})$.

Capacity. The capacity P is the maximum bandwidth (in Mbps) that one can send through the uplink of the ISP.

In addition to these three parameters, an additional parameter called the Committed Information Rate (CIR) is important to model. CIR represents the committed amount of billable traffic that must be sent through an ISP. The CIR is paid for in advance, whether or not it is used. CIRs are considered in the later part of the paper in section IV.

2) The Bandwidth Cost Minimization Problem: The optimization problem proposed here is part of a core component in a CDN that senses the incoming traffic requests and assigns them to servers in multiple ISPs. Typically, the routing is performed by resolving domain names using DNS, and the incoming traffic represents requests from thousands of nameservers around the world. Since each nameserver can be routed independently to an ISP, we assume that the traffic is splittable and assignable in any fashion to the ISPs[3].

The Internet bandwidth cost minimization problem is modeled as follows. The billing period (typically one month) is divided into M 5-minute time buckets. We model the incoming traffic as a sequence b_t, $1 \leq t \leq M$, where b_t is the average traffic (Mbps) in time bucket t. Each b_t represents the average traffic demand from end-users that must be served from the contracted ISPs at time bucket t. At any time t, a traffic routing algorithm partitions the incoming traffic b_t and assigns y_t^j Mbps to ISP_j such that $\sum_j y_t^j = b_t$. Further, it ensures that capacity constraints are met at each ISP_j and at each time $1 \leq t \leq M$, i.e., $y_t^j \leq P_j$, where P_j is the capacity of ISP_j.

An offline algorithm knows the entire time-ordered input sequence of traffic demands, $I = \langle b_t \rangle$, $1 \leq t \leq M$, for the entire billing period. It makes traffic routing decisions based on this complete knowledge. An online algorithm makes routing decisions at time t knowing only b_j, $1 \leq j \leq t$, i.e., knowing only the past and current values. Note that the incoming traffic b_t, the traffic assignments y_t^j, and capacities P_j are *integral* values in the units of bits per second.

In this paper, we study both offline and online algorithms for traffic management that optimize the total cost incurred in the network contracts for the billing period. We use the notion of competitive ratio to bound the cost $C_A(I)$ of an online algorithm A in terms of the optimal offline cost of

[3]This is a good first-cut approximation as most of the internet web traffic comes from a large number of nameservers, where each nameserver accounts for only a small fraction of the total traffic. Further, each individual web request induces only a small amount of traffic and can be served within a short period of time (milliseconds), as web objects are small on average (under 10KB).

$C_{OPT}(I)$. In particular, a deterministic online algorithm A is said to be c-competitive if there exists a constant α such that for all input sequences I, $C_A(I) \leq c \cdot C_{OPT}(I) + \alpha$. A similar competitive notion applies to randomized online algorithms where the *expected* value of the cost is used instead. Note that the competitive ratio guarantees derived for our online algorithms hold in the worst-case, irrespective of the behavior and (un)predictability of the incoming traffic.

II. The Offline Algorithm

In this section, we derive an optimal offline algorithm that routes traffic with minimum total bandwidth cost to ISPs with AVG or MAX contracts.

To start with, assume that we are given contracts from m MAX ISPs Max_i, $1 \leq i \leq m$, such that $C_{Max_1} \leq C_{Max_2} \leq \cdots \leq C_{Max_m}$. Define the threshold t_{Max_i} of an ISP Max_i to be the maximum traffic routed during the billing period through that ISP. The proof of the following Lemma is easy, and is omitted.

Lemma 1: There exists an optimal solution in which Max_i is not used in a time interval unless each ISP Max_j, $j < i$, has been used to its full capacity of P_{Max_j}.

Thus the greedy algorithm of using a cheaper MAX ISPs to its full capacity before using a costlier MAX ISPs routes traffic through m MAX ISPs with the least cost. As the cost is determined by the bucket with most traffic to be routed the time taken to calculate the cost of the optimal routing is $O(m \log m + M)$ (sorting is necessary).

Now we give a similar greedy algorithm for routing traffic when we have only AVG ISPs. Assume that we are given contracts from n AVG ISPs Avg_i, $1 \leq i \leq n$, such that $C_{Avg_1} \leq C_{Avg_2} \leq \cdots \leq C_{Avg_n}$. The proof of the following Lemma is easy, and is omitted.

Lemma 2: In any optimal solution, ISP Avg_i is not used in a time interval unless each ISP Avg_j, $j < i$, is used to its full capacity.

