

# Cross-Layer Design for Improved QoE in Content Distribution Networks

Muslim Elkotob  
MEDIA BROADCAST GmbH  
Kaiserin-Augusta-Allee 104-106  
10553 Berlin, Germany  
muslim.elkotob@media-broadcast.com

Karl Andersson\*  
Pervasive and Mobile Computing Laboratory  
Luleå University of Technology  
SE-931 87 Skellefteå, Sweden  
karl.andersson@ltu.se

## Abstract

In this article, we present the challenges faced by communication networks in delivering high-quality video content to mobile and stationary devices serving CDN users. Starting with Content Distribution Network types, an overview is given in order to show how this field develops on the research as well as on the commercial side. Research or academic CDNs, new aspects and features are tested to add scalability and functional features to known systems and prototypes. On the other hand, commercial CDNs serve alternatively customers that use broadcasting services. We also show which performance metrics such as quality of experience (QoE) and time to first byte (TTFB) best capture the dynamics of traffic and services in CDNs. The core of this paper is our proposed network architecture for CDN providers/operators. The architecture combines video multicast optimized trees and cross-layer coordination between the physical DWDM layer (L1) and network layer (L3) for achieving higher efficiency and lower latency values for live streaming and on demand (VoD) video. Due to the pilot implementation of the presented concept being limited in scale, we use simulations in order to perform proof-of-concept on a sufficiently large environment. Results show that there is a strong correlation between the TTFB and QoE metrics with the former taking on values as low as 75 msec in a national 3-tier network. Ultimately, our aim is to familiarize readers with the field of CDNs and also to help them see how network research especially in the architectural, protocol design, and cross-layer design help bring applications in this field to a quality level acceptable by a large community of users.

**Keywords:** Content Delivery Network (CDN), Peering, IP Transit, Video Streaming, Quality of Experience

In this article, we present the challenges faced by communication networks in delivering high-quality video content to mobile and stationary devices serving CDN users. Starting with Content Distribution Network types, an overview is given in order to show how this field develops on the research as well as on the commercial side. Research or academic CDNs, new aspects and features are tested to add scalability and functional features to known systems and prototypes. On the other hand, commercial CDNs serve alternatively customers that use broadcasting services. We also show which performance metrics such as quality of experience (QoE) and time to first byte (TTFB) best capture the dynamics of traffic and services in CDNs. The core of this paper is our proposed network architecture for CDN providers/operators. The architecture combines video multicast optimized trees and cross-layer coordination between the physical DWDM layer (L1) and network layer (L3) for achieving higher efficiency and lower latency values for live streaming and on demand (VoD) video. Due to the pilot implementation of the presented concept being limited in scale, we use simulations in order to perform proof-of-concept on a sufficiently large environment. Results show that there is a strong correlation between the TTFB and QoE metrics with the former taking on values as low as 75 msec in a national 3-tier network. Ultimately, our aim is to

---

*IT CoNvergence PRACTice (INPRA)*, volume: 1, number: 1, pp. 37-52

\*Corresponding author: Tel: +46-0910-585364

familiarize readers with the field of CDNs and also to help them see how network research especially in the architectural, protocol design, and cross-layer design help bring applications in this field to a quality level acceptable by a large community of users.

**Keywords:** Content Delivery Network (CDN); Peering; IP Transit; Video Streaming; Quality of Experience.

## 1 Introduction

After the number of video-capable mobile network devices grew and services got more personalized, the demand for content delivery in a timely manner in different formats grew as well. Content Delivery Networks in all their forms address those challenges in the area of media streaming. Web proxies preceded CDNs as a static storage and localized mechanisms, whereas CDNs focus on an overlay for multimedia content, intelligent caching, fast video content retrieval and adaptation, and selective customized coding and storage.

With the transition of classical broadcasted services such as Live-TV into Catch-up TV in online media portals, and with the emergence of Hybrid Broadcast Broadband TV (HBBTV) with connected TVs over the Internet, the area of Content Delivery Networking (CDN) continues to boom.

After content used to be stored on a single server, and then backed-up with mirror systems, it got more and more distributed until the topic of CDNs started to be researched and produced systematically dealing with issues such as content caching, fetching, storage, and adaptation. Figure 1 shows a classification of content delivery systems and CDN which is a specific and evolved type in that group.

With the transition of classical broadcasted services such as Live-TV into Catch-up TV in online

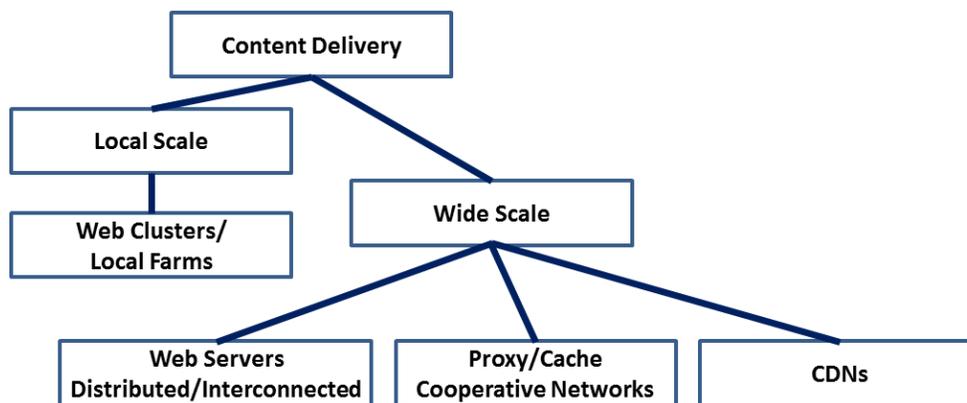


Figure 1: CDNs and Content Delivery in General

media portals, and with the emergence of Hybrid Broadcast Broadband TV (HBBTV) with connected TVs over the Internet, the area of Content Delivery Networking (CDN) continues to boom.

The paradigm shift towards stored and streamed content instead of broadcasting over classical media, as well as the personalization of real-time and multimedia services has been a key factor in shaping the development of content distribution networking. Content generation or acquisition, adaptation, and transport are key aspects in a CDN. Video traffic on a network, especially during a major event such as the soccer world cup where millions of users watch content live online would pose a challenge to CDN providers and this in turn is an aspect for network designers to focus on. Provisioning and dimensioning of a network, including the core and the access part have to take into account the requirements of CDN

traffic.

This article focuses on the technical challenges faced by a Content Delivery Network (CDN) provider in connection with innovative research-oriented ways to overcome them. Video streaming on its own poses performance challenges in IP networks due to the stringent requirements on performance parameters such as delay, jitter, etc. Furthermore, the burst-like pattern with which video traffic is generated causes the demand on bandwidth for video traffic to vary irregularly. A Content Delivery Network (CDN) provider has to cope with varying customer demands, assure acceptable quality of experience (QoE), and run a profitable business model.

