Delay and Mobility Aware Content Placement in Named -VEC

Sai Vishwak Gangam(2020202006)

UNDER THE GUIDANCE OF
Prof. Vignesh Sivaraman

CS9.412.PG-Semester Project

International Institute of Information and Technology, Hyderabad

May 16, 2022

CONTENTS

1 Introduction

With a growing society comes a need for growth in technology as well. The internet architecture prevailing today was developed in the year 1960's with the intention to interconnect few computing resources across geographically distributed user base. The solution brought us TCP/IP stack which is a communication protocol where hosts can speak by establishing communication pipes between them. But the demand for mobile traffic is growing exponentially and will soon surpass what can be supplied with traditional host centric architecture. The Cisco Visual Networking Index forecast predicts that by the year 2020, mobile and wireless devices will account for 71 percent of the total IP traffic, which leads to a seven-fold increase in mobile data traffic globally between 2017 and 2022.Part of the cause for mobile data traffic growth is the rising use of video traffic. By 2020, 82 percent of all IP traffic will be IP video traffic, which will be a fourfold increase from 2017.

To meet this ongoing demand Content Delivery Network(CDN) architecture is particularly in use.The basic concept of CDN is having a network of servers close to the network edge with up to date content to optimize availability and latency. As the video traffic demand increase, as well as becomes more mobile, the placing of network edge servers and non-feasibility of 100 percent replications of the main servers, CDN is just not sufficient enough for the growing demand due to many users requesting the same server for the same content, which causes traffic overload. Not only is this relating to mobile devices, but the increase in IoT devices will add additional difficulty to the performance and reliability of the CDNs.

These shortcomings motivated the research community to look for alternate architectures for the future Internet and Information-Centric Networking is one such alternative. The concept of ICN(Information-Centric Networking) defines a new communication model that focuses on what is being exchanged rather than which network entities are involved in exchange of information. From the ICN perspective contents are first class network citizen rather than hosts.ICN's primary objective is to shift the current host-oriented communication model toward a content-centric model for effective distribution of content over the network. ICN relies on location independent naming, in-network caching and name-based routing for effective distribution and access of content over network. In recent years this paradigm shift has generated much interest in the research community and sprung several research projects around the globe to investigate.The traditional network consists of hosts being named by their IP addresses, which is based on their location in the network. Therefore, since the name is bound to the address, and the address is based on location, the address will change when the physical location is changed. With a rise in mobile devices, this will be an increasing problem. With ICN, this is no longer a problem, because ICN does not request content based on the address of the host, but rather based on the content itself. Instead of needing to know a host, in ICN, just the name of the content will do.

As one key enabler of Intelligent Transportation System (ITS), Vehicular Ad Hoc Network (VANET) has received remarkable interest from academia and industry. The emerging vehicular applications and the exponential growing data have naturally led to the increased needs of communication, computation and storage resources, and also to strict performance requirements on response time and network bandwidth. In order to deal with these challenges, Mobile Edge Computing (MEC) is regarded as a promising solution. MEC pushes powerful computational and storage capacities from the remote cloud to the edge of networks in close proximity of vehicular users, which enables low latency and reduced bandwidth consumption. Driven by the benefits of MEC, many efforts have been devoted to integrating vehicular networks into MEC, thereby forming a novel paradigm named as Vehicular Edge Computing (VEC).

So our research is on improving quality of service for consumers by reducing content access delay in vehicular edge computing domain by integrating it with Information Centric Networking(ICN).

2 Background on ICN

To meet the requirements of future internet architecture various architectures have been proposed under the umbrella of ICN including Data Oriented Network Architecture(DONA), Content Mediator Architecture for Content Aware Networks(COMET),convergence,Mobility First,Scalable and Adaptive Internet Solutions (SAILs),Publish-Subscribe Internet Technology (PURSUIT), Nebula,Expressive Internet Architecture (XIA), ChoiceNet , Content-Centric Network (CCN), Named-Data Network (NDN) and green ICN.Figure 1 depicts the timeline of ICN projects.