Thus the greedy algorithm where in each interval a cheaper AVG ISPs is used to its full capacity before using costlier AVG ISPs routes traffic through n AVG ISPs with the least cost. We can find the most expensive AVG ISP that needs to be used in a bucket in $O(\log n)$ time by using binary search to search for the bucket capacity in an array of size n, whose k^{th} element is $\sum_{i=1}^{k} C_{Avg_i}$ for $1 \leq k \leq n$. As the ISPs need to be sorted by their cost, the total time taken to calculate the cost of the optimal solution is $O((n + M) \log n)$.

Now we consider the general case when we have *both* MAX ISPs and AVG ISPs. Assume that we are given contracts from m MAX ISPs Max_i, $1 \leq i \leq m$, such that $C_{Max_1} \leq C_{Max_2} \leq \cdots \leq C_{Max_m}$. Further, assume that we are also given contracts from n AVG ISPs Avg_i, $1 \leq i \leq n$, such that $C_{Avg_1} \leq C_{Avg_2} \leq \cdots \leq C_{Avg_n}$. Now we show that there exists an optimal solutions which has the same form as Figure 1.

Lemma 3: There exists an optimal solution such that in any time interval an AVG ISP is used only if all MAX ISPs are used to their respective thresholds for the billing period.

Proof: Start with any optimal solution where ISP Avg_i receives $x > 0$ units of traffic in a time interval, but some ISP Max_j is used less than its threshold by $y > 0$ units. By moving $\min\{x, y\} > 0$ units of traffic from Avg_i to Max_j, the total cost of ISP Avg_i does not increase while the cost of Max_j remains the same. Thus, the overall cost does not increase and we have an optimal solution which satisfies the given property. □

Thus there exists a dividing line (Figure 1) such that all traffic below this line is routed through MAX ISPs and all traffic above is routed through AVG ISPs.[4] Thus the problem can be broken into three parts - finding the optimal height h of the dividing line, routing traffic below the height h through MAX ISPs and routing traffic above the height h through AVG ISPs. The problem of routing traffic below (resp., above) the dividing line at height h through only MAX ISPs (resp., AVG ISPs) can be solved by greedy algorithms given above. The Max-Threshold h, defined to be the sum of the thresholds of the MAX ISPs, can be found by binary search using the following lemma.

Define $C_{Max}(h)$ (resp., $C_{Avg}(h)$) to be the total cost of routing traffic below (resp., above) the dividing line at height h through the MAX ISPs (resp., AVG ISPs) using the greedy algorithms given above and define $C'_{Max}(h^+)$ (resp., $C'_{Avg}(h^+)$) to be its right derivative at h i.e., $\lim_{\delta h \to 0^+}(C_{Max}(h + \delta h) - C_{Max}(h))/\delta h$. Let $C(h) = C_{Max}(h) + C_{Avg}(h)$ and thus $C'(h^+) = C'_{Max}(h^+) + C'_{Avg}(h^+)$.

Lemma 4: For all h_1,h_2 if $C'(h_1^+)$ and $C'(h_2^+)$ are well-defined and $x_1 - \sum_{i=1}^{n} P_{Avg_i} \leq h_1 < h_2 < \sum_{j=1}^{m} P_{Max_j}$, then $C'(h_1^+) \leq C'(h_2^+)$.

Proof: $C'_{Max}(h^+)$ is the cost of the cheapest MAX ISP that has not been used to its full capacity when the Max-Threshold is h. Thus $C'_{Max}(h^+)$ is defined wherever $C_{Max}(h)$ is defined, except when h is the sum of the capacities of the MAX ISPs. From Lemma 1, it follows that $C'_{Max}(h^+)$ is a non-decreasing function.

$C'_{Avg}(h^+) = -\sum_{i=1}^{M}$(cost of the most expensive AVG ISP used in the i^{th} interval when the Max-Threshold is h). Thus $C_{Avg}(h)$ is right differentiable wherever it is defined. From Lemma 2, it follows that $C'_{Avg}(h^+)$ is a non-decreasing function. The lemma follows as $C'(h^+) = C'_{Max}(h^+) + C'_{Avg}(h^+)$.□

Theorem 5: The cost of the offline optimal solution can be computed in $O(L(\log m + M \log n) + n \log n + m \log m)$ time, where m is the number of MAX ISPs, n is the number of AVG ISPs, M is the total number of intervals in the billing period and L is the number of bits required to represent the maximum amount of traffic sent in an interval.

Proof: $C(h)$ is a continuous function as both $C_{Max}(h)$ and $C_{Avg}(h)$ are continuous functions. From Lemma 4 and the

[4]Note that the algorithm could produce a solution that uses only the AVG contracts, if that is optimal, by computing the height h to be zero. In fact, that would be the case if the unit cost of the AVG contracts are significantly lower than the unit cost of the MAX contracts.