Basically a CDN provider has to acquire from a content server the multimedia content it will dis-

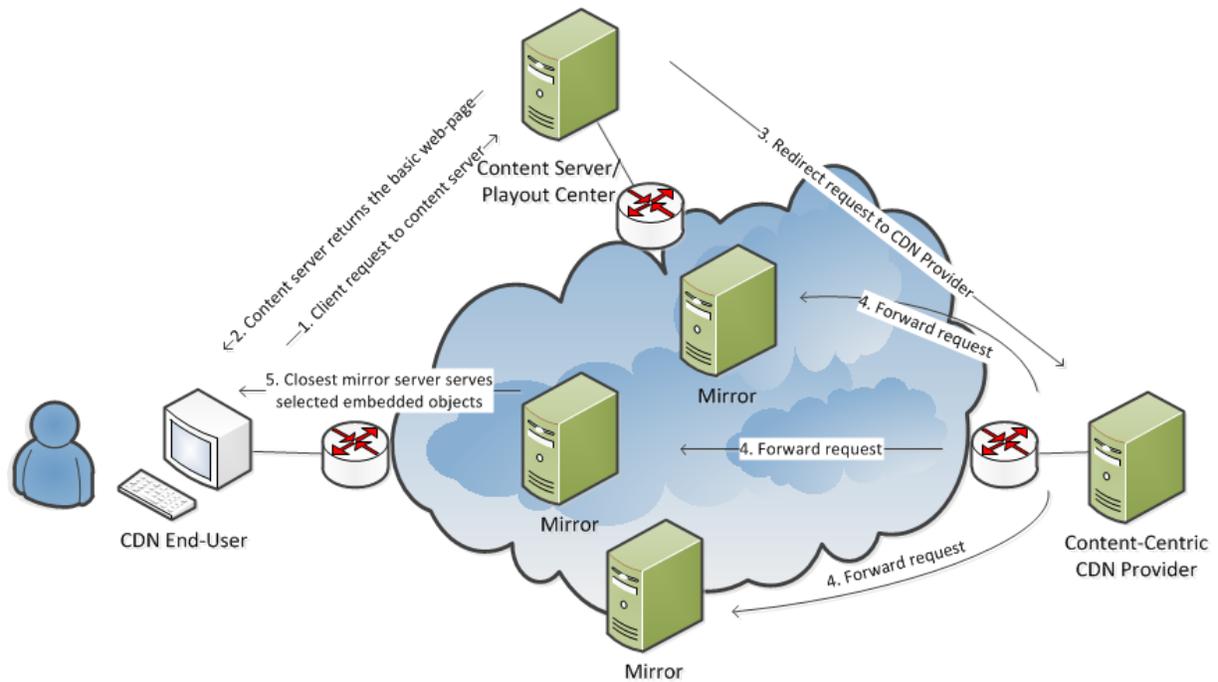


Figure 2: Basic principle of Content Distribution Networks

tribute. Such play-out centers are usually equipped with redundant servers for high reliability. Mirroring other servers down the communication chain is also common to ensure robustness and also allow load balancing at peak transmission moments (e.g. during important events where the demand on CDN content peaks). When clients request multimedia content from their closest contact point, the local caching servers are checked and if the page is not available (due to being pre-fetched and stored in the cache server), then the CDN network switches from short-tail content retrieval to long-tail content retrieval and requests the multimedia stream from the original location namely the play-out center in the backend.

We look at current CDN solutions of both formats academic (research oriented) and commercial. CDNs in the former category act as platforms for innovation in this area whereas CDNs of the latter category have to deliver a high QoE to a maximum number of customer devices with the available finite network resources. A commercial CDN is strongly governed by the underlying business model which tries to best capture the resources-revenues tradeoff. One of the major challenges for a content provider, service provider, or operator to launch a CDN is the initial investment for the basic infrastructure which includes the streaming servers with the content, the proxies, and the caching servers in addition to the network management software which governs the content exchange, fetching, trans-coding and streaming functionalities.

Network design that takes into account user-oriented aspects on the application layer and technical limitations and performance bottlenecks on lower protocol layers is cross-layer design that is specially adapted to the specifics of CDNs. We address this aspect step by step to give the readers an insight into how CDN network design is done.

The remainder of this paper is structured the following way: Section II covers related work, while Section III presents major challenges facing CDN operators. Section IV describes opportunities, while Section V presents our proposed solution. Finally, Section VI concludes the paper and indicates future work.

## 2 Content Distribution Networks in Academia and Industry

With content distribution networks being an important part of several domains and telecommunication services, we provide an extensive overview of the state of the art of CDNs. We cover video streaming platforms, major business model types, commercial CDN operators, and academic CDNs.

### 2.1 Video Streaming Platforms

There are several well-known video streaming platforms which act as content delivery networks. They usually buy large portions of capacity in core and metro networks (nationwide and internationally) and they also operate content servers on which the multimedia CDN content is produced, stored, cached, or processed. Tier-1 ISPs (Internet Service Providers) are typical owners of such platforms.

Examples are Akamai [3], Cogent [5], and Level3 [9]. For those Tier-1 ISPs, the core business is event-based CDN video streaming. Upon the occurrence of events in the areas of sports, politics, or any other areas, those CDN providers are able to accommodate and serve intermediate providers and end-customers with multimedia content because their networks are provisioned and dimensioned accordingly.

Recent studies have shown that the nature of demand for multimedia content has changed in terms of the peak-to-average data rate (PADR) by increasing from a factor of 2.9 to a factor of 6.5 according to a study conducted by the Swedish company Transmode [15]. For instance, a provider which has an average traffic load of 20 Gbps in its backbone per segment averaged over the whole network has to dimension its backbone in such a way to accommodate 6.5 fold for the case when events are to be covered which is around 130 Gbps in this case. This explains the trend for conglomerate and horizontal market growth trend based on expansion and acquisition pursued by Tier-1 ISPs. As the average traffic load grows, the peak traffic load to be covered for events in CDN mode is several fold of that capacity. Therefore, the asset or capital of each ISP that acts as a CDN is the peak capacity level it can accommodate.

### 2.2 Existing Business Models

The most popular existing CDN business models include the pay-per-view or transaction-based model and the flat-rate or all-you-can-watch model. The way content is accessed and the interactions among stakeholders within a CDN eco-system determine the type of underlying business model. The two main types are content-centric CDNs and access-centric CDNs.

Content-centric CDNs are models where content providers pay CDN operators to accelerate their own content through the network to reach end-user devices with high QoE. QoE is the key here because it is seen as the guarantee of retaining customers as opposed to having them switch to providers whose video quality is believed and perceived to be better. So those CDNs are sort-of QoE guarantors from a customer view point. What CDNs in this category do is to employ routing intelligence to guide user requests to the local servers (e.g. CDN caching servers). Akamai [3] is an example of a content-centric

CDN. When multimedia content is successfully tagged and guided throughout the network in terms of request (signaling) and data flow, the chances of achieving a high level of QoE to the customer's satisfaction becomes higher. When the CDN provider pays co-location fees, it achieves its profit and break-even point from re-seller revenues. Figure 3 shows the basic architecture of a content-centric CDN system and the interactions within.