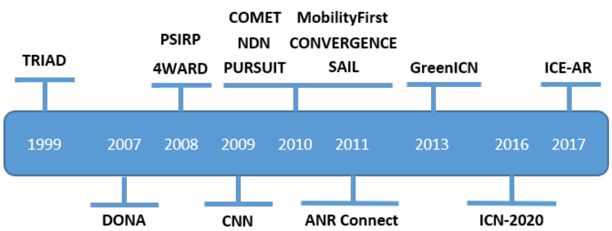


Figure 1: Timeline of ICN Projects

In this is section we will be discussing about working of Named Data Networking(NDN) as we are going to build our system model on that network architecture.

In NDN a client requests content by sending an interest packet into network with content name.NDN uses hierarchical naming which is very similar to URL.To request data objects, subscribers in an NDN network sends out an Interest packet. The Interest is sent to a Content Router (CR) and will be forwarded hop-by-hop across CRs. Each CR maintains three data structures: Forwarding Information Base (FIB), Pending Interest Table (PIT), and Content Store (CS). The FIB stores pairs of names to forwarding output direction. This is so that Interest packets can be forwarded to the correct place based on the requested content name. The PIT stores pairs of names to which interface requested the content. This is then used to backpropagate

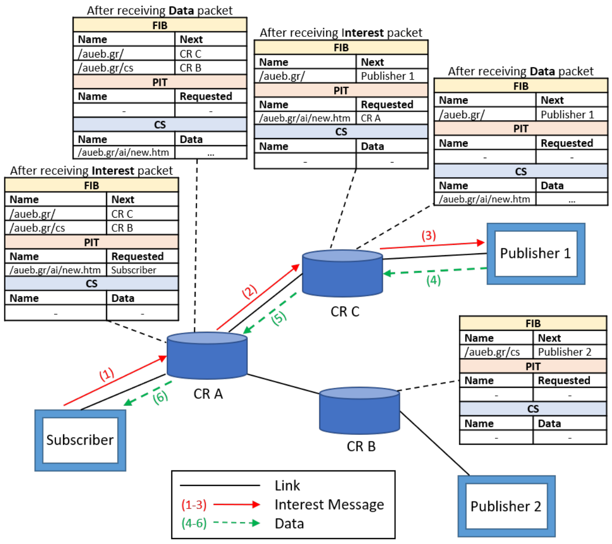


Figure 2: NDN ARCHITECTURE

data objects back to the subscriber. Lastly, the CS is, as the name suggests, storage for content that travels through the interface. This is a local cache. Once an Interest packet arrives at the CR, the FIB, PIT, and CS in the CR are updated.

3 Literature Survey

Nour et al.[1] has discussed about ICN cache placement scheme for IoT based traffic class. This paper talks about Near-ICN Cache Placement(NCP) scheme for IOT taking traffic class into consideration. NCP is designed to select optimal replica cache by minimizing three parameters:

- The delivery cost from original provider to the replica node
- Cost of caching the data in replica node
- The delivery cost from replica node to consumer

The objective function aim is to select the most demanded traffic class and place it close to the consumer which eventually reduces network delay and bandwidth. The proposed algorithm is executed in two phases, in first phase all the IOT nodes are sorted based on highest receive demands and the free cache memory.In second phase content placement is done on nodes that have the highest demands as well as they are far away from each other.This paper doesn't discuss or consider battery life of IOT Nodes(Energy Constraints), Mobility of IOT nodes and Content Freshness.

Asmat et al. [2] proposed a caching scheme for IOT contents known as Central Control Caching(CCC) scheme.This scheme avails the benefits of ICN in-network caching like reduction in content retrieval time, services separation of publisher and subscriber, better content accessibility, reduction in energy requirements,efficient bandwidth utilization. The proposed scheme divides the network into autonomous systems where all the autonomous systems are connected with a central network content controller(NCC) node. In the proposed scheme each node of a AS has a new structure along with FIB,PIT,CS known as CIT(Content Information Table) which makes node aware of the content location inside an AS(Autonomous system).The main aim is to minimize the hit to the IOT node for content in other terms energy has to be minimized. This paper also talks about Cache Data Freshness where content gets updated when the timer of the content in the routing table expires.This paper doesn't discuss about mobility of IOT nodes and delay requirements of consumers or subscribers.

