POLITECNICO DI TORINO
Repository ISTITUZIONALE

Energy consumption for data distribution in content delivery networks

Original
Energy consumption for data distribution in content delivery networks / Bianco, A., Mashayekhi, R., Meo, M.. - STAMPA. - (2016), pp. 1-6. (2016 IEEE International Conference on Communications, ICC 2016 Kuala Lumpur, Malesia Giugno 2016) [10.1109/ICC.2016.7511356].

Availability:
This version is available at: 11583/2647795 since: 2016-09-09T17:49:20Z

Publisher:
Institute of Electrical and Electronics Engineers Inc.

Published
DOI:10.1109/ICC.2016.7511356

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

06 October 2026

Energy Consumption for Data Distribution in Content Delivery Networks

Andrea Bianco, Reza Mashayekhi, Michela Meo
Politecnico di Torino, Torino, Italy
Email: {firstname.lastname}@polito.it

***Abstract*—A considerable percentage of worldwide electrical energy is consumed by information and communication technology. One significant element in this perspective are the data distribution systems via Content Delivery Networks (CDNs). We introduce a new model to compute the total energy consumption of CDNs which is based on a hierarchical Internet map and that takes into account the energy consumption needed to keep servers synchronized.**

The CDN is represented as a main server storing the whole data set and several surrogate servers, each caching a subset of the entire data set. Servers are located in a hierarchical three-tier network topology. We analyze the energy consumption trends as a function of the number of surrogate server. Results show that increasing the number of surrogate servers decreases the transmission delay but it does not always lead to decreasing energy consumption. Furthermore, the energy consumption profile as a function of the number of servers strongly depends on the ratio between the number of content requests and modifications. Finally, we show that the adoption of a hierarchical network model permits to highlight slightly different energy consumption trends with respect to those of standard "flat" network representation.

I. Introduction and Related Work

Information and Communication Technologies consume a considerable percentage of the total worldwide energy[1]. Internet accounts for one of the major, fast growing, energy consuming portion of the ICT consumption, and data centers and data distribution systems play a very significant role in this context[2]. Indeed, a vast amount of the currently used energy is consumed for data dissemination from a source to a number of users[3]. Therefore, energy management in data distribution systems is a hot research issue.

Content Delivery Networks (CDNs) are one of the most common data distribution systems. CDNs can be abstracted as a centrally managed pool of computing and storage resources, with high-speed Internet access. CDN sites are distributed at strategically chosen locations throughout the Internet or within ISP domains[4]. When considering data distribution, several energy contributions can be identified: storage energy, energy needed to move data from the server to the users, energy needed to keep the data structure synchronized among servers upon data modifications.

Several research activities studied the energy consumption of CDNs. In [4] and [5], the energy consumption of CDNs is computed, without considering synchronization energy consumption. The work in [2] studies energy distribution representing the Internet map through random graphs.

The novel contribution of this paper is including the synchronization energy consumption in the computation of the total energy consumption of CDN networks, while using a hierarchical Internet representation. We show that the contribution of the synchronization needed to propagate modifications from the main server to the surrogate server may become significant depending on the ratio between data modifications and user requests. As a consequence, we show that in some scenarios increasing the number of surrogate servers increases also the operational costs besides the capital expenditures.

II. Content Delivery Networks

We represent a CDN as a main server, named *primary server*, storing the whole data set, connected to several *surrogate servers* which are positioned on network edge, closer to end users. Surrogate servers store content depending on their cache size and on content popularity, possibly estimated among end users close to the surrogate servers. Storing contents in surrogate servers according to content popularity permits to save storage space, to reduce synchronization energy consumption, to reduce client download time and balance load among servers. It is well known that content distribution and management plays a fundamental role in a CDN. Indeed, a smart content selection through clever caching policies determines the efficiency of the CDN approach. The optimal placement of surrogate servers allows to provide high quality of service and low CDN prices [6]. Given a set of properly placed surrogate servers and a content to be delivered, an efficient content outsourcing strategy should be defined.

