International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-11 Issue-2, December 2021

A Proposed Architecture for Placement of Cloud Data Centre in Software Defined Network Environment

Mohit Mathur, Mamta Madan, Mohit Chandra Saxena

***Abstract:** Emerging technologies like IoT (Internet of Things) and wearable devices like Smart Glass, Smart watch, Smart Bracelet and Smart Plaster produce delay sensitive traffic. Cloud computing services are emerging as supportive technologies by providing resources. Most services like IoT require minimum delay which is still an area of research. This paper is an effort towards the minimization of delay in delivering cloud traffic, by geographically localizing the cloud traffic through establishment of Cloud mini data centers. The anticipated architecture suggests a software defined network supported mini data centers connected together. The paper also suggests the use of segment routing for stitching the transport paths between data centers through Software defined Network Controllers.*

***Keywords:** Cloud Computing, Software Defined Network, Traffic Engineering, Segment Routing, Latency, Data Center, Traffic Optimization, Mini Data centers.*

I.INTRODUCTION

A. Cloud Computing

Today, cloud computing has achieved its maturity. Cloud Computing is a service model that provide services to various types of customers, through network, independent of geographical locations of Cloud Service providers and cloud users. Cloud computing provides on demand services with shared pool of resources. Through Cloud Computing the vendors are providing services related to software, infrastructure and platform. The trend in the market shows a major shift towards cloud computing adoption as it cut down the cost of owing resources. The enterprises need not to worry about the maintenance. Besides this cloud computing carries many other advantages such as dynamic resource allocation, elasticity, scalability and pay per use. Cloud Computing is available in various modes like public Cloud, Private Cloud and Hybrid Cloud. Public Cloud provides services through public internet.

Manuscript received on December 20, 2021.
Revised Manuscript received on December 29, 2021.
Manuscript published on December 30, 2021.

* Correspondence Author

Mohit Mathur*, Associate Professor, Department of Information Technology, Jagan Institute of Management Studies (Affiliated to Guru Gobind Singh Indraprasth University), Delhi, India. Email: mohitmathur19@yahoo.co.in

Mamta Madan, Professor, Department of Computer Science, Vivekanand Institute of Professional Studies (Affiliated to Guru Gobind Singh, Indraprastha University), Delhi, India Email: mamta.vips@gmail.com

Mohit Chandra Saxena, Research Scholar, Computer Science & Engineering, Galgotias University, Noida (U.P), India. Email: mohit.chandra_phd20@galgotiasuniversity.edu.in

The services are either free or pay per usage. Examples of Public Cloud are Amazon Elastic Compute Cloud (EC2), Google App Engine. The private Cloud provides proprietary hosted services. The services of Private Cloud are limited to single organization's need. Examples include Amazon wen Services, Sales force Services. The third category i.e. hybrid cloud carry advantages of public as well as private cloud. It delivers the private cloud's high-security features, along with the fast connection and easy-to-access features of the public cloud.

B. Data Center Networking

Enormous scope Data Centers structure the center foundation support for the always growing cloud based Services. In this manner the exhibition and reliability attributes of Data Centers will fundamentally affect the versatility of these Services. Specifically, the Data Center network should be deft and reconfigurable to react rapidly to truly changing application requests and Service necessities. An examination of Data center topologies has been performed and these topologies were compared on the criteria. Data centers have Top of Rack (ToR) switches and these switches are interconnected to End of rack(EoR) switches which are associated by means of core switches. This methodology prompts critical transmission capacity oversubscription on the connections in the organization center, and incited a few scientists to propose substitute methodologies for flexible savvy network models. Cloud Data Centers today pack a great deal of computational force. For example, Amazon Data Center houses somewhere in the range of 50,000 and 80,000 servers with force utilization that is around 30-25 megawatts. These Data Centers consume space which is generally similar to a football field on the off chance that we mull over the all-out space involved by the computational assets themselves, the cooling region, and these sorts of things. Data Centers are accordingly giving scalable computing resources, for huge web scale Services. These Data Centers will in general be situated in topographically scattered regions. Frequently, they might be in distant regions for advancing on conditions like energy utilization and ensuring that we get monetary force. Those are a portion of the issues in the area of Data Centers. In the event that we look inside a Data Center, the network is one of the fundamental parts of the Data Center and these network fabrics associates the entirety of the servers that are inside a Data Center, and it additionally gives availability to the customers that are over here through the web.

The job of the routing fabric is both for internal communication among these servers inside the Data Center and external communication of these servers with clients.
The primary components of a data center are hundreds of thousands of servers. Large number of servers may be running some application for instance Gmail at the same time. Thus these servers need to communicate with each other rapidly; hence latency is significant in such cases.
For better performance commodity switches are used. Commodity Switches are routers and the Ethernet switches that are utilized and that is in the core of the wide area network. So those are the switches are commodity and accordingly it is modest, and thusly we can utilize that to build large scale network fabric. To provide elasticity to customer services, cloud providers operate multiple large data centers across the world at multiple geographical locations. These large data centers carry thousands of switches with thousands of servers with every server responding multiple requests from their clients using virtualization technologies.

C. Software Defined Networks (SDN)

Traditionally each network switch called a router, has its own control logic and they behave on the ground of information collected from the neighbors. Here the current state of network and the policy decisions like firewalls, access controls or bindings like DNS to IP and IP to MAC etc. were distributed. There is no single node having entire network view (Global). Every route work for both control plane (Route Computation) and data plane (Forwarding Packets) based on its own decisions. Any new network applications need to be building with hardware support which is quite difficult. This approach is inefficient to handle complex data centers with high density servers with multiple virtual machines running on them. Thus, the cloud providers have to consider the data center networks with a different perspective. A new technology called software defined networking is introduced to eliminate problems with traditional routers. The new paradigm called software defined Network (SDN), classify the data and control planes of a router into two different domains. SDN control plane is the one, responsible for intelligent decisions, and is controlled by a software program called controller. It performs various functions like route computation, forwarding these computations to another device switches. The Switches perform forwarding function [1].

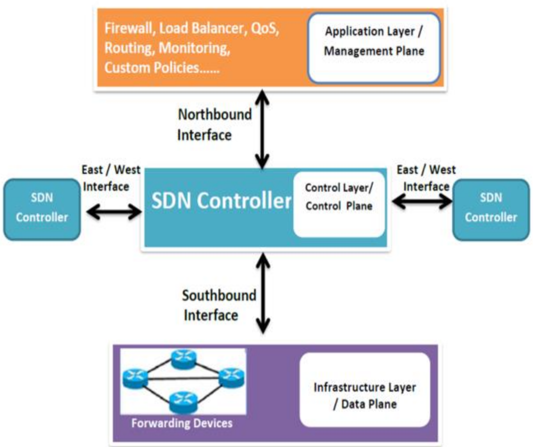

Figure 1. SDN Layered Architecture

The SDN architecture shown in figure 1, defines layered architecture having edge layer, SDN infrastructure layer and northbound application layer. The lowest layer i.e. edge layer is one in which servers are configured to handle the requests of the user. Various edge de--vices are involved in providing such services. The middle layer i.e. SDN infrastructure layer defines tradition SDN infrastructure including SDN controller and SDN Switches. The layer clearly defines the responsibilities of SDN controller and switches. The topmost layer defines northbound applications that actually define how control behaves. Northbound are then used to communicate. The purpose of these applications is optimal resource allocation, network management, security management, mobility management and many more. The communication received through northbound interface is then converted into command by controller and forwarded to data plane. In the multilevel architecture there is a need to define east/west interface also for communication. Sine a single controller faces limited scalability problem, a network can have multiple SDN domains, each controlled by an individual SDN Controller. Hence there is a requirement to make these controllers communicate, coordinate and share information with each other so that the network can provide optima routing and required quality of services. The SDN east-west interface defines communication between controllers of different domains. Controllers need to exchange information such as: Reach ability update: Exchange of reach ability information facilitates inter-SDN domain routing. Each flow may traverse through multiple SDN domains. Controllers of different domains exchange information related to routing paths, flow set up requests, Quality of Services, bandwidth, and software capabilities across domains. Standardization of east and west interface is still a infancy. No standard has been defined yet for east/west interface. Besides this, SDN also defines also address various technologies like software switch, virtualization. When SDN collaborate with optical network, it is called software Defined Optical Network (SDON) and provides all benefits of optical networking along with SDN. Another version of SDN called Software Defined Wireless Network (SDWN) is used for wireless network. Since Controller carry the global visibility, all innovations related to software can be implemented in controller, switches just obey controller instruction. All security capabilities like firewalls, authentication etc., Routing Capabilities, DNS and ARP bindings and even DHCP functionality can be implemented in the controller itself. The central controller can manage the entire network with dynamic programming. SDN and NFV (Network Function Virtualization) together provides more benefits. Through this a controller can optimize complex cloud network with dynamic network reconfiguration and resource provisioning. SDN empowers cloud computing with dynamic controllability because of centrally programmable interface i.e. controller. In real time Cloud Computing requires fast response to manage traffic.

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

A combination of SDN and Cloud can optimally rent visualized resources to cloud tenants. SDN can also monitor the status of resources being used. The SDN controller represents "Northbound Interface". An application in SDN translates the instructions into Open Flow Command or PCEP (Path computation Element Protocol) Instructions on its South bound Interface to configure switches. The overall functioning of a traditional open flow based SDN domain is as follows: All the switches in SDN domain have a flow table in which they store their internal state. The entries of flow table are compared with the packet headers field. Based on this switch takes action like sends the packet to certain port, drop the packet etc. When a switch recognizes a packet for which it does not carry any flow entry in its table, it forwards the packet to the controller. The controller then computes the optimal path for that packet flow and installs that path in all the switches along the path and returns packet to the switch to which it receives. All the packets that belong to same flow will follow the same path as they follow same rules.

