Evaluation of Free5GC Forwarding Performance on Private and Public Clouds

Yi Liu
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
liuy311@chinatelecom.cn

Qiaoling Li
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
liql16@chinatelecom.cn

Qingping Cao
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
caoqp@chinatelecom.cn

Zhilan Huang
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
huangzhil@chinatelecom.cn

Yangchun Li
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
liyc10@chinatelecom.cn

Yongbing Fan
China Telecom Research Institute
China Telecom Co.,Ltd.
Guangzhou, China
fanyongb@chinatelecom.cn

***Abstract*— There are differences in the way of public cloud and private cloud use network performance acceleration technologies (such as DPDK and SR-IOV), which will affect the performance of the 5GC network. This paper evaluates the difference in forwarding performance between AWS (public cloud) and the operator's private cloud by deploying Free5GC for testing. Under similar conditions, the 5GC forwarding performance of the private cloud using SR-IOV is higher than that of the public cloud, which is higher than that of the private cloud using DPDK.**

Keywords—5GC, public cloud, private cloud

I. Introduction

3GPP has designed Service-based Architecture (SBA) and Service-based Interface (SBI) in 5G SA. This design of 5G core (5GC) network architecture and interface facilitates the deployment of 5GC network based on cloud.

Since 2017, a series of tests and trials of deploying 5GC based on private cloud have been gradually carried out at home and abroad. These tests and trials are typically based on carrier-built private clouds.

Since June 2020, AT&T, Dish, Telefonica Germany and Swisscom have announced the migration of 5GC network to Azure or AWS cloud（Amazon Web Services cloud）. AT&T plans to migrate 5GC network to Microsoft Azure. DISH works with Nokia to deploy 5GC network on AWS public cloud, plans to be commercially availability in 2022 Q1. Telefonica Germany works with Ericsson to deploy 5GC ToB(To Business) network on AWS cloud. Swisscom has designated AWS as the preferred public cloud provider and plans to migrate OSS/BSS and 5GC networks to AWS cloud. Ericsson and Nokia also announced an alliance with Google Cloud and AWS to jointly build cloud-native 5GCnetwork solutions. With the gradual pilot and commercialization of 5G ToB, the industry shows a strong trend of deploying 5GC network based on public cloud.

The public cloud provides cloud resources that can be applied for on-demand, users can quickly deploy services on the public cloud. Deploying 5GC on the public cloud can reduce the cost of Network Construction and O&M (Operation and Maintenance) by using the cloud service provider's resources and services. However, the characteristics of the 5GC business, such as volume, business complexity, performance indicators, and service guarantee requirements, are quite different from the public cloud's service capabilities for IT applications. On the public cloud, customized acceleration technologies are generally used and are not disclosed to users. Users can directly use resources and services according to their needs and ignore the acceleration technologies.

The private cloud built by the user has stronger customization and enhancement features in terms of business, it is generally realized by commercial products or self-developed products. It can be enhanced for 5GC services, but it requires higher R&D(research and development) capabilities and O&M teams. When deploying services on the private cloud, user need to enable the acceleration technology configuration, such as DPDK(Data Plane Development Kit) and SR-IOV(Single Root I/O Virtualization).

Whether it is based on public cloud carrying 5GC or private cloud carrying 5GC, the functions and performance of 5GCnetwork deployed on cloud are always the key features to meet business requirements. R. Mohamed et al., in literature [1], conducts preliminary field trials for SA and NSA 5G networks based on a self-contained 5G base station. Literatures [2] carried out a taxonomy and survey of cloud services that are provided by three major cloud infrastructure vendors, [3] focuses on the performance testing of Web application systems in the cloud environment. [4] conducted an experiment study evaluated Amazon, Azure, Google, and IBM cloud infrastructure providers in terms of performance under Infrastructure as a Service (IaaS) perspective. [5] analyze the 5GC network capability requirements in various scenarios, and propose the traffic impact capability exposure method. [6] explores and verifies the feasibility of deploying virtual network elements (NEs) based on public cloud. However, none of the above studies have carried out research on the forwarding acceleration method and end-to-end forwarding performance of 5GC on public and private clouds.

The main commercial products of 5GC are from Huawei, Ericsson, ZTE and Nokia. Open source products include Free5GC, OAI-CN, Open5GS and FreePCRF, etc. Free5GC is an open source 5GC network project based on 3GPP R15. It has realized NEs of NRF, UDM, UDR, AUSF, NSSF, AMF, PCF, SMF and UPF, and is keeping on improving them. Free5GC can be deployed on the public or private cloud to simulate various 5GC scenarios.