Content outsourcing can be chosen among cooperative push-based, cooperative pull-based, or non-cooperative pull-based approaches. In cooperative push-based approaches, content is pushed to surrogate servers from the primary server. The primary server keeps a mapping between content and surrogate servers, and each request is preferentially directed to the closest surrogate server. Only if the request cannot be satisfied by the surrogate server, it is directed to the primary server. In non-cooperative pull-based approach, client requests are always directed to their closest surrogate server. If there is a cache miss, surrogate servers pull content from the primary server. The cooperative pull-based approach differs from the non-cooperative approach because surrogate servers cooperate to get the requested content in case of a cache miss.

III. The Proposed Model

The model we propose is peculiar in two dimensions: the way in which we model the Internet interconnection graph and the fact that we consider the energy cost of content

distribution among servers. The real Internet map is difficult to be estimated, due to the dynamic nature of the Internet, of its huge size and of its hierarchical and administrative-based structure, that has an impact on data distribution policies. To define a graph as close as possible to the real Internet layered ISPs architecture while keeping it simple and treatable, we choose a three tier model. For what concern the data distribution approach, we consider also the synchronization energy consumption in the total energy consumption computation. The synchronization energy is the energy consumed to propagate modified contents from the main server to all surrogate servers that need to host the content.

A. Internet map

As sketched in Fig.1, in the Internet, three types of ISPs can be identified. We label as Tier 3 the edge portion of the network, where edge routers connect to end users. Edge routers are located in Points of Presence (PoPs), which are on one side connected to the Internet via border routers, and on the other side to Customer Edge routers or Subscribers Edge routers. which connect end users to the Internet. Tier 3 ISPs are connected to Tier 2 ISPs by border routers. Tier 2 ISPs typically provide regional or national interconnection among PoPs. As such, "close-by" Tier 3 ISPs may be connected to the same Tier 2 ISP, being able to communicate independently of the core, Tier 1, network. Tier 1 networks are typically responsible of long-distance, international connections.

B. Surrogate server placement

We model a typical CDN network with one primary server that maintains all the contents in the data set, and a variable number of surrogate servers. The main reason to use surrogate servers is to move the content to network edge, closer to end users to improve end-user quality of experience. Several strategies for surrogate placement can be envisioned. In this paper, we use a somehow fair strategy according to our three tier topology model.

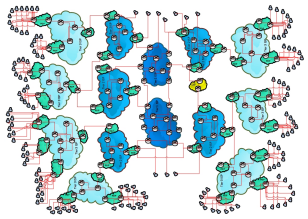

Fig. 1. Three layer Internet map

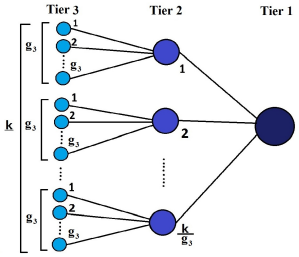

Fig. 2. Simplified map

Fig. 2 reports a simplified representation of the Internet map to better understand the placement strategy adopted in our model. We assume that a core network is available to provide global connectivity and that T_2 Tier 2 ISPs connect Tier 3 ISPs to the core. S surrogate servers, randomly positioned in Tier 3 ISPs are available, with $S < T_2$. We assume that at most one surrogate server exists in each Tier 3 ISP connected to the same set of Tier 2 ISPs. Therefore, for each Tier 2 ISP there is at most one surrogate server placed in one of the Tier 3 ISPs connected to that Tier 2 ISP. Obviously, the content provider is aware of the surrogate server positioning.

C. Content outsourcing

As previously described, content outsourcing can be chosen among cooperative push-based, cooperative pull-based, or non-cooperative pull-based approaches.Authors in [7] compared cooperative content replication and non-cooperative case in a joint optimization problem in CDNs. In this paper, cooperative push-based approach is considered, in which content is pre-fetched to the surrogate servers, being pushed by the primary server. In this scheme, the primary server keeps a mapping between content and surrogate servers, and each request is directed to the closest surrogate server. If a content is not hit i any surrogate server, the request is directed to the primary server. It is also important to note that when a content is modified, the modifications are propagated from the primary server to all surrogate servers that host that content.