D. Segment routing

Segment Routing is a type of source routing technique in which the source node specifies the route of packet towards destination. In Comparison, the other Routing techniques the path is chosen on the basis of destination. In Segment Routing the source node sends a packet with segments (An ordered list of instructions). This segment is embedded into packet header. In traditional segment Source Routing, the route is defined by source router. The path is not chosen on basis of destination as other Routing does. In Segment routing, a node guide a packet through segments which are list of instructions in ordered form. A segment is nothing but an instruction, may be topology or service based. The packet header contains the full instruction for path through the network. The segments are actually MPLS headers on IPv4 packets. Stack of MPLS labels are Instructions that are followed. The label at the top is processed by each router and the instruction in the label is followed to forward the packet. The packet reaches its destination after all labels are processed by routers on the way. The Segment Routing (SR) reduces complexity of control and user planes. Each router in a Segment routing network is identified by a SID (Segment Identifier). There are two types of SID defined in the SR RFC, global SID and local SID. The Global SID is advertised throughout the network and while local SIDs is local to the router. The Local SIDs is used to take decision within the router. For Example, selection of the exit interfaces for the traffic. There are two different types of segments: Global Segments-They calculate the shortest path using protocols like IS-IS or OSPF. Local Segments: They are used for hop by hop traffic diversion. As global segments have already directed the path, applying new route for each hop is an option in segment routing. With traditional MPLS traffic engineering using RSVP-TE all routers along the engineered route must maintain a state meaning that these routers must update with information about the end to end path and nodes. One of the major hindrances in network scalability is complexity of maintaining router states.

II.RELATED WORK/LITERATURE SURVEY

Pier Luigi Ventre, Mohammad Mahdi Tajiki, Stefano Salsano, Clarence Filsfilsdisused the SRv6 architecture for Traffic Engineering, Service Function The authors describe the advantages of the SRv6 technology with an SDN based approach in backbone networks. The authors also performed an evaluation of some performance aspects of our architecture and of the different variants of the Southbound APIs and analyzed the effects of the configuration updates in theSRv6 enabled nodes. Pier Luigi Ventre, Stefano Salsano, Marco Polverini, Antonio Cianfrani, Ahmed Abdelsalam, Clarence Filsfils, Pablo Camarillo, Francois Clad presents a tutorial and a comprehensive survey on SR technology, analyzing standardization efforts, patents, research activities and implementation results. The authors identified 8 main categories during our analysis of the current state of play: Monitoring, Traffic Engineering, Failure Recovery, Centrally Controlled Architectures, Path Encoding, Network Programming, Performance Evaluation and Miscellaneous. The paper then generate a report of experiences from survey work. Haisheng Yu ,Keqiu Li, and Heng Qi describes the use of distributed controllers to control forwarding devices in Software Defined Network architecture to solve the issues of scalability and load balance. It defines ASLB (active controller selection load balance), which proactively selects appropriate controllers for load balancing and minimize packet processing delays. The authors built a system and evaluated it on a physical platform. Results show that ASLB is much better than the static allocation scheme in terms of minimizing latency, bandwidth utilization, and throughput. Mohammad Ashrafi, Faroq AL-Tam, and Noelia Correia. This work focuses on the placement of controllers in software-defined networking architectures. A mathematical model is developed to place controllers under multicontroller switch-controller mapping, where a switch can be assigned to multiple controllers. The proposed model is shown to be effective and resilient under different failure scenarios while, at the same time, taking latency and scalability into consideration. DING Shichang, LUO Xiangyang,YE, YE Dengpan, LIU Fenlin discuss various IP geolocation algorithm and suggests Modified of IP Geolocation Algorithms Based on Rich-Connected Sub-Networks. They also experiment the performance of modified algorithm for China's Internet. Malvinder Singh Bali, Shivani Khurana discusses domains of latency from cloud to end user including, intra Cloud Latency, Network Latency, Processing Delay. They also highlighted the effect of latency on domains of Cloud network and also presented the service disruption effect due to DDoS attack on cloud network with simulated results. Shanhe Yi, Cheng Li, Qun Li, provides a survey that discusses the definition of fog computing and similar concepts. It also discussed various issues related to Fog Networking, Quality of Services, Interfacing and programming model, Computational Offloading accountability, billing and Monitoring, Provisioning and Resource Management., Security and Privacy.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Tim Verbelen, Pieter Simoens, Filip De Turck, Bart Dhoedt, provides a survey on cloudlets. They also discussed Cloudlet Architecture and evaluation along with a prototype implementation, showing the advantages and capabilities for a mobilereal-time augmented reality application.

Adam Wolfe and Paul Lu worked on reducing the communication overhead caused by Virtual Machine.They also proposed use of VDE networking for reducing total latency.

Ajith Singh and Hemalatha provide a survey on finding causes of latency for different geographical locations. They also compared latency caused due to use of different browsers .Different locations like cybercafé and university campus were used to check the latency caused in accessing Google Docs.

Qia o feng Qin, Konstantinos Poularakis, George Iosifidis, LeandrosTassiulassuggests a multi-controller edge system to reduce delay overheads. They suggest placement of controllers in the edge network. They also provided approximation solutions using linearization and supermodular function techniques.

Fan Zhao, XIangyangLuo, Yong Gan, ShoodiZu, Qingfeng Cheng, Fenlin Liu provides a survey on various IP geolocation based methods for judging users geolocation based on IP address. They classify IP geolocation methods as City level geolocation and street level geolocation and there classifications. They further propose a geolocation method based on identification in routers and local delay distribution similarity.

YehiaElkhatib, Barry Porter, Heverson B. Ribeiro, Mohamed fatenZhani, JunaidQadir, Etienne Riviere[40] provides a study on feasibility and readiness of microclouds for fog computing. They provided experiments using Rasberry Pi for hosted applications in microclouds.

Gilles A S Keupond Jo, Nogbou G. Anoh, Joel C Adepo, Souleymane Oumtanaga discussed the reactive and proactive approach of routing in SDN. They also provide experimental evaluation of proactive approach. They proposed a hybrid approach based on algorithm backpressure and provide tests on Ryu controller version 3.2 to calculate latencies.

III.PROBLEM FORMULATION

A. Growing Traffic

Considering the Public cloud which is most popular flavor of cloud computing offers various benefits to organization including lower cost, scalability, performance etc. One of the most important concerns when we look at public cloud is rapidly increase in the network traffic patterns due to sharp increase in number of connected devices and advent of technologies like IoT. According to a report generated by Deep Packet Inspection in April 2020, the Global Cloud traffic increase from 1.36% in Feb 2020 to 1.83% in April 2020 of the total internet. According to another statistics published by statistica.com the cloud data traffic is expected to grow to 19.5 zetta bytes by end of 2021.The predicted number of IoT devices is around 8.74 billion in 2020 and expected to increase 25.4 billion by 2020. According to prediction these IoT devices will be having a very highly dense network approximately 1 million devices per kilometer. This leads to very huge amount of data storage and processing with extremely high network performance. These increasing patterns raise an alarm on the current infrastructure which is affecting the network performance and needs to be upgraded. The tradition centralized data Center are far away from end users and cannot provide ultra-low latency and high bandwidth connectivity. They are not suitable for increasing traffic requirements due to emerging technologies like IoT. According to a report published by CBinsights 175 zettabytes of data expected by 2025. The current bandwidth available is not fair enough to accommodate the future requirements. The Problem is the Network being less focused part by cloud providers in the past. According to Cisco Annual Internet Report (2018-23) the expenditure on public cloud services and infrastructure is expected to grow to nearly $500 billion in 2023, a Compound Annual Growth Rate (CAGR) of 22.3%, according to market research firm International Data Corp. For last few years the data center traffic remained centralized and static.

B. Limitations of existing Mega datacenters

Today billions of devices are connected to mega data centers that provide various cloud services over internet. A matter of concern in cloud computing is continuous rise in the number of data hungry customers and applications. Due to huge amount of data produced by consumers, the trend of data center construction gets affected. The number of such mega data center is increasing day by day with the growth of cloud services. The acceptance of these monolithic data centers is dangerous. Despite advancements of technologies, the requirements like timely delivery and meeting Quality of Services (QoS) are not up to the mark. Though these huge data centers provides hyper scaling, but hype scaling does not involve expanding data centers to multiple floors. Moreover, the energy consumed by mega data center is quite high. According to the census done by National Resources Defense Council (NRDC) the energy consumed by these mega data centers will be 140 billion kWh by 2020. Besides energy consumption these mega data centers possess other problems like latency in delivery, Complex cabling and maintenance. Another issues related to data, is to keep data in local geographical boundaries because of security issues and Regulatory constraints.

C. Data center Geographical Location as a factor to Latency

Global Scale Large data Centers consist of tens to hundreds of thousands of servers. Several Cloud providers' carries large, centralized data centers, but this approach carries some issues. The main key issue with these data centers is Response Time. The Geographical distance from Data Centers to end user increase latency, this is a key challenge for some applications and services. It may include Video Conferencing, Real time Gaming etc. Though Content Delivery Networks were developed near to end users to resolve the issue but still they can keep static content.

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Geographical location of data center plays an important role in the latency during communication of client and server. Here are the ping results of some cloud data centers. The ping is performed from New Delhi location having IPv4 Address 172.16.11.218 performed on 19.10.2020 between 10 AM to 3 PM IST.

CloudPing.info

Amazon Web Services™ are available in several regions. Click the button below to estimate the latency from your browser to each AWS™ region.

