A stock options metaphor for content delivery networks

Elias Vathias1 · Stathes Hadjiefthymiades1

Accepted: 31 December 2021 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
We focus on the timely issue of Content Delivery Network (CDN) resource management. We introduce a multi-stage scheme that leverages solutions from the capital market. The mechanism of Stock Options (SOs) is used to address a potential scarcity of resources, not adequately addressed by other predictive mechanisms previously introduced by the authors. Using a Predictive Reservation Scheme (PRS), network resources offered by a CDN are monitored through established techniques (Kernel Regression Estimators) in a given time frame. Next, a Secondary Market (SM) significantly reduces resource waste by allowing the fast exchange of unused (remaining) resources granted to Origin Servers (OSs). This exchange occurs either by implementing socially optimal practices or by allowing automatic electronic auctions at the end of the day (EoD) or shorter time intervals. Finally, we further enhance our Load Prediction Mechanism (LPM); SOs are purchased and exercised, depending on the lack of resources at the EoD. As a result, OSs may acquire resources (if required) at a standard price. The effectiveness of the proposed stock market-based CDN resource management framework further improves.

Keywords Content Delivery Network · Origin Server · Meta-CDN · Forecasting · Resources · Socially Optimal Practices · Automatic Electronic Auctions · Capital Markets · Stocks · Secondary Market · Stock Options

Elias Vathias
evathias@di.uoa.gr

Stathes Hadjiefthymiades
shadj@di.uoa.gr

1 Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece
1 Introduction

Nowadays, the adoption of CDNs is standard practice for many content-generating entities. The population of OSs is constantly increasing with content that becomes multimedia-richer, “heavier,” and, structurally, far more complex. In this evolving ecosystem, we aim to rationalize the reservation and use of CDN resources from both the OS (client) and the CDN perspectives. The use of CDN resources needs to be carefully planned in the space-time domains. The CDN gradually shifts its attention to more fine-grained resource management schemes that minimize overheads (unused capacity, etc.). Schemes and techniques for more accurate resource claims are of common interest as they positively impact the involved clients (OSs) and the CDN operator/provider. Resource utilization can be maximized and, thus, advance the pertinent economy in total (i.e., improve the position of both sellers and buyers). CDN providers can deliver better services to a broader audience at more competitive prices. Furthermore, improving resource utilization may result in the reduction of network congestion.

We adopt an improved framework for the methodical management of CDN resources. Specifically, we borrow concepts and techniques from the capital market [37, 58, 62, 63]. We treat the CDN resources (assets) as (capital) stocks. The stockholders are the content-generating organizations (OSs). According to the client needs for resources, stocks are purchased dynamically, aiming to sustain the experienced visitor load efficiently. The CDN monitors the incoming traffic of each client and learns the temporal distribution of their resource use (bandwidth) to protect against resource over/under-use. This knowledge is of utmost importance for the rationalization of CDN resource management. Using the proposed LPM, the CDN then establishes load predictions for the immediate future to reserve resources for the OS well beforehand and, thus, cope with the fluctuating load. Resources can be claimed from the CDN but also traded among clients (Fig. 1). A tool for trading CDN resources, which improves the resilience of the LPM to forecasting failures, is the capital SM. CDN resources are traded in the SM by the actual proprietors (i.e., clients that have already acquired resources but experience load incompatible with their expectations/predictions).

This paper introduces another tool that further improves the forecasting and resource trading mechanisms’ resilience, the capital SOs. According to CDN predictions, each OS buys SOs at prices and for the duration specified by models such as that of Black-Scholes (BS) [37], Barone-Adesi & Whaley [3], Bjerkusund & Stensland [8], Ju & Zhong [38], and Binomial and Trinomial Trees [19]. These SOs are exercised if and when needed (i.e., if an OS runs out of resources at the end of the billing period and cannot locate any available surplus in the SM).

Overall, this resource management scheme (Fig. 1) is essential for CDN providers that “lease” finite resources (bandwidth) from other operators. The latter stakeholders (higher-level entities) constitute a meta-level in the CDN ecosystem (meta-CDN). We access the performance of the proposed scheme using extensive real traces taken from high traffic OSs, such as chambers of commerce, universities, popular sports content websites, and e-commerce websites. Our findings show that the introduction of the SOs enables the proposed framework to further “absorb” forecasting failures.
and optimize CDN resource management. We compare the proposed framework with other plans currently being used in the market. We argue that this capital market metaphor is a well-suited option for structuring the modern World Wide Web (WWW) ecosystem.

The paper is organized as follows. Section 2 discusses the related work and Sect. 3 presents the architecture of the proposed framework. Section 4 describes current CDN pricing policies on the market. Section 5 deals with the simulations we choose to perform, describes their related parameters and discusses the corresponding metrics and results. Finally, Sect. 6 concludes the paper and discusses future work.

2 Related Work

Internet resource pricing is a crucial factor for the efficient allocation of Internet resources and the determinant of profit [65]. CDNs [10, 15] have to deal with the cost of interconnection and traffic on their networks. At the same time, they are trying to increase their revenues by choosing effective pricing strategies that have to do with the efficient allocation of Internet resources.

In view of the above, there are three primary pricing models: flat pricing, usage-based pricing, and congestion pricing.

Taking a look back at the early days of the Internet, users utilized a small number of network resources. Thus, Internet Service Providers (ISPs) [49, 50, 56], aiming to attract many users, adopted a flat-rate billing, charging users based on access costs. Flat pricing was easy to implement and also stimulated network usage. However, drawbacks emerged as the increase of network content, in conjunction with the lack
of incentives for efficient network resource usage, resulted in increased traffic and the overall network performance degradation. To address this, usage-based pricing [17, 21, 31] was proposed. The main idea was that if the charge were usage-based, and since Edell and Varaiya [22] showed that users are susceptible to pricing, a fairer and more efficient use of resources would be feasible. Among the problems that had to be addressed were the privacy issues in processing audits and the charge of non-expected traffic, i.e., advertisements, spam, etc. For example, the 95th percentile pricing became an industry standard. In this pricing model, the peak flow within 5% of the total time (36 h per month) is free of charge.

However, network traffic continued to increase, exacerbating congestion. As a result, the aforementioned pricing became more complex, leading to a relatively dynamic pricing model, congestion pricing [16, 29, 40, 42, 43, 48, 49, 61, 68], which has been studied extensively. Congestion pricing dynamically sets prices that can reflect approximate real-time network resource usage and, especially when the network is busy, encourages shifting the traffic from peak time to non-peak time, reducing congestion.

Here it is worth noting that, in the CDN context, congestion reduction is not a key goal of pricing. Content providers (CPs) subscribe to CDN services to overcome, among other things, Denial-of-Service (DoS) Attacks or Flash Crowds. Moreover, the traffic of various CPs is unlikely to surge at the same time. So CDNs can temporarily adjust their infrastructure to handle the traffic spike and improve the availability of content. As a result, the research on congestion pricing, while relevant, does not directly affect the CDN.

With regard to the pricing of the different models applied, pricing mechanisms can be categorized into two types: best-effort and Quality of Service (QoS)-enabled. In the best-effort type, users are charged according to access rate or resource usage. In the QoS-enabled type, ISPs tend to serve different data streams with different QoS and price levels. Generally, for best-effort networks, pricing is always done at the edge of networks and incurs a lower overhead cost, while QoS-guaranteed services involve a higher audition and accounting cost. Priority-based pricing was first proposed by Cocchi et al. [13, 14] to perform service layering with the corresponding pricing. Another well-known QoS differentiation proposal was Odlyzko’s Paris Metro Pricing [55], which divides the network into subnets and charges them differently. QoS-guaranteed network architectures (e.g., IntServ and DiffServ) and their corresponding pricing mechanisms have been widely studied [18, 20, 25, 30, 59, 64].

There are two main models for determining the appropriate price levels regarding network pricing methods: system optimization (models) and strategic optimization (models). System optimization models are mainly based on optimization theory like the concept of the Network Utility Maximization (NUM) framework proposed by Kelly [41], which is the initial work of Internet system optimization, as well as other works on network utility maximization [12]. Strategic optimization models, i.e., considering the strategic behaviors of others when setting prices or making other decisions [4, 5], are based on non-cooperative games [53, 54].

Following those previously mentioned, several recent studies have been conducted in various aspects of caching and CDN technologies, including resource management and pricing.
2.1 Caching

In the field of multi-level caching, it is observed that some clients exhibiting “aggressive” behavior tend to monopolize the cache disk space, thus, enjoying high hit rates. At the same time, other clients are confined to restricted disk space and suffer the eviction of “important” objects, thus, experiencing numerous cache misses and underperformance. Alyfantis et al. [2] proposed a framework that discourages monopolizing the cache disk space by a minority of clients while rewarding clients who contribute to the overall hit rate by using more disk space. Hosanagar et al. [33, 35] find that the adoption of traditional best-effort caching will decrease as OSs move towards dynamic content and simultaneously seek accurate business intelligence regarding website usage. They argue that CDNs can play a vital role in intermediating between OSs that seek the benefits of edge delivery and ISPs that can install servers at the network’s edge. OSs can enjoy the benefits of edge delivery of content without incurring the costs of best-effort caching.

2.2 ISPs – CPs Business Relations

In the field of business relations between ISPs and powerful CPs, Hu et al. [36] described a marketplace for content (digital media) distribution involving both fog/edge and remote cloud computing and storage resources. They conclude that “a small side-payment encourages to use the edge cloud, and caching in the edge cloud benefits the subscribers” [36]. Moreover, Gaivoronski et al. [28] developed several Stackelberg-type game-theoretical models to analyze the relations between a powerful CP and ISPs. They conclude that (a) paid content peering can be mutually beneficial to CPs and ISPs even when the CP has the market power to force the connectivity provider to accept the free content peering and (b) in the case of competition in Internet provision, the paid peering can bring a competitive advantage for the ISPs, which have such agreements with CPs.

2.3 Distributed Group of Nodes

Another standard model for studying caching proxies, CDN, and Peer-to-Peer (P2P) technologies is that of a distributed group of nodes, where each of them uses the storage capacity to create copies of objects, either through replication (permanent copies) or through caching (temporary copies) and render them available to local and remote users. In the replication case, the authors in [45] propose a two-step local search (k-parameter) algorithm that protects nodes from mismanagement. In the case of caching, the authors in [44] propose detection, addressing, and adjusting mismanagement mechanisms.

2.4 Content Delivery Networks

In the CDN area, when several Service Classes (SCs) with different QoS are offered to the publishers, the authors in [24] discussed a simple differentiated service type architecture for CDNs. They proposed a pricing scheme to complement this archi-
tecture and provide fair service to the subscribed publishers (OSs). They also have investigated and suggested methods to determine the optimal pricing of these services, the optimal allocation of resources between the services, and the optimal number of services to be offered.