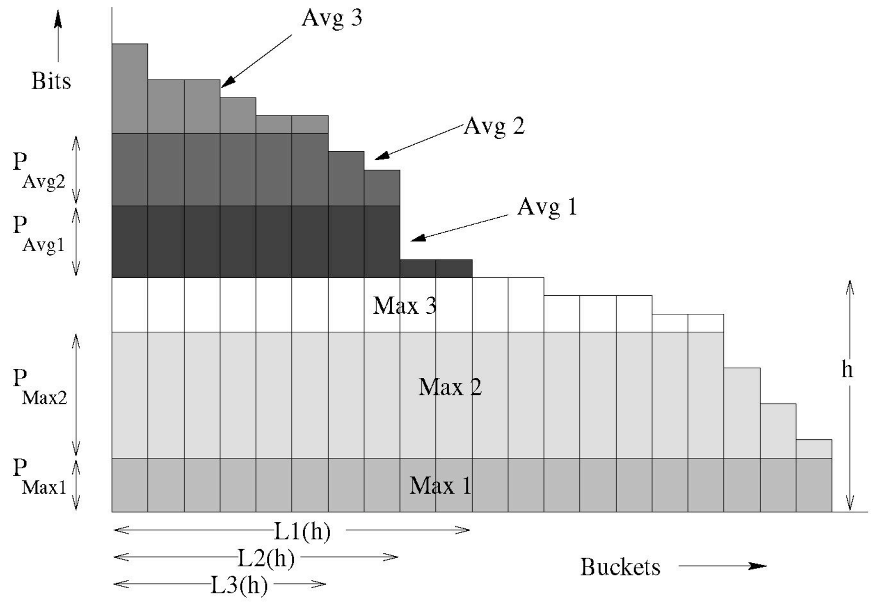

Fig. 1. The structure of an optimal offline solution

fact that $C(h)$ is continuous it follows that $C(h)$ reaches its minimum whenever $C(h^+)$ changes from being non-positive to being positive. We use binary search over all values of h to find the h such that $C(h^+)$ changes sign. There are at most 2^L possible values for h. So there are at most $\log(2^L) = L$ steps. At each step of the binary search we need to calculate $C'(h^+)$ (using $C(h)$) and this can be done in $O(\log m + M \log n)$ time. Thus we can calculate the cost of the optimal routing in $O(L(\log m + M \log n) + n \log n + m \log m)$ time.□

III. Online Algorithms

We provide both deterministic and randomized optimal online algorithms for the problem of routing traffic through AVG and MAX ISPs with minimum cost with competitive ratios of 2 and $\frac{e}{e-1}$ respectively. Note that an online algorithm at time t knows the current and past traffic values, $b_1, b_2, \cdots, b_t$, but does not know future traffic values $b_{t+1}, b_{t+2}, \cdots, b_M$.

A. Optimal Deterministic Online Algorithm

In this section, we present a 2-competitive deterministic online algorithm A that routes traffic through AVG and MAX ISPs. Assume the time-ordered sequence of traffic demands is $I = \langle b_1, b_2, \cdots, b_{M-1}, b_M \rangle$. At a given time interval t, the online algorithm A does the following:

1) Runs the offline algorithm OPT described in Section II on the input $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots, 0 \rangle$. That is, runs the offline on a prefix of the input assuming all future time intervals have zero traffic.
2) Routes the current traffic b_t in the same manner as OPT.

First, we show that the Max-Threshold, i.e., the sum of the thresholds incurred in the MAX contracts, can only increase with time as we progress through the month.

Lemma 6: Let h_t be the Max-Threshold of OPT on input $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots 0 \rangle$. Then, for all $1 \leq t \leq M-1$, $h_t \leq h_{t+1}$.

Proof: Assume $h_t > h_{t+1}$. The cost of routing the traffic $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots 0 \rangle$ with a Max-Threshold of h_t is less than or equal to cost of routing the same traffic with a Max-Threshold of h_{t+1}. As $b_{t+1} - h_t < b_{t+1} - h_{t+1}$ the contribution in the total cost of routing the $t+1^{th}$ interval traffic above the Max-Threshold through the AVG ISPs with a Max-Threshold of h_t is less than or equal to the same with a Max-Threshold of h_{t+1}. Thus with a Max-Threshold of h_t we can route the traffic $\langle b_1, b_2, \cdots, b_t, b_{t+1}, 0, 0, \cdots 0 \rangle$ with the same or lower cost than with a Max-Threshold of h_{t+1}. This contradicts the fact that for no Max-Threshold of $h > h_{t+1}$ can we route the traffic $\langle b_1, b_2, \cdots, b_t, b_{t+1}, 0, 0, \cdots 0 \rangle$ with the same or lower cost. Hence proved by contradiction.□

Theorem 7: The competitive ratio of the deterministic online algorithm A is 2.

