JUNIPER NETWORKS | Engineering Simplicity

White Paper

HOW TO MEASURE PERFORMANCE OF A VIRTUAL FIREWALL ACROSS PUBLIC CLOUDS

Deploying security in the public cloud must account for performance and scale of the virtual firewall

©2020, Juniper Networks, Inc.

TABLE OF CONTENTS

©2020, Juniper Networks, Inc.

EXECUTIVE SUMMARY

Network solution architects designing hybrid cloud networks must choose the right type of virtual instance from the many available in public cloud domains to ensure the experience is not compromised as users transition to the public cloud.

This user experience hinges on throughput and latency—two key parameters that need to be measured and used when planning and deploying the network. Unfortunately, network performance in the public cloud has its ups and downs. For various reasons—most out of the administrator's control—the public cloud infrastructure does not offer the same predictability and visibility that most network and security administrators get with their in-house deployments.

The problem is exacerbated by a lack of reliable and consistent tools for measuring performance and scalability in the public cloud. Contrast this with on-premises networks, where a number of third-party hardware and virtual tools from many vendors are available to measure the performance and scalability of the network and security equipment being deployed.

This white paper explores the key parameters that administrators need to understand in order to accurately measure the performance of a virtual firewall. Essential among them is an understanding of the virtual firewall's architecture. Instances offering a clear demarcation between the forwarding plane and the control plane must be preferred over those that do not have this separation.

This paper also examines the Juniper Networks® vSRX Virtual Firewall, which features unmatched performance per core, provides additional flexibility when deployed in the public cloud by utilizing all available vCPUs and memory, increasing the performance and scale of key measurement parameters as the capacity of the selected instance type increases.

Finally, this paper provides guidance on the methods used to measure the throughput of virtual firewall instances in public cloud environments.

Introduction

Digital transformation is requiring organizations to deploy their workloads across multiple clouds, including on-premises clouds spread across multiple geographical locations and, increasingly, in public clouds. As network engineers plan the transition to the public cloud, they are faced with difficult decisions about how to design the network in the cloud while providing the necessary security and visibility.

Public cloud providers such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) advertise the reduced burden of maintaining the underlying host machines and the security of the host operating system (OS) among the many benefits of moving to the public cloud. However, public cloud deployments also introduce their own complexities, one of which is throughput considerations and the resulting user experience.

©2020, Juniper Networks, Inc.

Designing the Cloud Environment

One of the prime considerations in the design phase is the consistency of the end-user experience, regardless of where the workload resides. Consistency is the result of reliable latency and network throughput. There is a huge need for repeatable and dependable methods for measuring the performance of a virtual machine (VM) hosted in the public cloud. The design process, however, is extremely challenging because of:

- The multitude of instance types offered by the cloud provider, each with its own cost and performance implications
- The changes in the expected performance of networks deployed in the public cloud

Contrast this with the private cloud or on-premises data centers, where all hosts and the network are under the purview of a professional conducting the measurement. Let's take a look at the different environments for performance tests between the public and private cloud:

Workload Deployment Across Multicloud

"Multicloud" is typically used to describe enterprises that deploy their workloads and applications in different cloud environments, ranging from private on-premises data centers to public data centers. These public cloud deployments can be from multiple providers based on the requirements and capabilities of each.

Table 1. Public and Private Cloud Environments

Public Cloud	Private Cloud
It is hard to manage specific details of the deployment, such as non-uniform memory access (NUMA) socket and network interface card (NIC) mapping where deployed instances are abstracted.	Physical servers, connectivity, and the deployment specifics for the guest OS can be controlled.
BIOS settings are nonnegotiable; hence, optimal settings for throughput based on the virtual instance are hard to achieve. This means that measured performance is suboptimal.	Since the servers are in-house, all settings for optimal throughput—such as hyperthreading, huge page sizes, etc.—can be set and controlled.
Interconnectivity between two VM instances is not clear. Are they hosted on the same server? Are they on different physical hosts? How many switches are being traversed in the path? All of these are either unknown or difficult to ascertain.	Every detail of the topology that impacts performance measurement is under the administrator's control and is therefore easy to manage and change to achieve high throughput.
Traffic generator tools such as iPerf3 and NTttcp themselves are a function of the virtualization settings, similar to the device being measured for throughput.	The traffic generator can be a physical device such as those provided by Keysight, Spirent, and others, guaranteeing that the packets are sent at a prescribed rate.

As shown in the table, the challenges of measuring throughput in the public cloud can lead to a potentially poor network design based on inaccurate expectations of network performance.

High-throughput virtual firewalls and routers do not perform the same when deployed in the public cloud. Some of this degraded performance can be blamed on the Infrastructure as a Service (IaaS) network. For instance:

- The AWS cloud claims a maximum throughput of up to 100 Gbps[1] with the c5n.18xL instance type. It offers 72 vCPUs to achieve this throughput.
- Microsoft's Azure cloud offers up to 30 Gbps[2] of network bandwidth for some of its high vCPU count general purpose (x86) server instances.
- Google Cloud caps the throughput at 32 Gbps[3] for some of its standard instance types.