2022 IEEE Cloud Summit | 978-1-6654-5113-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/CLOUDSUMMIT54781.2022.00008

978-1-6654-5113-0/22/$31.00 ©2022 IEEE
DOI 10.1109/CloudSummit54781.2022.00008

In this paper, we conduct performance test of Free5GC on public and private clouds. The configurations of the resources

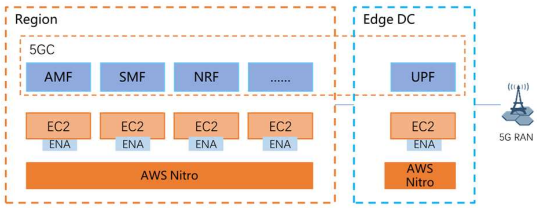

Fig. 1. 5GC architecture on Public cloud (AWS)

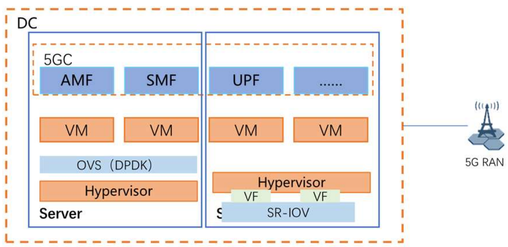

Fig. 2. 5GC architecture on Private Cloud

are similar on both clouds for fair comparison. Based on the testing result, we analyze the advantage and disadvantage of these two solutions.

The method in this paper simulates the end-to-end environment, which can be closer to the actual commercial scenario. It uses professional instruments to simulate network traffic, and measures the traffic rate of users using 5G under the simulation situation, more realistically reflect the performance impact of private cloud and public cloud acceleration technologies.

II. BACKGROUND

A. 5GC cloudifying technology architecture

We choose AWS as the public cloud example, and build private cloud by ourself. The 5GC Deployment Architecture on AWS is shown in Fig. 1. VM specifications cannot be self-defined on AWS, and different VM specifications mean different network performance. Therefore, VM specifications need to select properly. On AWS public cloud environment, we focus on the cloud services/capabilities required to deploy the 5GC, and the compatibility of the 5GC with the public cloud.

The 5GC Deployment architecture on private cloud is shown in Fig. 2. We use x86 servers to build private cloud environments, running DPDK-accelerated OVS (OVS + DPDK) or SR-IOV acceleration to provide high performance network for virtual machines (VMs). We also choose to enable NUMA affinity and anti-affinity, CPU pinning, and memory huge page.

5GC forwarding performances on public or private cloud are affected by many factors, such as VM specifications, network performances, network acceleration technology, 5GC NE's compatibility with the cloud environment, and 5GC NE deployment scheme. Testing a 5GC system relies on complete and sophisticated end-to-end performance tests by using meters.

B. Network performance acceleration technology

AWS uses self-developed Nitro hardware to provide an Enhanced network (EN). VMs access EN through an Enhanced Networking Adapter (ENA), and EN uses SR-IOV to provide high-performance networks. Performance optimization of most VM instances on AWS is provided in the cloud resource layer by default, such as memory huge page and NUMA binding. We only need to configure the VM Guest OS to use the VM. There is no explicit configuration requirements in the cloud resource layer.

The VM was created by AWS AMI and has its own ENA driver with high network bandwidth and performance.

On the private cloud, VMs can access the network by using either a vSwitch or a passthrough NIC. The network forwarding performance of vSwitch is poor due to the nature of software forwarding. Data Plane Development Kit (DPDK) is a network driver developed by Intel. It can be used in vSwitch to greatly improve network forwarding performance. SR-IOV virtualizes a physical NIC into multiple VFs, and the VF passthrough to VM to achieve network performance close to that of a physical NIC.

On private cloud there are two sets of environments respectively for DPDK-accelerated OVS and SR-IOV to test and compare.

C. Test requirements and indicators

ETSI - ES 203 539 [7] has defined the metrics and measurement methods for the energy efficiency of functional components of NFV environment. In this literature, the concepts of Userful Output are defined, and maximum capacity of the system under test (SUT) is depending on the different functions. SUT is not confined to some NFV scenarios. So in ES 203 539, Userful Output may have different meanings for different systems under test.

In 5GC network scenario, Useful Output can be a measure of throughput for user-side functionality, such as bandwidth (BPS), packet rate (PPS), and so on. Useful Output can also be the measurement of the capacity of the control plane, such as Subscribers (Sub), Simultaneously Attached Users (SAU), or PDU session number. In 3GPP TS 28.554[8], end-to-end KPI (Key Performance Indicators) of 5G system is defined, including Mean registered Subscribers, Registered Subscribers of network through UDM, PDU session Establishment, etc. These can be considered useful outputs for 5GC.

In summary, 5GC systems are large and complex. Performance indicators can not be measured by a single dimension/metric, but rely on complete and complex end-to-end performance testing based on meters. This paper verifies the forwarding performance of 5GC user plane and presents it with relatively intuitive performance indicators. This environment used to verify the 5GC forwarding performance is constructed based on instrument simulation terminal, base station and Data Network and carrying out surrounding test.

III. EXPERIMENTAL DESIGN

A. Test method

In order to test and evaluate the forwarding performance of 5GC on public cloud and on private cloud environment, this paper designs a test method based on the following principles:

- Align configurations of the public cloud and private cloud: Ensure that the 5GC environment configurations of public and private cloud are the same as possible or comparable;
- Compare different network acceleration technologies: Test the network performance acceleration technology specific to the public cloud (such as AWS's ENA) and private cloud (DPDK and SR-IOV) to compare their performance;
- Adopt a close-to-real user traffic model: refer to the actual traffic model of users in the existing network to simulate common user Internet access scenarios such as user access to video websites, instant messaging software (QQ), Weibo, etc., covering HTTP, UDP and other network protocol traffic , which can reflect not only the traffic forwarding capability of the user plane, but also the packet loss of the network;
- Intuitive and visible result data presentation: Present the 5GC forwarding performance result data from the perspective of the user side, the final result data is the total user upload and download speed, and the packet loss in the corresponding situation.