In similar research, the authors in [34], using analytical models, addressed the optimal pricing of the offered services and studied how external factors, such as the cost of bandwidth and security issues affect pricing. For example, they found that (a) declining bandwidth costs will negatively impact CDN revenues and profits, (b) CDNs will have to lower prices in light of increasing security concerns associated with the content distribution, or they will need to invest in developing and deploying technology to alleviate the security concerns and (c) larger CDN networks can charge higher prices in equilibrium, which should strengthen any technology-based economies of scale and make it more difficult for entrants to compete against incumbent firms.

Hosanagar et al. [32] study the optimal pricing for a monopoly CDN. They find that traditional usage-based pricing plans should entail volume discounts when subscribing CPs have similar traffic burstiness levels, but that volume discounts can prove suboptimal when traffic burstiness is highly heterogeneous. Moreover, they find that profitability from a percentile-based pricing plan can be substantially higher than traditional usage-based billing.

The authors in [23] investigated the maximization of the benefits gained by OS and CDN. Among the results of their research, they found that (a) as the CDN lowers its price, it receives higher interest from the publishers, (b) lowering the price more than a specific level reduces the revenue, because the CDN has a limited cache space and the requests of publishers cannot be completely satisfied, (c) the surrogate revenue is maximized when the total publisher demand is equal to the CDN cache space, (d) surrogates need not be very close to the users and (e), while the system optimum investment maximizes the total of the publisher utilities, it reduces some of the publishers' individual utilities.

Besides “classic” single CDNs, various other solutions are offered to the OSs: Cloud CDNs (CCCDNs) combine resources (compute, storage, and network that are available as web services) with CDN services. Broberg et al. [9] introduced a system that uses storage cloud resources and deploys as many surrogates in user-requested locations as their storage and transfer budget allows, keeping them active as long as sufficient budget remains. Then they measure the utility of content delivery via this system and use it to devise a request-redirection policy that ensures high-performance content delivery [57].

Multi-CDN solutions are viewed as an overlay of existing, individual CDNs, offering more Points of Presence (POPs), optimal reach and redundancy, which are essential if one wants to reach an international or worldwide audience. For example, the authors in [1] refer to Netflix that employs a blend of data centers and CDNs for content distribution and propose a solution that can improve the average bandwidth by more than 50%.

Often, OSs deliver their content to multi-CDN through the use of content brokers. Mukerjee et al. [51] argue that brokers have been shown to invalidate many traditional delivery assumptions (e.g., shifting traffic invalidates short- and long-term
traffic prediction) by not communicating their decisions to CDNs, which can have unintended consequences, including higher costs. Then they analyze these problems [52], examine the design space of potential solutions, and find that a marketplace design (inspired by advertising exchanges) potentially provides an interesting trade-off.

Private CDNs (P-CDNs) are used by companies with high bandwidth and throughput needs or in-house network capacity (such as telecommunication companies), or high security and compliance requirements (such as payment or government companies). P-CDNs do not share resources with other customers. Canali et al. [11] present the deployment of P-CDN by a company with a networking infrastructure that is outsourced to a third party. Their findings suggest that:

- the introduction of a P-CDN can increase by more than eight times the number of clients that can access the multimedia data at the highest quality,
- there is a significant trade-off between the performance of the P-CDN in terms of media quality for the users and the number of edge servers deployed, and
- as the cost of edge servers may be high for a single CP, the parameter for deciding if a company branch is large enough to be chosen to host an edge server remains a critical decision for the tuning of the model.

Telco CDNs are telecommunication companies or telecommunication service providers that are licensing CDN platforms or merge with existing CDN companies and take advantage of the lower bandwidth costs. However, owning a telecommunication network alone is not enough for a Telco CDN to enjoy market benefits. Lee et al. [47] analyze the strategic interactions between CDNs and Telco CDNs and study the conditions that can lead to alliances among Telco CDNs. They also provide evidence that if a telco CDN appropriately manages to offer better service quality exploiting its competitive advantages (e.g., joint traffic engineering and content distribution), market benefits are possible. The authors in [46] show that there are cases under which the potential for full resource pooling and revenue sharing among a Telco CDN federation is beneficial, although in most cases, resource pooling on its own brings more benefits to each individual telecom operator. Frangoudis et al. [27] propose an architecture for on-demand service deployment over a Telco-CDN, where the CDN resources are leased dynamically in different regions based on the customer demand. Spagna et al. [60] offer design considerations for building Telco CDNs with a slight focus on mobile networks. They discuss such questions as cache placement, request routing, and content outsourcing.

In Virtual (or Virtualized) CDNs (vCDNs), virtual caches are deployed dynamically (as virtual machines or containers) in physical servers distributed across the provider’s geographical coverage. A vCDN can be more cost-effective than a CDN running on dedicated infrastructure since virtual machines better utilize server resources. Frangoudis et al. [26] propose an architecture for CDN as a Service (CDNaaS), which allows CPs to order and deploy the vCDN surrogate servers in ISPs. ISPs leverage CDNaaS to receive the CP requests, orchestrate the resources, and deploy the surrogate server functionality on available infrastructure.
P2P CDNs are based on P2P technology and offer lower cost and high-quality services by leveraging the massive fragmentation of idle resources in edge networks. Xu et al. [66] propose and analyze a novel hybrid architecture that integrates CDN- and P2P-based streaming media distribution and significantly lowers CDN capacity reservation cost without compromising the media quality delivered. Yin et al. [67] evaluate the performance of LiveSky, a commercially deployed hybrid CDN-P2P live streaming system, using data from these real-world deployments and argue that such a hybrid CDN-P2P system (a) provides quality and user performance comparable to a CDN and (b) effectively scales the system capacity when the user volume exceeds the CDN capacity.

2.5 Auctions in Resource Management

Regarding auctions in the resource management area, the authors in [39] propose an auction approach to dynamically allocate the wireless spectrum in a SM to enable its better use. The authors in [6] propose ‘Progressive Second Price’ auctions to allocate network bandwidth based on the demand and willingness to pay from competing users who decide their bids based on their accurate valuation. The authors in [7] consider a two-level hierarchical business model for selling bandwidth. In this model, a single vendor allocates bandwidth to intermediary service providers, who in turn sell their assigned shares of bandwidth to their customers at the lower level.

2.6 Our Contribution

The proposed framework uses (a) load monitoring, modeling, prediction, and (b) capital market instruments (SM) from the area of corporate finance, offering a fine-grained resource management scheme. Such a scheme reduces overheads and allows the CDN to plan resource needs carefully. We compare in detail the efficiency of a set of pricing plans and corresponding resource allocations.

The novelty of this paper, in relation to our previous work [62, 63], is in the adoption of another capital market instrument, namely the SOs. The costs of acquiring the SOs are calculated a priori. Then, depending on possible failures of our LPM, SOs are purchased to acquire resources at advantageous prices in the future. These contracts are exercised if and when needed (i.e., if an OS runs out of resources at the EoD and cannot find the total amount required through the SM). The efficiency of the proposed framework further improves.

3 Architecture of the Proposed Framework

This section recapitulates the architecture of the proposed framework so far, detailed in [62, 63]. We continue with the description of the SOs and how they are incorporated into the proposed solution. This architecture allows (a) CDN clients (OS) to rationalize their reservation, use, and exchange of CDN resources and (b) the CDN to also rationalize the reservation and use of resources in meta-CDN instead of blindly reserving bandwidth.
Overall, we adopt a PRS that involves four different (operational) aspects (Fig. 2). Such a scheme tries to accurately model the client’s (OS’s) load and use such information for prompt reservations of CDN resources. To implement these resource reservations within the CDN we adopt the financial instrument of stocks. Although the clients want to predict load as accurately as possible to avoid excessive or under-reservation, predictions cannot be 100% accurate. To face the PRS’s failures, we introduce the SM, where OSs can exchange resources at mutually (across OS) advantageous prices.

Finally, we introduce the SOs. Each OS can buy SOs at prices and a duration defined by models like the BS. The OS can exercise them if needed (i.e., if an OS runs out of resources at the end of the billing period and cannot find the needed resources through the SM).

### 3.1 LPM

The LPM itself is the first aspect of the PRS [62, 63] and is based on Kernel Regression Estimators (KRE) and other complementary mechanisms described as follows (Fig. 3):

- **The Transient High Load Detection Mechanism** (THLDM) detects transient high load created by phenomena like Flash Crowds (legitimate load) or DoS attacks (malicious load) and excludes the corresponding load from the next season’s forecasting.
- **The KRE** accurately models the OS workload over specific time frames and outputs a “refined” model of the anticipated load. This refinement involves a stepization of the derived (and constantly updated) function of the load over time.
- **The Inertia Region Detection Mechanism** (IRDM) modifies the KRE output to include negative step-like segments termed load inertia regions whenever the modeled load leaves local maxima.
- The **Deviation Early Detection Mechanism** (DEDM) monitors with high time granularity the LPM’s metrics about the actual load and adjusts the resources reservation for the next time unit (i.e., next day), modifying the KRE output for that specific time interval.

- The **Initial Resource Reservation Monitoring Mechanism** (IRRMM) applies only to the initial period where no previous actual data exists. The LPM’s result is binding as an upper limit regarding the next reservation of resources by the OS. This prevents the system from malicious users (OSs) who could reserve many resources in advance at small prices and, subsequently, sell them through the SM at a certain profit.

### 3.2 Stocks

To implement the predictive resource reservation within the CDN we adopt the financial instrument of stocks. In the context of our problem, stocks are units of bandwidth share traded at specific prices. Our stock-based modeling is premised upon the following: (a) CDN resources are not infinite i.e., the number of stocks that the CDN can trade at any price is countable, and (b) CDN charging for resource reservation/use relies on the volume/quantity and the planned use date/time. Urgent requests for resources are “penalized” at high prices. Reservations confirmed well beforehand are preferentially priced.

The CDN client, OS, buys stocks to immediately reserve resources at current prices available (Fig. 2). When an OS buys stocks, it actually buys the right to use specific resources (bandwidth) for a specific duration and at a specific price. These rights are easily exchanged via the SM (Figs. 1, 2, and 4). This feature and the overall ease of managing resources are the main reasons for choosing modern financial market instruments in the proposed framework design.
3.3 Secondary Market

The first line of defense for the PRS’s failures is the SM, where neither new financial instruments are issued nor do the issuers, such as corporations, raise new funds. Only investors can purchase from other investors [58].

In our PRS, we establish an SM for stock trading (Figs. 1, 2, and 4). When the reserved resources are significantly underused, the OS can sell a percentage of the (unused) resources to another OS through the SM. Thus, the OS will reduce or even eliminate the unused yet reserved resources (balancing the prediction failure) and improve resource use efficiency (usage relative to reservation). Conversely, when an OS’s reserved resources prove insufficient, the OS can seek extra resources through the SM. Thus, the OS avoids being penalized (overcharged) by the CDN for the needed resources (an implication of forecasting failures) and manages to “absorb” the unexpected load through the SM traded resources.