Wang et al. [3] mainly talks about maximizing the quality of experience of ICV (Internet Connected Vehicle). They have considered mobility and caching the contents to minimize the both transmission delay and propagation delay,also considered content popularity patterns. The idea they used is to divide the content into chunks and place it in different ICV (v2v communication) and in different SBS (v2B communication). Paper doesn't talk about energy requirements of ICV devices, nor the freshness of data present in cache.

Serhane et al. [4] discussed about energy aware cache placement scheme for IOT-based ICN networks. This paper focuses on in-network caching in wireless IOT to maximize the energy-efficiency.The proposed scheme aims to maximize the energy-saving by trading off between content transmission energy and content caching energy. For placement of content in cache nodes different metrics like Content Popularity Ratio, User Proximity Ratio, Content Caching Reward, Content Energy Reward is taken into consideration. The proposed algorithm is as follows:

- **Case 1**: If the content is popular and fetching energy is higher than caching energy,the node with maximum user proximity ratio is cached.
- **Case 2**: If the content is not popular and transmission energy is low than the caching energy,the node with maximum content caching reward is cached.
- **Case 3**: If the content is not popular and transmission energy is high than the caching energy than the node with maximum content energy reward is cached.
- **Case 4**:If the content is not popular and fetching energy is less than caching energy content is not cached.

Li et al. [5] discusses about energy efficient computation offloading in vehicular edge cloud computing environment. The authors proposed a resource block based model of the computation offloading service(COS) under the VECC framework.The model outcome is offloading decision (i.e whether to compute locally or offload to edge node) with the aim of minimizing overall energy usage. Two equilibrium points namely energy equilibrium point and time equilibrium point are developed for allocation of resource blocks to edge nodes to perform computation requirements of consumers. Later the offloading decision problem is reduced to 0-1 Knapsack Problem and a greedy algorithm is developed accordingly which allocate the highest available value density to users. This paper doesn't consider mobility of IOT Nodes. They just talk about energy saving and minimization of allocation of resource blocks.

Zhu et al.[6] developed an cooperative offloading algorithm which maximizes the energy efficiency for a vehicular edge computing network, while guaranteeing the shortest delay of worst case vehicle. Authors have exploited geometry and unidirectional nature of road where road is divided into different segments and each segment has VEC edge server associated with it. A two step optimization algorithm is proposed where in the first step optimization of server selection is done considering maximum delay achievable and in the second step energy efficiency is maximised maintaining the delay.This paper considers mobility constraints of vehicles by tuning frequency of edge server such that vehicles should get the computational service results before it leaving the coverage area.

Deepak et al.[7] proposed an optimization framework that can be used to deliver data/service to the connected vehicles such that bandwidth cost objective is optimized.The framework optimises two objective functions namely minimization of maximum bandwidth utilization over all edges and minimization of total bandwidth cost overall edges.Our bandwidth and mobility constraint model is inspired from this paper. Coming to bandwidth constraint minimum data an edge node can serve to a vehicle is determined considering jam density in the coverage area, velocity of vehicle and bandwidth offered by the edge node,vehicular traffic in the coverage area of edge is modelled using M/D/C/C queuing model.Mobility model is developed accordingly based on the fact that the time required to process the service and time taken to transfer the data between vehicle and edge is lesser than the time the vehicle remains in the coverage distance of edge.

Quevedo et al.[8] proposed a framework to meet data freshness requirements of consumers in ICN-IOT network.This paper has inspired us to construct data freshness model for our framework. In the consumer driven freshness approach consumers specify their particular requirements in terms of information freshness. This parameter will be interpreted by the Content Stores which will consider it to determine whether or not suitable information is available.