D. Cache management

Once the number of surrogate servers and their position is determined, we must define the cache management strategy and cache size in surrogate servers. We analyze three scenarios in which each surrogate server is assumed to have a storage capacity equal to a fraction of the total storage size, denoted B_{tot}, needed to host the entire data set. Obviously, the cache size strongly affects the hit probability, because the larger the cache size, the larger the hit probability. The strategy to choose the content pieces to be cached in each surrogate server is of fundamental important. For what concerns cache management, several research works discuss the best strategy to distribute data among different servers in CDNs. Tuncer et al. investigate in[8] lightweight strategies that can be used by the ISPs to manage the placement of contents in the various network caching locations according to user demand characteristics. Their proposed strategies depend on the volume and the nature of contents in the system. Baliga et al.[9] suggest that frequently used data are better to be replicated and kept close to end users, while rarely accessed data should be replicated less and kept in the primary server only. We analyze two caching strategies to distribute data replica in surrogate caches in this paper. The first one is a uniform distribution strategy in data replication: The data set to be stored in each cache is selected according to a uniform distribution among the whole data set stored in the primary server, i.e. without taking into account any popularity distribution. The second strategy is popularity-based. Contents are cached in the surrogate servers according to their global popularity following a Zipf distribution with parameter $\alpha = 0.8$. [10].

E. Assumptions and notations

The main assumptions and notations are as follow.

- There are S surrogate servers and one primary server. All surrogate servers are located in Tier 3 ISPs. The primary server hosts all M contents.
- Contents are of the same, fixed size B bits.
- All surrogate servers have the same storage capacity equal to a fraction S_C of the total storage capacity, B_{tot}, of the primary server.
- There are T_3 Tier 3 ISPs. Each Tier 3 is connected to n end users, for a total number of users $N = n \cdot T_3$.

Symbol	Default value	Description
S	–	number of surrogate servers
S_C	20%, 40%, 50%	cache size (surrogate servers storage capacity)
M	1000	total number of contents in the data set
B	10^6 bits	size of each content
t	6000 s	time period of the analysis
n_m	$S_C \cdot S$	number of replica for content m
r_m	100,1000	requests for content m
m_m	10,100	modifications to content m
H_{sd}^A	3	hops to fetch content from the same Tier 3 ISP
H_{sd}^B	14	hops to fetch content from the same Tier 2 ISP
H_{sd}^C	25	hops to fetch content from the core network
H_{ps}	-	hops from primary to surrogate servers
T_3	1000	number of Tier 3 ISPs
g_3	20	number of Tier 3 ISPs connected to Tier 2 ISP
T_2	50	number of Tier 2 ISPs
N	2,000,000	total number of end users
n	2000	number of end users per Tier 3 ISP
P_{st}	$7.84 \cdot 10^{-12}$ W	storage power consumption per bit
E_r	$1.2 \cdot 10^{-8}$ J/bit	router energy consumption per bit
E_l	$1.48 \cdot 10^{-9}$ J/bit	link energy consumption per bit
E_{sr}	$2.81 \cdot 10^{-7}$ J/bit	server energy consumption per bit

TABLE I. MAIN NOTATION AND PARAMETERS SETTING

- For each group of g_3 Tier 3 ISPs, there is one Tier 2 ISP. Thus, there are $T_2 = \frac{T_3}{g_3}$ Tier 2 ISPs.
- All Tier 1 ISPs are considered as the core network.
- Each user request is directed to the closest surrogate server that contains the chosen content, if any. Otherwise, it is directed to the primary server.
- The hit probability for each surrogate server, i.e., the probability that the requested content is hosted in the surrogate server is denoted by P_{hit} and depends on the cache management strategy. In the uniform caching strategy, the hit probability is equal to the server relative cache size S_C, the percentage of the data stored in the surrogate server with respect to the total storage in the primary server. In popularity-based caching policy following a Zipf distribution with parameter $\alpha = 0.8$, according to [10], the hit probability is equal to $P_{hit} = 0.82$. Of course, the primary server has hit probability equal to 1.
- When a content is modified, the primary server instantaneously propagates the modified content to surrogate servers based on the probability that each surrogate server hosts each content.
- The average path length in Tier 3 ISPs is set to 3 hops.
- Requests for each content m are generated according to a Poisson distribution with parameter r_m.
- Modifications of each content m are generated according to a Poisson distribution with parameter m_m.