The other type of CDN providers is the access-centric CDN type where the data flow for multimedia

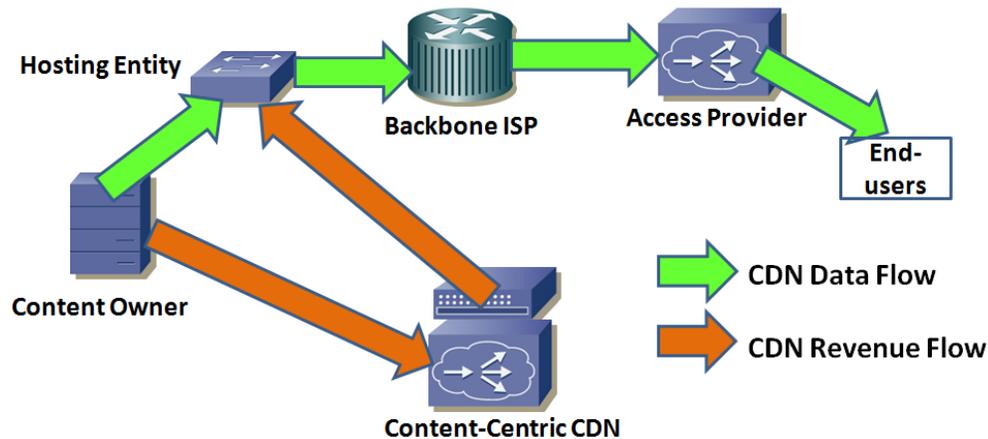


Figure 3: Content-Centric CDN Architecture and Flows

content is identical to its counter-case (the content-centric CDN model), but the revenue flow is from the access provider stakeholder towards the CDN provider. Access providers act as “fat-ISPs” where they provide not only access in terms of connectivity but also access to popular content delivered just as CDNs deliver their content. Subscribers of those fat-ISPs would then have access to popular multimedia content which is supplied by access-centric CDN providers to their respective access providers.

### 2.3 Commercial CDN Operators

This is the category that forms the current and future market and there are several established players such as Netflix [12], Lovefilm [10], MyVideo [11], etc. What characterizes and differentiates commercial CDN operators are their size, their library of content, their mode of operation, and their positioning within their eco-system besides stakeholders in the underlying network architecture. Netflix for instance is known as one of the CDN giants in North America due to its large library of multimedia content and also its ability to serve customers nationwide with a broad range of content. As the volume of multimedia content in the library grows and as the demand for quality of experience (QoE) at high levels increases or at least stays stable, CDN operators are faced by the need to invest in their network infrastructure. This includes leased lines or dark fibers in the backbone, caching and streaming servers, intelligent content fetching mechanisms from the core network, and even hardware and software mechanisms for QoE provisioning and maintenance on the access links. Commercial CDN providers belong to either of the two business models presented in the previous subsection namely: content and access centric CDNs.

### 2.4 Academic CDNs

Academic CDNs are video streaming platforms and content delivery networks which offer caching and streaming services for video content within a research-like environment. Those CDNs allow for experimenting with the purpose of technically improving network performance in terms of QoS and QoE but

they remain far from actually deployable networks for paying end-users within a commercially viable business model.

The most popular academic CDNs include: Globule (Vrije Universiteit Amsterdam) [8], FCAN (Flash Crowds Alleviation Network) [22], CoDeeN (Princeton University) [4], CoralCDN [6], and CO-MODIN (COoperative Media On-Demand on the InterNet) [21].

What academic CDNs have in common is their experimental nature for enhancing research in multimedia content distribution with quantitative measurable goals. Moreover, the deployment of experimental and research CDNs allows reaching architectural changes which provide significant performance boosts. This is necessary because the classical architectural models have started to reach their limitations and cannot cope with the growing volume of individual video stream (with the shift from SD to HD video) and collective volume as well. Several papers in the literature address this issue of the race between capacity growth and network architectural evolution. For instance in [23], Roy et al. introduce a metric called “network-cut exhaustion probability” and based on this parameter they determine upgrade requirements on the network. This constant methodological provisioning and dimensioning is an integral part of the CDN world as well. In [24], machine learning methods are proposed for classifying video traffic based on packet size in order to accommodate video streams in a better way and achieve acceptable performance (QoE) levels.

### 3 Major Challenges Facing CDN Operators

In this section, we briefly cover the key challenges which a CDN operator or provider faces when serving, acquiring or retaining customers as well as in regard to the dynamics and competition faced on the CDN market.

Streaming video, whether in play-out or live mode composes the major part of CDN traffic. This traffic is packed into streams and sold as services as described in the previous section. Subscribing customers are either broadband connection subscribers at home or mobile device owners to whom streaming video is available. One barrier is the acquisition of mobile device (e.g. smart phone) customers to use CDN services; the main challenge here being the cost as well as QoE as discussed below. Achieving and maintaining high QoE levels with the growing traffic volumes and increasing QoS challenges in the core and access networks is a barrier to CDN operator success. Furthermore, the major CDN type which is content-centric as described in Section II in this paper faces a major cost barrier incurred by peering and IP-transit expenses.

#### 3.1 Barriers for the Mobile Video Market

Mobile video traffic is growing with different standards allowing for sufficient bandwidth and terminals capable of trans-coding and buffering as well as supporting a broad range of streaming video formats. However, CDN operators face the challenge of covering up for their costs and achieving profits when it comes to customers with mobile devices because the cost barrier for mobile connectivity is a problem in many countries. Mobile broadband, whether 3GPP LTE (Long Term Evolution) [1], or EDGE (Enhanced Data Rates for Global Evolution) [2], or any other wireless access technology capable of transporting video streams.

In [16], the authors propose variable-rate video coding to adapt to channel conditions and achieve higher bandwidth efficiency. Lowering the transmission rate to adapt to weaker QoS conditions could work within some bounds, but when watching live streaming videos or even playback from an archived multimedia CDN library, such techniques do reach their limitations. This issue has also been addressed in [18] for interactive video streaming where the performance bottlenecks were identified and partly

alleviated. However in [18], various wireless technologies are used for higher bandwidth efficiency on their wireless devices for video streaming, whereas most business models involving CDN subscribers tie the user mostly to one access technology. Even when this is not the case, there has to be a default access technology, in mobile broadband mode in order to transport video streams the mobile device. The current price models where volume-based accounting is employed make CDN video services less attractive due to the price hurdle. On the other hand, the perceived quality or quality of experience (QoE) of video on small mobile terminals (e.g. smart phones) is not always sufficiently high enough to be adequate for CDN video streaming.