Proof: The total cost C_A of algorithm A equals the sum of the cost $C_{A,Avg}$ incurred in the AVG contracts and the cost $C_{A,Max}$ incurred in the MAX contracts. Note that the final threshold h_M of A equals the threshold h_{OPT} computed by the offline optimal algorithm OPT. Also by lemma 6 $h_M \geq h_t$ for all $t \leq M$. Therefore,

$$C_{A,Max} = C_{OPT,Max} \leq C_{OPT} \tag{1}$$

Let $C^t_{A,Avg}$ be the cost incurred in AVG ISPs by algorithm A during the first t time intervals. Let C^t_{OPT} be the total cost incurred by the optimal offline algorithm OPT when provided an input of $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots, 0 \rangle$. We prove by induction on t that $C^t_{A,Avg} \leq C^t_{OPT}$.

Base Case: When $t = 1$, algorithm A runs OPT on the first input and behaves identical to it. Therefore,

$$C^1_{A,Avg} = C^1_{OPT,Avg} \leq C^1_{OPT}$$

Inductive Case: Assume that the hypothesis is true until t. So $C^t_{A,Avg} \leq C^t_{OPT}$. As C^t_{OPT} is the optimal offline solution for the input $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots, 0\rangle$, we have $C^t_{A,Avg} \leq C^t_{OPT} \leq$ the cost of the optimal offline solution with Max-Threshold as h_{t+1} for the same input. The contribution in the cost of $C^{t+1}_{A,Avg}$ and C^{t+1}_{OPT} of sending part of the data in the $t+1^{th}$ interval through the AVG ISPs is the same. This is because in both cases only the data more than h_{t+1} is sent through the AVG ISPs. Adding this cost to the extremities of the inequality given above we get $C^{t+1}_{A,Avg} \leq C^{t+1}_{OPT}$. This completes the induction. Therefore,

$$C_{A,Avg} = C^M_{A,Avg} \leq C^M_{OPT} = C_{OPT} \qquad (2)$$

Thus, combining equations 1 and 2, $C_A = C_{A,Max} + C_{A,Avg} \leq 2C_{OPT}$ □

Theorem 8: The competitive ratio of 2 achieved by Algorithm A is the best possible for any deterministic online algorithm.

Proof: We first prove that the Ski Rental problem[11] is a special case of the traffic routing problem. Given a ski rental problem where the cost of renting a pair of skis is 1 and the cost of buying them is p, the optimal strategy when you ski k times is to buy skis in the beginning if $k \geq p$, and rent otherwise. Given an instance of the ski rental problem we create an instance of the traffic routing problem with one MAX ISP of cost p and one AVG ISP of cost M, where $M \geq k$ is the number of intervals in the billing period. The input traffic $b_t = 1$ bit, if $1 \leq t \leq k$, and zero for $k < t \leq M$. The capacity of each ISP is 1 bit/interval. Thus in an interval one can only send either 0 or 1 bits through an ISP. It is easy to verify that in the original ski rental problem the optimal solution is to buy skis if and only if the optimal solution for traffic routing problem is to use only the MAX ISP for routing the entire traffic. Similarly renting skis is optimal if and only if AVG ISP is used to route the entire traffic in the optimal solution. Also the optimal cost in both problems is the same.

If for any $\epsilon > 0$ if there exists a deterministic online algorithm with competitive ratio of $2 - \epsilon$ we can use it to get a $2 - \epsilon$ competitive deterministic online algorithm for the ski rental problem using the construction given above. This contradicts the fact that ski rental problem has a lower bound on the competitive ratio of a deterministic online algorithm of $1 + \frac{\lceil p \rceil - 1}{p}$ which $\longrightarrow 2$ as $p \longrightarrow \infty$[11], [12]. □

B. Optimal Randomized Online Algorithm

In this section we describe an $e/(e-1)$ competitive randomized online algorithm A_{Rand} which

1) Picks z between 0 and 1 according to the probability density function $p(z) = \frac{e^z}{e-1}$.
2) Routes the traffic using the deterministic online algorithm A_z.

If the time-ordered sequence of traffic demands is $I = \langle b_1, b_2, \cdots, b_M\rangle$ then at a given time interval t, the deterministic online algorithm A_z does the following:

1) Runs the offline algorithm $OPT(z)$ described in Section II on input $\langle b_1, b_2, \cdots, b_t, 0, 0, \cdots, 0\rangle$ but with the costs of all MAX ISPs multiplied by z.
2) Routes the current traffic b_t in same manner as $OPT(z)$.

