Scholars Journal of Engineering and Technology (SJET)
Abbreviated Key Title: Sch. J. Eng. Tech.
ISSN 2347-9523 (Print)
ISSN 2321-435X (Online)

A Unit of Scholars Academic and Scientific Society, India
www.saspublisher.com

A Method of Modeling a Behavior of NFV Infrastructure Based on Network Calculus Theory

Hassan Mohamed Muhi-Aldeen*
Telecommunication system department, Odessa National Polytechnic University, Odessa, Ukraine

Original Research Article

***Corresponding author**
Hassan Mohamed Muhi-Aldeen

Article History
Received: 03.01.2018
Accepted: 14.01.2018
Published: 30.01.2018

DOI:
10.21276/sjet.2018.6.1.3

Abstract: A new approach of modeling of NFV infrastructure based on network calculus theory is proposed. The proposed approach gives possibility to modeling and analysis "end-to-end" time and performance quality of service (QoS) characteristics that dependent on current state of infrastructure components. The modeling the characteristics of computing nodes of NFV infrastructure is determined by term of arrival, service and backlog curves. The calculation values of such network characteristics of the computing nodes as delay and processing speed is given; the equation, making it possible to estimate the "end-to-end" processing delay and queue length is formulated based on partial obtained results.
Keywords: NFV infrastructure, network calculus, computing node, service provision.

INTRODUCTION

NFV infrastructure (NFVI) [2, 3] is single platform for data processing, storage and transmission implemented through the interaction of physical and virtual network resources. The analysis of mechanisms and tools of QoS services provision modeling in NFVI has shown that the focus is on functional indicators of quality and structural analysis of NFV infrastructure elements that ensure the correct compilation and provision of services [1-3]. At the same time, as non-functional quality indicators (time characteristics, performance, reliability, etc.) are not defined, and for each NFV solutions can vary significantly. The complexity of predicting the amount of available virtual network resources at the time of service formation and launch, the migration and the dynamic nature of the services provided, the lack of information on the characteristics of the equipment on the client side, and the equipment supporting the NFV in the zones of service providers exacerbate this situation [4].

Analysis of the boundary values of the quality indicators is a key mechanism for evaluating the functioning of the NFV infrastructure. The analysis of boundary quality indicators consists in determining the maximum and minimum values of indicators such as delay, bandwidth, processing speed, and the analysis of the dynamics of their changes in the process of providing the service "end-to-end" [5, 6, 7]. Thus, in assessing the overall quality of service delivery through the NFV infrastructure, the following components can be distinguished: the quality of the virtual computing resources/nodes, the quality of the network infrastructure, the quality of the IaaS components (Database-like-Service, Load-Balancing-as-a-Service), the effectiveness of interaction between controls and orchestration with the elements of the virtual network infrastructure. A number of work [6, 8, 9] suggested the use of a mathematical model of a network calculus algebra for the estimation of boundary indicators of multi-service networks.

The theory of network calculus allows modeling the behavior of network elements in the terms of both a continuous and discrete time model. The network calculus tools allow you to formalize and investigate the dependencies of QoS parameters for both the composition of network elements and for each element individually, and determine how intermediate (point) and boundary values of quality indicators for the implementation model. Based on the above features, the theory of network calculus proposed to use as the basis of the method of analyzing QoS NFV solutions, as well as to determine the minimum requirements for virtual and physical elements of NFVI, the observance of which allows providing services with guaranty QoS level.

Analysis of service provision mechanism in NFV infrastructure

The process of services provision in NFVI has a number of significant differences that are due to the presence and interaction of physical and virtual components. Based on [1, 2] the mechanism of service provision initially modelling by forwarding virtual network function service forwarding graph (VNF Forwarding Graph) [2, 4] and call “end-to-end” service provision. The main goal of the VNF Forwarding Graph is to define interactions and links between an ordered set of NRF&S modules to the subsequent provision of services.

Set of sequential chain of NVFI elements interaction with indication of possible physical (PhC) and virtual channels (VC), as well as connection computation nodes/virtual machine (VM) involved in the process of service providing is result of forwarding graph generation. VM can be represents as an interface module for the implementation of network functions and services that acts based on number of rules and policies necessary to handle incoming and outgoing data streams. The choice of a particular rule or policy depends on the decision of the of management and orchestration level (MANO).