Operators try to achieve a balance between unicast and multicast streaming in their network in order to utilize resources (mainly bandwidth) as efficiently as possible and at the same time maintain performance levels for instance for retransmissions or adapting collective rates for a group of streams as demonstrated in [17]. Subscribers to flat-rate wireless broadband face the problem of service quality, since their traffic is not prioritized as “Premium” unless they pay additional charges and certain CDN providers also provide “Premium” content for additional charges. So watching premium content with premium quality (QoE) becomes an expensive venture for users and dims the growth chances of CDN providers on the mobile market. Ongoing research tries to address this challenge.

### 3.2 Achieving Sustainable Quality of Experience Levels

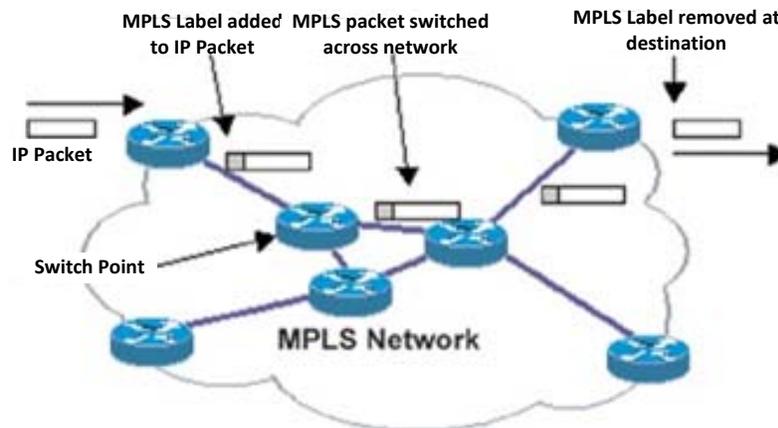


Figure 4: Multi-Protocol Label Switching in Backbone

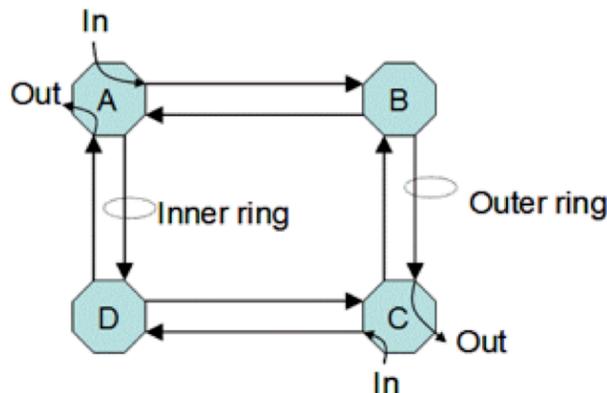


Figure 5: Ethernet Ring Protection in Backbone

**MPLS** (Figure 4): Most carriers and providers, including CDN providers try to maintain control over their own routing paths and thus deploy MPLS (Multi-Protocol Label Switching) and VPLS (Virtual Private LAN Services). When using either a Layer2 VPN service in the network or a L3 service-oriented MPLS service, scalability issues arise, especially when working with video traffic as in the case of CDN. The key to achieving high QoE in CDN networks. More details on this point are provided in Section V in this paper. Scalability issues arise because when using dynamic link metrics, as in line with the nature of CDN content can cause oscillations in the network and load the backbone links with the primary and backup paths almost to their limits [26]. On the other hand, MPLS-compatible schemes such as Profile-Based Routing (PBR) work out fine for CDN networks when the load is relatively low or moderate, but with larger loads, problems occur too. The key point here is that IP routing as well as the MPLS path determination in the default as well as backup cases becomes contra-productive because of the complex routing relationships on intermediate nodes resulting from the video packet fetching from CDN caches and irregular user requests.

**Native Ethernet** (Figure 5): As a response to the aforementioned challenges, we propose in this paper a cross-layer approach for CDN network operators and providers whereby CDN video traffic, whether multicast or unicast is packed into Native Ethernet Packets and transported over a L1 network consisting of leased lines or optical fibers (e.g. dark fiber, gray fiber) with the routing logic built into the protected wavelengths of the Reconfigurable Optical Add-Drop Multiplexers (ROADM)s in the optical network which are collocated with the backbone core network IP routers. This simplified design (with more information in Section V) allows bypassing many Layer3 core routers in the backbone and simply letting video streams exit at target locations, even in multicast mode using the EXMP (Ethernet Muxponder) technology where an example is available from [7]. The value-add is providing operators with connection-oriented transport services over native Ethernet based packet-optical networks, in this particular case for CDN operators.

### 3.3 Peering and IP-Transit

Both peering and IP-transit are key technical and monetary components in an operating CDN network. Those components can form hurdles to the launch or success of a CDN network or video platform because of the incurred costs are substantial in order to perform peering or IP-transit. The costs depend on the volume (e.g. committed, consumed, etc.) of multimedia traffic transferred or exchanged.

- **Peering:** Is when two ISPs or stakeholders make an agreement on a certain port to let the traffic destined to each of them pass through in order to reach its required destinations (end-users). Peering can be either public or private and it allows for packet exchange on a horizontal level (without crossing from a higher to a lower tier ISP or vice versa);
- **IP Transit:** Is the service where Internet traffic, and in this particular case CDN video streaming traffic, is transferred from one network to another, for further reach and connectivity. Seen on a larger scale, it is equivalent to having traffic of a small ISP or networked domain passed over to a Tier-2 and then eventually to a Tier-1 ISP to reach the global platform.

After discussing the major challenges faced by CDN providers and highlighting their significance, we present in the next section a quick insight into opportunities in deploying a CDN network.

Table 1: Important CDN Metrics

Metric	Typical range	Remarks
Packet Size	65-1518 bytes	Very wide range; influences the packets per second (PPS) frequency upon limited capacity. Higher for I-Frames and low for P-Frames and B-Frames
TTFB	50-350 msec	Time to First Byte
QoE MOS	2.81%	Quality of Experience Mean Opinion Score; theoretical range: 1-5; practical range in this model 3.5-4.2
Load	0-100%	It is possible to load the network with more than 100% of its link capacity
Utilization	0-1.0	It reflects the effective usage of link capacity in terms of goodput as opposed to the more generic metric reflecting the load on the network

## 4 Opportunities for CDN Networks

For successful CDN network design, deployment and eventually profitable management, several factors have to come together, and the factors as well as their interactions are a result of ongoing research in this field. In this section, the most important metrics which characterize CDN-based services are outlined and also the organic growth model of CDN networks is introduced and discussed as a strength of such a model.