The test methods and steps used in this paper are as follows:

- Select comparable VM specifications on the public cloud and private cloud environment respectively, and deploy Free5GC;
- The public cloud and the private cloud use network performance acceleration technology to improve VM performance (ENA for AWS, DPDK-accelerated OVS and SR-IOV for the private cloud);
- Spirent Landslide meters are used to model user traffic by emulating end users, base stations, and data networks(DN);
- Final result takes the average value of user upload and download speed and packet loss within 2 hours. The sum of the upload and download rates constitutes the throughput. The upload/download packet loss result is the percentage of the upload/download packet loss to the upload/download packets.

Limited by the measured object, instrument, environment and other factors, the test is carried out only for the forwarding performance (bps) and packet loss.

The test method in this paper is used to verify the forwarding performance of a set of 5GC network elements, can also be used for the test of multiple sets of 5GC networks.

B. *Test topology*

In this paper, we choose AWS as the public cloud, using the cloud resources and cloud services of AWS Ningxia Region to deploy Free5GC. We build the private cloud environment in the laboratory and deploy Free5GC based on the self-developed virtualization software with a variety of private cloud optimization methods.

The version 3.1.0 of Free5GC is used in this test. There are 9 network elements in current version. Each network element is deployed with 1 VM, and there is 1 VM for independent database and 1 VM for WebConsole. 11 VMs in total are used in this test. In addition, there are 3 VMs for Spirent 5GC test software, with version Landslide 21.0.5.

Free5GC is deployed according to the official documentation, but the default parameters of UPF's BUFBLK POOL, NUM OF 5G UE, PFCP XACT POOL

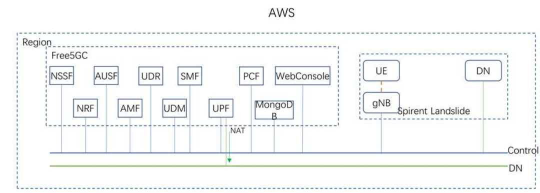

Fig. 3. Test Logical Topology on AWS

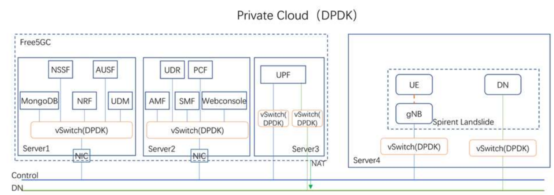

Fig. 4. Test Logical Topology on Private Cloud (DPDK)

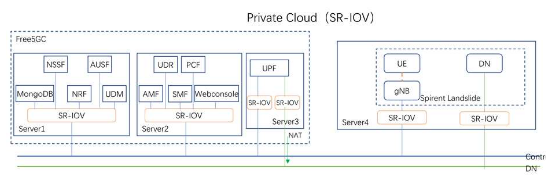

Fig. 5. Test Logical Topology on Private Cloud (SR-IOV)

and AMF's MaxNumOfPDUSessions are not enough to support this test (errors such as "pool is empty" will be reported when there is a large number of users), Therefore, the corresponding value is adjusted to satisfy the test. Other than that, Free5GC will not be optimized.

All VMs are deployed in the same Region (limited by the conditions, ignoring edge nodes for now) on AWS, each VM is configured with 2 network interfaces, one for Control Plane and the other for Data Plane (Spirent Landslide and UPF Connect to Data Network). The logical topology is shown in Fig. 3.

The private cloud environment uses 4 Sugon I620-G30 physical servers to deploy 5GC based on enhanced self-developed virtualization software, and the physical network links are all 10Gbps. 3 physical servers with 11 VMs to deploy Free5GC ; another physical machine deploys VMs to install Spirent 5GC test software(Landslide). In DPDK acceleration scenario, each VM has 2 vNICs, and the connection is enabled with DPDK-accelerated OVS, one vNIC is used for Control Plane and the other is used for Data Plane. The logical topology is shown in Fig. 4.

Similarly, in SR-IOV acceleration scenario, each VM has 2 SR-IOV VFs as vNICs, which are used to connect Control Plane and Data Plane respectively. The logical topology is shown in Fig. 5.

C. *VM specification selection*

The VM specifications and optimization settings on private cloud and AWS are shown in TABLE I.

TABLE I. COMPARISON OF VM SPECIFICATIONS

configuration	Private cloud		AWS		
	t1	*t2*	*c5.xlarge*	*c5.2xlarge*	*c5.9xlarge*
vCPU	4,CPU affinity, NUMA affinity	8,CPU affinity, NUMA affinity	4	8	36
vCPU model	Intel Xeon Silver 4114 CPU @ 2.20GHz	Intel Xeon Silver 4114 CPU @ 2.20GHz	Intel Xeon Platinum 8124M CPU @ 3.00GHz	Intel Xeon Platinum 8275CL CPU @ 3.00GHz	Intel Xeon Platinum 8124M CPU @ 3.00GHz
Memory (GB)	8,Huge Page	16,Huge Page	8	16	72
Hard disk(GB)	80,local disk, RAW	80,local disk, RAW	80,EBS	80,EBS	80,EBS
Network card model	Intel X520	Intel X520	——	——	——
Number of network interfaces and bandwidth	2vNIC, 10Gbps	2vNIC, 10Gbps	2 Network Interface, up to 5Gbs, ENA	2 Network Interface, up to 10Gbps. ENA	2 Network Interface, 10Gbps, ENA
Guest OS	Ubuntu 18.04		EC2 image(AMI): Ubuntu Server 18.04 LTS(HVM)		

