How Do Content Delivery Networks Affect the Economy of the Internet and the Network Neutrality Debate?

Patrick Maillé, Bruno Tuffin

How Do Content Delivery Networks Affect the Economy of the Internet and the Network Neutrality Debate?

Patrick Maillé[1] and Bruno Tuffin[2]

[1] Institut Mines-Télécom / Télécom Bretagne, 2, rue de la Châtaigneraie, 35576 Cesson Sévigné Cedex, France, patrick.maille@telecom-bretagne.eu
[2] Inria Rennes Bretagne Atlantique, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France, Bruno.Tuffin@inria.fr

Abstract This paper investigates the economic impact and strategies of Content Delivery Networks (CDNs), Internet actors that reduce the capacity needs in the backbone network and improve the quality perceived by users. We consider so-called push and pull models where the traffic is paid by the sender or the receiver, respectively, as well as the situation where the CDN is (vertically) integrated to, i.e., owned by, an Internet Service Provider (ISP). We then discuss the implication of CDNs into the network neutrality debate, another issue forgotten by researchers and regulators.

1 Introduction

Content delivery networks (CDNs) are large distributed systems of servers deployed within the Internet, aimed at putting data closer to end users to offer better quality of service (QoS) and availability [3]. The role of CDNs is to produce the highest performance for content accessed by users, some typically demanding applications being video delivery or gaming. They are therefore of interest for content/service providers, but also for network providers since reducing the load (congestion) on the network: the most popular pieces of content, being stored at the edge of the network, do not need to be downloaded from their source upon each request.

But the telecommunications economics literature barely addresses the role of CDNs, although they have become major actors of the Internet. While there has been a huge amount of work on CDNs from a technical point of view, the literature on the economics of CDNs is more limited. Existing works focus on the optimal pricing strategies of CDNs [8], or discuss if and how different CDNs within a network should cooperate when setting their caching strategies [1,4,5,7,9]. But the interactions between CDNs, ISPs, and CPs, have received little attention, while they have a major impact all the different actors.

This paper is a first step in the direction of modeling and analyzing the relations between CDNs and other Internet actors:

- We extend a previous model of ISP competition by including CDNs into the

picture, to study how they interact and what their best strategies are.
• We consider both the so-called *pull model* – where the ISP requesting content has to pay the ISP hosting it – and the *push model* – where the transfer fee is paid by the ISP hosting content to the one hosting end users; in addition the CDN can be an independent entity or be owned and managed by an ISP. For each case, we illustrate the impact of CDNs on ISPs' revenues and on fairness between CPs.
• We then discuss the impact of CDNs on the *network neutrality debate* [10,11,12], in which CDNs have also barely been addressed although they reduce traffic transit, thereby addressing in part the concerns that ISPs express in the debate.

2 Model

Extending the model in [6], we consider two ISPs (A and B), in competition for end users. We assume a continuum of content, whose total mass normalized to 1 represents the total volume downloaded per time unit (i.e., the mass of each piece of content is proportional to its popularity). Let x (resp., $1-x$) be the proportion of content that is *directly* connected to the Internet through A (resp., B).

At each ISP, we can have a CDN, indexed by the name of its host ISP, caching some of the content rooted at the other ISP to improve QoS and avoid potential transfer fees between ISPs. Each CDN can be an independent actor (ex: Akamai), or a service implemented by the hosting ISP (the CDN is then said *vertically integrated*).