### 4.1 Important CDN Metrics

One important metric, also present in the web-server world is the so-called "Time To First Byte" or TTFB. This metric indicates the time it takes from the instant a request in a CDN network is initiated until the first CDN multimedia byte is play-ready on the user terminal (playout buffer). Due to the fact that video files have a varying range of length, and since routing over the core network could take place in multi-path mode for better load balancing and network utilization, defining a generic metric which reflects CDN video transport time is practically impossible. Therefore, as a benchmark and as a part of a Service Level Agreement (SLA) between a CDN provider and its customers, the TTFB is taken as an indicator. Another important factor related to the video streams is the average packet size (or the content packet size if the stream is encoded that way). This parameter influences the behavior of the stream over the network and the scalability of QoS mechanisms employed on the individual links. In CDN networks, packet size and stream encoding and bundling are key features for link shaping and network performance modeling in both the access part and the core part. Resource modeling on the access link for accommodating many users and conforming to performance requirements is studied in a sample case study for vehicular networks with a 3/3.5G downlink (UMTS, EDGE [2]) in [19]. In Table 1 below, the metrics characterizing CDN models are summarized and explained.

### 4.2 Organic CDN Growth

Because CDN is the area where content management and QoS mechanisms on a closed domain network meet, the organic growth of such networks is possible. Stakeholders can interact and expand or share responsibility in way which allows expansion. For instance, an infrastructure provider (or operator) can open its network to a content provider to run its content and then both can expand in terms of content and coverage depending on what it is most beneficial to acquire.

The strict distinction between stakeholders is disappearing, and this lowers the price/cost barrier and allows for healthy expansion of CDN networks. For this reason, when establishing a CDN network, the assumption of linear-scale organic growth can be made as a basis for profit projection.

## 5 Proposed Architecture and Model Evaluation

Within the area of CDN networks, there are several algorithms for selecting content sources and distribution or caching points in order to optimally load the network and achieve the best possible video streaming quality on client terminals. Figure 6 shows the typical relative performance of the various CDN algorithms namely: the greedy algorithm, the hot spot algorithm, the tree algorithm, and the random algorithm. Depending on what type of topology is used, different algorithms perform differently. Here a random topology is taken as an example to show how the different algorithms perform compared to each other. The paper however, with its simulations and performance metrics analysis deals with tree-like topologies in the backbone and aggregation network and a random set of distribution points (i.e. end users in a B2C scenario and ISP peering points in a B2B scenario).

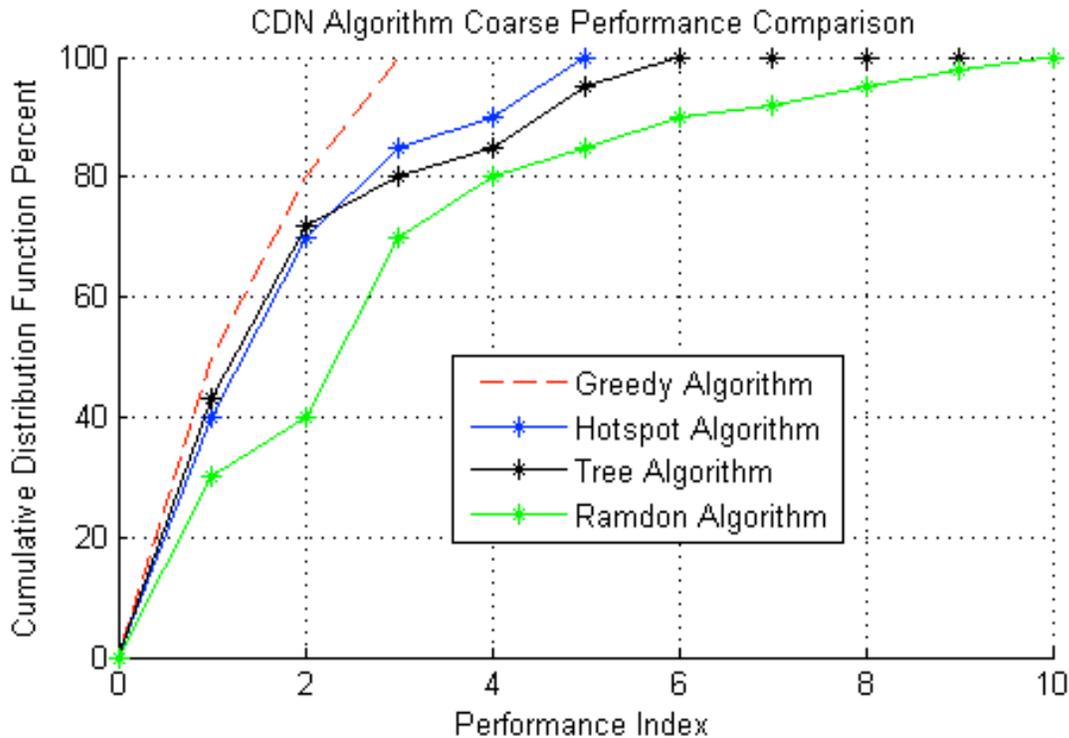


Figure 6: Algorithms in CDNs

We propose in this paper synchronizing the play-out centers which are the CDN multimedia content sources via broadband links such as 10Gbps leased lines (LL) or dedicated wavelengths on a dark fiber. Then the selection of Points of Presence (PoPs) among existing backbone node locations has to take place during network design while focusing on the following factors:

- Geographical factors: e.g. geo-blocking, distance from play-out centers, distribution and density of CDN users, etc.;

- Economic/Business factors: e.g. peering opportunities with last-mile carriers, access to end-customers via peering/IP-transit, etc.

