

Technical analysis of content placement algorithms for content delivery network in cloud

Suman Jayakumar¹, Prakash Sheelvanthmath², Channappa Baslingappa Akki³

¹Department of Computer Science and Engineering, Visvesvaraya Technological University, Belagavi, India

²University Institute of Engineering, Chandigarh University, Ajitgarh, India

³Department of Computer Science and Engineering, Indian Institute of Information Technology (IIIT) Dharwad, Dharwad, India

Article Info

Article history:

Received Jan 11, 2021

Revised Jun 15, 2021

Accepted Jun 30, 2021

Keywords:

Cloud computing

Content delivery network

Content placement algorithm

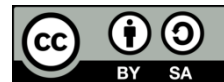
Origin server

Surrogate server

ABSTRACT

Content placement algorithm is an integral part of the cloud-based content delivery network. They are responsible for selecting a precise content to be re-positioned over the surrogate servers distributed over a geographical region. Although various works are being already carried out in this sector, there are loopholes connected to most of the work, which doesn't have much disclosure. It is already known that quality of service, quality of experience, and the cost is always an essential objective targeting to be improved in existing work. Still, there are various other aspects and underlying reasons which are equally important from the design aspect. Therefore, this paper contributes towards reviewing the existing approaches of content placement algorithm over cloud-based content delivery network targeting to explore open-end re-search issues.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Suman Jayakumar

Department of Computer Science and Engineering, Visvesvaraya Technological University

Belagavi, India

Email: jayakumarsuman@gmail.com

1. INTRODUCTION

With the advancement of cloud computing and its associated utility, there is an increased proliferation of resource leasing to construct a content delivery network [1]. Owing to the limited features of the content delivery network, the recent trend of work is majorly towards migrating to the cloud-based content delivery network which is mainly due to magnified traffic over large datacenters. The benefit of migrating to the cloud-based content delivery network is security, reliability, elasticity, flexibility, and scalability. Apart from this, the operational expenditure is significantly controlled by the cloud-based content delivery network, which leads to maximized adoption of both small and large-scale businesses with higher agility. The content management subsystem system focuses on two aspects, viz. to perform placement of surrogate server and what to select for content and to deliver them. The content placement selection is further carried out by two standard approaches of push-based and pull-based strategy [2]. The push-based content placement strategy is mainly cooperative in its form, while the pull-based strategy is again classified into cooperative and non-cooperative form. The concept is purely based on obtaining the contents by surrogate servers from the origin server. The content placement algorithm's effectiveness depends upon the cost factor by positioning the precise contents over the correct surrogate servers. The content placement algorithm must yield contents with the maximized resolution, which is also prone to unpredictability with reduced delay. At present, there are various content placement algorithms developed to be working over a cloud-based content delivery network.

This paper discusses various technical attributes and aspects to understand the effectiveness of existing content placement algorithms over the cloud-based content delivery network. The paper's organization is as follows: Section 1 discusses the criticality and significance of the content placement algorithm, followed by discusses the research problems and briefing of the proposed solution, while section 2 illustrates the proposed analytical model. Section 3 discusses simulation outcomes while section 4 summarizes the paper.

At present, there are three types of contents, viz. static and dynamic contents [3]. There are two types of content in the cloud-based content delivery network, i.e., data and metadata [4]. The conventional content delivery network is designed only for static, while a cloud-based content delivery network is designed both static and dynamic. The majority of the present era contents are user-generated and are essential multimedia files from the social network [5], [6]. They are immensely large in shape. At present, four essential factors are required to ensure its incorporation for the development of the content placement algorithm (CPA) viz: i) fault tolerance, ii) consistency, iii) latency, and iv) cost [7]. From a fault tolerance viewpoint, content availability is the prime aspect. From a consistency viewpoint, replica number and frequency of updating replica as the major objectives. The major components are round trip time, end-to-end delay, throughput, hop count, jitter, and geodesic distance from the latency viewpoint. The essential components are storage, bandwidth, and resource allocation [8]. Hence, these are mandatory to possess as an essential characteristic while developing such an algorithm. However, existing approaches don't have such inclusion in their design aspect, and hence they do not have full-fledged coverage of the essential problems. One essential thing to the node is that existing models do not consider any form of underlying networking system and hence fail to offer the claims of fault tolerance mainly, which should always be the core target. So, the content placement algorithm's significance is only high if all these four essential factors are retained in any design or to the maximum extent. Therefore, with content placement algorithm enhancement over the cloud-based content delivery network, upcoming applications cannot be supported. There is a need for more research work in this area. They demand more consistency, content availability, and fault tolerance features while working with the cloud-based content delivery network. Therefore, an effective and optimal design of content placement algorithm is required to offer a better quality of service and quality of experience for any service relayed to end-users in the cloud environment.