The SM’s adoption proves beneficial for all involved parties (the seller, the buyer and the CDN). The seller reduces its potential loss (or even makes a profit), the buyer adopts a cost-efficient scheme for unforeseen load, while the CDN increases its resource utilization indicator and improves its competitive position by offering lower prices to its clients.

3.4 SOs

Following the SM’s implementation, there may still be clients who lack resources at the EoD and who will, therefore, turn to the CDN to buy resources at penalty prices. At this point, we present an additional tool that complies with the SO philosophy. By
using SOs, each OS is additionally charged only with the cost of their acquisition, but at the same time, he gains access to cheaper resources.

A SO is a privilege sold by one party to another. It gives the buyer the right (but not the obligation) to buy (call option) or sell (put option) a stock at an agreed-upon price (strike price) within a specified period [37]. European SOs can only be redeemed at the expiration date. On the other hand, American SOs can be exercised any time between the date of purchase and the SO expiration date. This greater flexibility of American SOs (for the buyer) results in a higher risk for the seller. As a result, American SOs are more expensive than European SOs.

We choose to use the American SOs in the proposed framework so that each OS can exercise them whenever needed. However, how are the SOs integrated into the overall function of the proposed framework? In summary, the CDN records how many of its customers at the EoD are deficient in resources and how many resources are missing from each of them. Based on this information, forecasts are made for future failures, and call options are purchased, at the beginning of each day, for each OS for future needs. So, each OS pays the cost of acquiring SOs and, when resources are needed, it acquires the resources at the strike price simply by exercising SOs he has already bought.

At the end of each day, for each OS, there may be available: (a) SOs purchased at the beginning of the day, and (b) additional SOs that were purchased in the previous days which have not yet been exercised or expired. If the OS needs resources, it exercises a corresponding number of SOs and acquires resources at the strike price.

In detail, the situations that may occur at the EoD for each OS are as follows:

- Based on the remaining active SOs, the resources that can be purchased are more than or equal to the resources that need to be purchased by the OS. In this case, the necessary SOs are exercised, and the OS purchases the required resources at the strike price.
- Based on the remaining active SOs, the resources that can be purchased are less than the resources that need to be purchased by the OS. In this case, all available and active SOs are exercised, and the OS purchases the related resources at the strike price. The remaining resources, for which there is no available SO, are purchased directly from the CDN at a penalty price.
- No active SOs are available. In this case, all needed resources are purchased directly from the CDN at a penalty price.

### 3.5 Framework Implementation

The proposed framework can be implemented as software that runs on CDN, evaluates the recorded information from the previous season, estimates future workload for each OS, and enables the trading of resources between OS through SM. Regarding the simulation process, Matlab, C, and Java code were used. The total computation time (on a system with a Quad processor and 8 GB RAM) for emulating the operation of six OSs, one CDN, and one meta-CDN, for 52 weeks (1 year) is approximately 36 min.
4 CDN Market Pricing Policies

Currently, CDN offerings are significantly diversified in terms of their ubiquity, pricing strategies, network capabilities and security, content management, etc.

This section outlines the main pricing policies currently available in the CDN market for all possible traffic volumes (TrV). CDNs offer (a) Pre-Costed (PrC) plans, (b) Pay As You Go (PAYG) plans, (c) Mixed plans, and (d) Free plans. The pricing policies discussed below are differentiated in terms of the billing period, the minimum contract duration, minimum seasonal costs, and the possibility of sharing contracted resources amongst commonly owned OS.

4.1 PrC Plans

In PrC plans, a CDN announces specific SCs, each of which prescribes a predetermined maximum allowable TrV at a specific cost. The OS selects the desired SC based on its TrV estimate. If the actual TrV turns out to be less than or equal to the maximum allowed TrV of the selected SC, the OS receives no discounts on the agreed charges. Unused resources are usually not transferred to the next billing period. Nevertheless, some CDN (e.g., MaxCDN) allow the transfer and use of unused resources, just for the next billing period. If the actual TrV proves to be higher than the maximum allowable TrV of the selected SC, the OS pays a penalty for the extra resources.

4.2 PAYG Plans

In PAYG plans, the OS does not need to estimate its TrV for the next billing period. The OS pays according (a) to the actual resources consumed and (b) to the SC that its TrV belongs. As a result, the CDN does not know how many resources should be reserved for each OS for the next billing period.

This may be one factor that contributes to the fact that SCs of all possible TrV are usually much more expensive in PAYG plans than PrC plans, even in the same CDN. As an example, we mention CDN77, which in PAYG plans charges $0.049 per GB for 5,000–20,000 GB per month and $0.045 per GB for 21,000–60,000 GB per month. The same CDN, for its PrC monthly plans charges $0.02 per GB for 6,000–25,000 GB per month with a penalty value of $0.022 per GB and $0.0189 per GB for 25,000 to 100,000 GB per month with a corresponding $0.02 penalty value. We observe that the prices per GB in PAYG plans are at least twice as high as the corresponding PrC ones.

In PAYG plans, again, various SCs with scalable charges are offered to the OS. The more resources are used by an OS, the lower the GB is charged. There are two billing models currently applied in the market: Unified and staggered.

In the unified model, the OS is charged at the cost corresponding to the scale of its total TrV. For example, if the monthly TrV of the OS is 1,240 GB and the cost per GB for TrV, ranging from 1,000 GB per month to 10,000 GB per month, is $0.020 per GB, then the cost the OS is required to pay is 1,240 GB * $0.020 Per GB = $24.8.

In the staggered model, the OS is charged for the first SC’s TrV at the cost of the first SC, the second SC’s TrV at the cost of the second SC, etc. Thus, if the cost of up
to 10 GB per month is $0.035 per GB, from 10 to 100 GB per month is $0.030 per GB, from 100 to 1,000 GB per month is $0.025 per GB and from 1,000 to 10,000 GB per month is $0.020 per GB, then the cost the OS is required to pay for 1,240 GB actual TrV is 10 GB * $0.035 per GB + (100 − 10) GB * $0.030 per GB + (1,000 − 100) GB * $0.025 per GB + (1,240 − 1,000) GB * $0.020 per GB = $30.35.

It is evident that the staggered model leads to increased costs, for the OS, compared to the unified one. In this example, the staggered model of PAYG is 22.38% more expensive. The only case in which the staggered and unified models lead to equal costs is when the TrV of the OS remains within the limits of the 1st SC.

4.3 Free Plans

It is worth mentioning that there are CDNs such as Cloudflare, Incapsula, Photon (by Jetpack), Swarm CDN, Corn CDN (peer-to-peer MIT project), and jsDeliver, which also offer free plans. These plans include many of the advantages of a CDN service, such as page load speed, website security, Distributed Denial-of-Service (DDoS) attacks protection, and load balancing & failover.

With these free solutions, the small OSs have the opportunity to exploit CDN with apparent benefits.

This, in turn, results in the increased use of CDN technology. A percentage of OSs, who were reluctant to test the CDN services, eventually choose to go to a non-free plan. Gradually, CDN services cease to be a privilege only for the very large and economically efficient OSs.

4.4 Pipeline Pricing or Burstable Billing

All the above pricing models are based on aggregate usage. Some CDNs also offer Pipeline Pricing, which is sometimes referred to as Burstable Billing or 95/5 pricing. Amazon CloudFront calls it Reserved Capacity Pricing. This pricing model is based on the 95th percentile of traffic. Pipeline Pricing lets the OS consume all the needed bandwidth at an agreed maximum rate (i.e., 1 Gbps). The OS cannot exceed that agreed threshold for more than usually 5% of the month (1.5 days); otherwise, it pays the penalty.

5 Simulation and Results

We simulate the proposed framework to assess its performance. We also simulate the main pricing plans currently available in the market [62, 63]. We compare the results and present the effectiveness of the proposed framework in relation to that of the other solutions currently offered in the market.

The rest of this section is structured as follows: In Sect. 5.1, we present the basic design of the simulations we perform and describe in detail their essential parameters and characteristics. We also analyze the cost of TrV - Stocks, and SOs. Section 5.2 briefly reports the plans that we choose to simulate and present their essential characteristics. In Sect. 5.3.1, we report and compare the simulation results from the OS
A stock options metaphor for content delivery networks

point of view. In Sect. 5.3.2, we report and compare the simulation results from the CDN point of view. In Sect. 5.3.3, we define the efficiency (E) of each plan under consideration, present the results, and comment on them. Finally, in Sect. 5.4, we comment on the results as a whole.

5.1 Simulation Definition, Parameters, and Pricing

The simulations involve 6 OSs, 1 CDN, and 1 meta-CDN, are trace-driven, based on anonymized Apache web server cache logs of primary OSs, and cover a duration of 364 days (52 weeks).

More specifically, the logs have been obtained from the following CPs: two chambers of commerce portals (with business information and personalized services), a university website (with student information and personalized course and curriculum services), a popular sports content website (with details on race dates and results, personal drivers pages, national records, and statistics) as well as two popular e-commerce websites. In Fig. 7 of Sect. 5.3, we report the total actual TrV served by the CDN for each OS for the entire simulation period.

5.1.1 Billing Period and Pricing

An essential feature of the proposed framework is that pricing is calculated per day. In other words, the billing period lasts one day. On the other hand, the pricing is usually calculated per month in the market plans. To facilitate the simulations we perform, we assume that the billing period of the simulated market plans lasts 4 weeks or 28 days.

5.1.2 SCs, TrV Range and Accuracy

Another important parameter of the simulations we perform is the number of SCs. Initially, there were usually 3 SCs offered in the market plans: low, medium, and high. More recently, the CDNs began to offer more SCs. We assume that the more SCs are available for an OS to choose from, the more proportional, fair, and low pricing the OS can obtain. Therefore, we simulate two sets (with 3 and 6 SCs, respectively) in all the plans under review, thus doubling the simulation runs.

Following the general market trends and practices, in the case of 3 SCs, we set the following monthly TrV ranges: $P_1 <= 10^2$ GB per month, $10^2$ GB per month $< P_2 <= 10^4$ GB per month, $10^4$ GB per month $< P_3 <= 10^6$ GB per month and $10^6$ GB per month $< P_4$ (unlimited). Respectively, in the case of 6 SCs, the monthly TrV ranges are: $P_1 <= 10^2$ GB per month, $10^2$ GB per month $< P_2 <= 10^2$ GB per month, $10^3$ GB per month $< P_3 <= 10^3$ GB per month, $10^3$ GB per month $< P_4 <= 10^4$ GB per month, $10^4$ GB per month $< P_5 <= 10^5$ GB per month, $10^5$ GB per month $< P_6 <= 10^6$ GB per month and $10^6$ GB per month $< P_7$ (unlimited). The mapping of the above TrV in daily traffic is done by dividing these ranges by the number of days of the month (28 for the simulations). Thus, in the case of 3 SCs, the daily TrV range of SC $P_1$ is 3.57 GB per day and so on. In Tables 3 and 4, the first three columns show the detailed TrV ranges for 3 or 6 SCs and monthly or daily billing. As indicated in the
simulation results (see Sect. 5.3), in some of the plans under consideration, the number of offered SCs plays a crucial role in the cost that each OS pays and the utilization rate of reserved resources.