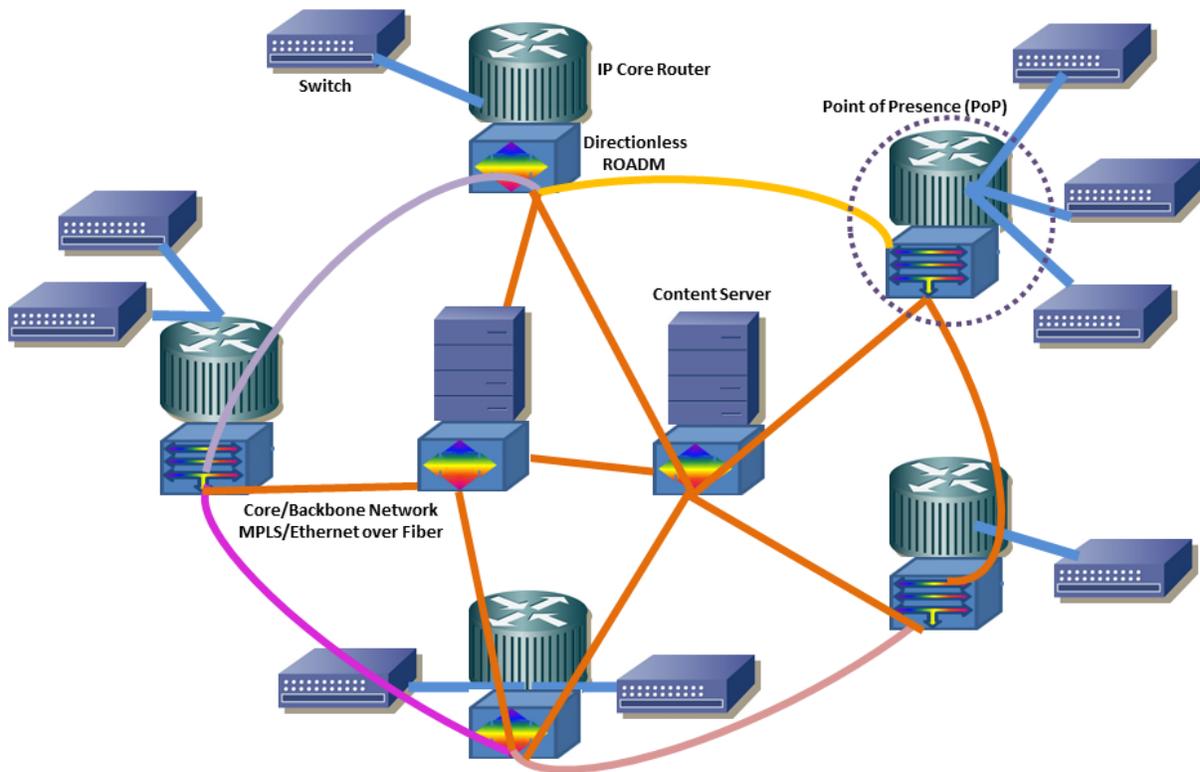


Figure 7: Proposed CDN Architecture

Most operators who own or run a leased backbone deploy MPLS (Multi-Protocol Label Switching) which combines Layer-2 and Layer-3 features with one or two alternative paths for traffic re-routing upon the failure of the default multi-hop path between two nodes.

The issue with MPLS is that it does not scale well with larger bandwidths and increasing loads in the network. With Standard Definition (SD) video traffic with streams in the order of several 100 Mbps data rates, MPLS could still scale fine in the backbone domain. However, with currently available High Definition (HD) streams with a bandwidth ranging between 1.6 and 3 Gbps, MPLS does not scale that well anymore; moreover, when using video in multicast mode, scalability becomes increasingly low.

A typical CDN network as shown in Figure 8 supports both mobile and stationary users in the access part, connected via 3G/LTU and Ethernet/xDSL respectively. The aggregation layer or tier in the network is key for performance management like load balancing and reliability, and for the core part where B2B peering and IP-transit points are installed, an optical backbone (e.g. ROADM based) is recommended to dedicate separate wavelengths to CDN traffic.

To evaluate the benefits of the proposed architecture, we built a simulation model using OPNET IT Guru [14] for the topology in Figure 6 with two play-out centers as sources of multimedia content and five points of presence (PoPs) where multimedia content is fetched from the source and cached besides being streamed into the last mile or access network segment. The data path shown in Figure 3 is the same as the one used in the simulation model. All connections whether between PoPs or play-out centers are 10 Gbps leased lines in bundles of 3 per segment capable of supporting the capacity (with overhead) of 30 Gbps per link. High Definition (HD) streams of video content are used to fill up the capacity of the

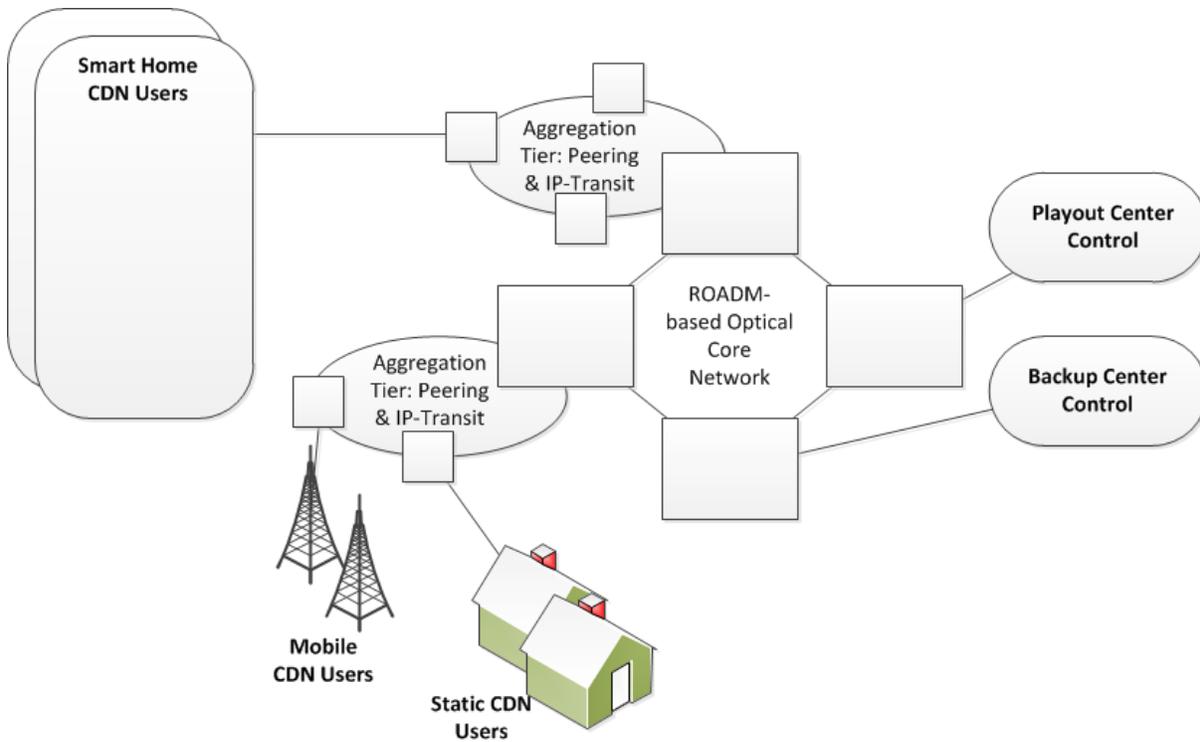


Figure 8: CDN Network High-level Architecture

links uniformly with discrete values of 30, 50, 70, 90 and 110 percent in both modes video on demand (VoD) and Live Streaming. For Live Streaming, short-tail traffic is used whereas for VoD, long-tail traffic is used in the simulation. The video packet size is kept at the maximum of 1518 bytes and the packet rate is throttled and adjusted so as to fill the capacity with the different line rates on the x-axis of the performance graphs (namely 30%, 50%, 70%, 90%, and 110% of the 30 Gbps capacity of the links). On the backbone, two alternatives on top of the optical (L1) equipment are used for performance comparison: Native Ethernet in Ethernet Protected Ring (EPR) mode and Multi-Protocol Label Switching. The main difference between those two underlying backbone mechanisms is that the former uses ITU-T G.8032 automatic protection switching by changing the transport direction on the ring upon a break or failure, whereas the latter uses one or two alternative switched-label paths to the default route upon breaks or extreme load.