Consider that there are a group of origin servers O as well as a group of surrogate server S in such a way that there is a higher number of origin servers O compared to the surrogate servers S . Considering such system consists of N number of contents, graph G can be formulated as $G=(V, E)$ where vertices V is union operation between O and S . The edge E is considered as directional communication link that connect server Φ_i with another server Φ_j .

The capacity of storage Cap_i is considered for all surrogate servers such that the value of i resides between 1 and $|S|$. The channel capacity of μ_{ij} is considered for all the links between the server with each data bearing the cost of τ_{ij} and duration of communication $d(e_{ij})$, active time of channel $a(e_{ij})$, and passive time of channel $p(e_{ij})$. All the content c_m is considered to possess a specific type t_m , size α_m , similarity factor β_{mn} obtained from different contents. Other parameters considered for modelling are an index of popularity γ_m , rate of refreshing r_m , metric for quality-of-service η_m . For all the contents c_m , every modelling considers formulating rate of request r_m , and cost for positioning the content c_m over a surrogate server is represented as ψ_{im} along with consideration of the rate of failure f_{im} . The prime problem will be to explore an optimal strategy for content placement considering that a precise number of replicated version of all the contents c_j along with their respective positioning over the surrogate serves in S with the condition that all the contents c_j should cater up the metric for quality of service η_j . Apart from this, it must also minimize the cost associated with the delivery and hosting of the contents and increase the utilization of the bandwidth and quality of service for all the end-users. This should be ascertained considering consistency, availability, and latency. Figure 1 highlights the problem of content placement, which shows a single origin server connected with the surrogate server that is distributed geographically using cloud-based content placement. The variable N is used to represent the content placement group over the origin server, which is being managed externally by the service provider of a cloud-based content delivery network. The figure also highlights the end-users' demands associated with the contents over different forms of surrogate servers. The prime target is to efficiently replicate the content and position the content over the surrogate server to offer seamless services with higher service quality.

This section discusses the existing approaches of content placement algorithms (CPA) applied over the cloud-based content delivery network (CCDN). It should be known that existing approaches of CPA mainly target controlling the operational cost and enhance the quality-of-service Wu [9], Jin *et al.* [10]. It also focuses on catering up the content popularity that is dynamically changing and understanding the access pattern of the contents associated with the end-user. The current development process in CPA also associates with the operational cost, which is connected with resource consumption over the cloud where the primary

entities are cost associated with content storage over a surrogate server and bandwidth consumption to extract the contents finally followed by updating operation Jia *et al.* [11] and Herbaut *et al.* [12]. Usually, it is found that the service provider of CCDN incurs this cost with the participation of third parties' cloud infrastructure. The studies in such direction are mainly meant for the reduction of such operational cost along with the assurance of the quality of service for end-users Frangoudis *et al.* [13] and Xie *et al.* [14]. Existing approaches towards the quality of service associated with CPA is mainly confined to either soft guarantee or hard guarantee for some specific metrics in quality of service Sung *et al.* [15], Xuan *et al.* [16]. There is a various representation of metrics associated with quality of service in the study of CPA associated with CCDN, e.g., round-trip time, hop count, jitter in delay, geodesic distance, end-to-end delay, and latency [17]. Existing studies have also noticed an inclination of the area of interest from the quality of service to the quality of experience as seen in Casas' work [18]. However, such categories of work are usually subjective or highly user-centric. However, there are also existing studies that claim that adherence towards such quality of experience is quite challenging, as reported in the work of Casas *et al.* [18] and Seufert *et al.* [19]. Hence, it was found that the quality of the service is not equivalent to the quality of experience owing to the adoption of different dependable parameters to meet the satisfaction of users concerning the services of CCDN.