The main parameters that affect the “end-to-end” QoS NFV during service provision are given in Table 1. The first column of the table contains the stages of the service provision life cycle. The following columns show the key quality of service parameters for each individual stage of life cycle [4, 7].

Table-1: Parameters that affect quality indicators within the life cycle of a service

№	Stage	QoS		
		Time indicators	Availability/Accuracy	Reliability
1	Initial organization/orchestration of service provision (analysis of statistical and resource allocation, preconfiguration of network components)	Delays in the process of allocating and reserving resources	Service configuration in based on SLA principles	The reliability of the resources allocation (the ratio of successful /unsuccessful delivery of services, the amount of data lost)
2	Virtual machines dedication and configuration	Computing resources time spent. Delaying in planning and distribution of tasks	Failure computing nodes synchronize and control components.	Premature calculation of performance/service delivery reliability
3	NFVI configuration	Delay in the process of virtual resources allocating to the formation of a virtual network infrastructure	Seamless convergence of virtual and physical network elements	Resource reservation
4	Virtual network operation/service provisioning	Data processing time (delay, jitter, etc)	Service ability coefficient, packet loss rate	Failure of network devices, both physical and virtual nature (VM)
5	Final orchestration (resources management)	Data processing time (delay, jitter, etc)	Service ability coefficient, packet loss rate	Failure of the compute node (VM) in the process of releasing resources

As shown in Table-1 one of the top quality indicators throughout the service provision life cycle is time indicators (delay) and reliability indicators. Time indicators can be divided by the following:

- A time of configuration and orchestration of NFVI, T_{orc} ;
- A time of new instance of service launch, T_{ins} ;
- A time of resource operation/service provision/migration time, T_{serv} ;
- A time of resource release and re-allocation of network resources, T_{rls} .

Time interval T_{orc} includes the time of sending the request by the user and receiving the request NFVM - T_{req} , free resources searching and efficiency of resources checking and allocation time T_{DOA} . $T_{orc} = T_{req} + T_{DOA}$

The launch new instance of service is considered as beginning of the useful phrase of service life cycle T_{ps}. The duration of the useful phase of life cycle includes the time of service instance launch, time of network configuration, the service provision time, and time of service finalization: $T_{ps} = T_{ins} + T_{NF} + T_{serv} + T_{rls}$.

Diagram of the procedure of orchestration and provision service in NFVI taking into account time indicators is shown in Fig-1.

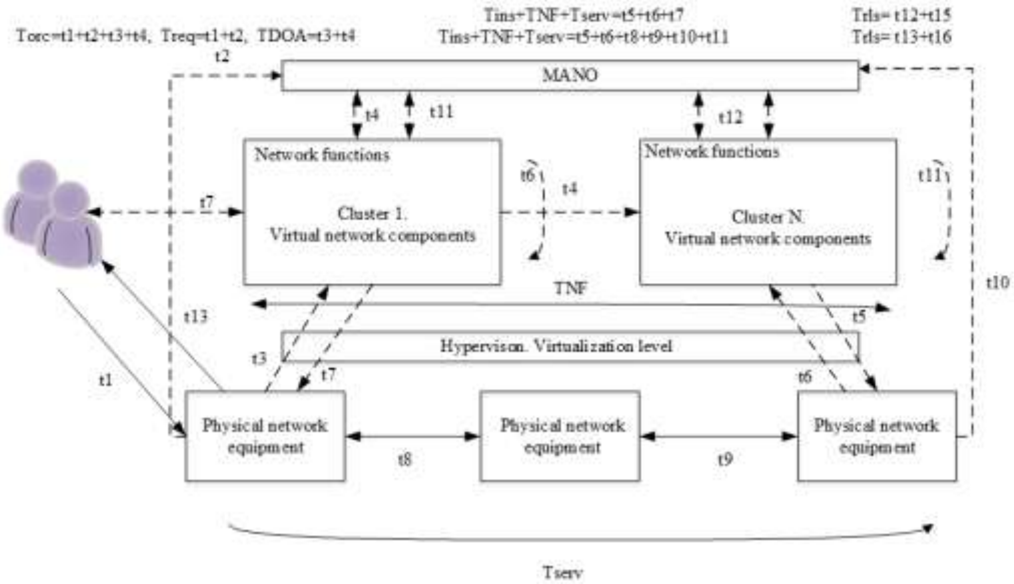


Fig-1: Diagram of service orchestration and provision

Delays have a primary impact on QoS. The analysis of service provision processes in NFVI showed that the greatest impact is caused by downtime delays due to insufficient amount of computing resources or migration of the service to another computing resource. The following reasons for delays is identified [5, 6]:

- Downtime caused by hardware or software failures;
- Service unavailability caused by characteristics of the computing node (virtual processor, virtual nic or virtual disk resources is not enough to provide the service);
- Duration of the service provision exceeds the maximum allowed downtime associated with the network characteristics of the infrastructure.

