Abstract
The structure of the Internet serves as a big "commoditizer" of all traffic. Therefore all data, be it time critical or not is transported at the same speed. However, recent trends in the internet are changing this structure. The practices of multi-homing and using content delivery networks reduce the commodity nature of data being transported and put terminating Internet service providers in a position to price discriminate against specific providers or types of traffic. We firstly formalize multi-homing and content delivery networks, we then derive end user prices for paid content and lastly show consequences of the modeled situation. We thus show how the two technologies to bypass crowded peerings change the Internet business model. Traffic which is sensitive to transport quality, such as business critical or delay sensitive traffic, will be paying higher fees to terminating ISPs.

1. INTRODUCTION

It is a trivial thought that content (a service, music or text) needs to be delivered to its consumer. However, apparently this problem is by and large being ignored in the web service hype. Implicitly in all service visions it is assumed that the Internet is there and available in appropriate quality. There is for example no work that the authors are aware of, which connects web service availability with the use of the underlying network infrastructure. The available literature on web service standards does consider quality of service of web services but only as a question of setting a protocol standard to exchange quality information between two nodes in a network [7, 23]. A first tentative investigation into the matter of availability of web services is provided by [16] and [17]. Even though [16] do not attribute the differences in download speed to connection quality, such an interpretation does not seem far-fetched, especially when considering the geographical disparities in the measured data.

The authors of this work believe that one key factor for the slow uptake of services (SaaS for example) by businesses is a result of the performance and reliability problems [15]; and these are in

1 Institut für Wirtschaftsinformatik, Universität St. Gallen
2 Institut für Technologie und Management, Technische Universität Berlin
large part due to today’s Internet. This work focuses on one specific aspect of quality of service of the Internet and analyses the economic impact of two quality assurance technologies on today’s pricing regime of the Internet.

Internet pricing is a vast field of research that has its roots in telecommunications. This paper departs somewhat from the “classical” literature on communications pricing by considering a special pricing problem present in today’s Internet that has not been covered by the extant literature. We focus on content oriented pricing of network services. In such a scenario Internet service providers (ISPs) with access to end users (EUs) can discriminate against content providers and charge higher prices for termination. This scenario departs from the idealized bill and keep regime that is often used as a basis for analysis of network pricing decisions. However, our scenario is quite realistic. It is commonplace that content providers (CPs) directly buy transit from terminating ISPs, thus effectively paying them for preferential access to end users. This is a viable strategy because by bypassing crowded peerings used by the CP’s original ISP, the CP gets higher quality access to EUs. This is due to the fact that peering points are a major bottleneck on the Internet. While there are usually few capacity problems present inside a single provider’s network, peerings are often at their capacity limit [19], [4]. For the purpose of this paper we call this practice multi-homing (MH). Content delivery networks (CDNs) are also a popular way to enhance the flow of information on the web. A CDN uses local caches to keep distributed images of content close to EUs thus bypassing peerings. The CDN effectively buys transit from the ISP that terminates the traffic.

The paper is structured as follows: First we explain the relevant entities of the current Internet that we need for a formal model. Then we present a formalized treatment of four scenarios that show how CDN and MH affect ISPs incentives to price traffic. Lastly we discuss consequences of our model and sketch out an agenda for further research.

2. STATUS QUO: THE MARKET FOR INTERNET CONNECTIVITY

2.1. Overview

Figures 1 and 2 show the structure of the Internet as commonly used to model the Internet [18], [12] and [21]. In figure 2 at the bottom, customers and content providers receive and send traffic, respectively. ISPs interconnect to exchange traffic to and from their customers (EUs or CPs). In figure 2 traffic could for example originate at the content provider and be passed on to its ISP 3. From there it is routed to ISP 1 to be terminated at end user (EU) one. The ISPs are fully meshed, each one interconnecting with all the others. It is a common approximation [12] that CPs (web sites) only send traffic and EUs only receive traffic.

2.2. Internet Service Providers

ISPs provide connectivity to Internet users who are either end users or content providers. The ISPs pay the market price $p_w$ to the ISPs for termination of traffic. Typically ISPs have no lack of bandwidth on their backbones and can provide quality assurance to traffic either through excess capacity or network management techniques commonly employed only within a single provider’s network [22]. Bottlenecks (Not meant in the sense of bottleneck facility as in [1, 2] may be present in the peering points and in the access network. We ignore possible problems due to constrained access bandwidth and concentrate on the peerings. All of the following ignores dynamic considerations of integration and foreclosure through the ISP.
2.3. Points of Interconnection

In figure 2 the circle with arrows in upper middle represents the point of interconnection where the three ISPs interconnect their networks. This is also called a “peering point”. There are two dominant modes of interconnection: Peering and transit. Peering is a settlement free agreement to exchange traffic while transit involves payment for exchanged data typically on a bit-rate per time scale. Typically peering agreements are used between ISPs of similar size while transit is paid from small ISPs to larger ISPs.