Region	Latency
US-East (Virginia)	1058 ms
US East (Ohio)	322 ms
US-West (California)	296 ms
US-West (Oregon)	319 ms
Canada (Central)	283 ms
Europe (Ireland)	183 ms
Europe (London)	166 ms
Europe (Frankfurt)	176 ms
Europe (Paris)	542 ms
Europe (Stockholm)	187 ms
Middle East (Bahrain)	72 ms
Asia Pacific (Hong Kong)	128 ms
Asia Pacific (Mumbai)	48 ms
Asia Pacific (Osaka-Local)	191 ms
Asia Pacific (Seoul)	1579 ms
Asia Pacific (Singapore)	103 ms
Asia Pacific (Sydney)	3065 ms
Asia Pacific (Tokyo)	646 ms
South America (São Paulo)	387 ms
China (Beijing)	201 ms
China (Ningxia)	431 ms
AWS GovCloud (US-East)	279 ms
AWS GovCloud (US)	313 ms

HTTP Ping

Figure 2. Ping Report for AWS services with cloudPing.info

As we can see in Figure 2 that the ping latency from AWS cloud is smallest for Asia Pacific (Mumbai, India) , It is nearest having the smallest geographical distance from source of ping i.e New Delhi.

Search:

#	GCP Region Name	Region Code	Mean	Median	Min	Max	Test 1	Test 2	Test 3	Test 4
1	Taiwan (Changhua County)	asia-east1	405 ms	390 ms	380 ms	460 ms	390 ms	389 ms	460 ms	380 ms
2	Hong Kong (Hong Kong)	asia-east2	355 ms	355 ms	352 ms	357 ms	357 ms	357 ms	352 ms	353 ms
3	Japan (Tokyo)	asia-northeast1	464 ms	457 ms	439 ms	502 ms	502 ms	474 ms	440 ms	439 ms
4	Japan (Osaka)	asia-northeast2	490 ms	486 ms	473 ms	516 ms	516 ms	489 ms	473 ms	482 ms
5	South Korea (Seoul)	asia-northeast3	478 ms	470 ms	461 ms	513 ms	513 ms	467 ms	461 ms	472 ms
6	India (Mumbai)	asia-south1	135 ms	124 ms	103 ms	188 ms	188 ms	142 ms	103 ms	105 ms
7	Singapore (Jurong West)	asia-southeast1	249 ms	250 ms	236 ms	258 ms	256 ms	258 ms	236 ms	244 ms
8	Indonesia (Jakarta)	asia-southeast2	301 ms	299 ms	269 ms	337 ms	312 ms	337 ms	269 ms	286 ms
9	Australia (Sydney)	australia-southeast1	552 ms	535 ms	[illegible]	629 ms	629 ms	555 ms	511 ms	514 ms
10	Finland (Hamina)	europe-north1	638 ms	627 ms	314 ms	986 ms	986 ms	342 ms	911 ms	314 ms
11	Belgium (St. Ghislain)	europe-west1	734 ms	727 ms	371 ms	1112 ms	1112 ms	371 ms	1082 ms	372 ms
12	UK (London)	europe-west2	736 ms	718 ms	366 ms	1141 ms	1141 ms	370 ms	1066 ms	366 ms
13	Germany (Frankfurt)	europe-west3	753 ms	742 ms	374 ms	1153 ms	1153 ms	380 ms	1104 ms	374 ms
14	Netherlands (Eemshaven)	europe-west4	762 ms	753 ms	382 ms	1160 ms	1160 ms	402 ms	1103 ms	382 ms
15	Switzerland (Zürich)	europe-west6	783 ms	767 ms	385 ms	1215 ms	1215 ms	385 ms	1146 ms	387 ms
16	Canada (Montreal)	northamerica-northeast1	571 ms	575 ms	278 ms	854 ms	854 ms	308 ms	842 ms	278 ms
17	Brazil (São Paulo)	southamerica-east1	[illegible]	[illegible]	397 ms	1206 ms	1206 ms	421 ms	[illegible]	397 ms
18	USA (Iowa)	us-central1	541 ms	539 ms	272 ms	816 ms	792 ms	272 ms	816 ms	285 ms
19	USA (South Carolina)	us-east1	609 ms	612 ms	308 ms	904 ms	904 ms	321 ms	903 ms	308 ms
20	USA (Northern Virginia)	us-east4	594 ms	599 ms	294 ms	883 ms	879 ms	319 ms	883 ms	294 ms
21	USA (Oregon)	us-west1	593 ms	701 ms	250 ms	721 ms	250 ms	721 ms	703 ms	[illegible]
22	USA (California)	us-west2	495 ms	486 ms	244 ms	762 ms	244 ms	[illegible]	762 ms	252 ms
23	USA (Utah)	us-west3	258 ms	247 ms	246 ms	290 ms	290 ms	246 ms	246 ms	248 ms
24	USA (Nevada)	us-west4	247 ms	246 ms	244 ms	250 ms	244 ms	244 ms	250 ms	248 ms

Figure 3. Ping Report for GCP services with www.GCPping.com

As we can see in Figure 3that the ping latency from GCP Region is smallest for Asia Pacific (Mumbai, India) , It is nearest having the smallest geographical distance from source of ping i.e New Delhi

Azure Ping Test (Latency)

Missing your favourite cloud provider or a specific region? E-Mail to Varun Agrawal (Varun@VarunAgw.com).
Compare ping (latency) for other cloud providers.

Note: Sorting will be enabled after you press Stop.
Search:

#	Azure Region Name	Region Code	Mean	Median	Min	Max	Test 1	Test 2	Test 3
1	United States (Central US)	centralus	327 ms	341 ms	291 ms	350 ms	291 ms	350 ms	341 ms
2	United States (East US 2)	eastus2	248 ms	249 ms	245 ms	250 ms	249 ms	245 ms	250 ms
3	United States (East US)	eastus	252 ms	245 ms	239 ms	271 ms	245 ms	239 ms	271 ms
4	United States (North Central US)	northcentralus	281 ms	284 ms	266 ms	293 ms	266 ms	284 ms	293 ms
5	United States (South Central US)	southcentralus	302 ms	298 ms	286 ms	323 ms	323 ms	298 ms	286 ms
6	United States (West US 2)	westus2	254 ms	248 ms	245 ms	270 ms	248 ms	245 ms	270 ms
7	United States (West Central US)	westcentralus	297 ms	292 ms	287 ms	311 ms	311 ms	287 ms	292 ms
8	United States (West US)	westus	269 ms	257 ms	257 ms	293 ms	293 ms	257 ms	257 ms
9	Canada (Canada Central)	canadacentral	295 ms	290 ms	284 ms	311 ms	284 ms	290 ms	311 ms
10	Canada (Canada East)	canadaeast	286 ms	270 ms	260 ms	329 ms	260 ms	270 ms	329 ms
11	Brazil (Brazil South)	brazilsouth	362 ms	362 ms	360 ms	364 ms	364 ms	362 ms	360 ms
12	Europe (North Europe)	northeurope	190 ms	186 ms	172 ms	212 ms	186 ms	172 ms	212 ms
13	Europe (West Europe)	westeurope	188 ms	185 ms	175 ms	204 ms	185 ms	175 ms	204 ms
14	France (France Central)	francecentral	171 ms	173 ms	166 ms	175 ms	166 ms	173 ms	175 ms
15	United Kingdom (UK South)	uksouth	175 ms	172 ms	168 ms	185 ms	172 ms	168 ms	185 ms
16	United Kingdom (UK West)	ukwest	192 ms	187 ms	185 ms	203 ms	185 ms	187 ms	203 ms
17	Germany (Germany North)	germanynorth	178 ms	179 ms	173 ms	181 ms	173 ms	179 ms	181 ms
18	Germany (Germany West Central)	germanywestcentral	192 ms	191 ms	179 ms	206 ms	206 ms	191 ms	179 ms
19	Switzerland (Switzerland North)	switzerlandnorth	176 ms	177 ms	169 ms	182 ms	169 ms	177 ms	182 ms
20	Switzerland (Switzerland West)	switzerlandwest	156 ms	157 ms	153 ms	159 ms	153 ms	157 ms	159 ms
21	Norway (Norway West)	norwaywest	195 ms	196 ms	182 ms	208 ms	182 ms	208 ms	196 ms
22	Norway (Norway East)	norwayeast	191 ms	192 ms	185 ms	196 ms	185 ms	196 ms	192 ms
23	Asia Pacific (East Asia)	eastasia	138 ms	133 ms	126 ms	156 ms	126 ms	156 ms	133 ms
24	Asia Pacific (Southeast Asia)	southeastasia	103 ms	102 ms	92 ms	114 ms	92 ms	114 ms	102 ms
25	Australia (Australia Central)	australiacentral	201 ms	203 ms	195 ms	205 ms	195 ms	205 ms	203 ms
26	Australia (Australia Central 2)	australiacentral2	4631 ms	235 ms	179 ms	13479 ms	179 ms	13479 ms	235 ms
27	Australia (Australia East)	australiaeast	196 ms	186 ms	185 ms	216 ms	185 ms	186 ms	216 ms
28	Australia (Australia Southeast)	australiasoutheast	191 ms	191 ms	175 ms	207 ms	175 ms	207 ms	191 ms
29	India (Central India)	centralindia	75 ms	85 ms	47 ms	92 ms	92 ms	85 ms	47 ms
30	India (South India)	southindia	65 ms	62 ms	56 ms	76 ms	76 ms	62 ms	56 ms
31	India (West India)	westindia	53 ms	51 ms	47 ms	61 ms	61 ms	51 ms	47 ms
32	Japan (Japan East)	japaneast	185 ms	182 ms	176 ms	196 ms	176 ms	196 ms	182 ms
33	Japan (Japan West)	japanwest	167 ms	171 ms	153 ms	176 ms	153 ms	176 ms	171 ms
34	Korea (Korea Central)	koreacentral	182 ms	186 ms	167 ms	193 ms	193 ms	167 ms	186 ms
35	Korea (Korea South)	koreasouth	156 ms	161 ms	144 ms	162 ms	144 ms	162 ms	161 ms

Figure 4. Ping Report for Azure Cloud services with cloudpingtest.com

As we can see in Figure 4 that the ping latency from Azure cloud is smallest for India , West India , It is nearest having the smallest geographical distance from source of ping i.e New Delhi

Oracle Cloud Ping Test (Latency)

Missing your favourite cloud provider or a specific region? E-Mail to Varun Agrawal (Varun@VarunAgw.com).
Compare ping (latency) for other cloud providers.