Therefore, the attention during the analysis of NFVI implementation should be given to the following type of delays:

1. Forwarding or network delays that characterize the user's network environment and can be calculated as the difference between the service provisioning time of the computing device (for example, the streaming video fragment was sent) and the recipient received the time the service (the video stream fragment was delivered): $D_{NF} = F(T_{NFVi}) - F(T_{VFo})$
2. Processing or computational resources delays in the of NFV are characterized by a time interval between stopping the service provision - T_{to} and resuming or restarting the service - T_{ti} : $D_{VM} = F(T_{ti}) - F(T_{to})$.

Diagram of service provision process including session start, downtime and recovery process is depicted on Fig-2.

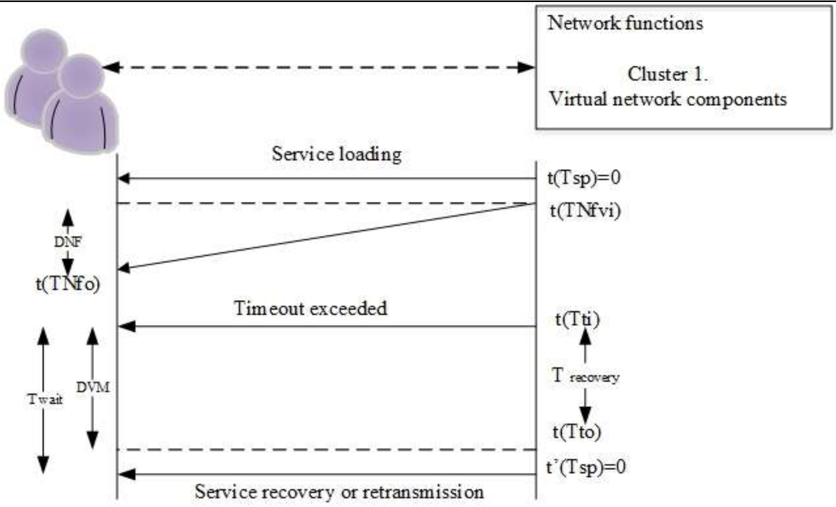


Fig-2: Service recovery scenario of service provision

Management and monitoring of NFV infrastructure by MANO needs to be guarantee for effectively services provision. Methods for assessing various QoS parameters can significantly vary and depend on the characteristics of network infrastructure elements.

Modeling and analysis of QoS NFVI indicators based on network calculus theory

In the process of analyzing and preventive estimation of QoS indicators characteristic for the implementation of NFV solutions, as for any other service-oriented architecture, a number of analytical methods can be used: system theory, queuing theory and network theory [8-10].

Elements of NFVI based on network calculus theory can be represented as follows:

1. Compute/handle nodes (different servers, VMs, routing and switching devices). Depending on the detail, the developer nodes can be represented by single elements or a combination of them (one VM or an entire cluster, a router, or a network domain).
2. Consumer nodes. Terminal equipment of service consumers, generating requests for the provision of services.
3. Data flow. For each flow, a arrival curve is defined that describes the transmission rate limitations and the possibility of its increase, as well as the sequence and rules of flow transmission through the handler nodes. For example, thread aggregation rules and service priority (IntServ, DifServ) [10, 11].

Modeling compute/handle nodes behavior

The behavior of each handler on the network is characterized by a service curve, which determines its performance, discipline of service, rules for allocating resources between threads. In accordance with features of the technology NFV-handlers nodes often act virtual devices: VM, routers, switches, firewalls, etc. Each node executes processing simultaneously a plurality of streams and providing multiple types of services. Therefore, the practical application of the description of the process of functioning of the handler by comparing the set of pairs of numerical values of the time-dependent arrival function (arrival curve) and the sending function is not possible.