Peering points are among the major bottlenecks of the Internet because it always takes both parties to agree to extend the capacity of a peering point. Since the telecommunications firms who own the Internet infrastructure tend to be not overly cooperative these peering points represent excellent opportunities for power play or generally uncooperative behavior. Peerings are frequently overloaded at peak times because one of the two parties involved has no interest in paying for extending its capacity [1] [5], [3], [8]. Ways for CPs to circumvent frequently overloaded (and therefore slow and packet dropping) peerings are multi-homing and the use of CDN services.

Transit in contrast to peering involves a payment from one ISP to the other for the delivery of traffic. With such an agreement a guaranteed bandwidth is bought. Due to strategic considerations the biggest networks only peer among themselves and charge smaller networks for sending traffic to them. Since small ISPs have to pay for sending traffic to larger networks which is necessary to reach the whole Internet they optimize their out-payments for transit fees by buying the least amount of bandwidth their users will tolerate.

2.4. Content Providers

Content Providers are for example websites that buy connectivity to the Internet from an ISP (possibly mediated through a hosting provider). Content providers are able to multi-home. They can buy connectivity for one fraction of their traffic from IPS1 and the rest from ISP 2. Furthermore they can host their service with any provider anywhere in the world giving them a very large set of ISPs to choose from. This creates competition amongst ISPs for CPs’ business and therefore CPs face a market price for Internet connectivity based on perfect competition. This price only includes un-prioritized traffic transportation across peering points. Canonical analysis [10-12, 14] usually assumes the following model of Internet payments: \[ EU \rightarrow ISP, ISP \leftrightarrow ISP, ISP \leftrightarrow CP \] (t=terminating, o=originating), ignoring where the CP gets funding from and emphasizing the analysis of the inter ISP settlement. Focusing on payments for content viewing, we model the payments flows according to the following scheme: \[ ISP, ISP \leftrightarrow ISP, ISP \leftrightarrow CP \leftrightarrow EU \]. We ignore payments from the EU to the terminating ISP for access to the Internet. Payments from the EU to the CP might be paid through watching advertisements or direct money exchange. The arrow with the a stands for the access charge that is paid from one ISP to the other for sending traffic to the receiving ISP[13]. p is the final price paid by the EU for viewing some content. \( p_a \) is the price paid from the CP to the ISP for reaching the EU. If the ISP receiving \( p_w \) cannot terminate the traffic it has to pay an access charge to another ISP able to terminate the traffic. If a CDN is involved in content delivery, the CP has to pay the cost of the CDN, too. Modifying the way CPs get access to the EUs and modifying payments accordingly will be the key part of this paper. The two variations we will consider are: multi-homing and CDN delivery. With MH, the terminating ISP is directly connected with the CP, while with CDNs a neutral third party mediates between CP and ISP. Under MH payment flows are \[ ISP \leftrightarrow ISP, ISP \leftrightarrow CP \leftrightarrow ISP \] and the originating ISP is cut out of the equation. With CDN delivery, the payments are: \[ ISP \leftrightarrow ISP, CDN \leftrightarrow ISP, CP \leftrightarrow CDN \].
2.5. End Users

Unlike CPs, EUs cannot divide their traffic amongst several ISPs and are immobile in the sense that they cannot chose their provider globally but need to chose among a small number of local ISPs. Therefore we consider a static scenario in which EUs are bound to their ISP, providing the ISP with a monopoly over terminating traffic to those EUs. Neither the competition among local ISPs for market share nor the dynamic considerations of consumers with switching costs [9] are considered here.

2.6. Multi-homing

Multi-homing is the common practice that CPs interconnect with several different ISPs. There are several instances of big content providers that are present in public peering points in order to gain direct access to ISPs with many EUs. Google for example is not just with one ISP in the USA but has its own transatlantic fiber capacity and is present in DE-CIX [6]. Besides the reliability aspect this practice allows the CPs to get direct access to an ISP’s network and thus be more in control over bandwidth agreements with that ISP.

Figure 1: Simplified model of the Internet. Figure 2: A tiered view of the Internet.

Figure 3: The Internet with MH. Figure 4: The Internet with CDNs.