Tab.I summarizes the notations used in the paper, as well as the values of the parameters to derive the numerical results.

F. Model formulation

We derive simple expressions to compute total energy consumption of the CDN modeled according to the above mentioned assumptions. The total energy consumption is the summation of four different components, namely storage, server, transmission and synchronization energy consumption. Storage energy consumption is the energy consumed to store the whole data set in all the servers. Server energy consumption is the energy consumed by the server for each received request to process it, get the content and send it. Synchronization energy consumption is the energy consumed to propagate modified content to the proper surrogate servers. Transmission energy consumption is the energy consumed to transmit content to the user that has requested it. We compute the total consumed energy in the considered measurement period t as the sum of four energy contributions derived in the next subsections:

$$E_{tot} = E_{storage} + E_{server} + E_{synch} + E_{tx} \tag{1}$$

Storage energy consumption

$$E_{storage} = \sum_m B n_m P_{st} t \tag{2}$$

where *B* is the content size, *t* is the time period in which the energy consumption is computed, n_m is the number of replica for content m and P_{st} is the storage power per bit.

Server energy consumption

$$E_{server} = \sum_m B r_m E_{sr} \tag{3}$$

where r_m is the number of requests for content m during period t, and E_{sr} is the server energy consumption per bit.

Synchronization energy consumption:

For any content modification, the content is propagated from the primary server to all surrogates hosting the content.

$$E_{synch} = \sum_m B m_m n_m \left[E_r \left(H_{ps} + 1\right) + E_l H_{ps}\right] \tag{4}$$

where H_{ps} is the average number of hops from the primary server to each surrogate server, m_m is the number of content modifications during period *t* for content m, E_r and E_l are the router and the link energy consumption per bit. When there is no surrogate server, $E_{synch} = 0$.

Transmission energy consumption:

When a user in a Tier 3 ISP makes a request, three different cases have to be considered, labeled as case A, B and C.

1: There is a surrogate server in the same Tier 3 ISP where the request is generated and the requested content is hit in that server. This happens with a probability denoted as P_A,

$$P_A = \frac{S}{T_2} \cdot \frac{1}{g_3} \cdot P_{hit} \tag{5}$$

2: There is not a surrogate in the same Tier 3 ISP where the request is generated, but there is a surrogate in the same Tier 2 ISP, and the requested content is hit in that surrogate server. The probability that this happens is denoted as P_B,

$$P_B = \frac{S}{T_2} \cdot \left(1 - \frac{1}{g_3}\right) \cdot P_{hit} \tag{6}$$

3: There is not a surrogate server neither in the same Tier 3 ISP nor in the same Tier 2 ISP, or there is a surrogate server in the same Tier 3 ISP or in the