Juniper Networks has a long history of offering high-performance routing, switching, and security solutions. This continued with the release of the vSRX Virtual Firewall, which provides high-performance firewall performance per vCPU core and increases the number of concurrent sessions as the amount of memory allocated to the vSRX increases. Even with the impressive performance of the vSRX on a standalone server, there are a number of

[1] For more information on AWS cloud performance claims: https://aws.amazon.com/ec2/instance-types/
[2] For more information Microsoft Azure Cloud performance claims: https://docs.microsoft.com/en-us/azure/virtual-network/virtual-machine-network-throughput and https://docs.microsoft.com/en-us/azure/virtual-machines/dv3-dsv3-series?toc=/azure/virtual-machines/linux/toc.json&bc=/azure/virtual-machines/linux/breadcrumb/toc.json
[3] For more information on Google Cloud performance claims: https://cloud.google.com/blog/products/gcp/5-steps-to-better-gcp-network-performance?hl=ml and https://cloud.google.com/compute/docs/machine-types

©2020, Juniper Networks, Inc.

considerations for measuring the performance of VMs in the public cloud:

1. The architecture of the firewall or software router product that is to be deployed must be well understood. Control and user plane separation is a fundamental architectural concept that lends itself to higher performance when compared to a monolithic architecture. Dedicated data plane vCPUs allow for traffic processing and forwarding, unburdened by control plane activities in a distributed architecture. The vSRX dedicates separate vCPUs for flow threads, ensuring that forwarding performance is maximized.
2. It is also important to understand the difference between scale and performance. Scale of any feature is usually a function of the memory available on the system. Performance is a function of the CPU. The architecture and speed of the processor determines the throughput capability of the device. The demarcation between scale and performance requirements will help scope the decision on the instance type from the IaaS marketplace.
3. The instance types available with the IaaS will need to be matched with the flavor of the VM selected for deployment. Not all VMs will utilize the vCPUs available to them. For example, the AWS c5.2xL offers eight vCPUs and 16 GB memory; if the VM chosen for deployment only uses four vCPUs and 12 GB memory, the instance will be underutilized. The vSRX does not have this limitation; it uses all the vCPUs and memory available in the instance across all public clouds.
4. When measuring throughput using instances in public clouds, the instance types of the traffic generators need to be chosen carefully. They must be capable of generating and sending traffic that is likely to stretch the limits of the device under test (DUT).

   For example, a two vCPU instance type running Linux should not be used to measure the performance of a 32 vCPU DUT. It is better to overprovision the traffic generators and receivers than underprovision them.
5. Across all public clouds, iperf3 and NTttcp are the preferred tools for throughput measurements. It always helps to read up on the documentation offered by IaaS providers looking for tips related to network throughput measurements.

   Parameters such as maximum segment size (MSS), warm-up time, ability to use all cores during traffic generation, and so on are explained in detail in such documents. The open-source tools iperf3 and NTttcp do not offer the flexibility and ease of use of a commercial third-party traffic generator tool, but they are used extensively and are the best available options for throughput measurements.
6. Iperf3 and NTttcp can also be used to measure latency in the network. Physical distance between the regions where the virtual private clouds (VPCs) are housed does have an impact on latency, just as it would in private cloud deployments. Most IaaS providers have multiple geographically diverse regions available to them. Choose the regions for your deployment based on the workloads and the sensitivity to latency when accessing these workloads. Finally, the creation of a test topology in the public cloud environment that closely mimics the final deployment scenario is highly recommended. There is no substitute for actually deploying the VM to be tested and running traffic through it.

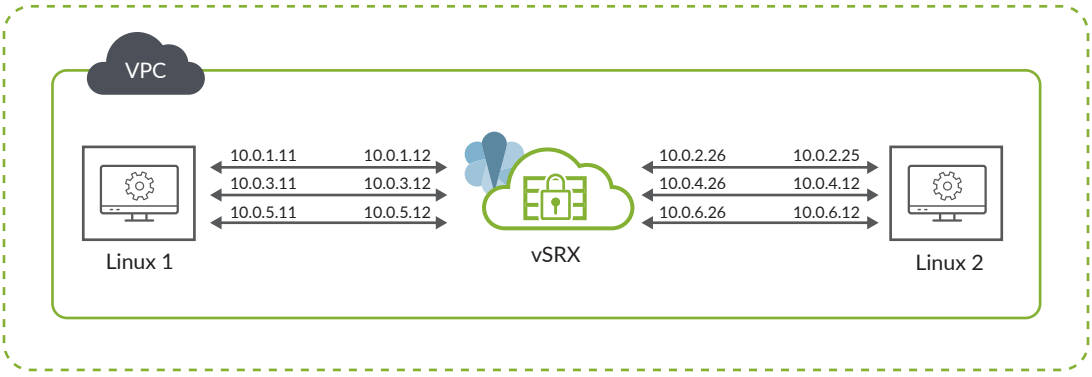

Figure 1: Sample topology for measuring performance with multiple interfaces

©2020, Juniper Networks, Inc.