2.7. Content Delivery Networks

CDNs consist of a network of servers that are distributed around the Internet within many ISPs’ infrastructures. A CDN takes content from a CP and caches it on those distributed servers, which has the effect that content is brought closer to the EU without passing through peerings. The CDN then delivers the content it is caching from the mirror site to the EU. By using the services of a CDN a CP does not need to multi-home with every possible network. We assume that CDNs do not make any profit and charge efficient prices.
3. MODELING THE EFFECTS OF MULTI-HOMING AND CDN

A key feature of the Internet is that all traffic is being treated (more or less) equally. Methods for discriminating against specific types of traffic are expensive and are not used aggressively. While this makes the Internet very egalitarian, it is a problem for content which needs assured performance parameters. Video streaming for example can only tolerate a certain package loss and Internet gaming requires assured low delays. Since the Internet cannot assure constant quality levels, methods to bypass the main bottlenecks are used by commercial CPs: CDN and MH. Both technologies “de-commoditize” Internet traffic because the source of traffic becomes a business partner.

In the following we will analyze the four combinations of market form faced by the CP and strategic behavior of the ISPs. The columns of table 1 represent the possible actions the terminating ISP can take: With MH it can perfectly discriminate against the known source of traffic and extract profits; with CDNs it can discriminate against all traffic that requires higher quality and has demonstrated this through the use of a CDN (self selection). The rows represent the market situation the CP is facing. Under competition many CPs are serving a market; with a monopoly, only one CP serves a specific market with one product. Cross product/market elasticities are assumed to be zero. The cells of the table show what the ISP will do to extract the maximum revenue from the CP.

<table>
<thead>
<tr>
<th></th>
<th>Perfect / 1st degree price discrimination (MH)</th>
<th>Class of Service / 2nd degree price discrimination (CDN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Competition</td>
<td>1: ISP sets access price until monopoly price for content is reached</td>
<td>2: ISP sets monopoly price for termination, inefficient since ISP sets avg. price</td>
</tr>
<tr>
<td>Content Monopoly</td>
<td>3: ISP extracts all revenue from CP through two part tariff</td>
<td>4: ISP sets monopoly access price but cannot capture all rents from CPs (dbl. marginalization)</td>
</tr>
</tbody>
</table>

Table 1: The four possible situations with an access monopolist

3.1. Scenario 0: No Discrimination by ISP Possible

The situation when the ISP has no way of differentiating different traffic, and therefore cannot set different prices for different traffic sources and different quality classes, is not shown in table 1. It corresponds to the “normal” way traffic is exchanged on the Internet. The price for terminating traffic $p_w$ is equal to the ISP’s cost $c$. The charge $p_w$ which is levied by the terminating ISP feeds through to the originating CP as a rise in connection costs to be included in the charge to the EU. The CP therefore can consider the market price as given. This treatment ignores the considerations put forward by [12].

3.2. Scenario 1: Perfect Discrimination by ISP and Content Competition

Here the terminating ISP can perfectly discriminate against CPs while the CPs face a competitive market. Assuming perfect competition only efficient firms are active in each differentiated market and the ISP can perfectly discriminate against each market segment (i.e. treat a whole segment like one single CP). By setting its price $p_{w, i}$ charged to the CPs of segment $i$ equal to the monopoly price the ISP can extract monopoly profits while leaving no profits to the CPs. Each CP sets the competitive price $p_i$ in segment $i$ equal to its marginal costs. For simplicity we assume that
marginal costs of production of the CPs are zero. Therefore $p_i = p_{w,i}$. The ISP with marginal costs $c_i$ per market segment determines $p_{w,i}$ by solving the following problem:

$$\max_{p_{w,i}} \left[ (p_{w,j} - c_i)D_i(p_{w,j}) \right].$$  

For example suppose that the demand function in one of the market segments $i$ is given by $D_i(p_{w,i}) = 1 - p_{w,i}$. Then the target function of the ISP is:

$$\max_{p_{w,i}} \left[ (p_{w,j} - c_i)(1 - p_{w,j}) \right].$$  

Solving this gives for price and output quantity yields

$$p_{w,j} = \frac{1 + c_i}{2},$$  

$$D_i(p_{w,i}) = q_i = \frac{1 - c_i}{2} \text{ and}$$  

$$\Pi = p_{w,j}q_i - c_iq_i = \left(\frac{1 - c_i}{2}\right)^2.$$  

**Interpretation:** The ISP converts each competitive market-segment $i$ into perfect monopoly and extracts the maximum profit. The output quantity is reduced and the CPs make no profits. Note that this result would not hold with users being able to switch between ISPs. Switching users without switching costs would restore competitiveness of the market. Due to long term contracts, however, the ISP has some market power.