Press F11 to exit full screen

Note: Sorting will be enabled after you press Stop.
Search:

#	Oracle Cloud Region Name	Region Code	Mean	Median	Min	Max	Test 1	Test 2	Test 3
1	Australia East (Sydney)	ap-sydney-1	190 ms	187 ms	187 ms	197 ms	197 ms	187 ms	187 ms
2	Australia Southeast (Melbourne)	ap-melbourne-1	182 ms	180 ms	180 ms	185 ms	180 ms	180 ms	185 ms
3	Brazil East (Sao Paulo)	sa-saopaulo-1	404 ms	398 ms	397 ms	416 ms	397 ms	398 ms	416 ms
4	Canada Southeast (Montreal)	ca-montreal-1	264 ms	251 ms	250 ms	290 ms	250 ms	251 ms	290 ms
5	Canada Southeast (Toronto)	ca-toronto-1	272 ms	274 ms	265 ms	276 ms	265 ms	274 ms	276 ms
6	Germany Central (Frankfurt)	eu-frankfurt-1	172 ms	172 ms	171 ms	172 ms	172 ms	171 ms	172 ms
7	India South (Hyderabad)	ap-hyderabad-1	57 ms	57 ms	55 ms	60 ms	60 ms	57 ms	55 ms
8	India West (Mumbai)	ap-mumbai-1	46 ms	42 ms	40 ms	55 ms	55 ms	42 ms	40 ms
9	Japan Central (Osaka)	ap-osaka-1	304 ms	309 ms	293 ms	310 ms	293 ms	309 ms	310 ms
10	Japan East (Tokyo)	ap-tokyo-1	306 ms	305 ms	304 ms	310 ms	304 ms	305 ms	310 ms
11	Netherlands Northwest (Amsterdam)	eu-amsterdam-1	183 ms	179 ms	178 ms	192 ms	179 ms	178 ms	192 ms
12	Saudi Arabia West (Jeddah)	me-jeddah-1	256 ms	235 ms	233 ms	300 ms	235 ms	233 ms	300 ms
13	South Korea Central (Seoul)	ap-seoul-1	349 ms	348 ms	344 ms	354 ms	354 ms	348 ms	344 ms
14	South Korea North (Chuncheon)	ap-chuncheon-1	328 ms	324 ms	320 ms	339 ms	339 ms	320 ms	324 ms
15	Switzerland North (Zurich)	eu-zurich-1	179 ms	178 ms	175 ms	183 ms	175 ms	178 ms	183 ms

Figure 5. Ping Report for Oracle Cloud services with cloudpingtest.com

As we can see in Figure5 that the ping latency from Oracle cloud is smallest for India West (Mumbai, India),

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

It is nearest having the smallest geographical distance from source of ping i.e New Delhi Hence as er the above reports, we can see that in all cases, it is very much clear that the nearest
(Geographical Distance) datacenter from New Delhi (From where the ping test is performed) is Mumbai, India/ India west and the ping latency is the smallest among all regions in all cases from all cloud data centers. This makes it clear that Geographical Distance of Data center plays a very important role in latency.

D. Regulatory Restrictions

Legal constraints play a very important role in determining where regulated and private data can be stored for companies across the global. Usually the laws of local jurisdiction are applicable to the agreements. Geographical locations of data center is a hindrance to public cloud Computing. Sometimes a problem of conflicting legislation is observed which forces a company in one jurisdiction to violate the laws governing data center where data is stored in another jurisdiction. Thus One of the main challenge a company face when accepting cloud computing environment is determining where and in what jurisdiction the data can be placed. The trans-border data flow (Data flow from one jurisdiction to another) i.e from data origin to Cloud Data center may face problems as original jurisdiction in such case may be difficult or illegal at datacenter location. For example EUDPD(European Union Data Protection Directive prohibits storing of European citizen's personal information outside the Europe. Such restrictions means that European Companies must ensure its customers that the cloud services are not replicated or moved to datacenters at countries other than Europe.

IV. EXISTING SOLUTIONS– FOG, CLOUDLET AND EDGE COMPUTING

There are lots of alternatives emerging in the cloud market to reduce latency and reducing the proximity between user and data center. Fog Computing and edge computing are the two major alternates of reducing the proximity of user to its data.

A. Fog Computing

Fog Computing is a system between cloud and end user that extends cloud services to network edge. In fog computing performs by physical devices in real time because these fog devices are distributed I the network edge and close to the user. Several applications put in the fog devices that each of which includes several modules. Fog devices are near the edge of the network such as routers, access points, gateways etc. Thus Fog architecture is three layered architecture with Mobile-Fog-Cloud hierarchy shown in figure 6. The layer between Cloud and Mobile is the intermediate fog Layer which consists of Fog servers placed at the edge of each network. The fog servers are lightweight and virtualized servers having wireless connections. The Fog servers can keep pre-fetched large volume of data.

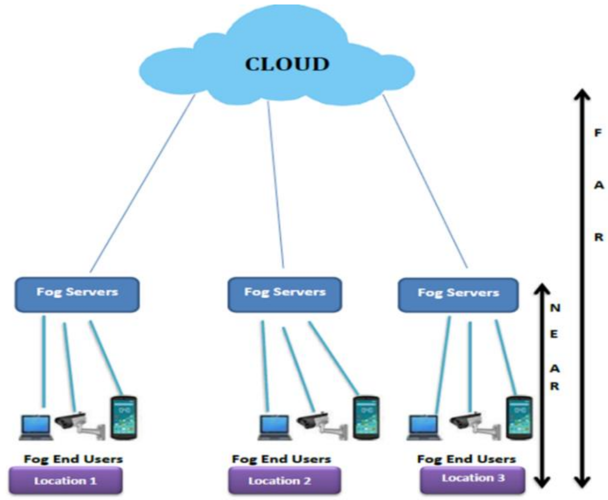

Figure 6. Fog Computing Architecture

Fog Computing suffer from major limitations firstly fog devices do not have enough sources to host all the modules of end users requirements, secondly the physical location takes away from anytime, anywhere, any data benefit of the cloud, thirdly resource management in edge architectures is very complex task, especially when a diverse set of services with different requirements need to be supported, fourthly there exists lot of security issues whilekeeping data on the intermediary devices. Moreover the cost of keeping this data is also a matter of concern, fifthly these fog devices will always be weaker when compared with the capacity of cloudservers, though fog computing is designed to reduce the latency for providing end user services however it may happen that delays may occur in finding the right fog device that meet all the requirements because a request may have to roam from one device to another until it reaches that fog device that fulfills the requirements.

B. Cloudlet

Today smart phones have been used in different aspects of our daily life. The enormous use of smart phones has shortened the battery life time as well as increases the need of storage and processing. Cloudlets emerged as to solve these problems. By offloading processing tasks and storage on cloud can reduce the power consumption and saves mobile storage. Cloudlet is a small cloud located very near to mobile user. Cloudlets are stateless, virtualized and resource rich servers on which mobiles can offload data and computations. Cloudlets need not be a fixed infrastructure; rather it may be formed dynamically with any device in Local Area Network with available resources. Though cloudlets have solved the problems of smart phones battery life, storage and latency of delivery. It still suffers from the following major limitations.

Firstly Cloudlets is a dynamic entity, one has to depend on service provider to provide such facility. Secondly there is still danger of cloudlet of going out of resources as its small and of limited resources. Thirdly Security is still am issue in dynamic cloudlets.

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

C. Edge Computing

Edge computing is a way to reduce latency by performing operations at the edge of destination i.e.closest logical location. It may be locally on a device closest to the destination. Sometimes, it may be closest data center. The key point of Edge Computing is that the processing must takes place off the network and on the Edge of Destination network. Edge Computing is getting popularity with the growth of Internet of Things as it helps in reducing latency for real time applications. Edge Computing also suffers from some limitations. The storage capacity requited for edge device is quite unpredictable in case of Edge computing. It also suffers from security challenges of these devices. Edge computing is primarily used only for analyzing the data. It requires additional advanced infrastructure at the edge.