Yonggong et al.[9] focussed on cache allocation problem which specifies how to distribute the cache capacity across routers under a constrained total storage budget for network. For formulation of optimal content placement problem authors have considered cache benefit which specifies how many intermediate nodes hit can be avoided if we cache the content at particular node and they have also considered content popularity such that more popular contents are given higher priority during allocation.

4 Analytical Model

Below is the table that describes parameters used for formulating objective function and respective constraints.

PARAMETERS	DESCRIPTION
E	Number Of Edge Nodes
F	Number Of Contents
V	Number Of Vehicles
Pop_f	Popularity Of Content f (between 0 and 1)
coverage_e	Coverage Length of Edge Node e(Km)
density_e	Number of Vehicles per Km under the coverage length of edge node e
$\text{band}_{e,v}$	Bandwidth(MBps) between edge node e and vehicle v
band_{e_1,e_2}	Bandwidth(MBps) between edge node e1 and edge node e2
size_f	Size(MB) of file f
$\text{t}_{e,f}$	Transmission delay of file f when a vehicle present in coverage area of edge e requests
$\text{tr}_{e_1,e_2,f}$	Transmission delay of file f when edge node e1 requests edge node e2.
vel_v	Velocity(m/sec) of vehicle v
$\text{coverage-time}_{e,v}$	Time(sec) a vehicle v will be under coverage area of edge node e
max-size_e	Maximum Cacheable size of edge node e
$\text{data}_{v,e}^{min}$	Minimum Data an edge node e can serve when a vehicle v requests
$\text{X}_{f,e}$	Decision variable whether file f is placed in edge node e or not

4.1 Parameters Formulation

$$t_{e,f} = \frac{size_f}{band_{e,v}/(density_e * coverage_e)} \tag{1}$$

Equation 1 represents the transmission delay model when a vehicle v under edge node e requests file f.

$$tr_{e1,e2,f} = \frac{size_f}{band_{e1,e2}} \tag{2}$$

Equation 2 represents the transmission delay between edge node e_1 and e_2 while transferring file f.

$$Data_{v,e}^{min} = \frac{band_{e,v}}{density_e * vel_v} \tag{3}$$

Equation 3 represents minimum data an edge node e can serve when a vehicle v requests.

$$coverage - time_{e,v} = \frac{coverage_e}{vel_v} \tag{4}$$

Equation 4 represents time a vehicle v will be under coverage area of edge node e under ideal conditions which is length of coverage area of edge node e divided by velocity of vehicle v

4.2 Objective Function

$$Minimize \quad \Sigma_{f\in F}\Sigma_{e\in E}(\, X_{f,e} * t_{e,f} + (\, 1 - X_{f,e}\,)(\, \Sigma_{e^1\in E}X_{f,e^1} * (\, t_{e,f} + tr_{e,e^1,f}\,)\,)\,) \quad (5)$$

Subject to Constraints:

$$\forall f \in F : \Sigma_{e\in E}X_{f,e} = 1 \quad (6)$$

$$\Sigma_{f\in F}Pop_f = 1 \quad (7)$$

$$\forall e \in E : \Sigma_{f\in F}size_f * X_{f,e} \leq maxsize_e \quad (8)$$

$$\forall e \in E \forall f \in F : size_f * X_{f,e} \leq data^{min}_{v,e} \quad (9)$$

$$\forall_{f\in F}\forall_{e\in E} : (\, X_{f,e}*t_{e,f}+(\, 1-X_{f,e}\,)(\, \Sigma_{e^1\in E}X_{f,e^1}*(\, t_{e,f}+tr_{e,e^1,f}\,)\,)\,) \leq coverage{-}time_{e,v} \quad (10)$$

Equation 5 describes minimization of overall delay considering every file and every edge node. Assumption is that all edge nodes are connected to each other. When file f is present in edge node e the delay will be $t_{e,f}$ otherwise the delay will be $t_{e,f} + tr_{e,e^1,f}$ where e^1 has the file f cached in it.