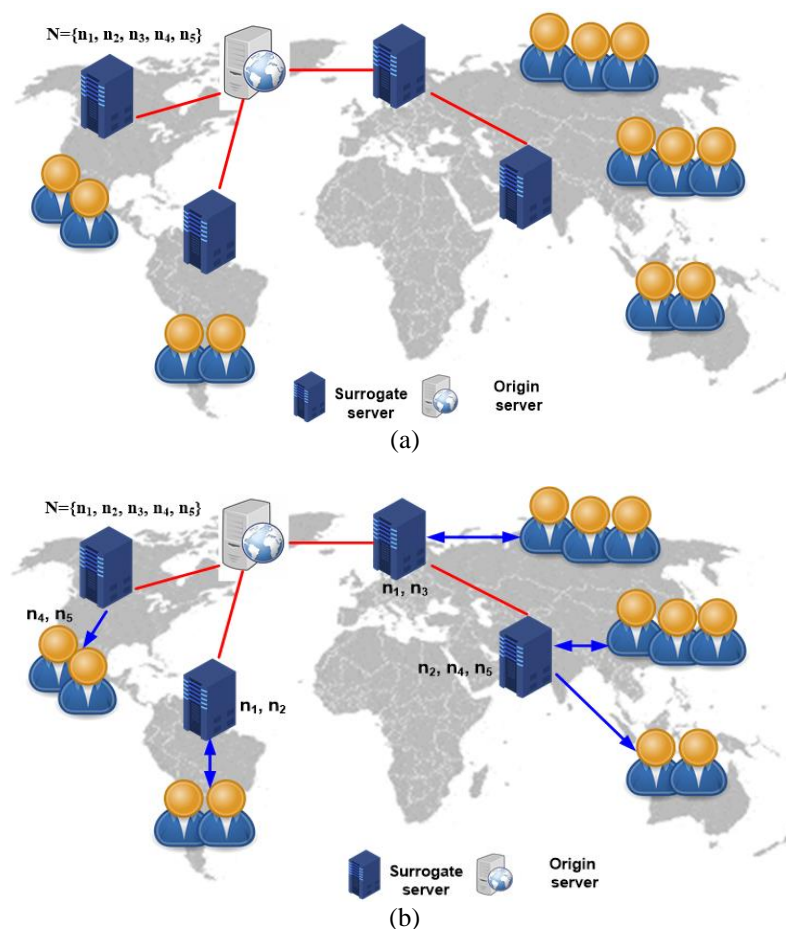


Figure 1. Associated problems of CP and the possibility of a solution: (a) problem scenario and (b) possibility of solution

Existing approaches of CPA are also focused on achieving an effective allocation of dynamic resources as well as better utilization of resources Padmavathi *et al.* [20] and Zhang *et al.* [21]. However, a closer look into such strategies towards resource-based approaches found that they are used to enhance the utilization of the resources that are already leased, and it doesn't perform any further leased resources. Although the term resource utilization is closely linked with the cost of operation, its usage differs from only reducing the cost of allocation of resources. Hence, both are not equivalent terms Liao *et al.* [22] and Yong [23]. Similarly, utilization of resources represents the transmission of maximum contents through the equivalent adopted surrogated server before adopting novel cloud resources. There is another similar form of

the term called resource provisioning in CPA over CCDN. This is related to adopt or release varied resources to cater to the dynamics of end-user demands Baccour *et al.* [24]. The cloud elasticity is leveraged using such an approach, and it has a direct contribution towards constructing CPA over CCDN. Therefore, these studies show certain standards for developing an efficient architecture of the CPA-based model, which are a cost of operation, quality of service, utilization of resources, and provisioning of resources as shown in Figure 2. There are various works towards developing a CPA in CCDN, mainly using a centralized greedy approach or distributed greedy approach. Existing studies also show that heuristics, non-convex optimization, Lyapunov optimization, and depth-first search are also used to implement the content's flow using either a push-based approach or pull-based approach as shown in Figure 3.

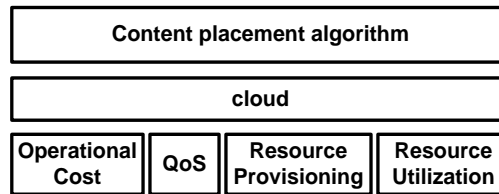


Figure 2. Dependable factors for efficient design of the CPA approach

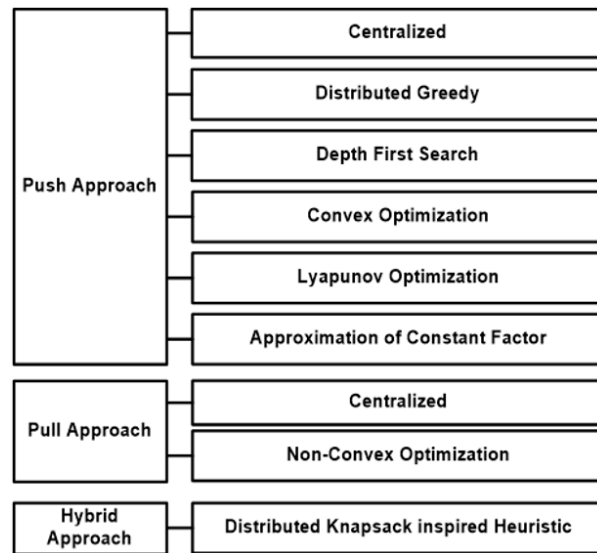


Figure 3. Taxonomies of CPA in CCDN

2. PROPOSED METHOD

From the Figure 3, it can be seen that the push and pull-based approach are the prime strategies for developing the CPA over CCDN. Following are the research work in both approaches:

2.1. Pull-based approach

This approach is mainly developed considering caching to ensure higher availability of contents and control the access delay of the contents. The entire operation is carried out on-demand, fetching the contents and targets for the surrogate server's optimal usage. The caching operation is carried out either by an en-route process, hierarchical process, and hybrid manner. In the process of en-route caching, the initial surrogate server manages the request of the end-user considering the route to the origin server from the end-user [25]-[29]. In hierarchical caching, the hierarchical levels are used to propagate the request until it meets the condition to be satisfied [30], [31]. Both approaches can be combined together to form a hybrid process that targets the availability of contents and minimizes access duration of contents [32]. Another pull-based approach called cooperative caching, which mainly orients about positioning the contents, includes performing a certain standard operation of consistency management, searching, and placement [33]. It was also seen that conventional CCDN deploys the caching for CPA development using the least recently used

replacement policy owing to its non-sophisticate design and implementation. There is also research work that uses the time-to-live factor for the development of replacement policy in caching. This is mainly used for the flexible management of CCDN for controlling losing contents over the network and the cost of operation. The concept of time-to-live-based caching is presented by Carlsson *et al.* [34] with a target to enhance the cost of cache storage. It is found capable of catering to both local and global requests. Hu *et al.* [35] have created a dynamic algorithm to reduce the storage cost and replication, and bandwidth. However, the study's outcome is found to offer a trade-off between contents service from either the CCDN node or source node. The issue of content placement problem was also studied by Jin *et al.* [36], where the aim was to reduce the bandwidth and storage cost. Adoption of the centralized system with enhanced DFS is investigated by Wang *et al.* [37] meant for cost-controlling resources. Guan and Choi [38] have implemented a technique using delay/latency reduction targeting towards increasing the cost of forwarding contents towards the surrogate server. The outcome shows a better degree of cooperation among the surrogate server. Rappaport and Raz [39] have presented a technique considering the constraint of the surrogate server's capacity to utilize capacity. Katsalis *et al.* [40] have presented a holistic method for minimizing operational cost considering content size variability. Another cost control approach was presented by Hu *et al.* [41] in the form of caching and differential provisioning using a greedy approach. The outcome shows significant supportability of service delivery on multi-path. The work of Chen *et al.* [42] has used pre-allocated greedy requests for controlling the cost of access control considering the constraint of quality of service based on distance. The outcome is found to adapt to all the patterns of request dynamically.

2.2. Push-based approach

Using this content delivery approach, the existing schemes perform the end-users allocation to the specific surrogate server that can cater to processing its request. In this approach, the mapping operation is carried out randomly. The input to the system is basically a group of surrogate servers to carry out content delivery. The existing system uses the prioritization of the surrogate server for a better quality of service. End-users' assignment is carried out to the surrogate server considering the distance factor based on geodesic distance, delay, hop count [43]. The existing system has also witnessed weight-based prioritization of surrogate servers. The ratio of channel capacity and storage unit is important for the surrogate server to optimize the storage cost and content delivery cost [37], [41]. The surrogate server's prioritization is carried out using path metrics parameters to minimize the cost of content delivery and facilitate a better quality of service [44]. The problem associated with assigning the surrogates can be mitigated by prioritizing them. Existing approaches have also witnessed dynamic techniques to offer better adaptability by re-provisioning resources [37], [35]. They are mainly meant for CCDN. The operational cost is also found to be controlled using a hybrid approach (push and pull) for CPA in CCDN with an idea of the implementation toward achieving a better quality of service [45], [46]. Apart from this, it should be noted that the assignment of surrogate server essentially demands centralized information; however, it was found that if the CPA techniques are designed in a distributed fashion, then they offer more resilient and fault-tolerant operation. Hence, most of the existing approaches don't support a long-term minimization of the operational cost after the contents are forwarded outside of the cloud.

3. STUDY OUTCOME

The prior section has discussed various standard techniques and methods for CPA over CCDN to adopt various methodologies towards addressing various problems. However, the majority of the problems are connected to cost control. Hence, it must have more insights into the core approaches and components used for design purposes in the existing system. A closer look into Figure 4 shows that the adoption of components differs for both the conventional approaches that essentially target cost minimization of content placement algorithm while working on CCDN. However, there are certain loopholes in the design methodology, which are required to be highlighted. Following are the observation being carried out while reviewing existing approaches:

- Resource provisioning is a significant contributor towards controlling operational cost irrespective of its form of short-termed or dynamic nature. In existing ideas about operational cost controlling-based approaches, the studies target studying the influence of cost over quality of service and investigating the impact of replicates over service quality on end-user. The open-end problem is mainly in assessing the influence of quality of experience over the quality of service associated with CPA design over CCDN.
- There is some dedicated work towards both push and pull approaches in designing CPA; however, they cannot be considered suitable for upcoming cloud system applications. The major open-end problems are assessing the impact of caching over the quality of service, understanding the impact of content prepositioning on the cost of operation, and investigating the cumulative gain from the existing push/pull method over content placement.