Note that A_1 is the deterministic online algorithm A given in section III-A. Define $C_{OPT}(z)$ to be the cost of the optimal offline solution with the same input but with the costs of all MAX ISPs multiplied by z. Let $C_{OPT,Avg}(z)$(resp., $C_{OPT,Max}(z)$) be the contribution in $C_{OPT}(z)$ due to the AVG (resp., MAX) ISPs. Similarly define $C_{A_z,Avg}$(resp., $C_{A_z,Max}$) to be the contribution in C_{A_z}, the total cost due to algorithm A_z, due to the AVG (resp., MAX) ISPs. Note that C_{A_z} and $C_{A_z,Max}$ are charged by the actual cost of the MAX ISPs but $C_{OPT}(z)$ and $C_{OPT,Max}(z)$ have a discounting factor of z. The proofs of the following two lemmas are similar to the analogous proofs of equations 1 and 2 in theorem 7.

Lemma 9: $zC_{A_z,Max} = C_{OPT,Max}(z)$

Lemma 10: $C_{A_z,Avg} \leq C_{OPT}(z)$

Lemma 11: For $0 \leq z \leq 1$, $C_{OPT}(1) - C_{OPT}(z) \geq \int_z^1 C_{A_w,Max} dw$

Proof: For any v such that $0 \leq z \leq v \leq 1$,

$$\begin{aligned} C_{OPT}(v) &= C_{OPT,Max}(v) + C_{OPT,Avg}(v) \\ &= vC_{A_v,Max} + C_{OPT,Avg}(v) \\ &\quad \text{(using lemma 9)} \\ d(C_{OPT}(v)) &= dv \cdot C_{A_v,Max} + v \cdot d(C_{A_v,Max}) \\ &\quad + d(C_{OPT,Avg}(v)) \qquad (3) \end{aligned}$$

Define $h(w)$ to be the Max-Threshold in the optimal offline solution ($C_{OPT}(w)$) when the cost of all MAX ISPs are multiplied by w. $h(w)$ is a non-increasing function of w. Also let C_{Max_w} be the original cost of the most expensive MAX ISP that was used in optimal offline solution $C_{OPT}(w)$ (or in C_{A_w}). As the actual cost of any MAX ISP used in the gap between $h(v+dv)$ and $h(v)$ would be at most C_{Max_v}, the increase in cost of MAX ISPs when Max-Threshold is increased from $h(v+dv)$ to $h(v)$ is at most $C_{Max_v} \cdot (h(v) - h(v+dv))$. Thus

$$\begin{aligned} -d(C_{A_v,Max}) &= C_{A_v,Max} - C_{A_{v+dv},Max} \\ &\leq C_{Max_v} \cdot (h(v) - h(v+dv)) \\ &= -C_{Max_v} \cdot d(h(v)) \qquad (4) \end{aligned}$$

The actual cost of any MAX ISP used in the gap between $h(v)$ and $h(v+dv)$ is at least $C_{Max_{v+dv}}$. Thus in the optimal solution when the cost of the MAX ISPs have been multiplied by v decreasing the Max-Threshold from $h(v)$ to $h(v+dv)$ decreases the cost due to the MAX ISPs by at least $vC_{Max_{v+dv}} * (h(v) - h(v+dv))$. The corresponding increase in the cost due to the AVG ISPs is at least this much, since $C_{OPT}(v)$ is the optimal cost. Thus,

$$\begin{aligned} d(C_{OPT,Avg}(v)) &= C_{OPT,Avg}(v+dv) \\ &\quad -C_{OPT,Avg}(v) \\ &\geq vC_{Max_{v+dv}} \\ &\quad \cdot(h(v)-h(v+dv)) \\ &= -vC_{Max_{v+dv}} d(h(v)) \end{aligned} \quad (5)$$

Substituting equations 4,5 in equation 3

$$\begin{aligned} d(C_{OPT}(v)) &\geq dv \cdot C_{A_v,Max} - \\ &\quad v(C_{Max_{v+dv}} - C_{Max_v})d(h(v)) \\ &= dv \cdot C_{A_v,Max} \\ &\quad -v \cdot d(C_{Max_v}) \cdot d(h(v)) \end{aligned}$$

Integrating v from z to 1 and using the fact that $C_{OPT}(v)$ is a continuous function and that the integral of the product of two differentials is 0, we get $C_{OPT}(1) - C_{OPT}(z) \geq \int_z^1 C_{A_v,Max} dv$. □

Corollary 12: $C_{OPT}(1) \geq \int_0^1 C_{A_w,Max} dw$

Theorem 13: The competitive ratio of the randomized online algorithm A_{Rand} is $e/(e-1)$