AWS has many instance types, which cannot be customized and correspond to the network bandwidth. According to the available instance types and the business characteristics of 5GC, three x86-based "computing intensive" instance type, c5.xlarge, c5.2xlarge and c5.9xlarge are selected in this test (in fact, "network acceleration" is more suitable, but it is temporarily unavailable in AWS Ningxia region). The first two specifications are common VM specifications, which meet the requirements of deploying Free5gc without causing too much waste of resources, but there is no obvious guarantee in terms of network bandwidth. The network bandwidth of these two types of VM has baseline bandwidth, and they need to compete for higher bandwidth, up to 10Gbps; The third specification far exceeds the requirements of Free5GC in terms of vCPU and memory, but 10Gbps is guaranteed in terms of network bandwidth. These three specifications can better evaluate whether different CPU cores and memory sizes have a great impact on the performance of Free5gc, as well as the network bandwidth performance of different AWS instance specifications.

According to official information, c5.xlarge c5.2xlarge and c5.9xlarge instances have the following specs: Custom 2nd generation Intel Xeon Scalable Processors (Cascade Lake) with a sustained all core Turbo frequency of 3.6GHz and single core turbo frequency of up to 3.9GHz or 1st generation Intel Xeon Platinum 8000 series (Skylake-SP) processor with a sustained all core Turbo frequency of up to 3.4GHz, and single core turbo frequency of up to 3.5 GHz.

In private cloud environment, Sugon I620-G30 physical server configuration is: 2-way 10 core Intel 4114 2.2GHz CPU, 384G DDR4 2666 memory, 2 * 1.2TB 10K SAS 2.5-inch 12Gb hard disk, 2G cache 8 port 12Gb RAID card (0/1/5) Raid card, 2 * 2 GE electric port, 4*10GE optical port.

VM specifications in private cloud environment can be aligned with AWS's c5.xlarge and c5.2xlarge, but cannot be aligned with c5.9xlarge. The c5.9xlarge specification has far exceeded the requirements of Free5GC, the laboratory use 8vCPU and 16G memory can also be compared with c5.9xlarge to a certain extent from the perspective of resource utilization and network bandwidth.

There are inevitable performance differences between private cloud server hardware and AWS hardware (platinum 8124m running single thread is about 22% higher than Intel 4114 [9] [10]), A preliminary horizontal comparison can be made when the VM specifications and the maximum bandwidth limit of the network are kept as consistent as possible.

D. *User traffic model*

The user traffic model refers to the service test model of 5GC in the telecommunications industry. User traffic is mixed traffic, including HTTP web, HTTPS, DNS, UDP and other services, with an average packet length of 600~700 bytes.

The test is based on the 5GC business process, including user registration online - establishing PDU - sending user data. The data flow starts to transmit after all users register and establish PDU. The specific test is carried out with 100, 500, 1000 and 2000 users (each user establishes one PDU) respectively, for 2 hours, to test the total user upload and download speed and the corresponding packet loss. The four user numbers can properly present the actual capacity of the service, and the maximum traffic model can also be close to the network bandwidth of 10Gbps in theory.

The configuration and flow model of Spirent test instrument are as follows:

- TS Data Gen Performance: MAX
- Test Activity: Capacity Test(with Data Traffic)
- Number of Subscribers：100/500/1000/2000
- Number of PDUs(1 PDU per subscriber) ：100/500/1000/2000
- Number of gNB：5
- Number of Network Host：2
- Traffic type: HTTP & UDP
- Traffic per subscribers:
  - TX: 2089kbps(UDP 567.2kbps, HTTP 1522.2kbps),422pps(UDP 175pps, HTTP 247pps)
  - RX: 3948kbps(UDP 1617.2kbps, HTTP 2331.6kbps),696pps(UDP 375pps, HTTP 321pps)
- Run Time：2 hours

IV. ANALYSIS OF EXPERIMENTAL RESULTS

The test result is shown in TABLE II. , including 18 data records, representing total user Upload Speed (Gbps), total user Upload Packet Loss, total user Download Speed (Gbps), and total user Download Packet Loss. Each record corresponding to one test combination of selected instance specification and user numbers for AWS and private cloud.

Each test combination consists of instance specification, acceleration technology (private cloud), and number of users.

For example, c5.xlarge-100 means the c5.xlarge instance type on AWS is used, with 100 users, , abbreviated as “x-1”, “x” stands for c5.xlarge (similarly “2x” for c5.2xlarge), “1” stands for 100 users (similarly “10” for 1000 users); “t1-DPDK-100” means the t1 VM of the private cloud environment, with 100 users, ,abbreviated as “t1-d-1”, “d” stands for DPDK acceleration (similarly “s” stands for SR-IOV), and “1” stands for 100 users.