- Another significant factor to observe is the content popularity, mining content, mobile content, and consistency/availability. Owing to the variable and fluctuating degree of content sensitivity, the existing content placement algorithm is not equipped to handle this sensitivity problem. Another open-end problem is associated with the content delivery to the mobile equipment where both quality of service and experienced cannot be retained. Availability, as well as consistency, has always been encountering the trade-off concerning the dynamic behavior of the user.
- None of the existing approaches towards CCDN has ever considered cost-effective redirection of request, directly affecting the quality of experience. It was also seen that there is a higher cost involvement of bandwidth and storage when redirection takes place between redirector and surrogate server.
- Existing approaches are found not suitable for the upcoming application. There is a growing interest in fog computing, which demands using a surrogate server with a non-conventional design. Moreover, there is also a need to test it for its supportability over the heterogeneous network, missing in existing studies.
- Finally, it should be noted that routes used for content delivery very often make use of different forms of access networks. At present, the studies carried out over the cloud environment have not considered this. It is essential to analyze the issues associated with controlling operation cost while using the heterogeneous network.

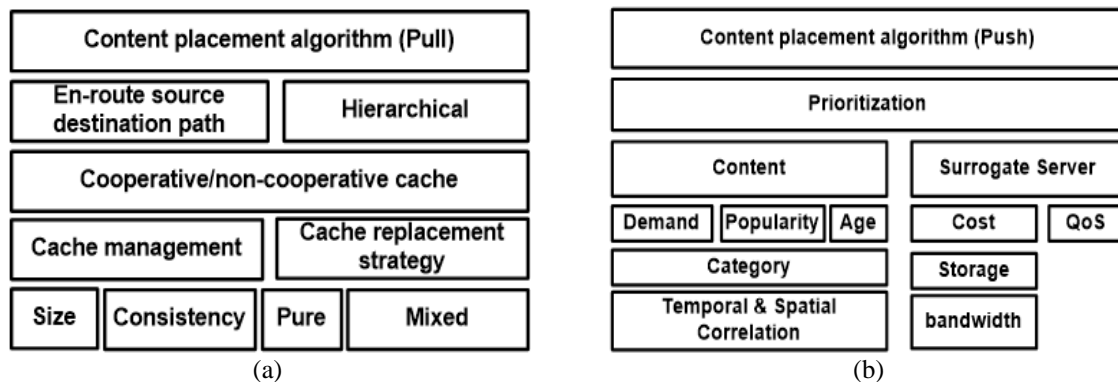


Figure 4. Core component adoption in existing content placement algorithms in CCDN: (a) essential components adopted for pull-based content placement algorithms, (b) essential components adopted for push-based content placement algorithms

4. CONCLUSION

This paper has discussed the existing approaches to content placement algorithms in the cloud content delivery network. The paper contributes towards discussing the design aspect of the existing approach and connects with the level of affectivity towards content dynamics. It is also discussed that an effective design of content placement algorithm will need better utilization of resources, quality of service maximization, minimizing operational cost, and performs resource provisioning. The future research work could be carried out towards addressing the existing problems by introducing a novel topology of content placement rather than reusing the legacy version in cloud CDN. The core target will be to develop a dynamic framework for CCDN where the focus will be towards optimizing the performance of quality of service, latency, traffic management, storage optimization.

REFERENCES




- [1] D. Sarkar, N. Rakesh and K. K. Mishra, "Content delivery networks: Insights and recent advancement," *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Wagnaghat, 2016, pp. 1-5, doi: 10.1109/PDGC.2016.7913113.
- [2] Q. Fan, H. Yin, Z. Jiang, H. Huang, Y. Luo, and X. Zhang, "Adaptive content management for UGC video delivery in mobile internet era," *Hindawi-Mobile Information*, vol. 16, 2016, doi: 10.1155/2016/3624860.
- [3] Z. Lu, J. Wu, Y. Huang, L. Chen, and D. Deng, "CPDID: A novel CDN-P2P dynamic interactive delivery scheme for live streaming," *2012 IEEE 18th International Conference on Parallel Distributed Systems*, Singapore, 2012, pp. 299-306, doi: 10.1109/ICPADS.2012.49.
- [4] S. Fati, P. Sumari, and C. W. Onn, "Content-aware resource allocation model for IPTV delivery networks," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 1, pp. 369-385, 2019, doi: 10.11591/ijece.v9i1.pp369-385.
- [5] M. Hussaini, S. A. Nor, and A. Ahmad, "Producer mobility support schemes for named data networking: A survey," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, pp. 5432-5442, 2018, doi: 10.11591/ijece.v8i6.pp.5432-5442.