Figure 1 represents the actors and their economic relations. Users are treated too as a continuum (individuals having a negligible impact) of total mass 1. The

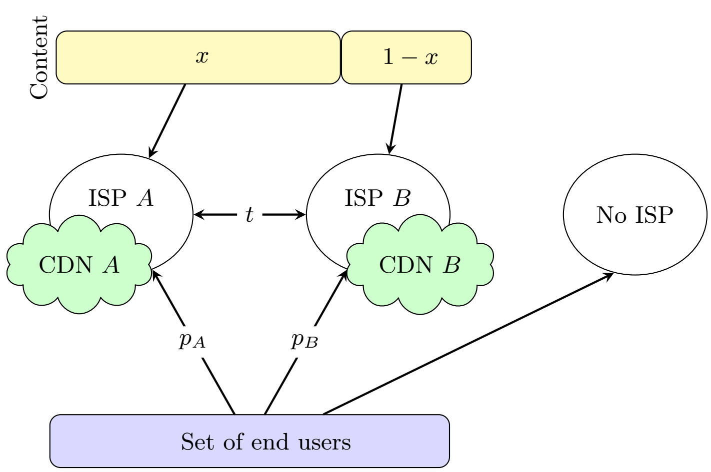


Figure 1. Relations between users, ISPs, and CPs

price p_A (resp. p_B) is the *flat-rate* subscription fee of ISP A (resp. B), and t is a unit price for the traffic transferred between ISPs: $t = 0$ corresponds to the classical peering agreement with no fee, while $t \neq 0$ means a paid transit.

Finally, only a portion of the downloaded content is cached by the CDNs, we denote those quantities by $y_A \leq 1 - x$ at CDN A and $y_B \leq x$ at CDN B.

Users' choice of ISP (if any) is through a standard *attraction model* [2]: the proportion of users choosing Option $i \in \{A, B, 0\}$ –0 meaning no subscription– is

$$\sigma_i(p_A, p_B) \stackrel{\text{def}}{=} \frac{(x_i/p_i)^\alpha}{(x_A/p_A)^\alpha + (x_B/p_B)^\alpha + 1/p_0^\alpha}, \tag{1}$$

where $\alpha > 0$ is a sensitivity parameter, p_0 represents the cost of not benefitting from any content, and x_i is the average perceived quality with ISP $i \in \{A, B\}$. Content is reachable with high quality $\gamma > 1$ (from the local ISP or its host CDN) or low quality 1 (from the distant ISP), yielding

$$x_A = 1 + (x + y_A)(\gamma - 1); \quad x_B = 1 + (1 - x + y_B)(\gamma - 1).$$

Each ISP $i \in \{A, B\}$ focuses on net benefits, stemming from:
1) end-users subscriptions, the corresponding incomes being proportional to the market share σ_i and the price p_i;
2) the potential gains/costs from hosting a CDN:

- if the CDN is an independent actor (Akamai for instance), we assume the CDN pays a per-unit-of-volume fee r to the ISP for delivering content from the CDN cache servers to the ISP users.
- If the CDN role is managed by the ISP itself (something increasingly happening), r is the per-unit-of-volume revenue the ISP gets from remote CPs for serving as a CDN. We also add a storing cost c_i depending on y_i, that can be considered convex since the most popular contents will be cached in priority.

3) And the potential transit fees between ISPs: content associated to ISP A (resp. B) and transferred to B (resp. A) being proportional to $x - y_B$ (resp. $1 - x - y_A$) and to ISP B (resp. A) market share, the net (possibly negative) amount $\Delta_{A,B}$ of traffic from A to B is

$$\Delta_{A,B} = (x - y_B)\sigma_B(p_A, p_B) - (1 - x - y_A)\sigma_A(p_A, p_B).$$

- In the *Pull model*, the transfer fee is paid by the ISP hosting end users to the one hosting content, i.e., the service is asked by end users. Let t be the per-unit-of-volume price. This leads to a revenue $t\Delta_{A,B}$ for A, from B (and an equal cost for B).
- *Push model*: the transfer fee is paid by the ISP hosting content to the one hosting end users, i.e., the service is asked by content providers. The same type of expression can be used with $t < 0$.

