G Model
SUSCOM-185; No. of Pages 10
ARTICLE IN PRESS
Sustainable Computing: Informatics and Systems xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Sustainable Computing: Informatics and Systems
journal homepage: www.elsevier.com/locate/suscom

On the energy consumption computation in Content Delivery Networks☆

Andrea Bianco, Reza Mashayekhi*, Michela Meo

Dip. di Elettronica e Telecomunicazioni, Politecnico di Torino, Italy

ARTICLE INFO

Article history:
Received 27 July 2016
Received in revised form 8 March 2017
Accepted 23 August 2017
Available online xxx

Keywords:
Content Delivery Networks
Energy consumption
Hierarchical Internet map
Synchronization energy consumption
CDN caching strategies

ABSTRACT

Data distribution via Content Delivery Networks (CDNs) is one of the most significant energy consuming sectors in the ICT area. A CDN system can be abstracted as a primary server housing the entire data set and several surrogate servers, each one caching a portion of the whole data set. We propose a new model to compute the total energy consumption of CDNs. The model is based on a hierarchical Internet representation, and includes the energy consumption needed to keep servers synchronized. We analyze the effect of synchronization and network topology on the total energy consumption of CDNs. Results reveal that increasing the number of surrogate servers reduces the transmission delay. However, this does not necessarily lead to a reduction of the total energy consumption. Furthermore, the effect of network topology and various caching strategies is described. Results show that CDN energy consumption strongly depends on the ratio between the number of content requests and content modifications, and considering a hierarchical network topology highlights slightly different energy consumption trends with respect to those of classical "flat" network representation.

1. Introduction

Nowadays energy consumption is becoming one of the biggest concerns throughout the world. One of the hottest energy consumption sectors is Information and Communication Technology (ICT) [1]. A considerable energy consuming portion of the ICT consumption is Internet with its dramatically fast growing trend. In this context, a very important role is played by data centers and data dissemination systems [2]. Delivering a vast amount of data from data servers around the world to a numerous number of users demands a considerable energy consumption [3]. Thus, energy management of data distribution systems is recently a hot research issue.

One of the most widely known data distribution systems is Content Delivery Network (CDN). 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]. Performing the process of data management and distribution in CDNs, energy is consumed in different sections due to different activities. These sections are: storage energy consumption, energy consumed by the servers to react to content requests, energy consumed to deliver the data to users, and finally the energy consumed to keep data servers synchronized upon data modification and insertion.

In this paper we propose a simple model for a deep understanding of CDNs energy consumption, that can be used to facilitate the design of more energy efficient solutions. One of the key features of the model is that it includes synchronization energy in the computation of the total energy consumption of CDNs. Through the model, we assess the effect of synchronization energy consumption when the ratio between the number of content modifications and insertions and content requests varies. We also consider the effect of the topology in the energy formulation to demonstrate that using a more realistic hierarchical topology representation can have a non-negligible impact. As a consequence, we show that in some scenarios increasing the number of surrogate servers increases also the operational costs, besides the capital expenditures. Finally, the trade off between energy consumption and data distribution delay is considered.

The paper stems from a previous preliminary contribution [5]. With respect to the previous version, the novelty of this paper consists of three aspects. First, the model is enhanced to take into account also new contents insertion; second, the model is validated through simulation; and, finally, additional scenarios are considered with different distributions of content popularity.

The paper is organized as follows. Section 2 discusses the related works. In Section 3 CDN architectures are briefly summarized, while Sections 4 and 5 present the new model to compute total

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author.
E-mail address: reza.mashayekhi@polito.it (R. Mashayekhi).

http://dx.doi.org/10.1016/j.suscom.2017.08.008
2210-5379/© 2017 Elsevier Inc. All rights reserved.

energy consumption of CDNs. In Section 6 the model is validated through simulation and in Section 7 performance results are presented and discussed. Finally, Section 8 concludes the paper and discusses possible future research activities.

2. Related works

Energy consumption of CDNs has been addressed in many previous research activities. In this work we target the aspects which have not been fully covered in previous works. The main differences are discussing the effect of a hierarchical topology instead of a random graph, and including synchronization energy consumption in the computation of total energy consumption of CDNs.