- [6] Cisco, "Cisco visual networking index: Forecast and methodology 2014–2019 white paper," *Cisco, Tech. Rep.*, 2015. [Online]. Available: http://s2.q4cdn.com/230918913/_files/doc_downloads/report_2014/white_paper_c11-481360.pdf
- [7] J. Sahoo, M. A. Salahuddin, R. Glietho, H. Elbiaze, and W. Ajib, "A survey on replica server placement algorithms for content delivery networks," in *IEEE Communications Surveys and Tutorials*, vol. 19, no. 2, pp. 1002–1026, 2017, doi: 10.1109/COMST.2016.2626384.
- [8] L. Sun, M. Ma, W. Hu, H. Pang, and Z. Wang, "Beyond 1 million nodes: A crowd sourced video content delivery network," in *IEEE MultiMedia*, vol. 24, no. 3, pp. 54–63, 2017, doi: 10.1109/MMUL.2017.3051522.
- [9] X. Wu, "Data sets replicas placements strategy from cost-effective view in the cloud," *Hindawi Publishing Corporation*, 2016, doi: 10.1155/2016/1496714.
- [10] Y. Jin, Y. Wen, and K. Guan, "Toward cost-efficient content placement in media cloud: Modeling and analysis," in *IEEE Transactions on Multimedia*, vol. 18, no. 5, pp. 807–819, May 2016, doi: 10.1109/TMM.2016.2537199.
- [11] Q. Jia, R. Xie, T. Huang, J. Liu, and Y. Liu, "The collaboration for content delivery and network infrastructures: A survey," in *IEEE Access*, vol. 5, pp. 18088–18106, 2017, doi: 10.1109/ACCESS.2017.2715824.
- [12] N. Herbaut, D. Négru, Y. Chen, P. A. Frangoudis, and A. Ksentini, "Content delivery networks as a virtual network function: A win-win ISP-CDN collaboration," In *IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6, doi: 10.1109/GLOCOM.2016.7841689
- [13] P. A. Frangoudis, L. Yala, and A. Ksentini, "CDN-As-a-service provision over a telecom operator's cloud," in *IEEE Transactions on Network and Service Management*, vol. 14, no. 3, pp. 702–716, Sep. 2017, doi: 10.1109/TNSM.2017.2710300.
- [14] G. Xie, Z. Li, M. A. Kaafar, and Q. Wu, "Access types effect on internet video services and its implications on CDN caching," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 5, pp. 1183–1196, May 2018, doi: 10.1109/TCSVT.2017.2652487.
- [15] J. Sung, M. Kim, K. Lim, and J. K. Rhee, "Efficient cache placement strategy in two-tier wireless content delivery network," in *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1163–1174, Jun. 2016, doi: 10.1109/TMM.2016.2543658.
- [16] L. D. Xuan, M. Seufert, F. Wamser, and P. T. Gia, "QoE aware placement of content in edge networks on the example of a photo album cloud service," *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*, 2016, pp. 443–448, doi: 10.1109/CCE.2016.7562677.
- [17] B. Briscoe *et al.*, "Reducing internet latency: A survey of techniques and their merits," *IEEE Communications Surveys and Tutorials*, vol. 18, 2014, doi: 10.1109/COMST.2014.2375213.
- [18] P. Casas, A. D'Alconzo, P. Fiadino, A. Bär, A. Finamore, and T. Zseby, "When YouTube does not work: Analysis of QoE-relevant degradation in Google CDN traffic," *IEEE Trans. Netw. Service Manage.*, vol. 11, no. 4, pp. 441–457, Dec. 2014, doi: 10.1109/TNSM.2014.2377691.
- [19] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hojfeld, and P. Tran-Gia, "A survey on quality of experience of HTTP adaptive streaming," *IEEE Commun. Surveys Tuts*, vol. 17, no. 1, pp. 469–492, 2015, doi: 10.1109/COMST.2014.2360940.
- [20] S. Padmavathi, P. K. Soniha, N. Soundarya, and S. Srimathi, "Dynamic resource provisioning and monitoring for cloud computing," *2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)*, Srivilliputhur, 2017, pp. 1–6, doi: 10.1109/ITCOSP.2017.8303132.
- [21] Z. Zhang, H. Xi and C. Song, "Dynamic optimal resource provisioning for VoD services under Amazon EC2's pricing models," *Proceedings of the 33rd Chinese Control Conference*, Nanjing, 2014, pp. 5527–5532, doi: 10.1109/ChiCC.2014.6895884.
- [22] J. Liao, J. Wang, B. Wu, and W. Wu, "Toward a multiplane framework of NGSON: A required guideline to achieve pervasive services and efficient resource utilization," in *IEEE Communications Magazine*, vol. 50, no. 1, pp. 90–97, Jan. 2012, doi: 10.1109/MCOM.2012.6122537.
- [23] Z. Yong, "Dynamical resource allocation controller research based on flow aware in content distribution networks," *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, Guangzhou, 2017, pp. 682–684, doi: 10.1109/CSE-EUC.2017.129.
- [24] E. Baccour, A. Erbad, A. Mohamed, M. Guizani, and M. Hamdi, "CE-D2D: Collaborative and popularity-aware proactive chunks caching in edge networks," *2020 International Wireless Communications and Mobile Computing (IWCMC)*, Limassol, Cyprus, 2020, pp. 1770–1776, doi: 10.1109/IWCMC48107.2020.9148355.
- [25] A. Gharaibeh, A. Khreishah, I. Khalil, and J. Wu, "Distributed online en-route caching," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 12, pp. 3455–3468, Dec. 2016, doi: 10.1109/TPDS.2016.2547396.
- [26] F. Berg, F. Dürr, and K. Roethermel, "Increasing the efficiency of code offloading through remote-side caching," *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Abu Dhabi, 2015, pp. 573–580, doi: 10.1109/WiMOB.2015.7348013.
- [27] J. Sahoo and R. Glietho, "Greedy heuristic for replica server placement in Cloud based Content Delivery Networks," *2016 IEEE Symposium on Computers and Communication (ISCC)*, Messina, 2016, pp. 302–309, doi: 10.1109/ISCC.2016.7543758.
- [28] X. Tang, H. Chi, and S. T. Chanson, "Optimal replica placement under TTL-based consistency," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 3, pp. 351–363, Mar. 2007, doi: 10.1109/TPDS.2007.47.
- [29] M. Ha and D. Kim, "On-demand cache placement protocol for content delivery sensor networks," *2017 International Conf. on Computing, Networking and Communications (ICNC)*, Santa Clara, CA, 2017, pp. 207–216, doi: 10.1109/ICNC.2017.7876128.
- [30] A. Araldo, M. Mangili, F. Martignon, and D. Rossi, "Cost-aware caching: Optimizing cache provisioning and object placement in ICN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 1108–1113, doi: 10.1109/GLOCOM.2014.7036957.
- [31] Y. Desmouceaux, M. Enguehard, V. Nguyen, P. Pfister, W. Shao, and É. Vyncke, "A content-aware data-plane for efficient and scalable video delivery," *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, Arlington, VA, USA, 2019, pp. 10–18.
- [32] M. Krishna, D.C. Bala, D. Gomes, J. P. Barraca, and H. Moreira, "Cloud-Centric Networking Using Virtualization and Resource Sharing," *In User-Centric and Information-Centric Networking and Services*, CRC Press, pp. 229–266, 2019.
- [33] P. Vashisht, R. Kumar, and A. Sharma, "Efficient dynamic replication algorithm using agent for data grid," *Hindawi-The Scientific World Journal*, 2014, doi: 10.1155/2014/767016.
- [34] N. Carlsson, D. Eager, A. Gopinathan, and Z. Li, "Caching and optimized request routing in cloud-based content delivery systems," *J. Perform. Eval.*, vol. 79, pp. 38–55, Sep. 2014, doi: 10.1016/j.peva.2014.07.003.
- [35] H. Hu *et al.*, "Community based effective social video contents placement in cloud centric CDN network," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Chengdu, China, Jul. 2014, pp. 1–6, doi: 10.1109/ICME.2014.6890134.
- [36] Y. Jin, Y. Wen, K. Guan, D. Kilper, and H. Xie, "Toward monetary cost effective content placement in cloud centric media network," in *Proc. IEEE International Conference Multimedia Expo (ICME)*, San Jose, CA, USA, 2013, pp. 1–6, doi: 10.1109/ICME.2013.6607582.