As Figure 8 shows, for relatively low loads, EPR and MPLS perform comparably for CDN video traffic in terms of TTFB (the time it takes to put the first video byte or packet on the player of the requesting client). However, for moderate loads of around 50%, EPR outperforms MPLS with 75 milliseconds TTFB times whereas MPLS delivers a TTFB of approximately 90 msec. For typical Service Level Agreement (SLA) bounds of 90 and 180 msec for premium and regular levels respectively, the EPR alternative for both VoD and live streaming remains within the acceptable interval even for higher loads as compared to MPLS which exits the acceptable SLA level already for a load of 80% whereas EPR does that for a loads larger than 92%.

For QoE modeling, the Mean Opinion Score for video is used; this metric is explained in Table 1 and some QoE example models for video are provided in [20]. For this model QoE is taken as the linear weighted sum of TTFB, delay bound, and packet loss rate (PLR) with all coefficients being equal and the result normalized to fit the MOS scale of 1-5. Both backbone technologies (ERP and MPLS) degrade in terms of QoE as the load is increased, however, due to its simpler operating mode, ERP degrades more

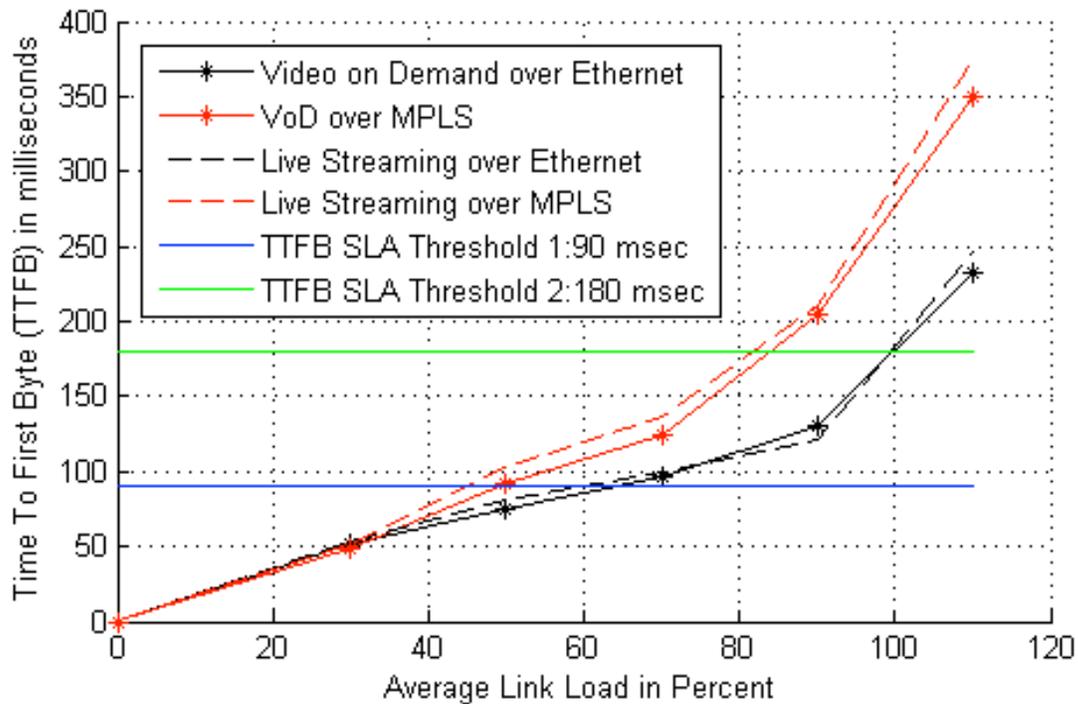


Figure 9: CDN TTFB Performance Ethernet vs. MPLS

slowly in MOS value and is thus able to deliver better perceived quality of CDN video even with loads in the range of 70-80%. An acceptable QoE MOS value is normally not lower than 3.6-3.7. An MPLS network drops below that QoE level and ceases to scale well with HD video streams already at loads around 60% whereas ERP maintains acceptable QoE values for loads up to almost 80% as observed in Figure 10.

## 6 Conclusion and Future Work

In this article we have presented an overview of existing CDN technologies and platforms and identified a current trend which has started to reach its limits when relying on a L2 and L3 based solution using MPLS in the backbone. Several algorithms are used within CDN networks for planning content delivery and other sub-operations such as fetching, coding, storage, and streaming. The choice of the appropriate algorithm or of the appropriate combination of algorithms depends on the density of the users, their distribution, and other system parameters which influence the performance of CDN networks. Moreover, using a cross-layer L1-L3 architecture and native Ethernet for video transport and broadcasting over the backbone delivers a better performance as simulation results show. The time to first byte (TTFB) metric used for web server performance also proves to be a viable metric for performance benchmarking in the CDN domain. The paradigm conveyed by this paper is a performance QoE (quality of experience) driven architectural design for getting more out of the network and sustaining growth and end-user satisfaction. Such an approach, once proven successful could become a trend-setter for further research in the area of multimedia content distribution and CDN networks. Load balancing within CDNs either of the same

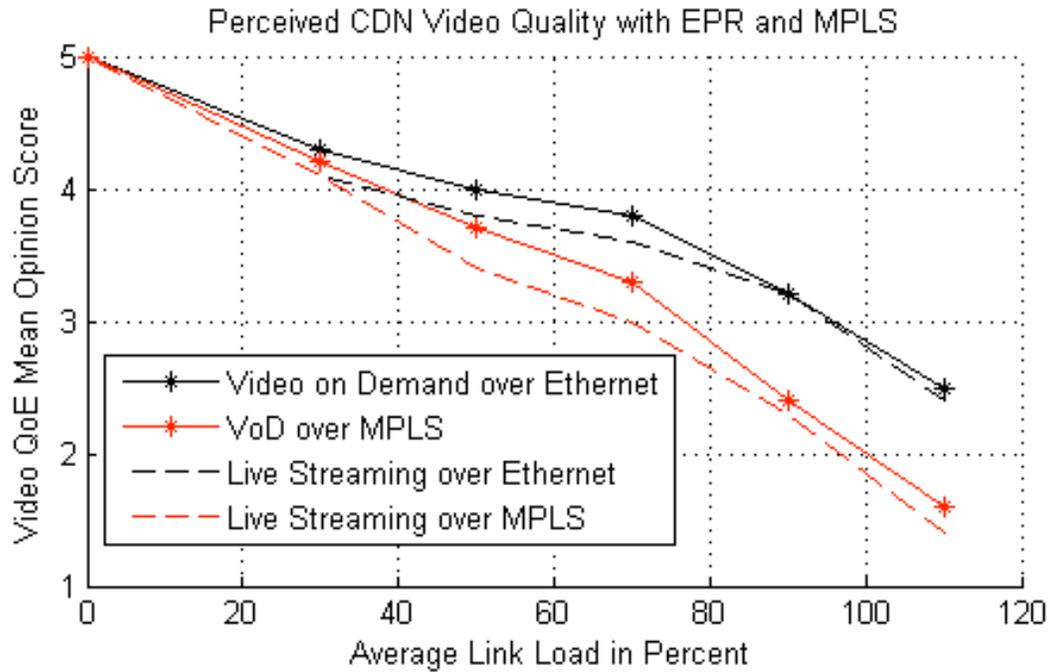


Figure 10: Quality of Experience in CDN with ERP and MPLS

type (e.g. stationary users or mobile users [25]) or of a mixed user base is an ongoing as well as future research item of interest for researchers and CDN providers as well.