Regrouping, the ISP utilities per time unit are:

$$
\begin{aligned}
U_A(p_A,p_B) &= \sigma_A(p_A,p_B)p_A + t\Delta_{A,B} + ry_A - c_A(y_A), \\
U_B(p_A,p_B) &= \sigma_B(p_A,p_B)p_B - t\Delta_{A,B} + ry_B - c_B(y_B).
\end{aligned}
\tag{2}
$$

We neglect the transfer cost of files from (remote) content providers to CDNs, since that transfer is done only once.

3 How much content to cache?

We focus here on the amount of cached content that benefits the most to the ISP hosting the CDN.

3.1 Pull model and independent CDN

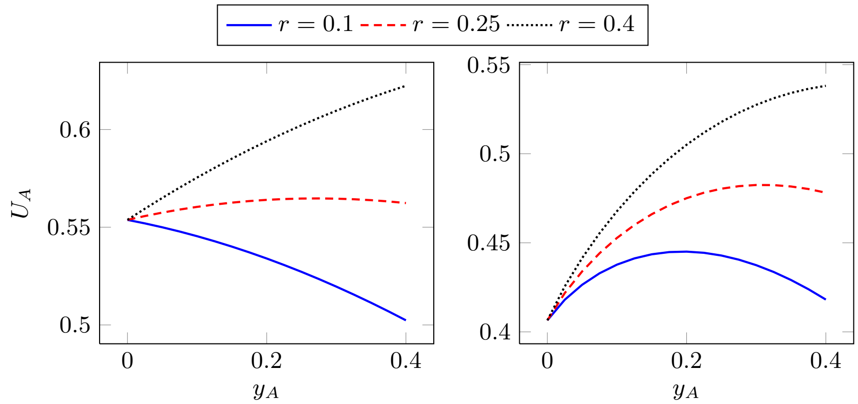
From (2), if ISP i hosts a CDN increasing its cached content: i) ISP i attracts more users because providing a better QoS (i.e., σ_i increases), ii) ISP i gains more from the CDN (ry_i increases with y_i), iii) ISP i pays less for transit because "pulling" less traffic. U_i thus increases with y_i, ISPs should then let CDNs cache as much content as possible.

3.2 Push model and independent CDN

Now the ISP still gains from CDN payments and increased attractiveness to users, but loses from transfer fees since its competitor will have to "push" less traffic. Figure 2(a) shows the values of U_A when y_A ($\in [0, 1-x]$) varies, for three values of r and with $\alpha = 1.5$, $x = 0.6$, $t = -1$, $y_B = 0.3$, $\gamma = 1.5$, $p_0 = 1$, $p_A = 1.3$, $p_B = 1.7$. Note that $c_A = c_B = 0$, since the storage cost is borne by the CDN. Depending on the parameters, it can be optimal to cache nothing, the whole content, or an intermediate amount, but we expect the optimal cached amount to increase with r, because the ISP's revenue from the CDN increases.

3.3 Pull model and integrated CDN

We take here the same values as for Figure 2(a), except that now $t = 1$ (pull model) and that the ISP bears storage costs, of the form $c_A(y_A) = (y_A)^{3/2}$. Because of that cost, we end up too with a trade-off. Figure 2(b) displays U_A in terms of y_A ($\in [0, 1-x]$). Note here that because $c'_A(0) = 0$, the derivative of the utility is positive at 0 (the other effects improve revenue), hence there are always incentives to cache some content. Again, when r is large enough, it becomes beneficial to cache all content.



(a) Push/independent (b) Pull/integrated

Figure 2. ISP A revenue versus cached amount y_A

3.4 Push model and integrated CDN

With the same set of values as for Figure 2(a), but with the storage costs $c_A(y_A) = (y_A)^{3/2}$, we studied the revenue U_A when y_A varies. The curves have a shape similar to those of Figure 2(a), we omit them due to space constraints. Even in the push model, an ISP can have an interest in using an integrated CDN if r is large enough.