Equation 6 describes a file f can be present at exactly one edge node.

Equation 7 specifies that sum of content popularity of all contents will be 1.Content Popularity of file f is number of requests a file f got divided by total number of requests of all files.

Equation 8 describes that for every edge node e sum of all sizes of file f present in e will not exceed maximum cacheable size of edge node e.

Equation 9 describes that for all edge nodes e and files f if a file f is present in edge node e than the size of file f should be less than minimum file size an edge node e can server when vehicle v requests.

Equation 10 describes the mobility constraint, when a vehicle v present in the coverage area e requests for a file f than it should get the file before vehicle leaves out of the coverage area of edge node e.

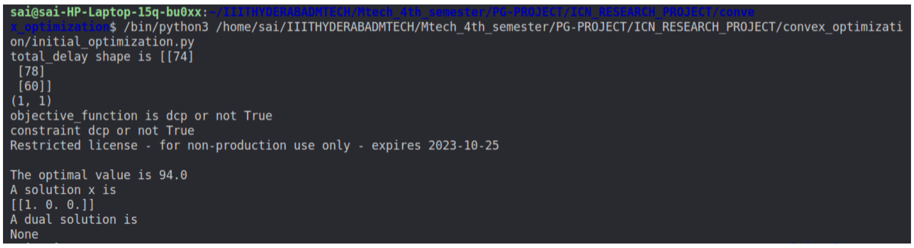
5 Proof of Concept

5.1 Example 1

Under ideal bandwidth,mobility and maximum size constraints assume there is one file f and 3 edge nodes e_1, e_2, e_3. Consider $t_{e_1,f} = 20$, $t_{e_2,f} = 30$,$t_{e_3,f} = 40$.
Now consider transmission delay from edge node e_1 to all other edge nodes is 2 sec.
Transmission Delay from edge node e_2 to all other edge nodes is 9 sec.
Transmission Delay from edge node e_3 to all other edge nodes is 5 sec.

If we place the file in the edge node e_1 total delay will be $20 + (30 + 2) + (40 + 2) = 94sec$ If we place the file in the edge node e_2 total delay will be $30 + (20 + 9) + (40 + 9) = 108sec$ If we place the file in the edge node e_3 total delay will be $40 + (30 + 5) + (20 + 5) = 100sec$

So its optimal to place the file in edge node e_1 with overall delay as 94 sec.



```
sai@sai-HP-Laptop-15q-bu0xx:~/IIITHYDERABADMTECH/Mtech_4th_semester/PG-PROJECT/ICN_RESEARCH_PROJECT/conve
x_optimization$ /bin/python3 /home/sai/IIITHYDERABADMTECH/Mtech_4th_semester/PG-PROJECT/ICN_RESEARCH_PROJECT/convex_optimizati
on/initial_optimization.py
total_delay shape is [[74]
 [78]
 [60]]
(1, 1)
objective_function is dcp or not True
constraint dcp or not True
Restricted license - for non-production use only - expires 2023-10-25

The optimal value is 94.0
A solution x is
[[1. 0. 0.]]
A dual solution is
None
```

Figure 3: CVXPY RESULT OF EXAMPLE1

5.2 Example 2

Consider Example 1 where there is one file with size 100 MB,now let us add maximum cacheable size of edge node constraint.
Assume maximum cacheable size of edge node e_1 is 50 MB
Assume maximum cacheable size of edge node e_2 ande_3 is 200 MB
Now though placing file on edge node e_1 is optimal but because of maximum cacheable size constraint the optimal solution is to place file on edge node e_3 with overall delay 100 sec.

```
objective_function is dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
Restricted license - for non-production use only - expires 2023-10-25

The optimal value is 100.0
A solution x is
[[0. 0. 1.]]
A dual solution is
None
```