- [37] F. Wang, J. Liu, M. Chen, and H. Wang, "Migration towards cloud assisted live media streaming," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 272-282, Feb. 2014, doi: 10.1109/TNET.2014.2362541.
- [38] X. Guan and B.-Y. Choi, "Push or pull? Toward optimal content delivery using cloud storage," *J. Netw. Comput. Appl.*, vol. 40, pp. 234-243, Apr. 2014, doi: 10.1016/j.jnca.2013.09.003.
- [39] A. Rappaport and D. Raz, "Update aware replica placement," in *Proc. Int. Conf. Netw. Service Manag. (CNSM)*, Zürich, Switzerland, Oct. 2013, pp. 92-99, doi: 10.1109/CNSM.2013.6727814.
- [40] K. Katsalis, V. Sourlas, T. Korakis, and L. Tassioulas, "A cloud-based content replication framework over multi-domain environments," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 2926-2931, doi: 10.1109/ICC.2014.6883769.
- [41] M. Hu, J. Luo, Y. Wang, and B. Veeravalli, "Practical resource provisioning and caching with dynamic resilience for cloud-based content distribution networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2169-2179, Aug. 2014, doi: 10.1109/TPDS.2013.287.
- [42] F. Chen, K. Guo, J. Lin, and T. La Porta, "Intra-cloud lightning: Building CDNs in the cloud," in *Proc. IEEE INFOCOM*, Orlando, FL, USA, 2012, pp. 433-441, doi: 10.1109/INFOCOM.2012.6195782.
- [43] C. Papagianni, A. Leivadreas, and S. Papavassiliou, "A cloud-oriented content delivery network paradigm: Modeling and assessment," *IEEE Trans. Dependable Secure Computing*, vol. 10, no. 5, pp. 287-300, Sep. 2013, doi: 10.1109/TDSC.2013.12.
- [44] M. Salahuddin, H. Elbiaze, W. Ajib, and R. Glitho, "Social network analysis inspired content placement with QoS in cloud-based content delivery networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2015, pp. 1-6, doi: 10.1109/GLOCOM.2015.7417793.
- [45] H. Li, L. Zhong, J. Liu, B. Li, and K. Xu, "Cost-effective partial migration of VoD services to content clouds," in *Proc. IEEE Int. Conf. Cloud Comput. (CLOUD)*, Washington, DC, USA, Jul. 2011, pp. 203-210, doi: 10.1109/CLOUD.2011.41.
- [46] G. Silvestre, S. Monnet, R. Krishnaswamy, and P. Sens, "AREN: A popularity aware replication scheme for cloud storage," in *Proceedings IEEE International Conference Parallel Distrib. Syst. (ICPADS)*, Singapore, Dec. 2012, pp. 189-196, doi: 10.1109/ICPADS.2012.35.