The accuracy of measurement of the TrV is 1 byte. This accuracy also affects the pricing accuracy of TrV and Stocks (see Sect. 5.1.7) and SOs (see Sect. 5.1.8).

5.1.3 Penalty Type

Following the general market trends and practices, we set two penalty types for the simulations: the high and the low. We set the high penalty (HP) price at 120% of the standard price of each SC and the low penalty (LP) price at 95% of the standard price of each SC. Note that in the “unlimited” SCs P4 (when 3 SCs are offered) and P7 (when 6 SCs are offered), no penalty is charged.

Both penalty types are simulated in all the plans under consideration, except the PAYG plans and the optimal proposed framework (see Sect. 5.2), due to the nature of their operation.

5.1.4 Unused Resources Exploitation (URE)

URE is the ability to transfer the unused resources to the next billing period. URE is enabled (and simulated) in some variants of our proposed framework, but not at the rest market plans (see Table 5).

5.1.5 SM and Auctions

At the EoD, double auctions are performed. Any OS with a surplus of resources (stocks) offers 100% of them for sale at a price greater than or equal to the one that it purchased them (either from CDN or through the SM). Also, any OS with a lack of resources (stocks) announces its intention to buy the stocks it needs at a price less than or equal to the standard price it can buy from the CDN.

In sellers-buyer pairs, where prices match, resources are sold in proportion to demand so that all buyers can benefit.

5.1.6 Forecasting

For the KRE (of the LPM), we choose a gaussian kernel smoothing function and a varying bandwidth from 0.8 to 5.1. Control time granularity indicates the rate at which measurements are taken, which is, in turn, used for forecasting. We take measurements every 24 h for the simulations, and the forecasting reference period is one week.

5.1.7 Pricing of TrV and Stocks

Initially, we consider that the costs reported below include all other CDN charges to OS, such as management fees and request and ingress costs. Moreover, the costs (prices) we adopt are currently used in the market, relate to the use (consumption)
per GB, and vary depending on the number (3 or 6) of the offered SCs. In Europe and America, prices are usually lower than on other continents.

In the case of 3 SCs, we set the following costs: P1: $0.0325 per GB, P2: $0.0225 per GB, P3: $0.0135 per GB, and P4 (unlimited): $0.012375 per GB. Respectively, in the case of 6 SCs, the costs are: P1: $0.035 per GB, P2: $0.030 per GB, P3: $0.025 per GB, P4: $0.020 per GB, P5: $0.015 per GB, P6: $0.012 per GB and P7 (unlimited): $0.011 per GB. The costs per GB of the 3 SCs are the averages of the corresponding costs of the 6 SCs. For example, in the 6 SCs, P1 costs $0.035 per GB (TrV range 0–10 GB per month), while P2 costs $0.030 per GB (TrV range 10–100 GB per month). When 3 SCs are available, P1 has costs ($0.035 + $0.030)/2 per GB (TrV range 0–100 GB per month). In addition, the SCs P4 and P7 are 9.09% cheaper than the previous ones (P3 and P6).

Regarding the pricing of stocks in the proposed framework, there is no differentiation: the cost is still based on TrV, and the cost per GB remains the same, regardless of the number of stocks corresponding to 1 GB. In the simulations of the proposed framework, we have adopted the ratio of one stock-one byte. The pricing accuracy of TrV and Stocks is 1 byte. Table 1 summarizes all the information discussed so far in Sect. 5.1.

### 5.1.8 Pricing of SOs

In this section, we comment on the pricing of SOs. The SOs are used only in some variants of the proposed framework (plans 6.a and 6.b) and not in the rest of the

| Table 1 General simulation parameters | Characteristic | Value(s) |
|---------------------------------------|---------------|----------|
| Meta-CDN                              | 1             |          |
| CDN                                   | 1             |          |
| OSs                                   | 6             |          |
| Duration of Simulation Period          | 364 days      |          |
| Billing Period                        | 1 day or 28 days |      |
| Number of SCs                         | 3 or 6        |          |
| Penalty Type                          | High (120% of Standard Price) and Low (95% of Standard Price) | |
| Accuracy of TrV Measurement           | 1 byte        |          |
| Accuracy of TrV, Stocks, and SOs Pricing | 1 byte      |          |
| Stock – TrV Ratio                     | 1 stock = 1 byte |      |
| Auction Type                          | Double        |          |
| Percentage of Unused Stocks traded in the SM | 100%     |          |
| Auction Sell Price                    | ≥ purchased price | |
| Auction Buy Price (bid)               | ≤ standard purchase price from CDN | |
| KRE Kernel Function                   | Gaussian      |          |
| KRE Kernel Bandwidth                  | Varying from 0.8 to 5.1 | |
| Control Time Granularity              | 24 h          |          |
| Forecasting Reference Period          | 7 days        |          |
simulated plans (see Sect. 5.2). These two plans are introduced for the first time in this paper.

For the pricing of European SOs, tools based on closed-form pricing equations of the BS model and their variants are commonly used. The BS model, also known as the Black-Scholes-Merton (BSM) model, estimates the variation over time of financial instruments (such as stocks or futures), assuming that these instruments will have a lognormal distribution of prices. For the pricing of American SOs, which we use in the proposed framework, many numerical techniques and approximations have been developed since they do not have closed-form pricing equations. Some of the most popular models are summarized below.

- **Barone-Adesi & Whaley.** This model [3] structures the value of American SOs into two parts. The first is the value of a European SO, and the second is the value of early exercise. The latter is given by partial differential equations, which Barone-Adesi & Whaley approximate with a quadratic equation (hence the method’s alternative name, the Quadratic Method).

- **Bjerksund & Stensland.** This model [8] was developed in 1993. The Bjerksund-Stensland model is a closed-form option pricing model used to calculate the price of an American option. The method is fast and computationally efficient. The Bjerksund & Stensland method is more accurate for long-dated options than the Barone-Adesi & Whaley method.

- **Ju & Zhong.** This model [38], first published in 1999, is more accurate than the quadratic approximation for short or long maturity options.

- **Binomial and Trinomial Trees.** Both models [19] involve three general steps: (a) A tree for SO prices is constructed. At each time step, the price can either go up or down (for binomial trees). Additionally, trinomial trees allow the SO price to remain the same at each time step. (b) The value of SO at maturity is calculated, and (c) The value of SO at any time before expiry is calculated through backward induction. Advantages of these models are: (a) They are easily understood and do not require complex calculations, (b) they can be quickly implemented, and (c) they can be modified to include dividends. Disadvantages are: (a) they do not produce exact option values (because of their discrete nature) and (b) constant volatility is assumed.

Evaluated SO Pricing Methods In the context of the proposed framework simulation, we initially examined three of the above methods for calculating SO costs: the BS (for European SOs, solely for price comparison purposes), the Barone Adesi & Whaley, and Binomial Trees methods for American SOs. In these three methods, the main parameters we use to calculate the cost of SOs are:

- **Time to Maturity (TTM).** Usual values for SOs are six months, yearly, or more extended. The simulation results revealed that a SO TTM of 60 days was sufficient to meet the OS’s needs almost entirely. Besides, CDN can vary this TTM, depending on the needs.

- **Strike price.** The strike price is defined as (a) in the case of HP plans, the standard stock cost, and (b) in the case of LP plans, the LP stocks cost (Tables 3 and 4).
A stock options metaphor for content delivery networks

- **Dividend.** We consider that the stockholders, i.e., the OSs, receive a zero dividend for their stocks.
- **Annualized volatility:** We consider a minimum possible volatility, 1%, for the stock cost in the future, just for the need to calculate the cost of SO. During the simulation, stock cost remains stable.

Figure 5 describes the cost of SO per different method in the HP plan (6.a), while Fig. 6 describes the cost of SO per different method in the LP plan (6.b). It should be noted here that the number of SCs is limited. As a result, the standard, LP, and HP selling prices and their corresponding SO prices are limited.
SOs calculated using the BS method are the cheapest. SOs calculated by the Barone Adesi & Whaley method are, on average, 2.62% more expensive than those of BS, while SOs calculated by the Binomial method are 2.56% more expensive than those of BS. We chose to use the Barone Adesi & Whaley method in the simulations we perform.

Table 2 summarizes the parameters and their values that we use to calculate the cost of SOs.

| Parameter                     | Value(s)                        |
|-------------------------------|---------------------------------|
| Style                         | American                        |
| Type                          | Call                            |
| Pricing Method                | Barone Adesi & Whaley           |
| HP Plans Strike Price         | Standard Stock Cost             |
| LP Plans Strike Price         | LP Stoke Cost                   |
| Dividend (Continuous Method)  | 0.02                            |
| TTM                           | 60 days                         |
| Annualized volatility         | 1.00%                           |

5.1.9 Comparative Assessment

Tables 3 and 4 summarize the costs of TrV, stock, and SO, per number of SCs and penalty type, that we use in our simulations.

5.2 Simulated Plans

We want to simulate the operation and measure the metrics of the PrC and PAYG plans described in Sect. 4. Moreover, we want to compare the metrics of these plans with the metrics of all variants of the proposed framework. The plans we choose to simulate, which are summarized in Table 5, are the following:

| SC              | TrV Range | TrV/Stock Cost | TrV/Stock Cost | HP SO Cost | LP SO Cost |
|-----------------|------------|----------------|----------------|------------|------------|
| P1              | 0–10^2     | 0–3.57         | 0.032500       | 0.030875   | 0.00027405 |
| P2              | >10^2–10^4 | >3.57–357.14   | 0.022500       | 0.021375   | 0.00018973 |
| P3              | >10^4–10^6 | >357.14–35,714.28 | 0.013500   | 0.012825   | 0.00011384 |
| P4              | >10^6      | >35,714.28     | 0.012375       | -          | -          |

Table 3 Cost and TrV details - 3 SCs

Table 4 Cost and TrV details - 3 SCs

SOs calculated using the BS method are the cheapest. SOs calculated by the Barone Adesi & Whaley method are, on average, 2.62% more expensive than those of BS, while SOs calculated by the Binomial method are 2.56% more expensive than those of BS. We chose to use the Barone Adesi & Whaley method in the simulations we perform.

Table 2 summarizes the parameters and their values that we use to calculate the cost of SOs.

5.1.9 Comparative Assessment

Tables 3 and 4 summarize the costs of TrV, stock, and SO, per number of SCs and penalty type, that we use in our simulations.