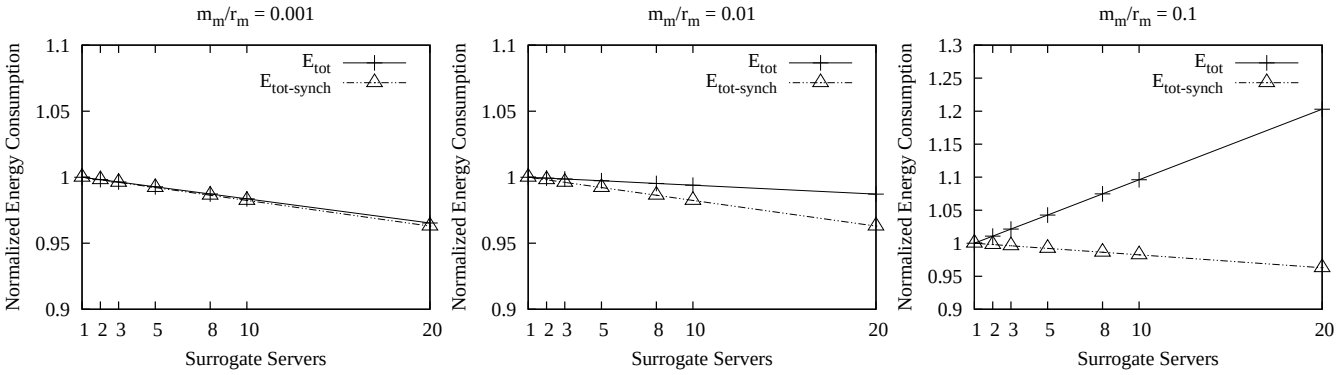

Fig. 3. Energy consumption with/without synchronization. Uniform caching. $S_C = 40\%$. $\frac{m_m}{r_m} = 0.001$ (left), $\frac{m_m}{r_m} = 0.01$ (middle), $\frac{m_m}{r_m} = 0.1$ (right)

same Tier 2 ISP but the requested content is not hit in that server. In this case, the requested content should be fetched through the core network. The probability that this happens is denoted as P_C:

$$P_C = 1 - (P_A + P_B) \tag{7}$$

Thus, the transmission energy consumption when there are S surrogate servers is:

$$E_{tx} = P_A \sum_m Br_m[E_r(H^A_{sd} + 1) + E_l H^A_{sd}] + P_B \sum_m Br_m[E_r(H^B_{sd} + 1) + E_l H^B_{sd}] + P_C \sum_m Br_m[E_r(H^C_{sd} + 1) + E_l H^C_{sd}]$$

where H^C_{sd}, H^B_{sd}, and H^A_{sd} are the number of hops to fetch the content from the core network, from a surrogate in the same Tier 2 ISP but not the same Tier 3 ISP, and from a surrogate in the same Tier 3 ISP.

IV. Results

We wish to assess the impact of the synchronization energy consumption in CDNs. Different scenarios are considered, varying: the number of surrogate servers from 1 to 20, the cache sizes at surrogate server ranging from 20%, 40% and 50% of the whole data set. The ratio between the number of modifications and the number of requests for each content ($m_m/r_m = 1, 0.1, 0.01, 0.001$. Two caching policies, uniform caching and popularity-based caching, are adopted. Finally, we take into consideration the trade-off between the total energy consumption of the system and the content access delay, measured as the average number of hops to hit the content.

A. Uniform caching policy

Fig.3 shows that, for the uniform caching policy and a cache size $S_C = 40\%$, the total energy consumption without considering synchronization energy consumption (denoted as $E_{tot-synch}$ in the plots) always decreases when adding surrogate servers. This is not always true if considering the synchronization energy consumption, which becomes relevant when the number of modifications for each content m_m is significant. Indeed, increasing the ratio with which contents are modified, increases also the synchronization energy consumption. As a result, the total energy consumption (E_{tot}) may increase when adding surrogate servers.

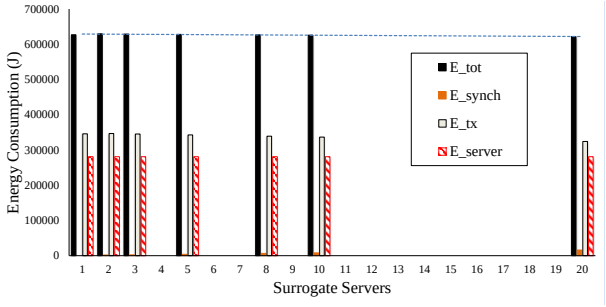

Fig. 4. Energy consumption components. Uniform caching. $S_C = 40\%$, $m_m/r_m = 0.01$, $m_m = 10$, $r_m = 1000$

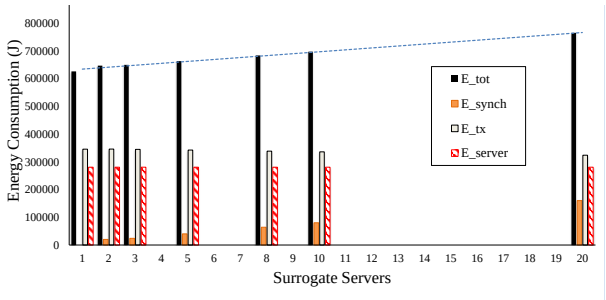

Fig. 5. Energy consumption components. Uniform caching. $S_C = 40\%$, $m_m/r_m = 0.1$, $m_m = 100$, $r_m = 1000$

Further analyzing the figures, it is also interesting to observe the effect of r_m, the number of requests for each content, on the total energy consumption. By increasing the number of requests for each content, both transmission energy consumption and server energy consumption increase. However, adding surrogate servers decreases the total energy consumption because contents get closer to users.