Proof: Define $P(z) = \int_0^z p(w)dw$. Then

$$\begin{aligned} C_{A_z} &= C_{A_z,Max} + C_{A_z,Avg} \\ &\leq C_{A_z,Max} + C_{OPT}(z) \\ &\quad \text{(by lemma 10)} \\ &= C_{A_z,Max} + C_{OPT}(1) \\ &\quad - \int_z^1 C_{A_w,Max} dw \\ &\quad \text{(by lemma 11)} \end{aligned}$$

$$\begin{aligned} E[C_{A_{Rand}}] &= \int_0^1 C_{A_z} p(z) dz \\ &\leq C_{OPT}(1) + \int_0^1 C_{A_z,Max} p(z) dz \\ &\quad - \int_0^1 p(z)\left(\int_z^1 C_{A_w,Max} dw\right) dz \\ &= C_{OPT} + \int_0^1 C_{A_z,Max} p(z) dz \\ &\quad - \int_0^1 C_{A_w,Max}\left(\int_0^w p(z)dz\right) dw \\ &= C_{OPT} + \\ &\quad \int_0^1 (p(z) - P(z)) C_{A_z,Max} dz \end{aligned}$$

$$\frac{E[C_{A_{Rand}}]}{C_{OPT}} \leq 1 + \frac{\int_0^1 (p(z) - P(z)) C_{A_z,Max} dz}{\int_0^1 C_{A_z,Max} dz} \quad \text{(by corollary 12)}$$

Setting $p(z) = \frac{e^z}{e-1}$ and $P(z) = \frac{e^z-1}{e-1}$ in the above result we prove the theorem. □

Theorem 14: The competitive ratio of $e/(e-1)$ achieved by Algorithm A_{Rand} is the best possible for any randomized online algorithm.

Proof: As in theorem 8 we use the fact that this problem is a generalization of the ski rental problem. The ski rental problem has lower bound on the competitive ratio of a randomized online algorithm of $e'_p/(e'_p - 1)$ where $e'_p = (1 + \frac{1}{p-1})^p$ when p, the ratio of the cost of buying to the cost of selling, is an integer. The algorithm which achieves this is similar to the randomized online algorithm for the snoopy caching problem[12]. Also $e'_p/(e'_p - 1) < e/(e-1)$ but tends to $e/(e-1)$ as p tends to ∞.

If for any $\epsilon > 0$ if there exists a $e/(e-1) - \epsilon$ competitive randomized algorithm for this problem then by the construction in theorem 8 we get a $e/(e-1) - \epsilon$ competitive randomized algorithm for the ski rental problem. A contradiction. □

IV. Extensions

In this section we consider two different extensions to our results. In Section IV-A we consider the notion of Committed Information Rate (CIR) and in Section IV-B we consider 95^{th} percentile contracts.

A. Committed Information Rate (CIR)

Committed Information Rate (CIR) represents the committed amount of billable traffic that must be sent through an ISP. The CIR is paid for in advance, whether or not it is used, at a rate that is usually lower than the rate for the traffic sent above the CIR. Since the cost for the CIR is independent of the usage it can be assumed to be 0 without loss of generality in all further analysis. Let $x_1 \geq x_2 \geq \cdots \geq x_M$ be the average traffic within each of the M 5-minute intervals during the billing period, placed in descending order. For an AVG contract, the bill for the month is $C_{AVG} * (\sum_i x_i/M - CIR_{AVG})$ if $\sum_i x_i/M \geq CIR_{AVG}$, otherwise it is 0 , where C_{AVG} is the unit cost and CIR_{AVG} is the CIR. For a MAX contract, the bill for the month is $C_{MAX} * (x_1 - CIR_{MAX})$ if $x_1 \geq CIR_{MAX}$, otherwise it is 0, where C_{MAX} is the unit cost and CIR_{MAX} is the CIR. Similar cost function exist for a 95^{th} percentile contract.

We derive offline algorithms for routing through ISPs with CIR by first considering routing through MAX ISPs alone and then AVG ISPs alone and finally when we have both as in Figure 2. The main difference between the algorithms for ISPs with CIR and ISPs without CIR is in the greedy algorithm for routing traffic through AVG ISPs alone.

Assume that we are given contracts from m MAX ISPs Max_i, $1 \leq i \leq m$, with unit cost C_{Max_i}, capacity P_{Max_i} and CIR CIR_{Max_i} ($\leq P_{Max_i}$), such that for all $j < i$, $C_{Max_j} \leq C_{Max_i}$. The proofs of the following Lemma is easy, and is omitted.

Lemma 15: There exists an optimal solution in which Max_i is not used more than its CIR in a time interval unless each ISP Max_j, $j < i$, has been used to its full capacity of P_{Max_j} and all MAX ISPs have been used at least to their CIR.