V.SUGGESTED MODEL -GEOGRAPHICALLY DISTRIBUTED SOFTWARE DEFINED LOCAL MINI CLOUD DATACENTERS

Geographical distance between client and the data center play a major role in traffic delivery. The more the geographical distance the more will be the hops and more will the problems like latency, loss etc. Moreover, in today's scenario where Mobile Computing and IoT (Internet of Things) Applications are now being realized, very low or no latency in delivery is the need of the hour. Many solutions have been provided till date toreduce/ eliminate the problems related to geographical distance between the client initiating a request and the data center. As discussed in the previous section, few solutions include Fog Computing or edge Computing. However, the technologies like Fog Computing /Edge computing suffers from the many limitations Thus besides many other solutions like Fog / Edge computing, one solution is proximity of end user data center. The more the geographical distance the more hops and more will be the latency as each hop require queuing, other type of delays. If this traffic is localized to mini data centers located at geographical proximity to users, the burden of core network can be relieved and bandwidth of core networks can be optimized. The overall latency can only be reduced by deploying Cloud severs closer to these devices locally which reduces the number of hops. We introduce Mini Data Centers which are within the proximity to end user in its local geographical area. The details about Mini Data centers are discussed in the next section. Thus localizing the cloud traffic prevents the traffic to cross through core networks and hence saving the bandwidth. Moreover, mini data centers solve the problems like congestion in core networks and manage the traffic locally with fewer efforts. Another benefit of mini data centers is that they can be easily scaled as per the traffic demands dynamically with the growing traffic. In case of outages, the nearby data center can be used to fulfill the request. The servers at mini data centers can process information for IoT like services locally. Besides processing, they are faster in data collection and analysis. This provides faster way to access your own data and stored in allocation with security provided of your own choice of protocols. Our proposed Geographically Distributed Cloud Mini Data center architecture contains Cloud Mini Data centers, Cloud Mega Data centers, SDN Controller, SDN Switches, SDN Gateway Controller and Central Traffic Engineering Server. Another challenge is the heterogeneous transport infrastructure with various links and IP Paths offering different SLA's or no SLA for reaching these data canters from the user location. This gives rise to the need of SLA based application aware routing. The SDN Controller, in this scenario, must collect the link state information from the network, process the same and understand the SLA's each path can offer, and provide the best path to specific applications based on their needs. These paths should be switched accordingly by the controller to maintain the promised SLA for application. The network can be sliced into various paths providing different SLA's and the application traffic can be marked with a color to identify it and map it to the path providing the right SLA.

A. Components Of The Architecture

Cloud Mini Data Centers

The architecture proposes multiple mini data centers with proximity of the end users. These Cloud mini data centers are deployed within a zone. Each zone contains at least one mini data center responsible for serving local customers. To build highly available services and applications, each region should have one Mega Data center.The replica of each mini data center is in mega data center of that region. This automatically brings up the cloud services for a zone whose mini data center services are not available, i.e. during downtime.These mini data centers are interconnected with other data centers (of same operator or other). The end users are connected to these Mini data centers either through ISP's through exchange between data center. We are considering mini datacenter to have approximately<50000 servers for our research purpose.

Cloud Mega DataCenters

For our research, we consider Mega Data Centers consisting of 200,000+ servers. These data centers are located in various regions throughout the globe. They are connected to mini datacenters of its region and mega data centers of other regions centers. They require long haul connectivity.

SDN Controller

Each mini data center has one or more SDN controller in which all programming regarding routing, management etc. is designed for that data center. This controller controls a set of switches within the mini data centers and can communicate with its peer SDN Controllers of other mini data centers via fiber optic links and Backbone network.
We will later discuss the communication in detail. In SDN-based architecture, each in- dependent sub-system (e.g. DC or optical transport network) is equipped with a software entity referred as controller. A controller has the complete view and control over its own domain, i.e. the pool of physical and virtualized resources in the domain.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In a Data center environment, the SDN controller might be managing the DC fabric based on VXLAN – EVPN architecture or preparatory architectures but Segment routing might not be starting from the end node. While most of the Linux Operating systems have evolved to be able to support the Prefix-SID required for Segment routing, there is still the need for other OS to comply so that an end to end SR path can be stitched from with-in the datacenter itself.

SDN Switches

It was discussed earlier that SDN switches are the actual forwarding device at the core of network. They are the hardware entity and work as per decisions of SDN Controller.

SDN Controller Gateway

The SDN Controller Gateway deals with physical and virtual network resources to establish connectivity between various distributed cloud providers virtual machines. SDN Controller gateway is connected to SDN Controller Gateway of other domains via Backbone network. The SDN Controller Gateway is responsible for providing connectivity between various resources for users without looking at type of connectivity (connectivity can be dedicated or public Internet). The SDN Controller Gateway ensures the connectivity non-invasive so that it conserves the configuration of cloud network. Broker at SDN Controller Gateway decomposes the user request by splitting algorithm (Expressed as resource graphs and sub graphs). The broker relates the sub graphs with SDN controller Gateway. The SDN Controller Gateway controls and configures SDN Controller to establish an inter data center connectivity to compose the overall graph corresponding to the original user request.

The central Traffic Engineering Server

The central Traffic Engineering Server controls the entire network. It keeps the detailed information of SDN Switch hardware. These servers were replicated throughout the globe. Each server is dedicated for a logical “Autonomous System” that share some set of IP prefixes. Each server is connected to its peer servers of other Autonomous Systems via SDN gateways. SDN Gateway abstracts details of flows and routes depending on the technology and architecture used. For Open flow based architecture, the gateways abstract the Open Flow and switch hardware from the central TE server while the IP routes and link states are abstracted from the central controller in the PCEP based architecture. The SDN Network is designed in such a way that it will support all traditional Routing protocols. Hence, it can work with both SDN enabled and non SDN devices. The central Traffic Engineering server employs two major services Traffic Measurement and Traffic Management. The Traffic measurement looks for monitoring, measuring and getting network status in SDN. Network status information includes the current topology connect ion status, ports’ status (up or down), various kinds of packet counters, dropped packet counters, utilization ratios of link bandwidths, end-to-end net- work latency, end-to-end traffic matrices and so on. The status of the network helps in validating the current state of the network and predicting the future traffic trends by analyzing packets. BGP-LS are the protocol which is used for collecting the link states and routing information from the network and passing it on the controller for making intelligent decisions. BGP-LS also enable the measurement of KPI values and update the controller on various Links and their associated KPI. The information thus received by the controller is used to decide the best path to accommodate the traffic and associated needs. This avoids future congestion in the network and improves network efficiency. We divide research work on traffic measurement into three directions: parameters of network measurement, a generic measurement framework, traffic analysis and prediction. Traffic management mainly studies how to manage and schedule network traffic based on the status of network information provided by the traffic measurement technology, to satisfy end user requirements of network applications, such as Quality of Services (QoS).

B. Inter Data Center Connectivity

Previously Data center interconnect (DCI) refers to the networking connect ion be- tween of two or more geographically distributed data centers. But for our sake its inter- connection between Data centers of one domain (i.e. Autonomous system) and connection of these Mini data centers with Mega data center of that region and further connection between geographically far mega data centers. We classify the communication domain in Software defined cloud network using segment routing into three categories – Intra Mini Data Centre Domain, Intra Mega Data Centre Domain /Inter Mini Data Centre Domain and Inter Mega Data Centre Domain. The connectivity between these domains is stitched either through BPG-LU (RFC-3107) or by a Segment routing technology called ODN (On-demand Next-Hop).

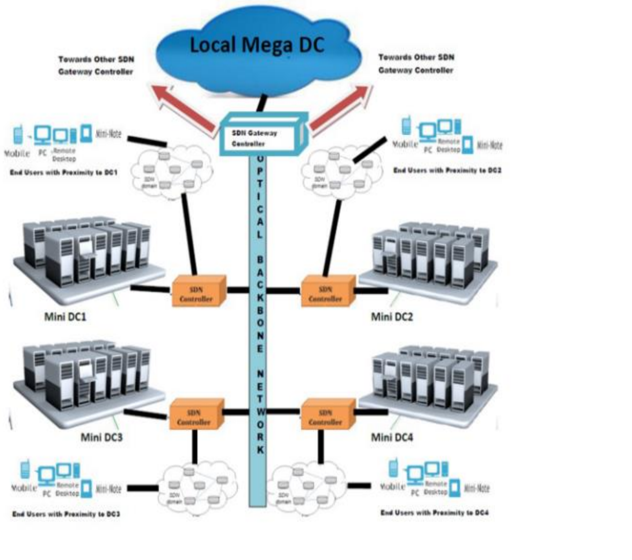

Figure 7. Inter Mini Data Center AS Architecture

C. Intra Mini Data Center Domain

The domain Consist of a local SDN Controller with SDN switches within a mini data center. The domain provides a view of connectivity and topology for a mini data center.

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

D. Intra Mega Data Centre Domain / Inter Mini Data Center Domain

The domain consists of several mini data center along with a local mega data center. The domain consists of Local SDN Controllers connected with each other as well as with a gateway controller of the domain via backbone network of that domain. A mini data center may be connected to backbone networks (many). Traffic to or from these mini data centers may be sent to or from any of these backbone networks. It provides edge devices for communication between different mini data centers. Figure 7 shows the interconnection of Mini Data centers of a geographical domain, connection to End Users, and to local mega data center via backbone. The communication can be described as Inter Mini Data Centre Communication (between Mini Data Centers of one domain) and Mini data center to Mega data center of one geographical domain. Inter-Mini DC data network consists of a set of Mini DCs interconnected through one or multiple optical fibers by an optical transport network. In SDN-based architecture, each independent subsystem (e.g. DC or optical transport network) is equipped with a software entity referred as controller. A controller has the complete view control over its own domain, i.e. the pool of physical and visualized resources in the domain. The network orchestra tor is a software component that coordinates resources over different domains and the operations through different controllers [14]. In a distributed Mini DC network, the main role of the orchestrator is the coordination of communication amongst geographically confined Mini DCs, taking into account information about the physical infrastructure. The communication amongst mini DCs is done using SDN Controllers of these mini DC's. The controller to Controller communication is most important part of this subsystem. Controllers are connected via optical fiber links to each other within a geographical area. We can call that area an autonomous system. This autonomous system ends at a SDN Gateway controller which takes/brings the traffic from other Autonomous systems. The SDN Gateway Controller has a full view of various SDN Controllers in its domain and other adjacent gateway controllers via SDN Traffic Engineering server of its Autonomous system [13]. The various information Exchanged between SDN Controllers include their topology, Network events like which link is down, bandwidth request information from user like allocate xyz GB, Various QoS requirements and the infrastructure status.