3.5 Caching and its impact on revenues

Figure 3 *(top)* plots the optimal cached amount y_A and the corresponding utility U_A in terms of x, for each of the four situations, adding the curve of U_A when no CDN exists. Note it is not relevant to compare the numerical values between integrated and independent: r is taken the same (a cost c_A is added to the integrated case), but it represents a price paid by the CDN in the independent case and a price paid by CPs in the integrated case. Remark in Figure 3 *(top)* that for the push/integrated case, y_A is 0 or close to 0. In the push/independent case, caching a non-negligible proportion is better for small x but the optimal cached part quickly goes to 0 as x increases. As seen before, all B content (i.e., $y_A = 1-x$) is cached in the pull/independent case. For the pull/integrated model the optimal y_A increases with x when $y_a < 1-x$. Note that U_A increases with x in the pull cases, but the trend is to decrease in the push ones, suggesting there is no interest for ISPs to attract CPs for the latter. Moreover the no-CDN curve is very close to the integrated one in the push case, because y_A^{opt} is close to 0. Caching does not yield a large revenue improvement in the pull/integrated case either (while it does for the independent case), but the gain can be significant with different parameters. Figure 3 *(bottom)* displays the same metrics, but when the price r varies. For small values of r, in both push cases there is no incentive

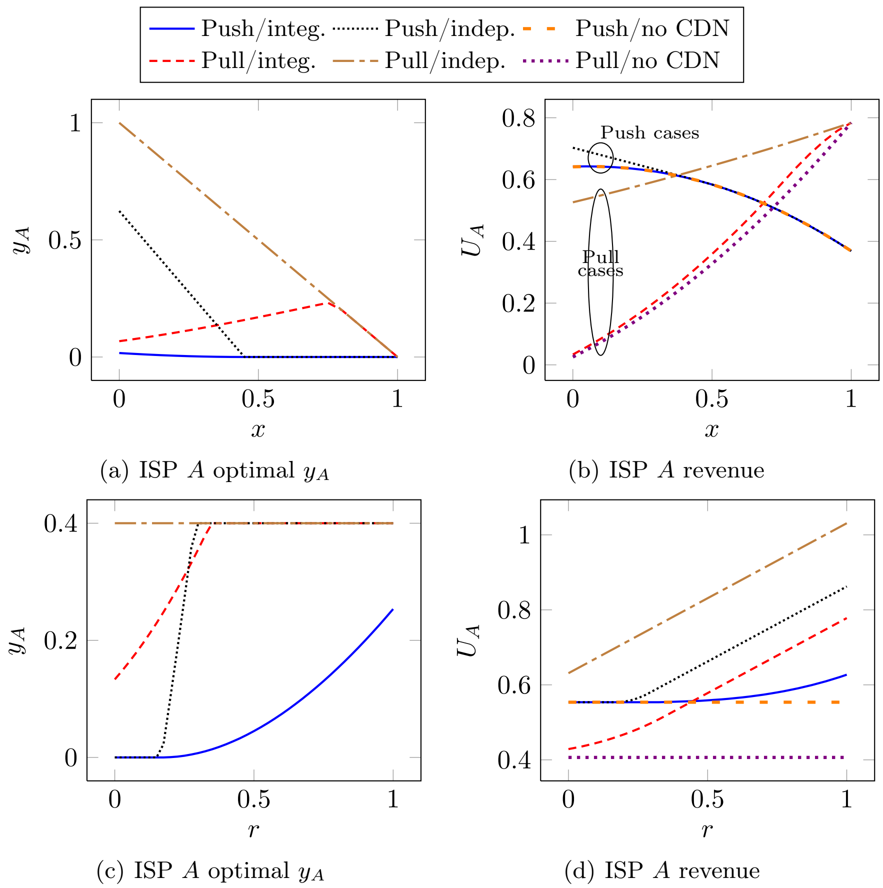


Figure 3. ISP A optimal y_A and utility U_A

to cache content, because gains are not sufficient, but above a threshold caching becomes beneficial. The revenue logically increases in r too (in cases with CDN).