5.2 Simulated Plans

We want to simulate the operation and measure the metrics of the PrC and PAYG plans described in Sect. 4. Moreover, we want to compare the metrics of these plans with the metrics of all variants of the proposed framework. The plans we choose to simulate, which are summarized in Table 5, are the following:
5.2.1 PAYG Plans

We simulate two PAYG plan variations (Sect. 4.2), one with the unified billing model (plan 4.b) and a second with the staggered one (plan 4.a). In these plans, no penalty is applied due to the nature of their operation.

5.2.2 PrC Plans

We simulate two variations of PrC plans (Sect. 4.1), one with a HP (plan 2.c) and a second with a LP (plan 2.d). We assume that the seller correctly estimates the TrV of the first month and chooses the appropriate SC in both these plans. Finally, we assume that, for each month that follows, the OS selects the SC that contains, within its TrV range, the actual TrV requested during the previous month. We note that such a selection does not necessarily imply the optimal (lowest) cost for the OS (Sect. 4.3 of [63]).

5.2.3 Optimal Cost PrC Plans

We also simulate two variations of optimal (minimum) cost PrC plans (Sect. 4.1), one with a HP (plan 2.a) and a second with a LP (plan 2.b). In these simulations, we assume that each OS, during each month, selects the optimal SC that results in the lowest cost (Sect. 4.3 of [63]). In this way, we want to determine the optimal (lower) cost for each OS if we assume that every OS can guess the optimal SC for every month for the actual resources served. Finally, we want to determine the margin for improvement that a CDN has by properly informing each OS to select the optimum SC during every billing period.
5.2.4 Proposed Framework Plans

We simulate the proposed framework and record its functionality and metrics by activating, in successive stages, its distinct individual mechanisms. More specifically:

a) First, we activate the forecasting operation (plans 3.1.a and 3.1.b).

b) Then we allow the exploitation of the unused resources, either for 28 days (plans 3.2.a and 3.2.b) or with no time limit (plans 3.3.a and 3.3.b).

c) Then we enable the exchange of stocks through the SM (plans 5.a and 5.b).

d) Finally, we activate the use of the SOs (plans 6.a and 6.b).

Plan 6.a is the successor to plan 5.a (where the penalty price of the resources is higher than the standard price), and plan 6.b is the successor to plan 5.b (where the penalty price of the resources is lower than the standard price). Plans 6.a and 6.b are the new plans we are simulating in relation to [63].

It is evident that there are no economic reasons for using SOs in a LP plan since the additional resources that each OS may need at the EoD can be obtained at a penalty price lower than the standard price. Why, then, do we also simulate this 6.b plan? The first reason is that we want to test and confirm the proposed framework as a whole, to observe and compare the performance of all the plans, including that of plan 6.b. The second reason is even more significant: Many times, in the markets, on a broader field of applications, customers acquire SOs to also ensure that they can access resources in the future. For these cases, therefore, we want to measure the performance of plan 6.b. Please note that, at the end of each day, ending SOs are deactivated and are no longer available the following day.

All the variants are simulated with HP and LP prices for 3 and 6 SCs. In total, we simulate 20 different variants of the proposed framework.

We assume that each OS orders resources close to the actual TrV during the first seven days of simulated time. Mechanisms for unused resources and SM are available from day 2 and day 1 of the first week, respectively. From the following week (second), the LPM, as well as SOs, are activated.

5.2.5 Optimal Proposed Framework Cost - Ideal Plan

This simulation (plan 1) calculates, for comparison purposes, the minimum cost that an OS will pay using the proposed framework. This will be experienced if the LPM is 100% accurate or if the SM succeeds in absorbing all LPM failures. One might say that the proposed framework’s optimal cost resembles a unified PAYG plan in which pricing is calculated per day rather than per month (high billing granularity). As discussed below and shown in the results, the billing period plays an essential role in the derived costs.

Summing up, we simulate the plans as indicated in Table 5.

5.3 Results

We begin the presentation of the simulation results by reporting the total actual TrV served by the CDN for each OS for the entire simulation period (Fig. 7). The total
actual TrV covers a wide range of TrV and helps us better study the behavior of the plans under consideration.

We then structure the simulation results into two sections: the OS pertinent results and the CDN results.

### 5.3.1 OS - Results

In this section, we study the performance of different plans from the point of view of the OS. Table 6 shows the OS’s metrics results for the entire simulation period when 3 or 6 SCs are offered.

#### 5.3.1.1 OS - Average Total Cost and Weighted Average Cost per Served Resources

Next, we comment on the average total cost (Table 6) that an OS is required to pay, for the entire simulation period, according to each simulated plan.
As we mentioned in [62, 63], up to plans 5.a and 5.b (Table 5) and before the introduction of plans 6.a and 6.b where the SOs are used, plan 5.b showed the best performance, achieving the closest (lowest) value to the Ideal plan (plan 1).

By activating the use of SOs and referring to 6 SCs and HP plans, the average total cost of plan 6.a is only 2.97% more expensive than Ideal plan 1, while the average total cost of plan 5.a (before the activation of SOs) is 4.91% more expensive. Thus, the activation of SOs further reduces the average total cost of the OS by 1.94%. For the case of 3 SCs, the activation of SOs further reduces the average total cost of the OS by 1.90%.

Respectively, in the LP plans (5.b and 6.b), the activation of the SOs slightly increases the average total cost of the OS by 0.09% (in the case of 6 SCs) and by 0.11% (in the case of 3 SCs). This price increase, essentially, stems from the cost of the SOs that the OS is required to buy. As we described earlier, we expected this plan (6.b) to be more expensive than 5.b, and we evaluated it mainly for confirmation purposes. Such a plan can be beneficial to an OS in cases where it needs to secure the availability of resources in the future. Nevertheless, the cost of 6.b is lower than all other plans (except 5.b) under consideration.
Here it is worth noting that the average total OS cost of plan 6.b is lower than the average total OS cost of plan 6.a. This is because the strike price for the purchase of resources (in the case of SO exercise) of plan 6.b is 5% lower than the corresponding strike price of plan 6.a. As a result, at the EoD, the extra resources purchased in the plan 6.b, are 5% cheaper than the ones in plan 6.a.

Summing up, if an **OS selects a plan with a lower-than-standard penalty price**, the best (lower cost) plan is 5.b. Plan 6.b is a bit more expensive (but can be used in case of a requirement to ensure the availability of resources in the future), and PAYG plans 4.b and 4.a follows. Plan 2.d is much more expensive.

If an **OS selects a plan with a higher-than-standard penalty price**, the proposed framework’s best (lower cost) plan is 6.a, and plan 5.a comes next. PAYG plan 4.b is about 0.34% less expensive than plan 6.a while plan 4.a is about 10% more expensive than 6.a. Finally, plan 2.c is much more expensive.

Lastly, the 3 SCs are always more expensive than the 6 SCs, in all plans under consideration.

Overall, the framework we propose allows the CDN to provide the requested services at an OS, at a cost close to the ideal, without requiring the OS to estimate and select the volume of the upcoming traffic. This cost is even lower (plans 5.b and 6.b) than the cost of PAYG plans (which in the real market is at least twice as high) and clearly lower than the cost of PRC plans.

Moreover, the negligible computational overhead for determining the SO disposal value is $O(n \cdot m)$, where $n$ denotes the number of SCs and $m$ the different possible maturity periods. In this paper, the number of SCs considered is 3 and 6 (see Tables 3 and 4, respectively), while a single maturity period is allowed (i.e., $m = 1$). The above figures a number of 6 and 12 SO valuations (calculations), respectively.

The above observations also apply to the OS Weighted Average Cost per Served Resources ($/per GB) metric (Table 6).

5.3.1.2 OS - Average Total Reserved Resources This section presents the average total resources that the OS reserves and pays (Table 7) to serve its needs, according to each plan, for the entire simulation period.

The introduction of the SOs (and hence of the plans 6.a and 6.b) does not change the image we described at [62, 63].

In PrC plans 2.c and 2.d, the resources reserved and paid for by the OS are significantly higher than the resources ultimately required and served to the OS customers. Obviously, with PrC plans, the OS pays for large amounts of resources that it does not need. On the contrary, in the optimal PrC plans 2.a and 2.b, reserved resources are very close to the required resources.

In PAYG plans, 4.a and 4.b, theoretically, no resources are pre-reserved for each OS. OSs consume whatever resources they need and pay only for what they have used. Therefore, we consider that reserved resources of PAYG plans correspond to 100% of the required resources. However, how can a CDN pre-reserve resources and serve the OS without knowing its consumption or consumption prediction? As discussed before, this is probably one of the main reasons that a CDN in the market offers PAYG plans with more than twice the price per GB compared to the PrC plans.

In plans 5.a, 5.b, 6.a and 6.b of the proposed framework, the reserved resources correspond to 100% of the required resources, regardless of the number of SCs. Also,
the OSs only pay for the resources they need. In addition, in relation to the PAYG plans, the CDN has a very good view of future demand and can, therefore, charge lower costs per GB and be very competitive with other CDNs that use PAYG or PrC plans.

Lastly, the Penalty type does not affect resource reservation at all in any of the plans considered.

### 5.3.2 CDN – Results

In this section, we study the performance of different plans from the point of view of the CDN. Table 8 shows the CDN’s Metrics results for the entire simulation period when 6 SCs are offered, while Table 9 shows similar results when 3 SCs are offered.

Which are the CDN related metrics that are worth examining? CDN cost is primarily due to the purchase of resources from the Meta-CDN. This cost also includes the CDN management cost, as mentioned earlier. CDN revenue arises from the sale of resources to the OSs. CDN profit results from its revenue minus its expenses. These three metrics are also calculated per served Resources.

In PrC plans, at each billing period, the CDN buys all the resources that each OS request. If some OSs have surpluses and some deficits in each billing period, the CDN can first carry out an internal redistribution of resources. If the total surplus of resources is not enough, then and only then, the CDN needs to resort to Meta-CDN to buy additional resources with a penalty price. Any OS that is short of resources is charged with a penalty for its entire deficit and not just for the deficit percentage that was not covered by the internal redistribution.
A stock options metaphor for content delivery networks

In PAYG plans, we do not know the algorithm that dictates when and how many resources will be purchased from Meta-CDN. They may perform blind Resources reservations from the Meta-CDN. Alternatively, assuming that the CDNs have a tracking mechanism for the total resources consumed per billing period, they may reserve for the next billing period as many resources as were consumed in the previous one. A third alternative is to reserve an extra percentage of resources (over-dimensioning) in relation to what was consumed in the previous billing period. In any case, by not knowing the upcoming load, the CDN is exposed to increased costs. Such costs are passed on to the OS. This is confirmed by the cost of PAYG plans in relation to the cost of PrC plans in the market. This is why the Cost and the Cost per Served Resources prices of 4.a and 4.b plans are theoretical, they do not correspond to reality, and we only study them for comparison purposes.