Lemma 15 gives us an optimal greedy algorithm for routing traffic through MAX ISPs alone. First use the CIRs

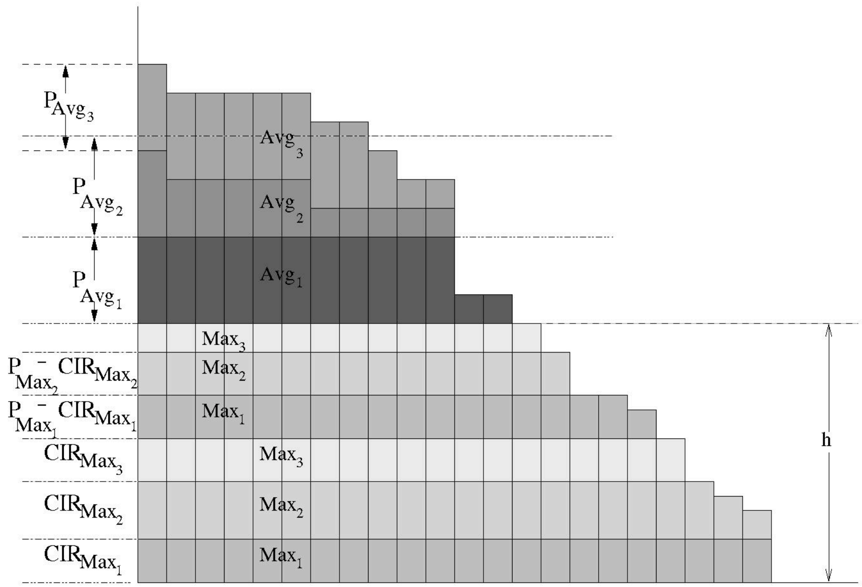

Fig. 2. The structure of an optimal offline solution when the ISPs have a CIR

of all the MAX ISPs and then route the remaining greedily by using cheaper ISPs to their full capacity before using costlier ISPs. This greedy algorithm also takes at most $O(m \log m + M)$ time to calculate the minimum cost of routing through m MAX ISPs.

Similarly there exists a greedy algorithm for routing traffic through n AVG ISPs. Assume that we are given contracts from n AVG ISPs Avg_i with CIR CIR_{Avg_i} and capacity P_{Avg_i}, $1 \leq i \leq n$, such that $C_{Avg_1} \leq C_{Avg_2} \leq \cdots \leq C_{Avg_n}$. Start with the costliest AVG ISP(Avg_n) and first route in each interval all the traffic $(b_k - \sum_{j=1}^{n-1} P_{Avg_j})$ that cannot be routed through cheaper ISPs due to capacity constraints. In case the AVG ISP Avg_n has not been utilized to its CIR then route more traffic to fully utilize the CIR such that the maximum remaining height of all intervals is minimized except for those intervals where this is not possible as the ISP Avg_n has been utilized to its full capacity. The remaining traffic is routed through ISPs $Avg_1, Avg_2, \ldots, Avg_{n-1}$ using the same procedure but now starting with ISP Avg_{n-1}, the next costliest.

Theorem 16: The greedy algorithm given above routes traffic through n AVG ISPs with minimum cost.

Proof: Given any solution we prove that it can be transformed into the greedy algorithm solution without increasing the total cost. Start with the costliest AVG ISP Avg_k used differently from the greedy solution. All the costlier ISPs Avg_i, $i > k$ and the traffic routed through them are not considered as the allocation of these costlier AVG ISPs is not changed.

If the total amount of traffic routed through ISP Avg_k is more in the greedy solution then one can route more traffic through Avg_k in the optimal solution as average traffic is less than the CIR and hence it is free. If the total amount of traffic routed through ISP Avg_k is less in the greedy solution then all of this traffic is being paid for as it is above the CIR. This traffic can be routed through other ISPs Avg_i, $i < k$ which have the same or lower costs.

In case the total amount of traffic routed through ISP Avg_k is the same then one can redistribute the traffic such that the traffic routed in all intervals through the ISP Avg_k is the same without changing the total cost. There exists a interval r (resp., s) in which more (resp., lesser) traffic is routed through ISP Avg_k in the greedy solution. More traffic is routed in interval r through ISPs $Avg_1, Avg_2, \ldots, Avg_{k-1}$ than in interval s as in the greedy solution we are trying to make the traffic routed through them the same. Thus for some x and some $l < k$ in interval r one can route x more units of traffic through ISP Avg_l instead of Avg_k and in interval s one can route x more units of traffic through ISP Avg_k instead of Avg_l. This transformation does not change the total cost. By repeatedly using this transformation one can convert the optimal solution into the greedy solution without changing the cost. □

Once the intervals and the AVG ISPs are sorted, the traffic that has to be sent through the most expensive AVG ISP due to capacity constraints can be calculated in $O(M)$ time. Routing more traffic to fully utilize the CIR of the most expensive AVG ISP can also be done in $O(M)$ time. This is done by keeping track of the heights at which the buckets in which more traffic has to be routed through the most expensive ISP changes. Thus routing traffic through n AVG ISPs with CIR can be done in $O(nM + M \log M + n \log n)$ time.