BIOGRAPHIES OF AUTHORS






Suman Jayakumar    received her B.E in Information Science and Engineering in 2007 from Vidya Vardhaka College of Engineering, Mysuru. She pursued Masters in Technology in Network and Internet Engineering from Sri Jayachamarajendra College of Engineering, Mysuru, in the year 2010. Ms. Suman Jayakumar is currently pursuing a Ph.D. in Computer Science at Visvesvaraya Technological University (VTU), Karnataka. She has over 12 years of experience in the IT industry and Academia. Her research interest includes Cloud Computing, Virtualization, and Software-Defined Networking. She can be contacted at email: jayakumarsuman@gmail.com.



Prakash Sheelvanthmath    has 24 years of experience in Training & Development, Education management, IT Infrastructure Management, both in academic and in the IT Training industry. Dr. Prakash S. has obtained a B.E degree in Electronics and Communication Engineering, M.Tech in Computer Science Engineering, and a Ph.D. degree in Computer Science on 4G wireless communications. He was chief editor for CSI Adhyan magazine for the year 2017-18. Dr. Prakash.S is the recipient of the "VMware Educator Excellence award" from VMware Inc. He is the recipient of the EMC Academic Excellence award for his five years of consistent service in Evangelizing third platform technologies like Cloud Computing, Big Data, and the Internet of Things. He received the BEST EDUCATOR AWARD from Salesforce. He has delivered talks at many Faculty development programs & chaired as an evaluator in many technical conferences. He has published many papers in International Journals and Conferences. He is currently serving as Professor and Executive Director, University Institute of Engineering, Chandigarh University. He can be contacted at email: prakash.hospet@gmail.com.



Channappa Baslingappa Akki    received his Bachelor of Engineering in Electrical & Electronics Engg, from University Visvesvaraiiah College of Engg., Bangalore in 1982, India and Master of Technology (M.Tech) in Computer Science, as well Doctor of Philosophy (Ph.D.) in Electronics & Computer Enggineering in the year 1990 and 1997 respectively from IIT (Roorkee), India. He has over 35 years of experience in both academics and industry. Dr. Akki has 40+ technical papers published/presented in National/International Journals/Conferences. His field of research is on Wireless Technologies, Big Data Analytics, and Security. He is currently serving as a Professor in the Department of Computer Science and Engineering at the Indian Institute of Information Technology, Dharwad. He is also serving as the Registrar at the Indian Institute of Information Technology, Dharwad, IIIT Dharwad, India, He can be contacted at email: akki.channappa@gmail.com.