In [4,6], energy consumption of CDNs is computed, without considering synchronization energy consumption. In [7], authors propose a novel load adaptation technique for Caching Points which not only enhances content download rate but also reduces transmission energy consumption through random sleep cycles. They also ignore the synchronization energy consumption in CPs. The authors of [8] also propose a multiphase sleep-wake mechanism (MSWM) for CDN to achieve energy saving. Synchronization energy consumption is also not considered in this work.

The work in [2] studies energy distribution representing the Internet map through random graphs. Random graphs are used to analyze data distribution systems for networks with different structures, where the main network parameters can be synthetically controlled. However, topologies derived from random graphs represent a fairly simplified model. Moreover they are not taking into account the hierarchical structure of the Internet, which may have a non-negligible impact on the energy consumption related to data distribution. In this paper, we model the Internet map according to a three-tier model which represents the hierarchical architecture of today Internet.

In [9,10], the authors highlight the effect of exploiting a distributed data center controlled and managed by an ISP. They discuss the usage of a hierarchical network topology to find the best placement of data caches and data centers to decrease the energy consumption, whereas we propose a hierarchical Internet topology model to realistically compute the CDN energy consumption.

Among three major CDN content outsourcing approaches, namely cooperative push-based, cooperative pull-based, or non-cooperative pull-based, previous works targeted different approaches. For instance authors in [11] compared cooperative content replication and non-cooperative case in a joint optimization problem in CDNs, while we utilize cooperative push-based approach in our model.

For what concerns cache management in CDNs, several research works discuss the best strategy to distribute data among different servers in CDNs. Tuncer et al. investigate in [12] 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. [13] 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 consider three different caching strategies in this work: *Static*, *Least Recently Used (LRU)*, and *Least Used (LU)*.

3. Content Delivery Networks

A CDN is represented by a main server, named *primary server*, which stores the entire data set and is connected to several *surrogate servers*. Surrogate servers are positioned at the network edge, closer to end users. Surrogate servers store contents in their cache based on the caching policy and content outsourcing strategy exploited in the network. Contents to be stored in surrogates can be chosen uniformly among the whole data set, or according to the content global popularity. Storing contents based on their popularity allows to save storage space, to balance load among servers, to reduce transmission energy consumption and consequently to reduce client download time. Content distribution and management in CDNs plays an important role. Indeed, the efficiency of the CDN approach can be determined by a smart content selection through clever caching strategies. The optimal placement of surrogate servers permits to provide high quality of service and low CDN prices [14]. Having a set of surrogate servers properly placed in the network and a smart content selection policy, 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. Indeed, the primary server keeps a mapping between surrogate servers and the contents stored in each one. Therefore, on each content request, the request is directed to the closest surrogate server hosting that content. Only if the request cannot be accessed by any of the surrogate servers, it is directed to the primary server. In non-cooperative pull-based approach, client requests are always directed to their closest surrogate server. On each miss, surrogate servers pull the missed content from the primary server. The difference between cooperative and non-cooperative pull-based approaches is that in cooperative approach surrogate servers cooperate to get the requested content in case of a cache miss. It means that the surrogate with a cache miss, pulls the content from one of the other surrogates instead of the primary server.

According to the above descriptions, in the next section we present our model.

4. System and model assumptions

In this section we propose a model to compute energy consumption of Content Delivery Networks [5]. The novelty of this model is in two dimensions. Firstly, the model includes synchronization energy into the total energy consumption; synchronization energy is the energy consumed to keep all surrogate servers updated through propagating the modified contents from the primary server to all surrogate servers housing that content. The second novelty of the model is that it considers a network topology that represents the Internet map in more realistic way with respect to what is usually done. The real Internet map is difficult to be represented, due to a number of factors: the dynamic nature of the Internet, its huge size and its hierarchical and administrative-based structure, that has an impact on data distribution policies. We represent the Internet topology with a three-tier model representing the layered ISPs architecture. In what follows, we provide details about our model, the topology, surrogate server management, and system assumptions.

4.1. Internet map

In the Internet, three types of ISPs can be identified. In Tier 3 ISPs, which are the edge portion of the network, edge routers connect end users to the Internet. They are located in Points of Presence (PoPs), which are on the 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. Border routers connect Tier 3 to Tier 2 ISPs. 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. Tier 1