The algorithm for routing traffic optimally through m MAX ISPs and n AVG ISPs with CIR is similar to the algorithm in section II when the ISPs did not have any CIR. This is true because the proofs of Lemmas 3 and 4 still hold. The only change is that the greedy algorithms given above are used to route traffic below (resp., above) the dividing line of height h through MAX ISPs (resp., AVG ISPs). Similar to the proof of Theorem 5 we can prove that the cost of the optimal offline solution can be calculated in $O(L(\log m + nM) + n \log n + m \log m + M \log M)$ time.

B. 95^{th} Percentile Contracts

Including network contracts that charge based on the 95^{th} percentile of the traffic renders finding the optimal offline solution NP-hard. In fact, the problem of determining whether a given input can be routed using only the free traffic (i.e., using only 5% of the intervals for each ISP) of a set of 95^{th} percentile ISPs is already NP-Hard. The proof involves a straight forward reduction from the Bin Covering Problem [7], which is known to be NP-complete in the strong sense, to this problem, and is omitted.

Theorem 17: Finding whether one can route the entire traffic with zero cost in a system consisting of n 95^{th} percentile ISPs is NP-Complete in the strong sense.

V. Concluding Remarks

An important contribution of this paper is that it opens up the algorithmically rich and practically important area of bandwidth cost optimization for CDNs and multihomed enterprises, using realistic contract models. This paper is but a first step into this area of research, and many open questions for future research remain. An immediate next step is to implement the algorithms in this paper and empirically study their behavior for actual traffic traces and network contracts derived from a CDN. Devising near-optimal online algorithms under the right adversarial model for AVG and MAX contracts with CIR is a problem of great importance for future work. Further, devising a suitable definition of approximation and finding good approximation algorithms for 95^{th} percentile contracts is another interesting avenue for future investigation. Finally, an important and critical avenue for future research is to introduce the notion of performance and extend our model and algorithms to simultaneously optimize both cost and performance. We believe that the current work provides a number of insights into reaching this final objective.

Acknowledgements

The work of Micah Adler is supported in part by the National Science Foundation under the NSF Early Career Development Award CCF-0133664 and an NSF Award CNS-0325726. The work of Ramesh Sitaraman is supported in part by a National Science Foundation award under grant number CNS-0519894.

References

[1] J. Dilley, B. Maggs, J. Parikh, H. Prokop, R. Sitaraman, and B. Weihl. Globally distributed content delivery. IEEE Internet Computing, September 2002, pp. 50–58.

[2] A. Akella, B. Maggs, S. Seshan, A. Shaikh, and R. Sitaraman. A Measurement-Based Analysis of Multihoming. Proceedings of the 2003 ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM), August 2003.

[3] Routescience. http://www.routescience.com.

[4] Internap: Premise-based Route Optimization. http://www.internap.com/products/route-optimization.htm.

[5] Limelight. http://www.limelightnetworks.com.

[6] Akamai. http://www.akamai.com.

[7] A. B. Assmann, D. J. Johnson, D. J. Kleitman, and J. Y.-T. Leung. On a dual version of the one-dimensional bin packing problem. *Journal of Algorithms*, 5(4):502–525, 1984.

[8] D. R. Dooly, S. A. Goldman, and S. D. Scott. TCP dynamic acknowledgment delay: Theory and practice. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pages 389–398, New York, 23–26 1998. ACM Press.

[9] R. Fleischer. On the Bahncard problem. *Theoretical Computer Science*, 268(1):161–174, 2001.

[10] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP- Completeness.* W.H. Freeman, 1979.

[11] A. R. Karlin, C. Kenyon, and D. Randall. Dynamic TCP acknowledgment and other stories about e/(e-1). In *ACM Symposium on Theory of Computing*, pages 502–509, 2001.

[12] A. R. Karlin, M. S. Manasse, L. Rudolph, and D. Sleator. Competitive snoopy caching. *Algorithmica*, 3(1):70–119, 1988.

[13] D. D. Sleator and R. E. Tarjan. Amortized efficiency of list update and paging rules. *Commun. ACM*, 28(2):202–208, 1985.

[14] A. Borodin and R. El-Yaniv. Online Computation and Competitive Analysis. *Cambridge University Press*, 1998.

[15] D.K. Goldenberg, L. Qiu, H. Xie, Y. R. Yang, and Y. Zhang. Optimizing Cost and Performance for Multihoming. *ACM SIGCOMM*, Portland, Oregon, 2004.