Overall, the Download Speed is higher than the Upload Speed, which matches the user traffic model. The Download Packet Loss is basically lower than the Upload Packet Loss, which also conforms to the working mechanism of the Client-Server mode and the HTTP+UDP mixed traffic mode. After the upstream HTTP+UDP packet loss, the download traffic will control the transmission speed according to the packet loss thus alleviate the Download Packet Loss.

TABLE II. TEST RESULT

Cloud	combination abbreviation	test combination	upload speed(Gbps)	upload packet loss	download speed(Gbps)	download packet loss
AWS	x-1	c5.xlarge-100	0.19	0.02150%	0.31	0.00000%
	x-5	c5.xlarge-500	0.49	3.19160%	0.98	2.93652%
	x-10	c5.xlarge-1000	0.57	18.95854%	1.05	20.81703%
	x-20	c5.xlarge-2000	0.54	59.86822%	0.73	49.76065%
	2x-1	c5.2xlarge-100	0.18	0.02244%	0.30	0.00000%
	2x-5	c5.2xlarge-500	0.74	0.20366%	1.33	0.10279%
	2x-10	c5.2xlarge-1000	0.84	2.16766%	1.93	1.77251%
	2x-20	c5.2xlarge-2000	0.79	43.67198%	1.74	22.48665%
	9x-1	c5.9xlarge-100	0.18	0.02188%	0.29	0.00000%
	9x-5	c5.9xlarge-500	0.85	0.03542%	1.39	0.00000%
	9x-10	c5.9xlarge-1000	1.32	0.10080%	2.60	0.00046%
	9x-20	c5.9xlarge-2000	1.55	16.32732%	3.74	1.46873%
lab	t1-d1	t1-DPDK-100	0.19	0.01925%	0.31	0.00248%
	t1-d5	t1-DPDK-500	0.43	0.02152%	0.89	5.75343%
	t1-d10	t1-DPDK-1000	0.60	0.76911%	1.02	36.23947%
	t1-d20	t1-DPDK-2000	0.34	73.61922%	0.43	4.68441%
	t1-s1	t1-SR-IOV-100	0.19	0.02216%	0.31	0.00000%
	t1-s5	t1-SR-IOV-500	0.88	0.03211%	1.42	0.00015%
	t1-s10	t1-SR-IOV-1000	1.50	0.07236%	2.66	0.02424%
	t1-s20	t1-SR-IOV-2000	0.75	50.98940%	1.11	0.38345%
	t2-d1	t2-DPDK-100	0.19	0.01998%	0.31	0.00314%
	t2-d5	t2-DPDK-500	0.42	0.08001%	0.90	6.23223%
	t2-d10	t2-DPDK-1000	0.59	1.94263%	1.00	36.41853%
	t2-d20	t2-DPDK-2000	0.29	76.70284%	0.33	5.63156%
	t2-s1	t2-SR-IOV-100	0.18	0.02347%	0.31	0.00000%
	t2-s5	t2-SR-IOV-500	0.87	0.03219%	1.41	0.00030%
	t2-s10	t2-SR-IOV-1000	1.50	0.07280%	2.66	0.00052%
	t2-s20	t2-SR-IOV-2000	1.25	24.79809%	1.75	1.66278%