Fig.3 (left and middle graphs) reports the total energy consumption of CDN system in the time period of the analysis t=6000 s for a ratio between number of content modifications (m_m) and number of requests to a content (r_m) equal to 0.001 (0.01). In both cases, the total energy consumption with and without synchronization energy consumption decreases,

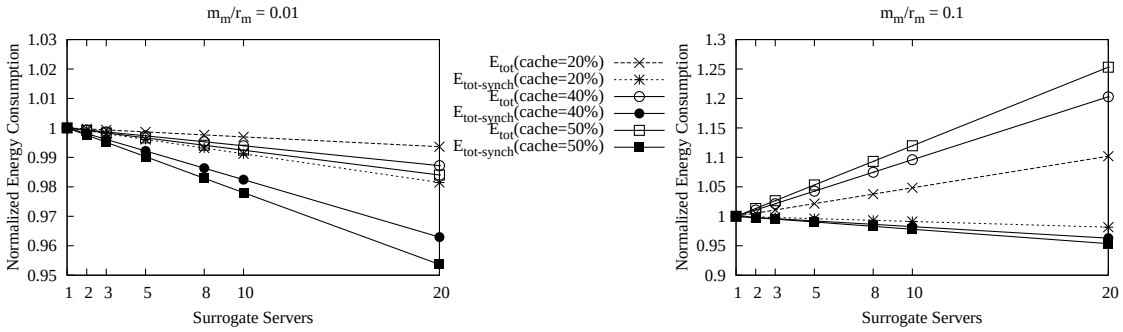

Fig. 6. Total energy consumption with/without synchronization. Uniform caching. $m_m/r_m = 0.01$ (left), $m_m/r_m = 0.1$ (right)

being the synchronization energy negligible with respect to server and transmission energy consumption. When the ratio m_m/r_m increases to 0.1 or to 1 (see. Fig.3(Right)) considering synchronization energy consumption becomes important. Thus, above a given frequency of modifications with respect to requests, increasing the number of surrogate servers increases the total energy consumption of a CDN network.

Figs.4 and 5 represent the breakdown of energy consumption for each component: total, synchronization, transmission and server energy consumption. The storage energy consumption is not reported, being negligible if compared to the other components. Indeed, the total amount of data in the system is assumed to be equal to $M = 1000$ content pieces, each $B = 10^6$ bits. The power consumption of servers to store a single bit of data is equal to $P_{st} = 7.84 \cdot 10^{-12}$ W [5], which is low in comparison with other energy components.

Fig.6 shows the total energy consumption with and without considering synchronization for uniform caching policy and the three different cache size scenarios. Regardless of the cache size, the importance of the synchronization energy consumption becomes evident only for a specific set of parameter values (see Fig.6(Right)). The cache size has an effect on the absolute values of energy consumption but not on the general trend as a function of the number of surrogate servers.

B. Popularity-based caching policy

Fig.7 shows that total energy consumption with and without considering synchronization energy consumption for popularity-based caching policy (Zipf distribution) to some extend leads to same results as in uniform caching strategy. The effect of the ratio between number of content modifications and requests is also critical in this strategy. It is seen that by increasing the number of modifications, after a threshold, the total energy consumption starts to increase. That is the effect of synchronization energy consumption. In the following subsection the exact threshold of this ratio for both uniform and popularity-based strategies are computed and presented.