The accumulation function $P(t) \in R$ expresses the dependence of the amount of data transferred by the node on the time of its operation, provided it is fully loaded, is called the productivity function of this node.

The delay function of handler node $D(t) \in S$ is an accumulative function that describes the dependence of the total amount of data processed and sent by the handler on the time of operation. The send function also depends on the characteristics of the function $A(t)$ and the properties of the handler.

Handler nodes have an accumulation function that results in the accumulation of a certain amount of data inside the handler node. The data processed to data stored ratio in handle node during the time of service provision (the lifetime of the service) is determined by backlog. The resulting curve, which specifies the amount of processed data for a certain period of time by certain handle node is represented by backlogged function (backlog curve).

Service curve is used for evaluation of changes in the QoS parameters. Service curve $\sigma(\tau)$ is an incremental function that characterizes the deviation of the current value of incoming data $S(t_i)$ at handler node from the value of the amount of data received at the previous moment $S(t_i - \tau)$.

For each flow an arrival curve *A (t)*, delay curve *D(t)*, service curve $\sigma(\tau)$ and backlog curve $b(t)$ are defined. The arrival curve describes the transmission rate limitations and the possibility of its increase, as well as the sequence and rules of flow transmission through the handler nodes.

The value of the backlog (t) at time t cannot be less than the difference between the amount of data received by the handler S, starting from an arbitrary moment and the amount of data that it could process during this time cannot be negative:

$$\forall t_i \le t : b(t) \ge [(A(t) - A(t_i)) - (P(t) - P(t_i))]^+ \qquad (1)$$

or

$$b(t) = \sup_{t_i \le t}\{[(A(t) - A(t_i)) - (P(t) - P(t_i))]^+\} \qquad (2)$$

$P(t)$ imposes a number of restrictions on the relationship between the arrival function A (t) and the sending function D (t), which can be specified as follows:

$$\begin{aligned} D(t) &= A(t) - \sup_{t_i \le t}\{[(A(t) - A(t_i)) - (P(t) - P(t_i))]^+\} \\ D(t) &= \inf_{t_i \le t}\{A(t) - [(A(t) - A(t_i)) - (P(t) - P(t_i))]^+\}. \\ D(t) &= \inf_{t_i \le t}\{A(t) + (P(t) - P(t_i))\} \end{aligned} \qquad (3)$$

Equation (3) allows to estimate the values preventively sending function, for each node in the network regardless of the number of streams.

The performance of the handler node depends on the type and value of data including the following:
p is a peak speed;
r is an average speed;
b is a changes in the intensity of the input data stream,
M is a maximum packet size,
R is a maximum processing speed of the data stream
Q is a processing rate of the individual stream,
q is a volume of the input buffer.
Set *(p, r, M, b)* indicates the main components of data flow specification T-SPEC [4, 7], set *(R, q, Q)* indicates the individual characteristics of handle node.

The fragment of the network infrastructure, for which the transmission of aggregated flows is characteristic, is shown in Fig-3.

Each data stream (R1-R3) has its own characteristics and is set by an individual load curve ($\sigma_1(t_i)$, $\sigma_2(t_i)$, $\sigma_3(t_i)$).

In the case of aggregation, the total multithread that contains multiple threads belonging to the same class of service is determined by generalized characteristics and should be described by a generalized load curve. The type of the arrival curve depends on the specification of the traffic class and is characterized by generalized indicators (max*p*, *maxr*, *maxM*, *maxb*): $\sigma_{AG}(t) = max_{n=1,2,3}\, pt + max_{n=1,2,3}\, b$

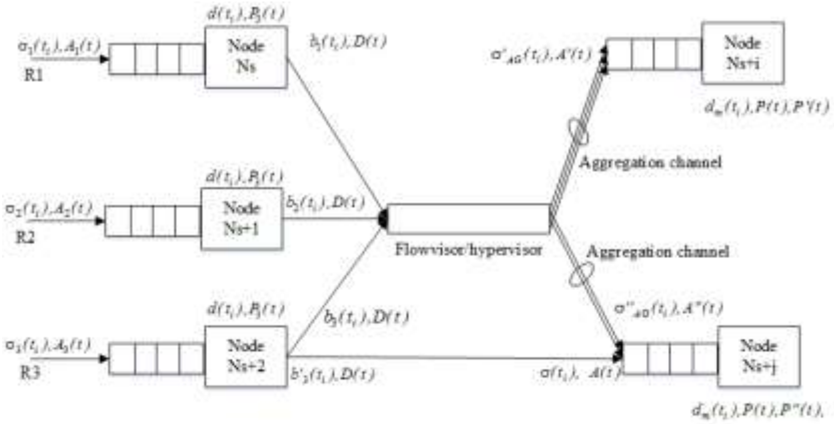


Fig-3: Structural scheme of data provision in NFVI

Each data flow has its service priority, which is defined in the ToS field [2, 3, 8]. The processing speed of the flow node by the flow handler, depending on its priority, can be estimated using the following equation:

$$r_i = \frac{\phi_i R_N}{\sum_{j=0}^{n} \phi_j} P_N(t), \tag{4}$$

where ϕ_i is weight (priority) of each stream serving by handle node, $P_N(t)$ is current performance of handle node, R_N average processing speed.

The value of delay is calculated by following equation:

$$d_i(t) = \frac{sM_i + \frac{\sum_{i=1}^{n} b_i - M_i}{r_i} r_i^{+}}{R_N}, \tag{5}$$

where s is total number of packets specific type of traffic transmitted during the time interval t, M_i is size i-th packet.

EXPERIMENTAL RESULTS

In order to analyze the adequacy of the use of a pro-assisted mathematical apparatus, an experimental study was conducted. The fragment of a network with support of technology NFV is considered. Provision of services with guaranteed quality in this fragment is achieved using mechanisms for reserving resources RSVP [11] (Fig-4).

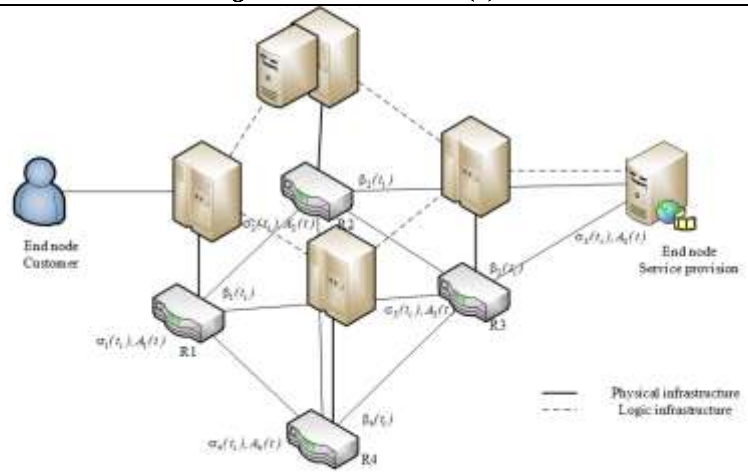
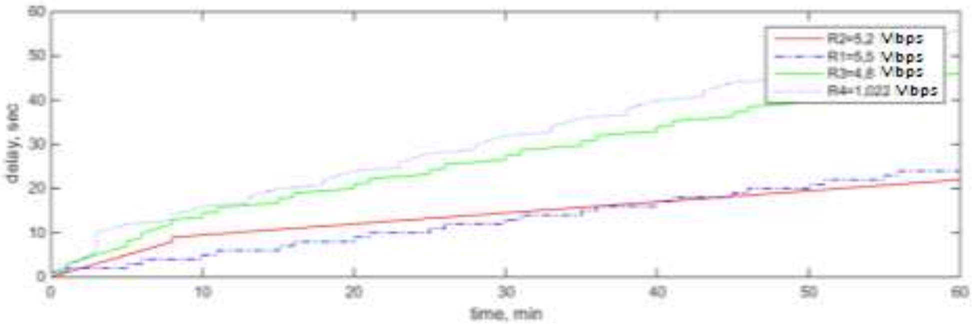