When performing the test (and later, during deployment), it is important to ensure that the vCPUs in the VM are being efficiently utilized. This can be achieved by:

1. Using multiple interfaces and multiple IP addresses/ports, so the receive-side scaling feature is able to spray the packets to multiple forwarding vCPUs to ensure they are all being used.
2. Monitoring the CPU utilization of the data plane vCPUs to ensure that the best possible performance is achieved. Sometimes, the virtual firewall may be capable of processing more packets and yielding higher throughput, but is limited by the bandwidth offered by the instance type selected. For example, Google Cloud Platform (GCP)[4] caps the throughput at 2 Gbps per vCPU and at 32 Gbps per instance type for the n1-standard series of instances. AWS[5] lists the network throughput available for each instance, as does Azure[6].
3. Setting the maximum transmission unit (MTU) and maximum segment size (MSS) to high values and ensuring they are consistent across availability zones (AZs) and regions.
4. If measuring IPsec performance, then the IPsec tunnels need to be distributed across different vCPUs to ensure that all the vCPUs are effectively utilized in a multiple vCPU VM instance.
5. Ensuring the tests are run multiple times on different days and different times of the day so a range of values can be established. The more information available during the design phase, the better.

vSRX Virtual Firewall

The vSRX offers unparalleled next-generation firewall (NGFW) security, including intrusion prevention system (IPS), malware protection, app control, and on-demand threat detection to safeguard applications and workloads hosted in the public clouds. The virtual firewall supports communications security with Secure SD-WAN and SD-LAN for secure segmentation between workloads.

The vSRX delivers unmatched performance per core. With the ability to use software-based receive side scaling, the vSRX ensures that all vCPUs available in the instance are being used for processing the packets traversing it.

The vSRX's automated provisioning capabilities allow network and security administrators to quickly and efficiently provision and scale firewall protection to meet the dynamic needs of cloud environments. By combining the vSRX with the power of Junos Space® Security Director or Contrail® Service Orchestration, administrators can significantly improve policy configuration, management, and visibility into both physical and virtual assets from a common, centralized platform.

vSRX is available in the marketplaces at AWS, Azure, GCP, and IBM cloud as both bring-your-own-license (BYOL) and pay-as-you-go (PAYG) offerings. If the vSRX is on your short-list of virtual firewalls for your cloud deployment, try the free trial today.

Try vSRX on Microsoft Azure at https://azuremarketplace.microsoft.com/en-us/marketplace/apps/juniper-networks.vsrx-next-generation-firewall-payg?tab=Overview.

Try vSRX on Amazon Web Services at https://aws.amazon.com/marketplace/pp/B01NAUWN0G?qid=1597078076428&sr=0-1&ref_=srh_res_product_title.

Try vSRX on Google Cloud Service at https://console.cloud.google.com/marketplace/details/juniper-marketplace/vsrx-next-generation-firewall.

Conclusion

Measuring firewall throughput in the public cloud environment can be a significant challenge. Fluctuations are typical from day to day, week to week, or even year to year. By following the best practices outlined in this paper, you will be able to design a network that is high-performing and resilient, while ensuring security in the public cloud.

[4] For more information on bandwidth per instance type on GCP: https://cloud.google.com/compute/docs/network-bandwidth
[5] For more information on bandwidth per instance type on AWS: https://aws.amazon.com/ec2/instance-types/
[6] For more information on bandwidth per instance type on Azure: https://docs.microsoft.com/en-us/azure/virtual-machines/dv3-dsv3-series

©2020, Juniper Networks, Inc.

About Juniper Networks

Juniper Networks brings simplicity to networking with products, solutions, and services that connect the world. Through engineering innovation, we remove the constraints and complexities of networking in the cloud era to solve the toughest challenges our customers and partners face daily. At Juniper Networks, we believe that the network is a resource for sharing knowledge and human advancement that changes the world. We are committed to imagining groundbreaking ways to deliver automated, scalable, and secure networks to move at the speed of business.

Corporate and Sales Headquarters
Juniper Networks, Inc.
1133 Innovation Way
Sunnyvale, CA 94089 USA
Phone: 888.JUNIPER (888.586.4737)
or +1.408.745.2000
Fax: +1.408.745.2100
www.juniper.net

APAC and EMEA Headquarters
Juniper Networks International B.V.
Boeing Avenue 240
1119 PZ Schiphol-Rijk
Amsterdam, The Netherlands
Phone: +31.0.207.125.700
Fax: +31.0.207.125.701

JUNIPER NETWORKS | Engineering Simplicity

Copyright 2020 Juniper Networks, Inc. All rights reserved. Juniper Networks, the Juniper Networks logo, Juniper, and Junos are registered trademarks of Juniper Networks, Inc. in the United States and other countries. All other trademarks, service marks, registered marks, or registered service marks are the property of their respective owners. Juniper Networks assumes no responsibility for any inaccuracies in this document. Juniper Networks reserves the right to change, modify, transfer, or otherwise revise this publication without notice.