4 Fairness concerns

We now focus on the impact of CDNs on CPs: Is there a difference of treatment among CPs according to the ISP they are attached to? To answer, we use the (aggregated) perceived quality Q_i of content associated to ISP i, that is γ for the customers of ISP i and for those of $j \neq i$ if the content is cached in j's CDN, and 1 otherwise. This gives:

$$Q_A = \gamma\sigma_A + \gamma\sigma_B y_B/x + \sigma_B(1 - y_B/x)$$
$$Q_B = \gamma\sigma_B + \gamma\sigma_A y_A/(1 - x) + \sigma_A(1 - y_A/(1 - x)).$$

Note that for any set of parameters, quality does not depend on the type of model (push/pull, integrated or not). Taking the values (when fixed) of Section 3.2 (adding $y_A = 0.2$), we display Q_A and Q_B in Figure 4 when x and y_A vary.

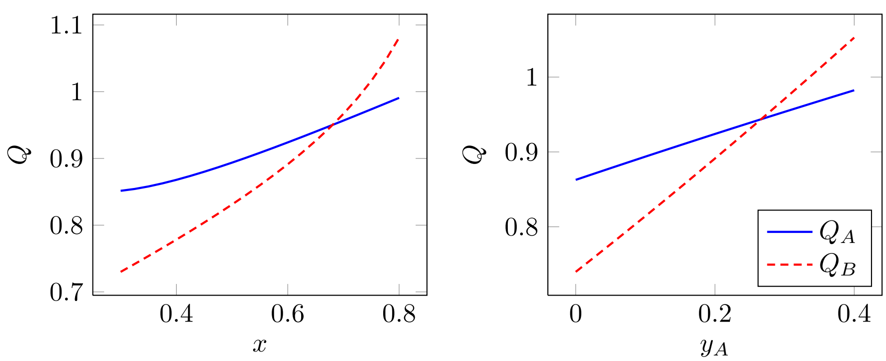


Figure 4. Content quality depending on host ISP

We observe a difference of treatment depending on the proportion of content hosted by ISPs (note that we have considered heterogeneous parameters). When x is small here, content hosted in B perceived better (because almost all content of B is cached in A, the y_i being fixed, while after some point the quality in A becomes better. This illustrates the impact of CDNs on perceived quality, and that CPs may be disadvantaged if associated with a given ISP. Figure 4 *(right)* shows that as the cached amount in A increases, the quality Q_B of content in B increases, but Q_A also increases (even if to a lesser extent), because σ_A increases.

5 Discussion: CDN and neutrality

The question of how CDNs impact the network neutrality debate hasn't been addressed much, but introducing them into the picture reshapes the economic relationships between actors. Several remarks on that matter can be made from our model:

1) For the pull model, the ISP has interest in accepting the operation of an external CDN, even without charging it. In all other cases, there are conditions on the set of parameters for this to occur, but the ISP often has an interest in hosting a CDN (integrated or not).

2) Hosting a CDN induces less traffic within the network, thus reduces the pressure to increase capacities.

3) The CDN operation fees are paid by CPs (to improve their QoS), which somehow is a shift in payment with respect to the side payments requested by ISPs in the neutrality debate. But this fee can hurt innovation exactly as side payments would: can regulators accept that and still prevent ISPs from imposing side payments?

4) Only the most frequently downloaded content is cached by CDNs. If the service is proposed to all (distant) CPs, this makes a differentiation between CPs, which is in a sense against the neutrality principle, because packets are not treated the same. One can also argue that it makes a difference of treatment between local and distant CPs. In any case, the CDN activity requires attention, and its conformity depends on the definition of neutrality.

5) The case $r = 0$ (no payment) would not create any neutrality problem. But as we have seen, this is acceptable for the ISP only for the pull model with an external CDN.

6) But we assumed here that CPs always prefer to use the CDN services, something to be further studied.