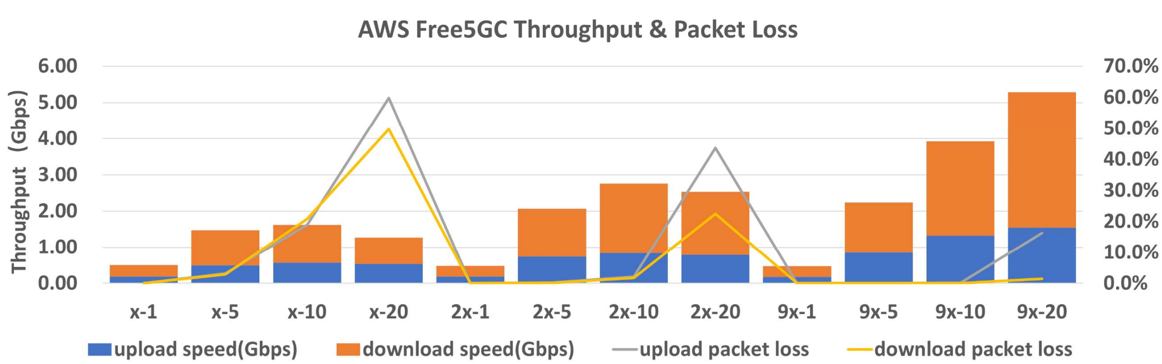

Fig. 6. AWS Free5GC Throughput & Packet Loss

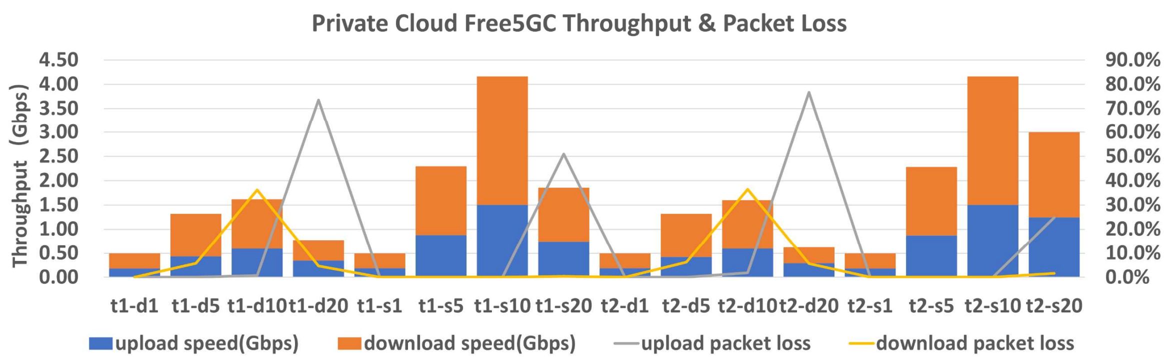

Fig. 7. Private Cloud Free5GC Throughput & Packet Loss

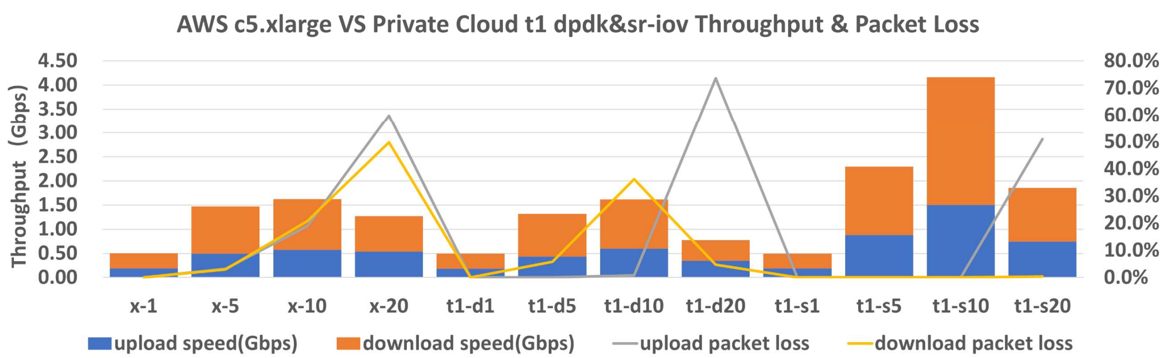

Fig. 8. AWS c5.xlarge VS Private Cloud t1 DPDK & SR-IOV Throughput & Packet Loss

A. *The impact analysis of AWS VM instance types on Free5GC performance*

The performance of Free5GC on AWS is shown in Fig. 6. Throughput is the sum of user upload and download speed. The forwarding performance of Free5GC shows an obvious growth trend with the increase of VM instance type. From 4vCPU 8G memory to 8vCPU 16G memory to 36vCPU 72G memory, the average forwarding performance is improved by 51.71% and 116.31% respectively. The packet loss decreases significantly as the size of the VM instance increases. Since there is a strong correspondence between the instance type of AWS VM and network bandwidth, the bandwidth of c5.xlarge (4vCPU 8G memory) is "up to 5Gbs", the bandwidth of c5.2xlarge (8vCPU 16G memory) is "up to 10Gbs", and the bandwidth of c5.9xlarge (36vCPU 72G memory) is "10Gbps", the increase in instance type also brings improvements in network bandwidth and performance.

B. *The impact analysis of private cloud VM specification and acceleration technology on Free5GC performance*

The performance of Free5GC on private cloud is shown in Fig. 7. The performance in the case of SR-IOV network acceleration is higher than that of DPDK OVS, with an average increase of about 154.07%.Different from AWS, the impact of the VM specification on performance is not obvious, when using the same network acceleration technology (such as DPDK OVS), the average difference is less than 0.5%. But in the case of 2000 users and with SR-IOV, the throughput of t2-s20 is improved by 61.51%compared to that of t1-s20. Unlike AWS, private cloud VM specifications are not directly related to network bandwidth. Based on this analysis, under the same network bandwidth guarantee, the impact of VM

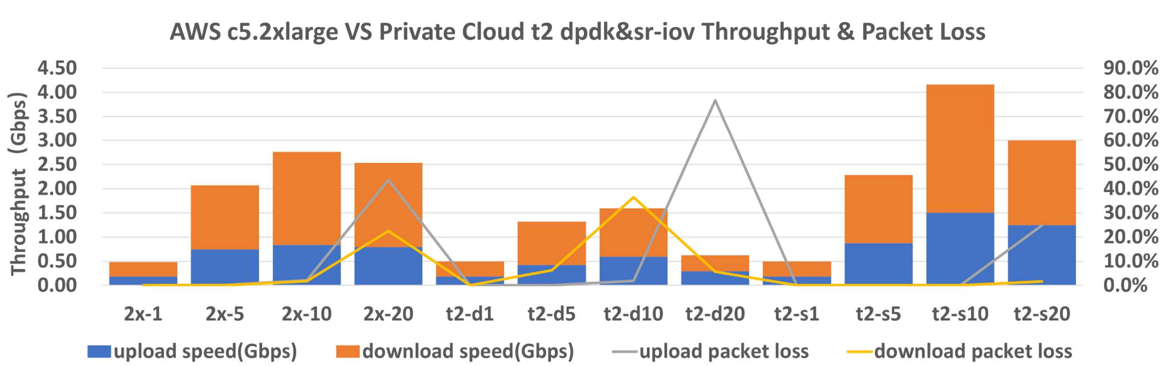

Fig. 9. AWS c5.2xlarge VS Private Cloud t2 DPDK & SR-IOV Throughput & Packet Loss