In the proposed framework, unlike previous plans, the CDN monitors resources demand. Therefore, using the modeling and forecasting mechanisms as well as the

| Table 8 CDN - results (6 SCs) |
|---|
| Plan | Resources Served (GB) | Resources Reserved (GB) | Revenue ($) | Revenue per Served Resources ($ per GB) | Cost ($) | Cost per Served Resources ($ per GB) | Profit ($) | Profit per Served Resources ($ per GB) |
| 1  | 153,174  | 153,174  | 2894.18   | 0.018895   | 1378.57 | 0.009000 | 1515.61 | 0.009895 |
| 2.c | 153,174  | 558,977  | 9816.47   | 0.064087   | 5017.68 | 0.032758 | 4798.79 | 0.031329 |
| 2.d | 153,174  | 558,977  | 9802.53   | 0.063996   | 5017.68 | 0.032758 | 4784.85 | 0.031238 |
| 2.a | 153,174  | 164,802  | 3766.88   | 0.024592   | 1444.80 | 0.009432 | 2322.13 | 0.015160 |
| 2.b | 153,174  | 164,802  | 3514.72   | 0.022946   | 1393.62 | 0.009098 | 2121.11 | 0.013848 |
| 4.a | 153,174  | 153,174  | 3284.16   | 0.021441   | 1378.57 | 0.009000 | 1905.59 | 0.012441 |
| 4.b | 153,174  | 153,174  | 2969.86   | 0.019389   | 1378.57 | 0.009000 | 1591.29 | 0.010389 |
| 5.a | 153,174  | 153,174  | 3036.40   | 0.019823   | 1378.57 | 0.009000 | 1656.83 | 0.010815 |
| 5.b | 153,174  | 153,174  | 2959.72   | 0.019323   | 1378.57 | 0.009000 | 1587.69 | 0.010365 |
| 6.a | 153,174  | 153,174  | 2980.05   | 0.019455   | 1378.57 | 0.009000 | 1599.20 | 0.010440 |
| 6.b | 153,174  | 153,174  | 2962.52   | 0.019431   | 1378.57 | 0.009000 | 1584.54 | 0.010345 |

| Table 9 CDN - results (3 SCs) |
|---|
| Plan | Resources Served (GB) | Resources Reserved (GB) | Revenue ($) | Revenue per Served Resources ($ per GB) | Cost ($) | Cost per Served Resources ($ per GB) | Profit ($) | Profit per Served Resources ($ per GB) |
| 1  | 153,174  | 153,174  | 3035.10   | 0.019815   | 1378.57 | 0.009000 | 1656.53 | 0.010815 |
| 2.c | 153,174  | 3,492,781 | 51008.17  | 0.333008   | 31433.40 | 0.205214 | 19574.77 | 0.127794 |
| 2.d | 153,174  | 3,492,781 | 50994.85  | 0.332921   | 31433.40 | 0.205214 | 19561.45 | 0.127707 |
| 2.a | 153,174  | 173,584  | 4515.21   | 0.029713   | 1448.72 | 0.009458 | 3102.50 | 0.020255 |
| 2.b | 153,174  | 173,584  | 4194.35   | 0.027383   | 1402.84 | 0.009158 | 2791.51 | 0.018224 |
| 4.a | 153,174  | 153,174  | 3480.59   | 0.022723   | 1378.57 | 0.009000 | 2102.03 | 0.013723 |
| 4.b | 153,174  | 153,174  | 3158.59   | 0.020621   | 1378.57 | 0.009000 | 1780.03 | 0.011621 |
| 5.a | 153,174  | 153,174  | 3226.53   | 0.021064   | 1404.84 | 0.009172 | 1821.69 | 0.011893 |
| 5.b | 153,174  | 153,174  | 3143.53   | 0.020523   | 1372.00 | 0.008957 | 1771.53 | 0.011565 |
| 6.a | 153,174  | 153,174  | 3168.95   | 0.020689   | 1380.82 | 0.009015 | 1788.13 | 0.011674 |
| 6.b | 153,174  | 153,175  | 3146.74   | 0.020544   | 1377.97 | 0.008996 | 1768.78 | 0.011547 |
SM and SOs (to address forecasting failures), together with the internal redistribution of any surpluses and deficits at the EoD, significantly reduces the need for extra resources with a penalty price.

5.3.2.1 CDN - Total Cost and Total Cost per Served Resources Regarding the metric of total CDN cost, as we mentioned in [62, 63] up to plans 5.a and 5.b and before the introduction of plans 6.a and 6.b that utilize the SOs, plan 5.b had the best performance (Fig. 8), having a cost even lower than the cost of the Ideal plan (plan 1), while plan 5.a had 1.8% higher cost than the Ideal plan. By activating the SOs, plan 6.a (that corresponds to plan 5.a) is just 0.15% more expensive than the Ideal plan, while plan 6.b (that corresponds to plan 5.b) is less expensive than the Ideal plan but more expensive than plan 5.b due to the cost of SOs.

Moreover, the total cost of CDN is very high in PrC plans 2.c and 2.d, while in optimal PrC plans, 2.a and 2.b decreases significantly. In PAYG plans 4.a and 4.b, the cost is identical to that of the Ideal plan.

However, how can the costs of plans 5.b and 6.b be lower than the cost of the Ideal plan? The primary reasons for this are the following: (a) a percentage of the resources is ordered at a penalty price from the CDN, (b) the total resources acquired are equal to the resources served to the OS, and (c) the LP price (plan b) is lower than the standard price.

Finally, the total CDN cost is affected by the number of SCs, only in PrC and optimal PrC plans. In conclusion, the proposed framework performs very well and allows CDN to serve OS at a very low cost.

The same observations also apply to the total CDN cost per served resources metric (Tables 8 and 9).

5.3.2.2 CDN - Profit to Cost Ratio Figure 9 presents the profit to cost (PtC) ratio for each plan under consideration. Evidently, a high ratio indicates a very efficient case.

The lowest PtC ratio is observed in PrC plans 2.c and 2.d, regardless of the SCs offered. This, coupled with the fact that the cost for the CDN (Figs. 8 and 9), as well as the OS cost (Table 6) of these two plans, is the highest, makes them the worst choice, since a CDN has to invest significantly, charge customers substantially and earn moderately.

The optimal PrC plans 2.a and 2.b present the highest PtC ratio with relatively low costs (Figs. 8 and 9) for the CDN. However, they are quite expensive for the OS (Table 6), so they are not the best choice for them (OS). We also note that optimal PrC plans describe the theoretical optimum of PrC plans. To approach this optimum,
an OS needs to be adequately informed by the CDN after each billing period and
accurately estimate the traffic it will receive in the future.

The results of the PAYG plans differ from one another: **plan 4.b** (unified charge)
has a lower PtC ratio than 2.a and 2.b plans, but it also has low costs for both CDN
(Figs. 8 and 9) and OS (Table 6). We, therefore, believe that it is an appropriate
choice for both the CDN and the OS. **Plan 4.a** (staggered charge), compared to plan
4.b, has a higher PtC ratio, the same cost for the CDN (Figs. 8 and 9) but a higher OS
cost (Table 6). This is an important reason (higher OS cost) not to be preferred by the
OS and, therefore, not widely offered by the CDN. This is why we consider plan 4.a
to be less favorable than plan 4.b.

Regarding the proposed framework, plans 5.a, 6.a, 5.b and 6.b have similar PtC
ratios to that of plan 4.b (Fig. 9), with PtC ratios of plans 5.a, 6.a and 5.b being
slightly better. Taking into account the facts that (a) plan 4.b has higher OS costs
(Table 6) than plans 5.b and 6.b, (b) plan 4.b has higher CDN costs (Figs. 8 and 9)
than plans 5.b and 6.b and (c) plan 4.b has no information about the upcoming traffic,
we consider the proposed framework the best option overall. Our scheme serves
at the lowest possible cost, sells at the lowest price (except the price of the Ideal plan,
which we are studying for comparison purposes), and its PtC ratio is noteworthy, as
it exceeds 115% of the CDN cost for 6 SCs and 128% of the CDN cost for 3 SCs.

**5.3.3 Efficiency**

We define the efficiency (E) of each considered plan as the distance (Euclidean)
of specific metrics of the plan from the corresponding metrics of the Ideal plan
(Table 10). These metrics are (a) the OS average cost per served resources ($ per
GB), (b) the CDN cost per served resources ($ per GB), and (c) the CDN profit per
served resources ($ per GB), as defined in Sect. 5.3.1 and 5.3.2. The shorter the dis-
tance, the better the efficiency (closer to that of the Ideal plan).

\[ E_i = |V_{ideal} - V_i|, i \in \{2.a, 2.b, 2.c, \ldots\} \]

\( V_{ideal} \) denotes the vector representation of the considered metrics for the Ideal plan,
while \( V_i \) denotes the same information for the considered plans.
E. Vathias, S. Hadjiefthymiades

When 6 SCs are offered, the LP plans 6.b and 5.b of the proposed framework have the lowest distance and, therefore, the highest efficiency. Next comes plan 4.b (PAYG with Unified Charge), and then the HP plans 6.a and 5.a of the proposed framework. The remaining plans have a low efficiency. In the case where 3 SCs are offered, the above remarks also apply. It is worth noting that the efficiency of plans with 6 SCs is significantly higher than the efficiency of the corresponding plans with 3 SCs.

Focusing on the proposed framework (Fig. 10): When 6 SCs and HP plans are offered, the activation of the SOs (switching from plan 5.a to plan 6.a) reduces the Euclidean distance (and therefore improves efficiency) by approximately 35.19%. In the LP plans, the activation of SOs (switching from plan 5.b to plan 6.b) improves efficiency by approximately 1.35%. When 3 SCs and HP plans are offered, the activation of the SOs (switching from plan 5.a to plan 6.a) improves efficiency by approximately 14.50%. In the LP plans, the activation of SOs (switching from plan 5.b to plan 6.b) does not affect efficiency.

### Table 10 Efficiency of plans (E)

| Plan | 6 SCs | 3 SCs |
|------|-------|-------|
| 2.a  | 7.77E-03 | 1.50E-02 |
| 2.b  | 5.67E-03 | 1.19E-02 |
| 2.c  | 5.54E-02 | 3.89E-01 |
| 2.d  | 5.53E-02 | 3.89E-01 |
| 4.a  | 3.60E-03 | 5.42E-03 |
| 4.b  | 7.03E-04 | 2.44E-03 |
| 5.a  | 1.21E-03 | 2.96E-03 |
| 5.b  | 6.44E-04 | 2.34E-03 |
| 6.a  | 7.85E-04 | 2.53E-03 |
| 6.b  | 6.34E-04 | 2.34E-03 |

**Fig. 10 Efficiency of the proposed framework**

When 6 SCs are offered, the LP plans 6.b and 5.b of the proposed framework have the lowest distance and, therefore, the highest efficiency. Next comes plan 4.b (PAYG with Unified Charge), and then the HP plans 6.a and 5.a of the proposed framework. The remaining plans have a low efficiency. In the case where 3 SCs are offered, the above remarks also apply. It is worth noting that the efficiency of plans with 6 SCs is significantly higher than the efficiency of the corresponding plans with 3 SCs.