E. Inter Mega Data Center Domain

The domain consist of SDN Gateway controllers for each mega datacenter connected to SDN gateway controllers of other domain and central traffic engineering server of a particular geographical area via backbone network[15].The domain provides global view of the Cloud network. It provides edge devices to communicate between different mega data centers. Figure 8 depicts the broader perspective of the architecture with interconnection of Mini datacenters with their local Mega datacenters, interconnection of mega datacenters and with central traffic engineering server using Optical fiber backbone network via SDN gateway controller. The SDN Gateway controller is responsible for controlling local SDN Controllers. The local controllers of the domain are responsible for making decision of forwarding the packet to the end user or to the SDN Gateway Controller. SDN Gateway controller provides specific policies or forwarding rules. These rules are then forwarded to local controllers. The local controller then store these policies or rules and make decisions accordingly. For SDN Gateway controller these local controllers are treated as agents. Thus, the overall services were provided with the cooperation between SDN Gateway controller and local controllers.

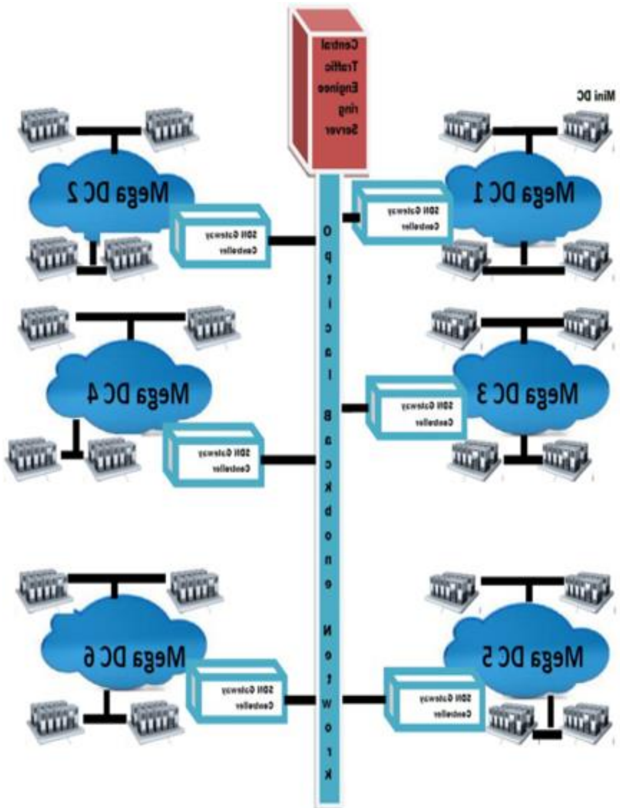

Figure 8. Inter Mega data center Communication and Interlinks of software De- fined Mini Data centers Regions, Availability Zones, and Local Zones

Cloud computing resources are hosted in multiple locations world-wide. These locations are composed of Regions, Availability Zones, and Local Zones.

Region:

Each geographical area is identified as a separate region and each region may have multiple isolated locations called Availability Zones.vA region is an independent geographical area. Each Region consists of one Mega Data center along with many mini data centers located in each zones. A region also consist multiple zones. A region must be an area with round trip time network latency < 1ms.Regional SDN controller must have detailed topological knowledge of its region and abstract information about the SDN Controller in nearby regions. It is a type of SDN controller Gateway. Each Region is completely independent.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Availability Zones

Zone is Cloud Services deployment area. Zone should be considered as single failure domain within a region. Each zone consists of at least one mini data center to provide cloud services.

Though it usually have many Mini Data Centers. Each Mini Data Centerslies in a local zone.The cloud services within a zone must be replicated at regional domain in mega data center of that region. Each zone is controlled by an SDN controller carrying the topological information of its zone as well as zones of that region.

Local Zones

Local Zones are the Nearest Location to end user. It consists of a single Mini Data Center where all cloud services local to that location are placed. The Local Zonal experience of a end user is based on Zone downtime, Zone latency and availability of resources.

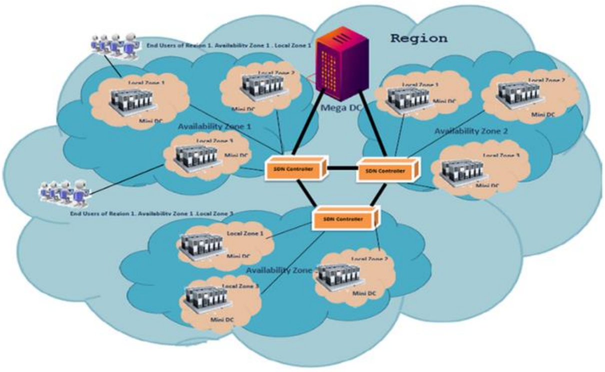

Figure 9 An Abstract View of local Zones, Availability Zones and a Region for suggested Architecture

Segment Routing for domains

For Routing information we suggest segment routing. Since most of the traffic generated is local for the mini data centers in a geographical domain, segment Routing is the best option for our geographically distributed mini data center architecture. Combining Segment Routing with SDN controller expands the set of available use cases and allows operators to gain the maximum set of benefits from the source routing technology. Centralized controller can actively collect and monitor the SDN topology network traffic engineering changes with global view of network. BGP Link state is the protocol which is utilized to collect and transport the link state information and the path KPI to the controller. Path Computation element protocol (PCEP) can then be used to present near real time views of topology deployment. The Path Computing Element (PCE) can be contained with a vendor's SDN controller, through which network commands can be given and acted upon. In traditional per flow Routing SDN controller directly interacts with each node in the traffic path. This can cause scalability issues as large amount of Routing states need to be maintained. In segment Routing the route information and configuration is done at the border of the network. The proposed architecture combines benefits of SDN and Segment Routing. A major challenge with traditional SDN is that SDN switch requires larger flow table as compared to tradition network. Maintaining flow table is highly expensive. Segment Routing also solves this problem. Segment Routing does not require maintaining large number of forwarding rules at each device and hence eliminates delays caused by network devices. The controller needs not to add forwarding rules to each individual device along the path. The controller need not redistribute the state to switches. OSPF can be used for communication between SDN and segment routing domains. A segment Routing link is kept in the table in a domain. In addition, an adjacency segment is kept which suggest actions that a packet transfers to egress nodes. SDN controller applies sequence of labels as combination of node segment and adjacency segment to the packet header. Thus Segment Routing carries benefit of scalability and minimizing delays. A major drawback of Segment Routing is that the packet size increases, which reduces available bandwidth. The combination of segment Routing with centralized PCE allows traffic engineered paths to be defined across multiple domains defining paths from metro networks to core networks, connecting data centers to WANs. Thus segment Routing in SDN domain pro- vides Increased Network Efficiency, Boosted network Performance, Increased Opex, efficiency through automation, Assured reliability and network up time, End to end Class of service assurance. The complete source Routing is classified into two categories. Intra Domain Traffic Routing and Inter Domain Traffic Routing. The traffic need to be steered within Intra Mega data center domain is called Intra domain Traffic steering. Here interconnection between the different mini data centers or between mega and mini data center is chosen. The traffic is steered within the Mega data center domain. Here the focus for Routing is on load balancing, achieving QoS and avoiding degradation of resources. Inter data center traffic is classified as Inter domain traffic routing. Here focus is on the optimal path across any network or across any collection of networks. The points of interconnections between the networks need to be selected. The composite end to end path comprise steering in source domain, choice of source domain exit point, steering cross backbone network of that domain, choice of network interconnections, choice of destination domain entry points and steering in destination domain. ODN (On demand Next Hop) is the technology which is powered by the SR and PCE architecture. This technology is used to stich the path between different domains, while connecting multiple data centers. Each domain is controlled by a specific controller or the PCE. The PCE of each domain can talk to each other to get the global view of the complete network. Based on this global view, the next hop address can be programmed on the bordering routers to connect to the neighboring domains. A packet crossing various regions during its journey, it may use different technologies, may have different administrative controls. Considering different administrative control is particularly important because operator of one region may not be willing to share in- formation about their networks and may be reluctant to allow third party to control their network. Using the Suggested Cloud Mini Data center architecture we consider how to get packet from source to the destination,

Retrieval Number: 100.1/ijeat.B33261211221
DOI: 10.35940/ijeat.B3326.1211221
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

we need to check the domain of the destination, the exit point from source domain, the entry point of the next domain towards destination, The path from next domain till the destination in the same way.

The local controllers and the gateway controllers play an important role in the whole procedure. It can be accomplished with the coordination amongst them. When, MPLS technology is used the labels may be distributed within a single domain using SDN Controller. Source Routes within a domain may be expressed label stacks by the controller of that domain. As in the figure the local controllers are connected to gateway controller via backbone network and further gateway controllers are connected to other gateway networks via backbone networks of their domain. The controllers may cooperate and share the information of different degrees. The optimization of path globally can achieve using the central traffic engineering server. Central traffic engineering server carry complete the domain networks information, backbone networks and end to end topology. It is responsible having knowledge of path computations and for issuing the necessary commands. The controller gets traffic engineering and topology information through BGP-LS. The local controller provides topology information to respective gateway controller and gateway controller stores respective central traffic engineering server. The route from source to domain egress controller is left to the controller of source domain and the route from destination gateway controller to the destination is left to the controller of destination domain. Controller from each domain is responsible for finding the best path to next domain but does not have knowledge about the best exit point from local domain. To solve this we suggest a technique in which each network advertises connectivity across the network about adjacent networks to controllers instead of advertising full topology. This technique is suitable for end to end domain interconnect where backbone network is under different controls from domains this information may be provided by the gateway controllers of these domains. For non SDN domains the packet can be forwarded by forwarding router based on locally determined additional set of labels that define the path to next hop as in the traditional networks. In this case the source puts a stack of labels in the packet so that it can reach its destination.