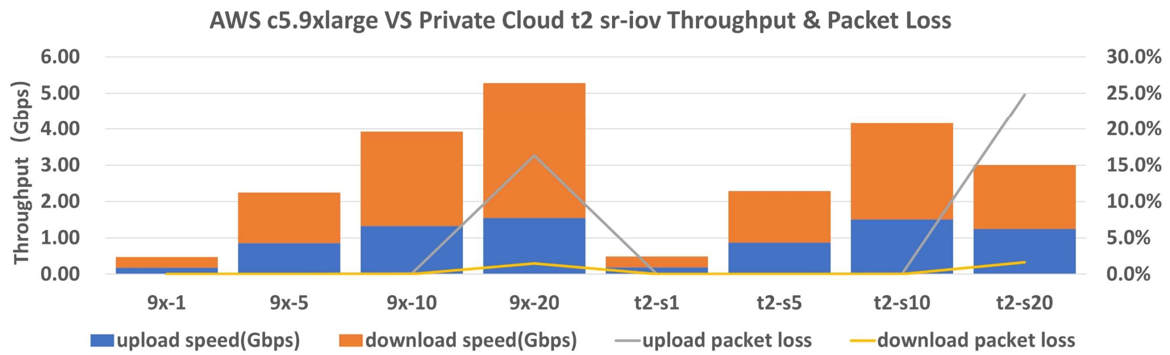

Fig. 10. AWS c5.9xlarge VS Private Cloud t2 sr-iov Throughput & Packet Loss

specifications on forwarding performance is small, Only when a large number of users (corresponding to large forwarding performance requirements) are achieved and there is still space for improvement in the network acceleration technology (such as t1-s20 and t2-s20), upgrading the VM specifications is effective.

C. *Performance analysis of Free5GC on AWS and private cloud*

The performance of Free5GC with c5.xlarge on AWS and t1 DPDK and t1 SR-IOV on private cloud is shown in Fig. 8. The performance of c5.xlarge is slightly higher than t1 DPDK, but the performance of t1 SR-IOV is significantly higher than that of c5.xlarge.

The performance of Free5GC with c5.2xlarge on AWS and t2 DPDK and t1 SR-IOV on private cloud is shown in Fig. 9. On AWS, the performance of the VM is improved after the VM instance type is increased. On private cloud, the VM instance type has little impact on performance. Therefore, the performance of c5.2xlarge is obviously better than that of t2 DPDK, which is about 108% higher on average. But the performance of t2 SR-IOV is still higher than that of c5.2xlarge, about 20% higher on average. And the packet loss of t2 SR-IOV is also lower than that of c5.2xlarge.

The performance of Free5GC with c5.9xlarge on AWS and with t2 SR-IOV on private cloud is shown in Fig. 10, c5.9xlarge performs better under large number of users. Below 1000 users, the performance difference between c5.9xlarge and t2 SR-IOV is less than 6%. In the case of 2000 users, the gap is reflected. The throughput of c5.9xlarge can still go up, while the performance of t2 SR-IOV has dropped significantly, which is lower than that of 1000 users. The throughput of c5.9xlarge is 43.23% higher than that of t2 SR-IOV. The packet loss is also similar. There is no significant difference below 1000 users. The c5.9xlarge has an advantage with 2000 users.

In general, with the same conditions, there is no insurmountable gap in the performance of Free5GC on public cloud and private cloud. Both have their own advantages and disadvantages:

- On public cloud (AWS), performance is affected by VM instance type and network bandwidth. The higher the instance type / network bandwidth, the higher performance. However, the increase in instance type also has an upper limit on the improvement of forwarding performance, which will also lead to higher usage costs (price, waste of resources, etc.). the actual use requires a reasonable selection of instance type;
- On private cloud, different acceleration technology have a greater impact on forwarding performance, and SR-IOV acceleration significantly improves performance, compared with DPDK, it has increased by more than 70%, up to 300%..
- Under the same conditions, the 5GC forwarding performance on the private cloud using SR-IOV is slightly better than that on the public cloud, with an average of 20.31% higher. In the case of high enough VM specifications, the performance of the public cloud can surpass the private cloud using SR-IOV by about 43%, but it also brings about 66% of the waste of resources.
- Compared with public cloud, the private cloud using SR-IOV has performance advantages, higher flexibility, and free configuration of resources. Public cloud has limited VM specifications, but public cloud resources can be used on demand, and user can ignore the underlying technical details, there is no need to configure various acceleration or binding means like a private cloud.

There are inevitable differences in the hardware of AWS and private cloud in this test, but the performance gap is not obvious when the VM size are the same and the network bandwidth is the same. Since AWS ENA uses SR-IOV by default, which is roughly the same as the private cloud SR-IOV technology, while AWS uses Nitro customized hardware optimization, and the private cloud optimizes the hypervisor layer for 5GC, the performance gap between the two is not large. However, there are different focuses between public cloud and private cloud.. The public cloud focuses on how to optimize the adaptation while using limited specifications (limited specifications, low compatibility with the business model), while the private cloud focuses on virtual Layer optimization and performance acceleration technology.