### 3.3. Scenario 2: Market Segmentation by ISP and Content Competition

This situation is similar to scenario one with a reduction in the ability to discriminate against CPs. With a CDN as mediator only CDN (i.e., quality sensitive) and non-CDN traffic can be distinguished. CDNs buy transit from the ISP who can thus differentiate between quality sensitive traffic from CDNs and ordinary traffic through normal peerings with other ISPs. The price taken by the competitive CPs depends on the price $p_{w}$ set by the ISP which acts like a monopolist. The ISP faces an aggregate demand for quality traffic. Assuming that the CDN segment is competitive, the ISP can set monopoly prices for CDN traffic which raises marginal costs of the CDN providers and thus costs for CPs. In the following treatment we assume the CDN costs to be fully absorbed into the marginal cost of the CP. For simplicity both are assumed to be constant and zero. Again the CPs are competitive in each market segment. In each of $n$ market segment the CPs have the target function

$$\max_i [ (p_i - p_{w})D_i(p_i) ]$$

where $p_i$ is the price charged to the EUs and $p_{w}$ is the uniform price for CDN traffic paid to the ISP. Different from scenario one, now the ISP cannot set different $p_{w,j}$ per market segment but just one averaged price for CDN traffic. Under perfect competition price equals marginal costs:

$$p_i = p_{w}.$$  

**Interpretation:** It is obvious form equation (6) that the ISP influences the pricing decisions of the CP’s through its price $p_w$. However, since the ISP can only set a a uniform price for all CDN traffic, it can only optimize across an aggregated $n$ demand function $D = \sum D_i$. Therefore prices in the EU market need not be those an integrated monopolist would have chosen. In this situation inefficiencies are present since the ISP charges an average price. Some segments profit from this by paying a price below the monopoly level, thus improving efficiency, while other segments pay more than the optimal monopoly price which might result in some markets not being served.
3.4. Scenario 3: Perfect Discrimination by ISP and Content Monopoly

The ISP can target the CPs individually and set individually profit maximizing prices. This situation seems prone to double marginalization with two cascaded monopolies. However, the ISP can anticipate this sort of inefficiency and avoid it by setting a usage price $p_w$ and then extracting all profits of the monopoly CPs through a fixed fee $A$. This way of pricing is known as franchising [20]. The total price set by the ISP is

$$T(q) = A + p_w q.$$  

(8)

Assuming that the ISP sets $p_w = c$ efficiently at the marginal cost of providing the service (the variable $q$ is the output quantity of the ISP), the CP chooses the optimal end user price $p$:

$$\max_p [(p-c)D(p) - A].$$  

(9)

This results in the monopoly EU price

$$p = \frac{1+c}{2}.$$  

(10)

By setting the fixed fee $A$ equal to the profit of the CP the ISP can now extract all profit without distorting the target function of the CP. The same result could have been achieved with a “tax” $\tau \Pi$ levied by the ISP.

*Interpretation:* The profit based charge $A$ does not alter the price paid by consumers but shifts profits from the CP to the ISP. This can be a problem since monopoly profits could be the reward for innovation and if those profits are taken away, innovation might not be profitable any more. This situation is good for static efficiency since the output decision of the CP is not changed by the ISP’s behavior. However, when considering the development over time, no matter how much of the CPs’ revenue is extracted by the ISP, the EU price for the content stays the same and thus there is no competitive pressure from EUs on the ISP.

3.5. Scenario 4: Market Segmentation by ISP and Content Monopoly

Each of $n$ CPs serves a monopoly market as in case three but the ISP cannot differentiate between those markets and can only optimize its revenue based on gross demand (as in case two) by all fully differentiated CPs. This corresponds to a situation know as double marginalization. Now the CPs (with constant marginal cost set to zero) set their price for content as a standard monopolist. Their target functions are:

$$\max_i [(p_i - p_w)D_i(p_i)]$$  

(11)

where $p_i$ is the price charged to the EU and $p_w$ is the price paid to the ISP.