The Overall Communication servicing data services can be defined as follows:

A packet originated from a source is first forwarded to local controller which pushes the segment header into the header as Controller carry complete topological information of its zonal local mini data center where data or compute services are located. The packet is then forwarded to the local mini data center with least latency from switches of that domain.

In case the services of local mini data centers are unavailable, the controller forward the packet to mega data center of that region, which carry the replica of that service request or to nearby zone mini data center of carrying compute services keeping load balancing and latency in mind in case packet require compute services.

When a request needs to be forwarded to domain outside a zone, global segments need to be added by the controller.

The benefit of segment routing is that the switches need not to keep any rule for forwarding a packet. It just inspects the segment header of the packet and forwards it to next switch. The controller also needs not to push rule into switches for various paths. This minimizes Controller switch packet exchange and hence reduces the traffic.

VI. CONCLUSION

Localizing data traffic geographically will have a great impact on current growing traffic trends. It saves networks bandwidth (throughout the globe) and improves quality of service and a solution to latency problems. Providing Mini Data Centers in a Local Zone localize the traffic and hence solve traffic related issues like latency, bandwidth and availability. This will also eliminate requirements of mediators such as content delivery networks, which adds a further layer of complexity in the network and incapable of delivering dynamic content. Localizing traffic will bring trust in people towards cloud. This also solves all regulatory restrictions related problems since the Mini Data Centers carry data of local personnel where local laws can be applied. Providing segment routing in SDN domains will provide additional benefits like Increased Network Efficiency, Boosted network Performance, efficiency through automation, Assured reliability and network uptime, End to end Class of service assurance. Our future work will look forward for designing a routing algorithm for finding best route for controller-controller communication, controller-switch communication and communication between Mini and Mega Data centers.

REFERENCES

1. Architecting Low Latency Cloud Networks, Arista White Paper, 2016 Arista Network
2. Adoption of Cloud Computing Hinges on Application Performance, white Paper 2015 ViaviSolutions
3. An Internap White Paper
4. White Paper, Pragmatic Network Latency Engineering Fundamental Facts and Analysis, RonyKay, Ph.DPresident/CTO Packet Networks
5. Jose Saldana, David de Hoz, JuliánFernández-Navajas, José Ruiz-Mas, “Small-Packet Flowsin Software Defined Networks: Traffic Profile Optimization”, Journal Of Networks, Vol. 10, No. 4,April 2015.
6. Hyunhun Cho , Jinhyung Park , Joon-Min Gil , Young-SikJeong and Jong Hyuk Park , “AnOptimal Path Computation Architecture for the Cloud-Network on Software-Defined Networking”,Mdpi journal-sustainability , ISSN 2071-1050, 2015.
7. Rui (Ray) Zhou,”Datacenter Network Large Flow Detection and Scheduling from the Edge”,Reading& Research Project - Spring 2014,
8. ShuhaoLiu_y, Hong Xu_, ZhipingCaiy, Low Latency Datacenter Networking: A Short Survey,Dec2013.
9. Kun-chanLan, John Heidemann,A measurement study of correlations of Internet flow characteristics,Computer Networks,Volume 50, Issue 1,2006,Pages 46-62,ISSN 1389-1286,https://doi.org/10.1016/j.comnet.2005.02.008.
10. Application Note, Alcatel Lucent” SDN Analytics For Elephant Flow Marking An InherentScalable Solution For The Enterprise”, 2014.
11. Derrick D’souza , Krishna PrabhuSundharan, SavithruLokanath, Vivek Mittal, “ ImprovingQoS in a Software-Defined Network”, Capstone Research Paper, April 30, 2016
12. Sara Bouchenak, Gregory Chockler, Hana Chockler, Gabriela Gheorghe, NunoSantos, Alexander Shraer, “Verifying Cloud Services: Present and Future” ACM
13. SelvarajKesavan, J.Jayakumar , “ Network Performance Analysis of Cloud BasedMultimedia Streaming Service” , International Journal of New Computer Architectures and theirApplications (IJNCAA) 4(3): 156-166 The Society of Digital Information and WirelessCommunications, 2014 (ISSN: 2220-9085)