Focusing on the proposed framework (Fig. 10): When 6 SCs and HP plans are offered, the activation of the SOs (switching from plan 5.a to plan 6.a) reduces the Euclidean distance (and therefore improves efficiency) by approximately 35.19%. In the LP plans, the activation of SOs (switching from plan 5.b to plan 6.b) improves efficiency by approximately 1.35%. When 3 SCs and HP plans are offered, the activation of the SOs (switching from plan 5.a to plan 6.a) improves efficiency by approximately 14.50%. In the LP plans, the activation of SOs (switching from plan 5.b to plan 6.b) does not affect efficiency.

### 5.4 Discussion

We summarize by answering whether each of the plans under consideration meets the requirements of the OS and the CDN. The OS’ needs include the low cost of buying, the optimized use of purchased resources, and not being obliged to predict the upcoming traffic. The CDN’s needs include the low service cost, a reliable forecasting traffic mechanism, the optimized use of purchased resources, a competitive sales
price compared to other CDNs, and the highest possible profit. Considering all these points, we argue that:

The PrC plans 2.c and 2.d have the most unfavorable behavior compared to the rest of the plans. The reason is that the CDN pre-commits many resources that, eventually, remain unused. This results in high costs, a relatively smaller percentage profit margin for CDN, and a high final cost for the OS. In addition, the OS is required to choose a plan before each billing period. We consider them the worst option for CDN and OS.

The theoretical optimal limit of the optimal PrC plan shows that there is much room for improvement in PrC plans, as long as the CDN informs the OS after each billing period about the SC that gives them the lower costs, as discussed in [62, 63]. If this happens, then the performance of PrC plans can approach the performance of plans 2.a and 2.b, which implies a significant reduction in CDN resource waste and costs, an increase in its profit rate, and a reduction in the final cost of the OS.

PAYG plans 4.a and 4.b exhibit an even better performance than optimal PrC plans, with even lower costs and zero unused resources, since they pay exactly what they use. Specifically, the 4.b plan (unified charge plan) is quite close to the metrics of the Ideal plan. However, how many resources are indeed reserved from the CDN before each billing period for each OS? The fact that the OS cost of PAYG plans in the real market is at least twice as high as that of PrC plans reinforces our hypothesis that CDNs do not have efficient LPMs for the upcoming traffic of each OS. This leads them to increased costs that are passed on to their customers.

On the other hand, the CDNs that insist on offering PAYG plans are likely to benefit from shorter billing periods, e.g., daily, as it happens in the Ideal plan (which is a PAYG plan with unified charge and daily billing period).

Plans 5.a and 6.a of the proposed framework approach PAYG theoretical performance, while plans 5.b and 6.b outperform them. OS is not required to predict the upcoming traffic or select a SC for the next billing period. CDN has a reliable predictive and failover mechanism for upcoming traffic per OS, resulting in very high resource utilization. The CDN’s cost is low, and its profit rate is satisfactory. Hence, the final OS cost is also low and, in some cases, even lower than the Ideal plan cost. We, therefore, believe that the proposed framework is the best choice for both OS and CDN.

Regarding the number of SCs and how it affects the plan’s metrics in question, in PrC plans 2.c and 2.d, 6 SCs give better results than 3 SCs in the Unused Resources and Cost fields. In optimal PrC plans 2.a and 2.b., the number of SCs has a smaller impact, not affecting the rest of the plans.

With regard to the penalty type and how it affects the performance of the plans in question, the LP gives lower costs in all plans except PrC (2.c and 2.d) due to many reserved resources that remain unused. The penalty type does not affect resource reservation.
6 Conclusions and Future Work

This paper presents a framework for managing CDN resources using tools from the capital market. We rely on our previous work [62, 63] that (a) deals with the efficient prediction of CDN resource use by OS and (b) uses a SM to redistribute resources among OS. We enhance the proposed framework by introducing the SOs capital instrument, simulating different plans and ways to redistribute the remaining resources among OSs.

We performed extensive simulations of all available plans in the market (Sect. 4). We also performed simulations of various variants of the proposed framework. In total, we simulated 34 plans (17 main plans with two sub-plans for 3 and 6 SCs each), covering 6 OSs with a wide range of TrV spanning 52 weeks.

We compared the simulation results from the point of view of OS and CDN independently. The results show that the proposed framework outperforms all the presented plans. Furthermore, by improving resource utilization, the proposed framework may favor the reduction of network congestion.

We believe that the proposed framework can also perform well on each of the individual “classic” CDNs, whether they operate autonomously (implemented as VCDN or CDNaas) or participate in other (multi-CDN or hybrid-CDN) formations. It is uncertain how the proposed framework will perform if brokers are involved since it depends on whether they allow or hide important information about the OSs’ resource utilization.

Apart from our current evaluation of the use of SOs, our plans for future work include (a) the evaluation of the exchange of unused SOs through the SM, (b) the investigation of the use of other estimators, and (c) the investigation of the possibility to use more than one estimator in parallel (estimator pool) and adopt a mechanism for the automatic selection of the more efficient one. Finally, we plan to investigate the co-existence of CDNs that use the proposed framework with other CDNs that do not and evaluate the metrics of the framework we propose for Meta CDNs and Cloud CDNs.

References

1. Adhikari, V. K., Guo, Y., Hao, F., Varvello, M., Hilt, V., Steiner, M., & Zhang, Z. L. (2012). Unreeling NetFlix: Understanding and improving multi-CDN movie delivery. 2012 Proceedings IEEE INFOCOM, Orlando, FL, 2012, pp. 1620–1628. https://doi.org/10.1109/INFCOM.2012.6195531
2. Alyfantis, G., Hadjiefthymiades, S., Merakos, L., & Kostopoulos, P. (2006). A distributed algorithm for sharing Web cache disk capacity. 12th International Conference on Parallel and Distributed Systems-(ICPADS’06), Minneapolis, MN, 2006, pp. 8. https://doi.org/10.1109/ICPADS.2006.12
3. Barone-Adesi, G., & Whaley, R. E. (1987). Efficient Analytic Approximation of American Option Values. The Journal of Finance, 42, 301–320. https://doi.org/10.1111/j.1540-6261.1987.tb02569.x
4. Başar, T., & Srikant, R. (2002). A Stackelberg Network Game with a Large Number of Followers. Journal of Optimization Theory and Applications, 115, 479–490. https://doi.org/10.1023/A:1021294828483
5. Basar, T., & Srikant, R. (2002). Revenue-maximizing pricing and capacity expansion in a many-users regime. In Proceedings. Twenty-First Annual Joint Conference of the IEEE Computer and Communications, New York, NY, USA, vol.1, pp. 294–301. https://doi.org/10.1109/INFCOM.2002.1019271
6. Bitsaki, M., Stamoulis, G. D., & Courcoubetis, C. (2005). A new strategy for bidding in the network-wide progressive second-price auction for bandwidth. In Proceedings of the 2005 ACM conference on Emerging network experiment and technology (CoNEXT ’05). Association for Computing Machinery, New York, NY, USA, 146–155. https://doi.org/10.1145/1095921.1095941

7. Bitsaki, M., Stamoulis, G. D., & Courcoubetis, C. (2006). An efficient auction-based mechanism for hierarchically structured bandwidth markets. Computer Communications, 29(7), 911–921. https://doi.org/10.1016/j.comcom.2005.08.012

8. Bjerksund, P., & Stensland, G. (1993). Closed-Form Approximation of American Options. Scandinavian Journal of Management, 9(1) S87–S99. https://doi.org/10.1016/0105-5221(93)90009-H

9. Broberg, J., Buyya, R., & Tari, Z. (2009). MetaCDN: Harnessing Storage Clouds for high performance content delivery. Journal of Network and Computer Applications, 32(5), 1012–1022. https://doi.org/10.1016/j.jnca.2009.03.004

10. Buyya, R., Pathan, M., & Vakali, A. (Eds.). (2008). Content delivery networks (Vol. 9). Springer Science & Business Media. https://doi.org/10.1007/978-3-540-77887-5

11. Canali, C., Corbelli, A., & Lancellotti, R. (2018). Designing a private CDN with an Off-Sourced Network Infrastructure: Model and Case Study. In 2018 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, 2018, 1–6. https://doi.org/10.23919/SOFTCOM.2018.8555766

12. Chiang, M., Zhang, S., & Hande, P. (2005). Distributed rate allocation for inelastic flows: Optimization frameworks, optimality conditions, and optimal algorithms. In Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, FL, vol. 4, pp. 2679–2690. https://doi.org/10.1109/INFCOM.2005.1498551

13. Cocchi, R., Estrin, D., Shenker, S., & Zhang, L. (1991). A study of priority pricing in multiple service class networks. In Proceedings of the conference on Communications architecture & protocols (SIGCOMM ’91). Association for Computing Machinery, New York, NY, USA, 123–130. https://doi.org/10.1145/115992.116005

14. Cocchi, R., Shenker, S., Estrin, D., & Zhang, L. (1993). Pricing in computer networks: Motivation, formulation, and example. IEEE/ACM Transactions on Networking, 1(6), 614–627. https://doi.org/10.1109/90.2666050

15. Courcoubetis, C., & Weber, R. (2003). Pricing Communication Networks Economics, Technology and Modelling. John Wiley. https://doi.org/10.1002/0470867175

16. Crowcroft, J., & Oechslin, P. (1999). Providing Internet access: What we learn from INDEX. IEEE Journal on Selected Areas in Communications, 13(7), 1162–1175. https://doi.org/10.1109/49.4144636

17. Daishan, M. H., & Verma, P. K. (2007). Resource Based Pricing Framework for Integrated Services Networks. Journal of Networks, 2(3), 36–45

18. Deutsch, H. P., & Beinker, M. W. (2019). Binomial and Trinomial Trees. In: Derivatives and Internal Models. Finance and Capital Markets Series. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-22899-6_9

19. Dovrolis, C., Stiliadis, D., & Ramanathan, P. (2002). Proportional differentiated services: Delay differentiation and packet scheduling. IEEE/ACM Transactions on Networking (TON), 10(1), 12–26. https://doi.org/10.1109/90.986503

20. Edell, R., McKeown, N., & Varaiya, P. P. (1995). Billing users and pricing for TCP. IEEE Journal on Selected Areas in Communications, 13(7), 1162–1175. https://doi.org/10.1109/49.4144636

21. Edell, R., & Varaiya, P. (1999). Providing Internet access: What we learn from INDEX. IEEE Network, 13(5), 18–25. https://doi.org/10.1109/65.793687

22. Ercetin, O., & Tassiulas, L. (2003). Market-based resource allocation for content delivery in the Internet. IEEE Transactions on Computers, 52(12), 1573–1585. https://doi.org/10.1109/TC.2003.1252853