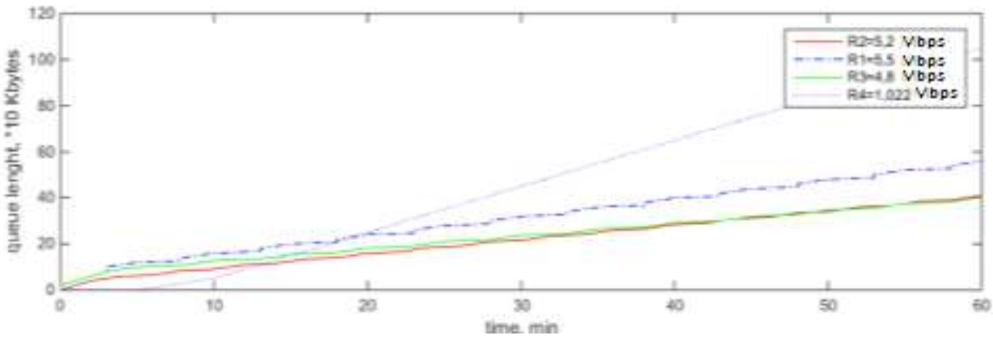
Fig-4: Fragment of experimental network

The handler nodes of experimental network are characterized by following: Tb_1 = 3,04 Gbps, Tb_2 = Tb_3 = 2,04 Gbps, Tb_4 =512 Mbps provided the maximum size of packets in the aggregation data flow M=1500 b; $d_{max1}(t)$ =600 ms, $d_{max2}(t)$ =620 ms, $d_{max3}(t)$ =612 ms; $d_{max4}(t)$ =1000 ms; R_1 =5,5 Mbps, R_2 =5,2 Mbps, R_3 =4,8 Mbps, R_4 =1,022 Mbps.

Graphs showing the time of the total delay and the waiting time in the queue, provided that the same discipline is maintained, are shown in Fig-5.



a)



b)

Fig-5: Processing (a) and handled/buffer (b) delays for different handle noodes

The delay value for the first node handler is 600 ms, with a minimum processing speed of 1,022 Mbps. Waiting in queue or delay time is 550 ms. The total end-to-end delay value is 1100 ms. The maximum delay occurs on node R4, thus, the given node does not allow to guarantee the proper quality of the provided services. In the process of reserving resources, the following route is being built: R1-R2-R3.

CONCLUSION

Analysis of the functional features of the MANO system throughout the service lifecycle has shown that time quality indicators such as delay and variations in its response time are critical. analyzing and modeling QoS NFVI indicators is a time consuming task. The existing methods of modeling and analyzing the behavior of multiservice systems do not allow to track the dynamic change in the characteristics of the NFV infrastructure. To simulate changes in network characteristics in the process of providing services to end users, it is proposed to apply the theory of network calculus. On the basis of the theory of network calculus, a number of definitions have been developed that allow the most complete modeling of the NFVI elements and the environmental conditions that affect their operation. Methods of analysis of NFVI behavior in two time models are proposed: continuous (load and service curves) and discrete (arrival and departure) time.

REFERENCES

1. Risso F. Network functions virtualization; 2012.
2. Xiang Z, inventor; FutureWei Technologies Inc, assignee. Methods and systems for managing capacity in a virtualized network. United States patent US 9,806,975. 2017 Oct 31.
3. Gabler H. ETSI-European Telecommunications Standards Institute. Nachrichtentechnische Zeitung (ntz). 1989;42:574-9.
4. Internet Research Task Force, Network Function Virtualization Research Group (NFVRG), https://irtf.org/nfvrg, 2015, Accessed: February, 03 2015.
5. Chung L, Nixon BA, Yu E, Mylopoulos J. Non-functional requirements in software engineering. Springer Science & Business Media; 2012 Dec 6.
6. Cloud Sourcing NFV VoIP and IMS. Hosted Model Delivers Superior Agility and Business Model Today [Online resource]/ Alianza Company. – 2015. Available at: http://www.alianza.com/hubfs/Collateral/Cloud_Sourcing_NFV_VoIP_-Alianza.pdf?t=1473364112924
7. Cisco corporation. Cisco Enterprise Network Functions Virtualization [Online resource]. Available at: http://www.cisco.com/c/en/us/solutions/collateral/enterprise-networks/enterprise-network-functions-virtualization-nfv/white-paper-c11-736783.pdf
8. Bouillard A, Jouhet L, Thierry E. *Service curves in Network Calculus: dos and don'ts* (Doctoral dissertation, INRIA).
9. Bouillard A, Junier A. Worst-case delay bounds with fixed priorities using network calculus. InProceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools 2011 May 16 (pp. 381-390). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
10. Cruz RL. A calculus for network delay. I. Network elements in isolation. IEEE Transactions on information theory. 1991 Jan;37(1):114-31.
11. Pana F, Put F. A Survey on the Evolution of RSVP. IEEE Communications Surveys & Tutorials. 2013 Jan 1;15(4):1859-87.