14. T. H. Szymanski, "Demonstration of an ultra-low latency energy efficient internet for cloud computing," 2014 IEEE 11th Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, 2014, pp. 1122-1123, doi: 10.1109/CCNC.2014.6994433."
15. Mohit Mathur, Mamta Madan, Kavita Chaudhary, A Satiated Method for Cloud Traffic Classification in Software Defined Network Environment, International Journal of Cloud Applications and Computing Volume 6 • Issue 2 • April-June 2016.
16. Yi Yaoa, Bo Shengb, Ningfang Mia, "A new packet scheduling algorithm for access points incrowded WLANs", Ad Hoc Networks. Elsevir.
17. M. J. Carmel Mary Belinda* and E. Kannan, "D2RCR-Dynamic Distributed Route Level Content based On-Body Routing in Distributed Cloud Environment", Indian Journal of Scienceand Technology, Vol 9(3), DOI: 10.17485/ijst/2016/v9i3/75971, January, 2016.
18. AlessioBotta , Walter de Donato, Valerio Persico, Antonio Pescapé, "Integration of Cloudcomputing and Internet of Things: A survey", Future Generation Computer Systems, ELsevir.,October 2015.
19. Cisco: Cisco visual networking index: Forecast and methodology, 2011–2016 (May 2012)
20. Tobias Hoßfeld, Raimund Schatz, Ernst Biersack, and Louis Plissonneau, "Internet VideoDelivery in YouTube: From Traffic Measurements to Quality of Experience" E. Biersack et al.(Eds.): Data Traffic Monitoring and Analysis, LNCS 7754, pp. 266–303, 2013, Springer-VerlagBerlin Heidelberg 2013.
21. Thesis, Lan Wang, "Traffic and Task Allocation in Networks and the Cloud",SubmittedtoElectrical and Electronic Engineering of Imperial College London and the Diploma of ImperialCollege London, january 2017.
22. Ahmed Abdelsalam, Michele Luglio, CesareRoseti, Francesco Zampognaro, "TCP Wave: Anew reliable transport approach for future internet", Computer Networks 112 (2017) 122–143, Elsevir
23. Dr Barry M Cook, C Paul H Walker, Low-Latency Packet Delivery In SpacewireNetworks,Session: Networks and Protocols , Long Paper 4Links Limited, UK
24. Roseti, Cesare& Kristiansen, Erling. (2008). TCP Noordwijk: TCP-Based Transport Optimized for Web Traffic in Satellite Networks. 10.2514/6.2008-5524.
25. Bin-Lei CaiRong-Qi Zhang , Xiao-Bo Zhou ,Lai-Ping Zhao ,Ke-Qiu Li ," ExperienceAvailability: Tail-Latency Oriented Availability in Software-Defined Cloud Computing", Journal OfComputer Science And Technology 32(2): 250–257 Mar. 2017. Doi 10.1007/S11390-017-1719- X
26. Adrian Taut Iustin A. Ivanciu Eduard Luchian Virgil Dobrota, Active Measurement Of TheLatency In Cloud-Based Networks, ActaTechnicaNapocensis Electronics AndTelecommunications, Volume 58, Number 1, 2017
27. White Paper, "Network Performance between Geo-Isolated Data Centers", 4th jan 2015.Neal Cardwell,Yucung Cheng, Eric Dumazet, "TCP Options for Low Latency: Maximum ACK
28. Delay and Microsecond Timestamps. IETF97: Seoul, Nov 2016.
29. MohitMathur. Elucidation of Upcoming Traffic Problems in Cloud Computing. Springer's (LNCS), CCIS series Vol. 90 in"Recent Trends in Networks and communications",ISBN978-3-642-14492-9 (Print) 978-3-642-14493-6 (Online), pages 68-79.
30. MohitMathur, Comprehensive Solution to Cloud Traffic Tribulations. , International Journal ofWeb Service Computing (IJWSC) ISSN: 0976 - 9811 (Online); 2230 - 7702 (print) Vol.1, No.2,December 2010 issue.
31. MohitMathur, MamtaMadan. Cloud Network Management Model- A Novel Approach tocloud Computing. IJCCSA, ISSN 2231-5853 (Online) 2231-6663 (Print) October 2014, Volume 4, Number 5.
32. MohitMathur, NitinSaraswat, Vijay B Aggarwal, Comprehensive Cloud Incremental Data-Application Migration – A Proposed Model for Cloud Migration, International Journal of ComputerApplications in Engineering Sciences, Volume III, Issue I, ISSN: 2231-4946, paper ID -IJCAESCSE-2013-009
33. White Paper , Network Performance Between Geo-Isolated Data Centers, Cloud Spectator,2015
34. upunKamburugamuve, Saliya Ekanayake, Milind a Pathirage, Geoferry Fox, "TowardsHigh Performance Processing of Streaming Data in Large Data Centers", Parallel and DistributedProcessing Symposium Workshops, 2016 IEEE International
35. PrabhjotKaur ,Reena Rani "Distributed and Cloud Computing Architecture" Imperial Journalof Interdisciplinary Research, Volume -3, Issue-2, 2017
36. Steen Larsen, ParthasarthySarangam, Ram Huggahalli "Architectural Breakdown of End toEnd Latency in TCP/IP Network", Computer Architecture and High Performance Computing,2007. SBAC-PAD 2007. 19th International Symposium
37. TingWang, ZhiyangSu, YuXiaBoQin,MounirHamdi "Towards Cost-effective and Low LatencyData Center Network Architecture" Computer Communications, elsevir Volume 82, 15 May 2016,Pages 1-12.
38. Qiaofeng Qin, KonstantinosPoularakis, George Iosifidis, LeandrosTassiulas "SDN Controller Placement at the Edge: Optimizing Delay and Overheads", IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, DOI : 10.1109/INFOCOM.2018.8485963
39. Fan Zhao, XIangyangLuo, Yong Gan, ShoodiZu, Qingfeng Cheng, Fenlin Liu, " IP Geolocation based on identification routers and local dely distribution similarity", Concurrency and Computation: Practice and Experience Published by John Wiley & Sons, Ltd., July 2018, https://doi.org/10.1002/cpe.4722
40. YehiaElkhatib, Barry Porter, Heverson B. Ribeiro, Mohamed fatenZhani, JunaidQadir, Etienne Riviere, " On Using Micro Clouds for deliver the Fog" , Published in: IEEE Internet Computing (Volume: 21 , Issue: 2 , Mar.-Apr. 2017), DOI: 10.1109/MIC.2017.35
41. DING Shichang, LUO Xiangyang, YEDengpan, LIU Fenlin Delay-Distance Correlation Study for IP Geolocation, Wuhan University Joual of Natural Sciences,2017, Vol.22 No.2, 000-000,DOI 10.1007/s11859-017-0000-0
42. Ruixiang Li , Yuchen Sun, Jianwei Hu, Te Ma, and XiangyangLuo, "Street-Level Landmark Evaluation Based on Nearest Routers" , Wiley, HindawiSecurity and Communication NetworksVolume 2018, Article ID 2507293, 12 pageshttps://doi.org/10.1155/2018/2507293
43. Shanhe Yi, Cheng Li, Qun Li," A Survey of Fog Computing: Concepts, Applications andIssues", Conference: Proceedings of the 2015 Workshop on Mobile Big Data (Mobidata '15) at Hangzhou, China, DOI: 10.1145/2757384.2757397
44. Gilles A S KeupondJo, Nogbou G. Anoh, Joel C Adepo, SouleymaneOumtanaga, " Hybrid Routing With Latency Optimization in ISDN Networks", Journal of Engineering Science and Technology, Vol. 14, , no.5(2019) 3062-3072.
45. Issued Paper, Pierre, National Resources Defense Council, America's Data Centers Consuming and Wasting Growing Amounts of Energy, February 06, 2015
46. M. S. Bali and S. Khurana, "Effect of latency on network and end user domains in Cloud Computing," 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), Chennai, 2013, pp. 777-782.
47. Adam Wolfe Gardon and Paul lu, "Low Latency Caching for Cloud –based Web applications," Department of Computer Science, University of Alberta, Edmonton, Alberta, Canada: Awalfe.Paul@cs.ualberta. , Sept. 16, 2011.
48. Ajith Singh and Hemalatha, "Comparative analysis of Low latency on different bandwidth and geographical locations while using cloud based applications," Head department of Software systems, Kalpagam university Coimbatore: IJAET ISSN: 2231-1963, Jan 2012.
49. Mohammad Heidari,"The Role of Modeling and Simulation in Information Security the Lost Ring", Springer, 1989, vol. 61.
50. Ankush Veer Reddy,"Usage of Opnet I.T tool to Simulate and test the security of cloud"(Project id-395) www.sci.tamucc.edu
51. Sonia and Satinder pal Singh."Analysis of Energy Consumption in Different types of networks For Cloud Environment" IJARCSSE Vol 2, Issue 2, Feb 2012 ISSN: 2277 128X
52. Tim Verbelen, Pieter Simoens, Filip De Turck, Bart Dhoedt, "Cloudlets: Bringing the cloud to the mobile user",Published in 2012 in Ghent Belgium by Ghent University Department of Information technology
53. Pier Luigi Ventre, Mohammad Mahdi Tajiki, Stefano Salsano, and Clarence Filsfils. 2018. SDN Architecture and Southbound APIs for IPv6 Segment Routing Enabled Wide Area Networks. <i>IEEE Trans. on Netw.and Serv. Manag.</i> 15, 4 (Dec. 2018), 1378–1392. DOI:https://doi.org/10.1109/TNSM.2018.2876251
54. P. L. Ventre et al., "Segment Routing: A Comprehensive Survey of Research Activities, Standardization Efforts, and Implementation Results," in IEEE Communications Surveys & Tutorials, vol. 23, no.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

1, pp. 182-221, Firstquarter 2021, doi: 10.1109/COMST.2020.3036826.
55. Haisheng Yu, Keqiu Li, Heng Qi, and Tom Chen. 2019. An Active Controller Selection Scheme for Minimizing Packet-In Processing Latency in SDN. <i>Sec. and Commun.Netw.</i> 2019 (2019). DOI:https://doi.org/10.1155/2019/1949343
56. Ashrafi, Mohammad & Al-Tam, Faroq&Correia, Noélia. (2019). Placement of Controllers in Software Defined Networking Under Multiple-Controller Mapping. KnE Engineering. 10.18502/keg.v5i6.7056.
57. Mamta Madan, A Tandon, Meenu dave Aug(2021)," Issues and threats in cloud network security" in Adavnces and application in mathematical sciences, Mili Publications, ISSN:0974- 6803vol no:20 , Issue no:10, pg no2077-2083 ,ESCi, Ugc-care
58. Mamta Madan, Nand P, Aneja j(2021), " Auto vectorization capabilities of the compilers , proceedings of 3rd International conference on computing , informatics and network. Lecture Notes in networks and systems, Vol 167, springer Singapore, https://doi.org/10.1007/978-981-15-9712-1_55
59. Mamta Madan, Madan R, (Sept 2013), " Optimizing Time Cost Trade off Scheduling by" International Journal of Application or Innovation in Engineering & Management (IJAIEM) , ISSN 2319 – 4847.
60. Mamta Madan, Madan R, (Aug 2013), " GASolver-A Solution to Resource Constrained Project Scheduling by Genetic Algorithm " in (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013 , ISSN : 2156-5570(Online).

AUTHOR PROFILE

Mr. Mohit Mathur, a true academician is an Associate Professor, Department of Information Technology, Jagan Institute of Management Studies, Delhi, India. He Completed his Graduation from Delhi University and Masters from National Institute of Electronics and Information Technology, Department of Electronics, Ministry of Information Technology, India.. He carries 20+ years of Rich Teaching Experience including teaching undergraduates and Post Graduate students including international students as well as professionals working in IT industry. He also carries experience of mentoring International students studying at Stratford University. His areas of Interests are Computer Networks/ Security/ Cloud Computing. He is Pursuing research on " A Technique for Optimal Cloud Traffic Delivery". His publication covers different aspects of cloud like security, traffic management, traffic classification, scalability, migration, software Defined Cloud Networks etc. He is a professional member of international organizations such as ISOC, AASCIT, IAENG, UACEE, IGI Global. He completed fellowship programme from Indian School of Internet Governance, inSIG2021 organized by ISOC India during November 2021. He also pursued summer faculty Research Fellowship Programme in the area of Cloud Computing at IIT Delhi during year 2018

Professor Mamta Madan is an accomplished professor of Computer Science. She started her journey in academics by graduating with B.Sc (Hons) Physics from Delhi University with excellent academic record and was awarded by National Graduate Physics Society Association at Delhi University for getting first position in the college. She further studied at Jamia Milia Islamia and holds dual Master degrees, in Computer Application (MCA) and Business Administration (MBA). Further in her pursuit of education she completed her Master of Philosophy (M.phil) and Doctoral research (P.hd.) in Computer Science from Banasthali Vidyapith – Rajasthan.Prof. Madan has over 16 years of experience in research and academics. She is associated with VIPS since inception. In recognition of her contributions in academics, Prof. Madan was honoured with the Best Teacher Award at VIPS in the year 2008-2009. During her tenure with VIPS, she also contributed significantly and spearheaded the training and placement cell for many years. Many placement drives were organised at VIPS under her guidance and students were placed in SAP, Infosys, Computer Science Corporation, HCL Technologies, Wipro, I Gate etc. Dr. Madan is also member of NAAC committee and Academic Council at VIPS.She is actively involved in research in the areas of artificial intelligence, software engineering, data mining and cloud computing. She is guiding different Ph.D students enrolled at various Indian universities. She is associated with many professional and research bodies like Central Board of Secondary Education, Computer Society of India etc. Her expertise goes well beyond the classroom, as she is in the panel of examiners at various universities and has evaluated numerous projects of computer science. She has published and presented many papers in National and International Journals of repute.

Mr. Mohit Chandra Saxena is a research scholar, pursuing Ph.D. in Computer Science & Engineering from Galgotias University with Software-defined Networks as the key area of research. He holds the position of Chief Technology Officer at Infinity Labs Ltd. A Public Limited company contributing to the technology upliftment into Software, Automation, Security, and SDWAN verticals. Mohit has 20 years of industry experience and has worked with Industry biggies like Cisco Systems, Orange, SNCH, etc. He holds an M.Tech in Computer Networking from Manipal Institute of Technology