Figure 4: CVXPY RESULT OF EXAMPLE2

5.3 Example 3

Now for the same above example add bandwidth schedulability constraints.
Assume minimum data size edge node e_1 can serve 100 MB
Assume minimum data size edge node e_2 can serve 200 MB
Assume minimum data size edge node e_3 can serve 50 MB
Now as the file size is 100 MB it cant be placed in edge node e_1 due to maximum cacheable size constraint and it cant be placed in edge node e_3 due to bandwidth constraint. So the optimal solution is to place file in edge node e_2 with overall delay of 108 sec.

```
objective_function is dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
constraint dcp or not True
Restricted license - for non-production use only - expires 2023-10-25

The optimal value is 108.0
A solution x is
[[0. 1. 0.]]
A dual solution is
None
```

Figure 5: CVXPY RESULT OF EXAMPLE3

6 Future Work

We will be extending this research by removing the exact one node cacheable constraint(i.e file can be cached at more than one edge node) and model file segmentation subject to some constraints. We will be considering freshness requirements of vehicles, energy requirements of vehicles and edge nodes in the future research.

References

[1] B. Nour, K. Sharif, F. Li, H. Moungla, A. E. Kamal and H. Afifi, "NCP: A near ICN Cache Placement Scheme for IoT-Based Traffic Class," 2018 IEEE Global Communications Conference (GLOBECOM), 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647629.

[2] H. Asmat, F. Ullah, M. Zareei, A. Khan and E. M. Mohamed, "Energy-Efficient Centrally Controlled Caching Contents for Information-Centric Internet of Things," in IEEE Access, vol. 8, pp. 126358-126369, 2020, doi: 10.1109/ACCESS.2020.3008193.

[3] C. Wang, C. Chen, Q. Pei, Z. Jiang and S. Xu, "An Information Centric Innetwork Caching Scheme for 5G-Enabled Internet of Connected Vehicles," in IEEE Transactions on Mobile Computing, doi: 10.1109/TMC.2021.3137219.

[4] Oussama Serhane, Khadidja Yahyaoui, Boubakr Nour, Hassine Moungla. Energy-aware Cache Placement Scheme for IoT-based ICN Networks. ICC 2021: IEEE International Conference on Communications, Jun 2021, Montreal (virtual), Canada. pp.1-6, ⟨10.1109/ICC42927.2021.9500341⟩. ⟨hal-03161227⟩

[5] X. Li, Y. Dang, M. Aazam, X. Peng, T. Chen and C. Chen, "Energy-Efficient Computation Offloading in Vehicular Edge Cloud Computing," in IEEE Access, vol. 8, pp. 37632-37644, 2020, doi: 10.1109/ACCESS.2020.2975310.

[6] Y. Zhu, T. Yang, Y. Hu, W. Gao and A. Schmeink, "Optimal-Delay-Guaranteed Energy Efficient Cooperative Offloading in VEC Networks," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9348017.

[7] D. Gangadharan, O. Sokolsky, I. Lee, B. Kim, C. -W. Lin and S. Shiraishi, "Bandwidth Optimal Data/Service Delivery for Connected Vehicles via Edges," 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), 2018, pp. 106-113, doi: 10.1109/CLOUD.2018.00021.

[8] J. Quevedo, D. Corujo and R. Aguiar, "Consumer driven information freshness approach for content centric networking," 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2014, pp. 482-487, doi: 10.1109/INFCOMW.2014.6849279.

[9] Wang, Yonggong Li, Zhenyu Tyson, Gareth Uhlig, Steve Xie, Gaogang. (2013). Optimal Cache Allocation for Content-Centric Networking. Proceedings - International Conference on Network Protocols, ICNP. 10.1109/ICNP.2013.6733577.

[10] S. Wang, Z. Zhang, R. Yu and Y. Zhang, "Low-latency caching with auction game in vehicular edge computing," 2017 IEEE/CIC International Conference on Communications in China (ICCC), 2017, pp. 1-6, doi: 10.1109/ICChina.2017.8330526.