V. CONCLUSION

From the test results, Free5GC performs better on VMs with higher instance type (also with higher bandwidth) on AWS. However, AWS's instance type lacks flexibility and selectivity, and can only be selected according to limited type, network performance instances type is not available in AWS BeiJing region and NingXia region. To ensure high network performance, higher instance type must be selected, which may lead to waste of resources. 5GC deployment needs to comprehensively consider resource allocation and performance requirements.

The performance of Free5GC on the operator's private cloud is also not bad. Even the hardware is not as good as AWS, with DPDK-accelerated OVS,t1-DPDK's performance can almost match the same size of AWS VM in 1000 users. In the case of using SR-IOV, it's performance surpasses the same size of AWS VM, and can also match the higher size of AWS VMs in 1,000 users. The operator's private cloud can effectively improve network performance and better carry 5GC and other services through customized optimization and using DPDK-accelerated OVS /SR-IOV.

This test has certain limitations. The Free5GC project has implemented functions, but has not optimized performance. In this test, Free5GC has not been optimized in a targeted manner, and only do some adjustments and modifications to ensure the smooth progress of the test. There is also a gap between the test environment and the actual 5GC commercial environment. There is also a clear gap between the AWS and the private cloud hardware. However, through horizontal comparison, we can still see the performance of Free5GC on both AWS and operator private cloud environments to a certain extent.

At present, commercial 5GC is still dominated by private cloud, and the private cloud case is relatively mature; the public cloud solution is still being explored, and there are few cases. This paper compares the performance differences between the two clouds by setting up a simulation environment, which can provide a simple reference for the deployment of 5GC on the public cloud.

On the public cloud, network performance acceleration configuration is not provided for users at the cloud resource layer, which is less flexible, but it also shields technical details, reduces the complexity of user configuration, and is more convenient to use. The private clouds have stronger autonomy, and user need to enable network acceleration configuration, which requires higher user team capabilities. In terms of performance, the deployment of large-scale 5GC is recommended to give priority to private cloud using SR-IOV. For small-scale 5GC services that do not care about the acceleration of underlying performance but have high performance requirements and need rapid deployment, it may be appropriate to consider deploying them on the public cloud. However, in commercial scenarios, 5GC deployment not only needs to consider performance issues, but also needs to consider many factors such as O&M, user data security, and system interconnection.

Private cloud hosting 5GC and public cloud hosting 5GC face similar but also different problems. This test is just a simple performance comparison. In practical applications, the 5GC deployment scale is larger, the architecture is more complex, and there are more details that can be optimized in terms of performance configuration. In the follow-up, we will continue to study the differences between private cloud and public cloud carrying 5GC, and further explore the differences in functions, performance, system connection, operation system, etc. of 5GC on private cloud and public cloud. Free5GC itself also has a lot of space for performance improvement, and the future can be expected.

References

[1] R. Mohamed, S. Zemouri and C. Verikoukis, "Performance Evaluation and Comparison between SA and NSA 5G Networks in Indoor Environment," 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), 2021, pp. 112-116, doi: 10.1109/MeditCom49071.2021.9647621.

[2] A. Alkhatib, A. Al Sabbagh and R. Maraqa, "Pubic Cloud Computing: Big Three Vendors," 2021 International Conference on Information Technology (ICIT), 2021, pp. 230-237, doi: 10.1109/ICIT52682.2021.9491680.

[3] H. Li, X. Li, H. Wang, J. Zhang and Z. Jiang, "Research on Cloud Performance Testing Model," 2019 IEEE 19th International Symposium on High Assurance Systems Engineering (HASE), 2019, pp. 179-183, doi: 10.1109/HASE.2019.00035.

[4] D. B. Oliveira, R. R. de Oliveira, R. F. Vilela, V. H. S. C. Pinto and R. N. Ungarelli, "Experimental Performance Evaluation Among Cloud Infrastructure Providers Under Different Load Levels," 2021 XLVII Latin American Computing Conference (CLEI), 2021, pp. 1-10, doi: 10.1109/CLEI53233.2021.9639943.

[5] L. Lin, B. Zhu, Q. Wang, L. Xu and J. Mu, "A Novel 5G Core Network Capability Exposure Method for Telecom Operator," 2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), 2020, pp. 1450-1454, doi: 10.1109/ISPA-BDCloud-SocialCom-SustainCom51426.2020.00217.

[6] D. Kumar, S. Chakrabarti, A. S. Rajan and J. Huang, "Scaling Telecom Core Network Functions in Public Cloud Infrastructure," 2020 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), 2020, pp. 9-16, doi: 10.1109/CloudCom49646.2020.00006.

[7] ETSI，"Measurement method for energy efficiency of Network Functions Virtualisation (NFV) in laboratory environment"， 1 June 2019, 7

[8] 3gpp, "3GPP TS 28.554 5G end to end Key Performance Indicators (KPI)", 2022, 9-42

[9] CPU114, Intel Xeon Platinum 8124M @ 3.00GHz, EB/OL, 2018, 2022, http://www.cpu114.com/cpu/3352/Intel+Xeon+Platinum+8124M+@+3.00GHz

[10] CPU114, Intel Xeon Silver 4114 @ 2.20GHz, EB/OL, 2017, 2022, http://www.cpu114.com/cpu/3095/Intel+Xeon+Silver+4114+@+2.20GHz