23. Ercetin, O., & Tassiulas, L. (2005). Pricing strategies for differentiated services content delivery networks. Computer Networks, 49(6), 840–855. https://doi.org/10.1016/j.comnet.2005.03.001

24. Fankhauser, G., & Plattner, B. (1999). Diffserv bandwidth brokers as mini-markets. Proceedings of the Workshop on Internet Service Quality Economics. MIT, US

25. Frangoudis, P. A., Yala, L., & Ksentini, A. (2017). CDN-as-a-service provision over a telecom operator’s cloud. IEEE Transactions on Network and Service Management, 14(3), 702–716. https://doi.org/10.1109/TNSM.2017.2710300
27. Frangoudis, P. A., Yala, L., Ksentini, A., & Taleb, T. (2016). An architecture for on-demand service deployment over a telco CDN. In 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, 1–6. https://doi.org/10.1109/ICC.2016.7510921

28. Gavironski, A. A., Nesse, P. J., & Erdal, O. B. (2017). Internet service provision and content services: paid peering and competition between internet providers. Netnomics, 18, 43–79. https://doi.org/10.1007/s11066-017-9114-x

29. Gibbens, R. J., & Kelly, F. P. (1999). Resource pricing and the evolution of congestion control. Automatica, 35(12), 1969–1985. https://doi.org/10.1016/S0005-1098(99)00135-1

30. Gupta, A., Stahl, D. O., & Whinston, A. B. (1996). An economic approach to networked computing with priority classes. Journal of Organizational Computing and Electronic Commerce, 6(1), 71–95, https://doi.org/10.1080/109019399609540269

31. Honig, M. L., & Steiglitz, K. (1995). Usage-based pricing of packet data generated by a heterogeneous user population. In Proceedings of INFOCOM’95, Boston, MA, USA, vol.2, pp. 867-874. https://doi.org/10.1109/INFONCOM.1995.515958

32. Hosanagar, K., Chuang, J., Krishnan, R., & Smith, M. D. (2008). Service Adoption and Pricing of Content Delivery Network (CDN) Services. Management Science, 54(9), 1579–1593. https://doi.org/10.1287/mnsc.1080.0875

33. Hosanagar, K., Krishnan, R., Chuang, J., & Choudhary, V. (2005). Pricing and Resource Allocation in Caching Services with Multiple Levels of Quality of Service. Management Science, 51(12), 1844–1859. https://doi.org/10.1287/mnsc.1050.0420

34. Hosanagar, K., Krishnan, R., Smith, M., & Chuang, J. (2004). Optimal pricing of content delivery network (CDN) services. 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the Big Island, HI, 2004. 10. https://doi.org/10.1109/HICSS.2004.1265480

35. Hosanagar, K., Krishnan, R., Chuang, J., & Choudhary, V. (2002). Optimal Pricing and Capacity Allocation in Vertically Differentiated Web Caching Services”. ICIS 2002 Proceedings. 44. https://aisel.aisnet.org/icis2002/44

36. Hu, X., Kesidis, G., Heidarpour, B., & Dziong, Z. (2020). Media delivery competition with edge cloud, remote cloud and networking. Netnomics 21, 17–36. https://doi.org/10.1007/s11066-020-09139-3

37. Hull, J. C. (2022). Options, futures and other derivatives, 11th Edition. Pearson, ISBN-13: 9780136939917

38. Ju, N., & Zhong, R. (1999). An approximate formula for pricing American options. Journal of Derivatives, 7(2), 31–40. https://doi.org/10.3905/jod.1999.319140

39. Kash, I. A., Murty, R., & Parkes, D. C. (2014). Enabling Spectrum Sharing in Secondary Market Auctions. IEEE Transactions on Mobile Computing, 13(3), 556–568. https://doi.org/10.1109/TMC.2013.17

40. Kelly, F. (1997). Charging and rate control for elastic traffic. European Transactions on Telecommunications, 8, 33–37. https://doi.org/10.1002/ett.4460080106

41. Kelly, F. P., Maulloo, A. K., & Tan, D. K. (1998). Rate control for communication networks: shadow prices, proportional fairness and stability. Journal of the Operational Research Society, 49(3), 237–252. https://doi.org/10.1057/palgrave.jors.2600523

42. Keon, N., & Anandalingam, G. A. (2005). A new pricing model for competitive telecommunications services using congestion discounts. INFORMS Journal on Computing, 17(2), 248–262. https://doi.org/10.1287/ijoc.1030.0062

43. Kunniyur, S., & Srikant, R. (2003). End-to-end congestion control schemes: Utility functions, random losses and ECN marks. IEEE/ACM Transactions on Networking (TON), 11(5), 689–702. https://doi.org/10.1109/TNET.2003.818183

44. Laoutaris, N., Smaragdakis, G., Bestavros, A., & Stavrakakis, I. (2006). Mistreatment in Distributed Caching Groups: Causes and Implications. Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications, Barcelona, Spain, 2006, 1–13. https://doi.org/10.1109/INFOCOM.2006.131

45. Laoutaris, N., Telelis, O., Zissimopoulos, V., & Stavrakakis, I. (2006). Distributed Selfish Replication. IEEE Transactions on Parallel and Distributed Systems, 17(12), 1401–1413. https://doi.org/10.1109/TPDS.2006.171

46. Lee, H., Duan, L., & Yi, Y. (2016). On the competition of CDN companies: Impact of new telco-CDNs’ federation. 2016 14th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), Tempe, AZ, 1–8. https://doi.org/10.1109/WIOPT.2016.7492919
A stock options metaphor for content delivery networks

47. Lee, H., Lee, D., & Yi, Y. (2014). On the economic impact of Telco CDNs and their alliance on the CDN market. 2014 IEEE International Conference on Communications (ICC), Sydney, NSW, 2950–2955. https://doi.org/10.1109/ICC.2014.6883773

48. MacKie-Mason, J. K. (1997). A Smart Market for Resource Reservation in a Multiple Quality of Service Information Network. Available at SSRN. http://dx.doi.org/10.2139/ssrn.975871

49. MacKie-Mason, J. K., & Varian, H. R. (1995). Pricing the Internet. https://econwp.uib.edu/econ-wp/comp/papers/9401/9401002.pdf

50. McKnight, L. W., & Bailey, J. P. (1997). Internet economics. Cambridge, Mass: MIT Press

51. Mukerjee, M. K., Bozkurt, I. N., Maggs, B., Seshan, S., & Zhang, H. (2016). The Impact of Brokers on the Future of Content Delivery. In Proceedings of the 15th ACM Workshop on Hot Topics in Networks (HotNets ’16). Association for Computing Machinery, New York, NY, USA, 127–133. https://doi.org/10.1145/3005745.3005749

52. Mukerjee, M. K., Bozkurt, I. N., Ray, D., Maggs, B. M., Seshan, S., & Zhang, H. (2017). Redesigning CDN-Broker Interactions for Improved Content Delivery. In Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies (CoNEXT ’17). Association for Computing Machinery, New York, NY, USA, 68–80. https://doi.org/10.1145/3143361.3143366

53. Myerson, R. B. (1997). Game theory: analysis of conflict. Harvard University Press. Cambridge

54. Myerson, R. B. (1997). Game theory: analysis of conflict. Harvard University Press. Cambridge

55. Odlyzko, A. (1999). Paris metro pricing for the Internet. In Proceedings of the 1st ACM conference on Electronic commerce (EC ’99). Association for Computing Machinery, New York, NY, USA, 140–147. https://doi.org/10.1145/336992.337030

56. Odlyzko, A. (2001). Internet pricing and the history of communications. Computer Networks, 36(5–6), 493–517. https://doi.org/10.1016/S1389-1286(01)00188-8

57. Pathan, M., Broberg, J., & Buyya, R. (2009). Maximizing Utility for Content Delivery Clouds. In: Vossen G., Long D.D.E., Yu J.X. (eds) Web Information Systems Engineering - WISE 2009. Lecture Notes in Computer Science, vol 5802, 13–28. Springer, Berlin. Heidelberg. https://doi.org/10.1007/978-3-642-04409-0_7

58. Ross, S. A., Westerfield, R., Jaffe, J. F., & Jordan, B. D. (2018). Corporate finance, Twelfth Edition. Published by McGraw-Hill Education. Dubuque

59. Semret, N., Liao, R. F., Campbell, A. T., & Lazar, A. A. (2000). Pricing, provisioning and peer- ing: dynamic markets for differentiated Internet services and implications for network interconnections. IEEE Journal on Selected Areas in Communications, 18(12), 2499–2513. https://doi.org/10.1109/49.898733

60. Spagna, S., Liebsch, M., Baldessari, R., Niccolini, S., Schmid, S., Garroppo, R., … & Awano, J. (2013). Design principles of an operator-owned highly distributed content delivery network. IEEE Communications Magazine, 51(4), 132–140. https://doi.org/10.1109/MCOM.2013.6495772

61. Varian, H. R., & MacKie-Mason, J. K. (1995). Pricing congestible network resources. IEEE Journal on Selected Areas in Communications, 13(7), 1141–1149. https://doi.org/10.1109/49.414634

62. Vathias, E., Nikolopoulos, D., & Hadjiefthymiades, S. (2016). A Capital Market Metaphor for Content Delivery Network Resources. 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), Crans-Montana, 101–108. https://doi.org/10.1109/AINA.2016.108

63. Vathias, E., Katsarou, E., & Hadjiefthymiades, S. (2017). A Secondary Market Metaphor for Content Delivery Networks. Netnomics. 18, 183–214. https://doi.org/10.1007/s10660-017-9120-z

64. Wang, X., & Schulzrinne, H. (2001). Pricing network resources for adaptive applications in a differentiated services network. In Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213), Anchorage, AK, USA, 2, 943–952. https://doi.org/10.1109/INFCOM.2001.916286

65. Xu, K., Zhong, Y., & He, H. (2014). Internet resource pricing models. Springer. New York. https://doi.org/10.1007/978-1-4614-8409-7

66. Xu, D., Kulkarni, S. S., Rosenberg, C., & Chai, H. K. (2006). Analysis of a CDN–P2P hybrid architecture for cost-effective streaming media distribution. Multimedia Systems 11, 383–399. https://doi.org/10.1007/s00530-006-0015-3
67. Yin, H., Liu, X., Zhan, T., Sekar, V., Qiu, F., Lin, C. … Li, B. (2009). Design and deployment of a hybrid CDN-P2P system for live video streaming: experiences with LiveSky. In Proceedings of the 17th ACM international conference on Multimedia (MM’09). Association for Computing Machinery, New York, USA, 25–34. https://doi.org/10.1145/1631272.1631279

68. Yuksel, M., & Kalyanaraman, S. (2003). Pricing granularity for congestion-sensitive pricing. In Proceedings of the Eighth IEEE Symposium on Computers and Communications. ISCC 2003, Kemer-Antalya, Turkey, 1, 169–174. https://doi.org/10.1109/ISCC.2003.1214118

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.