C. On the ratio between modifications and requests

Figs. 3-7 show that total energy consumption without considering synchronization energy consumption always decreases by adding surrogate servers. However, by considering synchronization energy consumption, when increasing the ratio between the content modifications and content requests, total energy consumption starts to increase when adding surrogates. Identifying this threshold would help network administrators in two directions. First, by providing guidelines for the number of surrogates that should be deployed. Second, the best caching strategy can be adopted to minimize the energy consumption. For the parameter set described in Tab.I, the threshold value for uniform caching policy is around 0.013, while it increases to around 0.028 for popularity-based caching. This is justified by the higher hit probability of the popularity based caching policy that permits to reduce the transmission energy consumption. Indeed, server, storage, and synchronization energy consumption are the same, regardless of the caching policies.

D. Youtube energy consumption

According to the statistics published by YouTube[11], over 6 billion hours of video are watched each month on YouTube, while 100 hours of video are uploaded to YouTube every minute. Taking these statistics into account, YouTube falls into the category in which the ratio between the number of content modifications and the number of requests for each content in a given period of time is close to $m_m/r_m = 0.001$. Thus, YouTube falls in the category of the CDN systems represented in Fig.3(Left), where by increasing the number of surrogate servers, the total energy consumption decreases even when considering the synchronization energy.

E. Prototype IP backbone network

In this section, we compare the results obtained by implementing the 3 layer model, with those of non-hierarchical Internet topology named prototype IP backbone network, presented in[5] and reported in Fig.8. To compute total energy consumption in this scenario, some modifications are required in previous formulas. Synchronization energy consumption and transmission energy consumption should be changed according to the number of hops in the topology. For this network, a semi-analytical approach is taken to compute H_r, which is the average number of hops between nodes. H_r is derived as a power law function of n_m (content replica). N_r is the number of nodes (number of routers considered in previous formulas).

$$H_r(n_m) = A\left(\frac{N_r}{n_m}\right)^{\alpha}$$

In a prototype IP backbone network, H_r is estimated as: $H_r(n_m) = 0.35\left(\frac{N_r}{n_m}\right)^{0.57}$. As mentioned before, storage

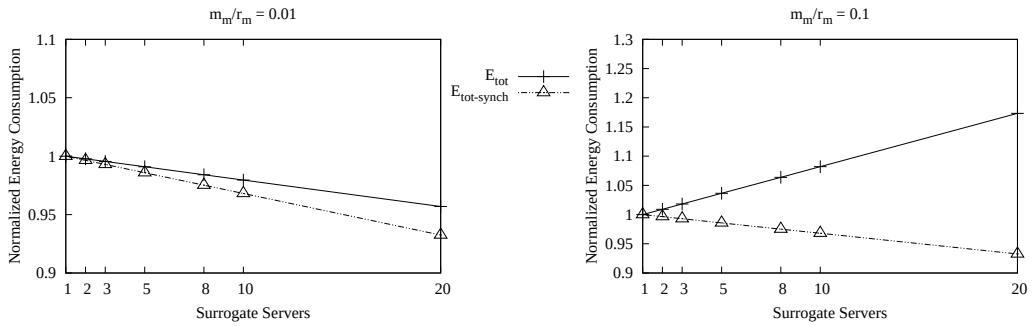

Fig. 7. Total energy consumption with/without synchronization. Popularity-based caching. Cache size =40%, $\frac{m_m}{r_m} = 0.01$ (left), $\frac{m_m}{r_m} = 0.1$ (right)

and server energy consumption are computed with the same formula, but synchronization and transmission energy consumption formulas should be modified according to:

$$E_{synch} = \sum_{M} Bm_m n_m \left[E_r \left(H_r(n_m) + 1\right) + E_l H_r(n_m)\right]$$

$$E_{tx} = \sum_{M} Br_m \left[E_r \left(H_r(n_m) + 1\right) + E_l H_r(n_m)\right]$$

where $H_r(n_m)$ is the average hops between any two nodes.

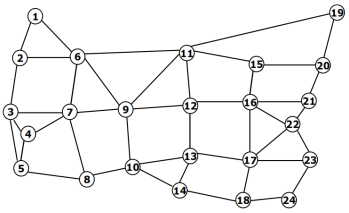

Fig. 8. The prototype IP backbone network, with $N = 24$ and $\Delta = 3.6$.