## Acknowledgment

This work was partially supported by the NIMO (Nordic Interaction and Mobility Research Platform) project [13] funded by the EU Interreg IVA North program.

## References

- [1] 3GPP (3rd Generation Partnership Project) Long Term Evolution (LTE). <http://www.3gpp.org/LTE>. accessed on March 15, 2013.
- [2] 3GPP EDGE (Enhanced Data rates for Global Evolution). <http://www.3gpp.org/specifications>. accessed on March 15, 2013.
- [3] Akamai CDN provider. <http://www.akamai.com>. accessed on March 15, 2013.
- [4] CoDeeN research CDN for PlanetLab. <http://codeen.cs.princeton.edu>. accessed on March 15, 2013.
- [5] Cogent CDN provider. <http://www.cogent.com>. accessed on March 15, 2013.
- [6] Coral Content Distribution Network research CDN. <http://www.coralcdn.org>. accessed on March 15, 2013.
- [7] Ethernet Muxponder (EMXP) Technology example. <http://www.transmode.com/press-releases/transmodebrings-mpls-tp-to-packet-optical-metro-networking>. accessed on March 15, 2013.
- [8] Globule research CDN. <http://www.globule.org>. accessed on March 15, 2013.
- [9] Level3 CDN provider. <http://www.level3.com>. accessed on March 15, 2013.
- [10] Lovefilm CDN. <http://www.lovefilm.com>. accessed on March 15, 2013.
- [11] MyVideo CDN. <http://www.myvideo.de>. accessed on March 15, 2013.

- [12] Netflix CDN. <http://www.netflix.com>. accessed on March 15, 2013.
- [13] NIMO: Nordic Interaction and Mobility Research Platform. <http://www.nimoproject.org>. accessed on March 15, 2013.
- [14] OPNET IT Guru Academic Edition Simulation Environment. <http://www.opnet.com>. accessed on March 15, 2013.
- [15] Transmode. <http://www.transmode.com>. accessed on March 15, 2013.
- [16] L. Al-Jobouri, M. Fleury, and M. Ghanbari. In *Proc. of the 18th International Conference on Signals, Systems, and Image Processing (IWSSIP'11)*, Sarajevo, Bosnia and Herzegovina. IEEE, June 2011.
- [17] L. Al-Jobouri, M. Fleury, and M. Ghanbari. Multicast and unicast video streaming with rateless channel-coding over broadband wireless. In *Proceedings of the 2012 IEEE Consumer Communications and Networking Conference (CCNC 2012)*, Las Vegas, Nevada, USA, pages 737–741, January 2012.
- [18] K. Andersson, D. Granlund, M. Elkotob, and C. Åhlund. Bandwidth efficient mobility management in heterogeneous wireless networks. In *Proc. of the 7th IEEE conference on Consumer communications and networking conference (CCNC'10)*, Las Vegas, NV, USA. IEEE, January 2010.
- [19] M. Elkotob. Architectural, service and performance modeling for an ims-mbms-based application. In *Proc. of the 2010 IEEE International Conference on Communications (ICC'10)*, Cape Town, South Africa. IEEE, May 2010.
- [20] M. Elkotob, D. Granlund, K. Andersson, and C. Åhlund. Multimedia qoe optimized management using prediction and statistical learning. In *Proc. of the IEEE 35th Conference on Local Computer Networks (LCN'10)*, Denver, Colorado, USA, pages 324–327. IEEE, October 2010.
- [21] G. Fortino, C. Palau, W. Russo, and M. Esteve. The COMODIN System: A CDN-based Platform for Cooperative Media On-Demand on the InterNet. In *Proc. of the 10th International Conference on Distributed Multimedia Systems (DMS'04)*, San Francisco, California, USA. IEEE, September 2004.
- [22] C. Pan, M. Atajanov, M. Hossain, T. Shimokawa, and N. Yoshida. FCAN: Flash Crowds Alleviation Network. In *Proc. of the 2006 ACM symposium on Applied computing (SAC'06)*, Dijon, France, pages 759–765. ACM, April 2006.
- [23] R. Roy and B. Mukherjee. Managing traffic growth in telecom mesh networks. In *Proc. of 17th International Conference on Computer and Communication Networks (ICCCN'08)*, Davis, California, USA, pages 1–6. IEEE, August 2008.
- [24] K. Takeshita, T. Kurosawa, M. Tsujino, M. Iwashita, M. Ichino, and N. Komatsu. Evaluation of HTTP Video Classification Methods Using Flow Group Information. In *Proc. of the 14th International Telecommunications Network Strategy Symposium (NETWORKS'10)*, Warsaw, Poland. IEEE, September 2010.
- [25] S. Wee, J. Apostolopoulos, W. Tan, and S. Roy. Research and design of a mobile streaming media content delivery network. In *Proc. of the 2003 International Conference on Multimedia & Expo (ICME'03)*, Baltimore, Maryland, USA, pages 5–8. IEEE, July 2003.
- [26] S. Yilmaz and I. Matta. On the scalability-performance tradeoffs in mpls and ip routing. In *Proc. of the SPIE ITCOM 2002: Scalability and Traffic Control in IP Networks*, Boston Miami, USA, July 2002.

## Author Biography



**Muslim Elkotob** was born in St. Petersburg, Russian Federation and received his M.Sc. degree in Electrical Engineering with specialization in Telecommunications from Technische Universität München (TUM), Germany in 2003 and his Ph.D. degree in Communications Engineering from Luleå University of Technology, Sweden in 2011. Muslim has worked as a research scientist and co-director of the competence center “Network and Mobility” at DAI Labs, in Berlin, Germany from 2004 to 2007. In 2008 and 2009 he was a research staff member in academia in Sweden. Since 2010, he is with Media Broadcast Germany, a part of the European TDF Group as senior expert and head of strategic network planning at the R&D department. His research interests include network resource management, performance modeling for multimedia applications, and self-healing and efficiency in large scale networks. He is a professional member of IEEE and ACM.



**Karl Andersson** received his M.Sc. degree in Computer Science and Technology from Royal Institute of Technology, Stockholm, Sweden, in 1993. After spending more than 10 years as an IT consultant working mainly with telecom clients he returned to academia and earned his Ph.D. degree from Luleå University of Technology (LTU) in 2010 in Mobile Systems. Following his Ph.D. degree Karl spent six months as a postdoc at Columbia University, New York, USA and was appointed Assistant Professor of Pervasive and Mobile Computing at LTU in 2011. His research interests are centered around mobility management in heterogeneous networking environments and mobile e-services. He is a senior member of IEEE.