Since the ISP can only 2nd degree price discriminate we will analyze the ISP’s problem based on average figures for price and demand from all CPs demanding quality (using CDN):

$$p = \frac{1}{n} \sum_{i=1}^{n} p_i, \text{ and}$$  

(12)

$$D = \frac{1}{n} \sum_{i=1}^{n} D_i(p_i).$$  

(13)

Assuming for simplicity that aggregate demand is linear: $D(p) = 1 - p$, the solution to this problem is similar to the treatment in section 3.2 with the $c_i$ replaced here by $p_w$ and yields optimal prices of $p = \frac{1+p_w}{2}$ and an optimal output $q = \frac{1-p_w}{2}$. Using this quantity, the optimization problem of the ISP with marginal output cost $c$ for CDN traffic is:

$$\max_{p_w} \left[(p_w-c)\left(\frac{1-p_w}{2}\right)\right].$$  

(14)
resulting in $p_w = \frac{1+c}{2}$ and $p = \frac{3+c}{4}$ for the final consumer price. Comparing the sum of the profits of ISP $\Pi_{isp}$ and CP $\Pi_{cp}$ we see that it is smaller than the profit $\Pi_{int}$ of an integrated monopoly provider:

$$\Pi_{isp} + \Pi_{cp} < \Pi_{int} \Leftrightarrow \frac{(1-c)^2}{8} + \frac{(1-c)^2}{16} < \frac{(1-c)^2}{4}.$$  

(15)

The end user price $p = \frac{1+c}{2}$ in monopoly is lower than the $p = \frac{3+c}{4}$ above (as long as $c < 1$ which makes sense because with $D = 1 - p$, $1 > p \geq p_w \geq c$ must be true if no losses are made and there is a positive demand).

**Interpretation:** On average across EU’s demand for the perfectly differentiated markets this situation is suboptimal since prices could be lower and revenues could be higher. In addition to this double marginalization problem, there exists a problem due to the averaging of price and demand of the different CPs. Since the price set by the ISP is targeted at the average CP, it will typically be either too high or too low. Thus there is a second source of inefficiency. ISPs could also opt for a different pricing model and charge (as in scenario 3) a fixed fee plus an efficiently set usage fee $p_w$. Thus inefficiencies due to the ISPs behavior would be removed. However, since the ISP can only set an average $A$ it is impossible to extract all rents. This would even make some market segments with low profits unattractive to serve since profits are too low to cover the fixed fee. Thus one has to chose between double marginalization reducing output in some segments and franchising resulting in some markets not being served at all.

### 3.6. Synopsis of the Four Scenarios

In all four scenarios welfare is below optimal and prices are above the competitive level. It is even possible as shown in case four that welfare is lower than in monopoly. The first general result therefore is that CDNs and multi-homing reduce welfare due to reducing the efficiency of price setting for data transport. This result is true if one ignores the welfare gains of being able to deliver higher QoS by the use of those technologies.

The second result which is common to all cases is that ISPs’ price setting reduces CPs’ profits. While this in itself does not need have a negative effect on welfare in a static environment it can be detrimental when considering monopoly revenues of CPs as the reward for innovation. If ISPs extract these profits, innovation might become unprofitable for CPs. Furthermore ISPs are able to exploit their access monopoly and create monopolies from otherwise competitive markets. The only precondition for this is a quality requirement of the service that does not allow the use of ordinary peerings.

### 4. CONCLUSIONS AND FURTHER RESEARCH

This work provides a new view on access pricing and quality of service on the Internet. Assuming that CDN and multi-homing are used to improve the quality of service provided to the EU we have shown how a “de-commoditization” of traffic enables the terminating ISP to charge more for termination. Our analysis shows that there exist incentives for ISPs to further degrade peering quality to attract more traffic to the more profitable segments. On the positive side, understanding multi-homing and CDNs as quality mechanisms opens up a whole new view on the quality of service debate. Standard approaches of QoS always require global carrier collaboration. All carriers have to agree on service classes and forward each other’s traffic with the appropriate service level. With multi-homing and CDN, an edge based solution to QoS is available that can deliver QoS for many applications on the Internet.
The presented work leaves open and poses many further research question that need to be addressed. On the technical side, it would be interesting to know whether CDN and MH can fully replace inter carrier agreements on quality parameters of traffic. Which quality mechanisms are necessary inside one carrier’s network to complement the peering bypass capability of CDN and MH with the ability to deliver to the EUs workstation? On the service oriented side, more work needs to be done to better understand the (non-) adoption of services and its connection with (un-) reliability and issues of the Internet. In an experimental set-up with one service hosted by CDN or MH and one service hosted within another network one could empirically validate the influence of peerings on user satisfaction.

There are many questions that we have not addressed. However, the text presents a new perspective on the QoS debate and on economic aspects of CDN and MH. Services hosted on the Internet are a case in point for the quality sensitivity of consumers and the lack of quality in the internet. We believe this paper will add an important building block to our understanding of QoS on the Internet and spark ideas to QoS enable the web.
REFERENCES