We believe that this work highlights the need for CDNs economic modeling in the Internet and for a deeper investigation of their impact on actors, fairness, and neutrality.

6 Conclusions and perspectives

This paper is a first step in the modeling and analysis of CDNs' economics, in interaction with end users, content providers and ISPs, in order to start discussing their influence into the network neutrality debate. We have presented several situations, depending on whether the CDN is vertically integrated with the ISP or not, combined with the so-called push and pull models, where the traffic is paid by the sender or the receiver respectively. We have determined how the CDN can compute the optimal amount data to be cached in each situation (which of course depends on the selected parameters) and the impact on revenue. The impact on the fairness in terms of treatment between the different content providers is then described.

We observed that revenue-oriented CDNs may treat content providers differently, which could go against principles of neutrality and hurt innovation. Again, conclusions are case dependent, since different parameter choices can lead to different conclusions.

As extensions of the work, we would like to investigate the viability of CDNs' business in neutral and non-neutral contexts, and to study the case of more complex topologies involving more ISPs, CDNs in competition, as well as more content providers. The case when ISPs pay for caching will also deserve attention.

References

1. Atzmon, H., Friedman, R., Vitenberg, R.: Replacement policies for a distributed object caching service. In: Proc. of CoopIS. pp. 661–674 (2002)
2. Ben-Akiva, M., Lerman, S.: Discrete choice analysis. The MIT Press, Cambridge, Massachusetts (1985)
3. Buyya, R., Pathan, M., Vakali, A. (eds.): Content Delivery Networks, Lecture Notes in Electrical Engineering, vol. 9. Springer Berlin Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-77887-5
4. Christin, N., Chuang, J., Grossklags, J.: Economics-informed design of CDNs. In: Buyya, R., Pathan, M., Vakali, A. (eds.) Content Delivery Networks, Lecture Notes in Electrical Engineering, vol. 9, pp. 183–210. Springer Berlin Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-77887-5_7
5. Chun, B.G., Chaudhuri, K., Wee, H., Barreno, M., Papadimitriou, C.H., Kubiatowicz, J.: Selfish caching in distributed systems: A game-theoretic analysis. In: Proc. of PODC'04. pp. 21–30. ACM (2004), http://doi.acm.org/10.1145/1011767.1011771
6. Coucheney, P., Maillé, P., Tuffin, B.: Network Neutrality Debate and ISP Inter-Relations: Traffic Exchange, Revenue Sharing, and Disconnection Threat. Rapport de recherche, Inria (Jan 2012), http://hal.inria.fr/hal-00659278
7. Hosanagar, K.: CDN pricing. In: Buyya, R., Pathan, M., Vakali, A. (eds.) Content Delivery Networks, Lecture Notes in Electrical Engineering, vol. 9, pp. 211–224. Springer Berlin Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-77887-5_8
8. Hosanagar, K., Chuang, J., Krishnan, R., Smith, M.: Service adoption and pricing of content delivery network (CDN) services. Management Science 54(9), 1579–1593 (2008), http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1080.0875
9. Hosanagar, K., Krishnan, R., Smith, M., Chuang, J.: Optimal pricing of content delivery network (CDN) services. In: Proceedings of the 37th Annual Hawaii International Conference on System Sciences (January 2004)
10. Lenard, T., May (Eds.), R.: Net Neutrality or Net Neutering: Should Broadband Internet Services be Regulated. Springer (2006)
11. Maillé, P., Reichl, P., Tuffin, B.: Internet governance and economics of network neutrality. In: Hadjiantonis, A., Stiller, B. (eds.) Telecommunications Economics - Selected Results of the COST Action IS0605 EconTel, pp. 108–116. Lecture Notes in Computer Science 7216, Springer Verlag (2012)
12. Wu, T.: Network neutrality, broadband discrimination. Journal of Telecommunications and High Technology 2(1), 141–176 (2003)