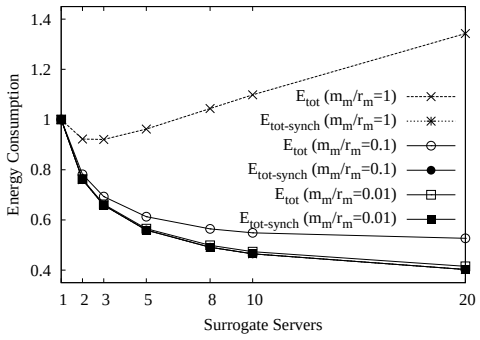

Fig. 9. Total energy consumption with/without synchronization, Prototype IP backbone network. $m_m/r_m = 0.01, 0.1, 1$.. Cache size = 50%

Fig.9 shows the results considering an IP backbone prototype instead of the three-layer topology. All other parameters are the same in both topologies. Differently from when considering the three-layer model, in the prototype IP backbone, the total energy consumption always decreases when the ratio between the number of content modifications and the number of requests for a content is equal to 0.1. The total energy consumption considering also the synchronization energy increases when the ratio becomes close to 1 , whereas this happened for a ratio 0.1 in the three-layer model. Finally, the energy consumption increases more than linearly when increasing the number of surrogate servers.

V. Conclusions

We introduced a new model to compute total energy consumption of CDNs. Two main extensions were introduced with respect to previous models. First, we use a three-layer Internet architecture instead of a "flat" graph. Second, we consider the synchronization energy consumption in the computation of the total energy consumption. Results show that the total energy consumption does not always decrease by adding more surrogate servers in a CDN network, depending on the ratio between the number of content modifications and the number of content requests. If this ratio overcomes a given threshold (around 0.013 for uniform and 0.028 for popularity-based caching policy), increasing the number of surrogate servers may increase the total energy consumption of the network, due to the increase in the synchronization energy consumption.

References

[1] C. Forster et al., "Understanding the Environmental Impact of Communication Systems", Ofcom final report, Apr. 2009.

[2] Uichin Lee, Ivica Rimac, Daniel Kilper, Volker Hilt, KAIST KSE, "Toward Energy Efficient Content Dissemination", IEEE Network ,2011.

[3] Uichin Lee, Ivica Rimac, Volker Hilt, "Greening the Internet with Content-Centric Networking", IEEE INFOCOM, 2012.

[4] A.Feldmann, A.Gladisch, M.Kind, C.Lange, G.Smaragdakis, F.J. Westphal, "Energy Trade-offs among content Delivery Architectures", Telecomm. Internet and Media Techno Economics Conf. (CTTE), 2010.

[5] Kyle Guan, Gary Atkinson, and Daniel C. Kilper, Ece Gulsen, "On the Energy Efficiency of Content Delivery Architectures", IEEE Internationl Conference on Communications (ICC), 2011.

[6] S.Sivasubramanian, G.Pierre, M.van Steen, G.Alonso, "Analysis of caching and replication strategies for Web applications". IEEE Internet Computing, 11 (1), pp. 6066, 2007.

[7] Kyongchun Lim, Yonghwan Bang, Jihoon Sung, June-Koo Kevin Rhee, "Joint Optimization of Cache Server Deployment and Request Routing With Cooperative Content Replication" IEEE International Conference on Communications (ICC), 2014.

[8] D.Tuncer, M.Charalambides, R.Landa, G.Pavlou, "More Control Over Network Resources: an ISP Caching Perspective", 9th International Conference on Network and Service Management (CNSM), Pages: 26 - 33, DOI: 10.1109/CNSM.2013.6727806, 2013.

[9] Jayant Baliga, Robert Ayre, Kerry Hinton, Rodney S. Tucker, "Architectures for Energy-Efficient IPTV Networks", OFC, 2009.

[10] Lee Breslau, Pei Cao, Li Fan, Graham Phillips, Scott Shenker, "Web Caching and Zipf-like Distributions: Evidence and Implications" IEEE INFOCOM, 1999.

[11] http://www.youtube.com/yt/press/statistics.html