Internet Engineering Task Force (IETF)
Request for Comments: 8172
Category: Informational
ISSN: 2070-1721

A. Morton
AT&T Labs

July 2017

Considerations for Benchmarking Virtual Network Functions and Their Infrastructure

Abstract

The Benchmarking Methodology Working Group has traditionally conducted laboratory characterization of dedicated physical implementations of internetworking functions. This memo investigates additional considerations when network functions are virtualized and performed in general-purpose hardware.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc8172.

Table of Contents

1. Introduction

The Benchmarking Methodology Working Group (BMWG) has traditionally conducted laboratory characterization of dedicated physical implementations of internetworking functions (or physical network functions (PNFs)). The black-box benchmarks of throughput, latency, forwarding rates, and others have served our industry for many years. [RFC1242] and [RFC2544] are the cornerstones of the work.

A set of service provider and vendor development goals has emerged: reduce costs while increasing flexibility of network devices and drastically reduce deployment time. Network Function Virtualization (NFV) has the promise to achieve these goals and therefore has garnered much attention. It now seems certain that some network functions will be virtualized following the success of cloud computing and virtual desktops supported by sufficient network path capacity, performance, and widespread deployment; many of the same techniques will help achieve NFV.

In the context of Virtual Network Functions (VNFs), the supporting Infrastructure requires general-purpose computing systems, storage systems, networking systems, virtualization support systems (such as hypervisors), and management systems for the virtual and physical resources. There will be many potential suppliers of Infrastructure

systems and significant flexibility in configuring the systems for best performance. There are also many potential suppliers of VNFs, adding to the combinations possible in this environment. The separation of hardware and software suppliers has a profound implication on benchmarking activities: much more of the internal configuration of the black-box Device Under Test (DUT) must now be specified and reported with the results, to foster both repeatability and comparison testing at a later time.

Consider the following user story as further background and motivation:

> I'm designing and building my NFV Infrastructure platform. The first steps were easy because I had a small number of categories of VNFs to support and the VNF vendor gave hardware recommendations that I followed. Now I need to deploy more VNFs from new vendors, and there are different hardware recommendations. How well will the new VNFs perform on my existing hardware? Which among several new VNFs in a given category are most efficient in terms of capacity they deliver? And, when I operate multiple categories of VNFs (and PNFs) *concurrently* on a hardware platform such that they share resources, what are the new performance limits, and what are the software design choices I can make to optimize my chosen hardware platform? Conversely, what hardware platform upgrades should I pursue to increase the capacity of these concurrently operating VNFs?

See <http://www.etsi.org/technologies-clusters/technologies/nfv> for more background; the white papers there may be a useful starting place. The "NFV Performance & Portability Best Practices" document [NFV.PER001] is particularly relevant to BMWG. There are also documents available among the Approved ETSI NFV Specifications [Approved_ETSI_NFV], including documents describing Infrastructure performance aspects and service quality metrics, and drafts in the ETSI NFV Open Area [Draft_ETSI_NFV], which may also have relevance to benchmarking.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Scope

At the time of this writing, BMWG is considering the new topic of Virtual Network Functions and related Infrastructure to ensure that common issues are recognized from the start; background materials from respective standards development organizations and Open Source development projects (e.g., IETF, ETSI NFV, and the Open Platform for Network Function Virtualization (OPNFV)) are being used.

This memo investigates additional methodological considerations necessary when benchmarking VNFs instantiated and hosted in general-purpose hardware, using bare metal hypervisors [BareMetal] or other isolation environments such as Linux containers. An essential consideration is benchmarking physical and Virtual Network Functions in the same way when possible, thereby allowing direct comparison. Benchmarking combinations of physical and virtual devices and functions in a System Under Test (SUT) is another topic of keen interest.

A clearly related goal is investigating benchmarks for the capacity of a general-purpose platform to host a plurality of VNF instances. Existing networking technology benchmarks will also be considered for adaptation to NFV and closely associated technologies.

A non-goal is any overlap with traditional computer benchmark development and their specific metrics (e.g., SPECmark suites such as SPEC CPU).

A continued non-goal is any form of architecture development related to NFV and associated technologies in BMWG, consistent with all chartered work since BMWG began in 1989.

3. Considerations for Hardware and Testing

This section lists the new considerations that must be addressed to benchmark VNF(s) and their supporting Infrastructure. The SUT is composed of the hardware platform components, the VNFs installed, and many other supporting systems. It is critical to document all aspects of the SUT to foster repeatability.

3.1. Hardware Components

The following new hardware components will become part of the test setup:

1. High-volume server platforms (general-purpose, possibly with virtual technology enhancements)

2. Storage systems with large capacity, high speed, and high reliability

3. Network interface ports specially designed for efficient service of many virtual Network Interface Cards (NICs)

4. High-capacity Ethernet switches

The components above are subjects for development of specialized benchmarks that focus on the special demands of network function deployment.

Labs conducting comparisons of different VNFs may be able to use the same hardware platform over many studies, until the steady march of innovations overtakes their capabilities (as happens with the lab's traffic generation and testing devices today).

3.2. Configuration Parameters

It will be necessary to configure and document the settings for the entire general-purpose platform to ensure repeatability and foster future comparisons, including, but clearly not limited to, the following:

- number of server blades (shelf occupation)
- CPUs
- caches
- memory
- storage system
- I/O

as well as configurations that support the devices that host the VNF itself:

- Hypervisor (or other forms of virtual function hosting)
- Virtual Machine (VM)
- Infrastructure virtual network (which interconnects virtual machines with physical network interfaces or with each other through virtual switches, for example)

and finally, the VNF itself, with items such as:

- specific function being implemented in VNF
- reserved resources for each function (e.g., CPU pinning and Non-Uniform Memory Access (NUMA) node assignment)
- number of VNFs (or sub-VNF components, each with its own VM) in the service function chain (see Section 1.1 of [RFC7498] for a definition of service function chain)
- number of physical interfaces and links transited in the service function chain

In the physical device benchmarking context, most of the corresponding Infrastructure configuration choices were determined by the vendor. Although the platform itself is now one of the configuration variables, it is important to maintain emphasis on the networking benchmarks and capture the platform variables as input factors.

3.3. Testing Strategies

The concept of characterizing performance at capacity limits may change. For example:

1. It may be more representative of system capacity to characterize the case where the VMs hosting the VNFs are operating at 50% utilization and therefore sharing the "real" processing power across many VMs.
2. Another important test case stems from the need to partition (or isolate) network functions. A noisy neighbor (VM hosting a VNF in an infinite loop) would ideally be isolated; the performance of other VMs would continue according to their specifications, and tests would evaluate the degree of isolation.
3. System errors will likely occur as transients, implying a distribution of performance characteristics with a long tail (like latency) and leading to the need for longer-term tests of each set of configuration and test parameters.
4. The desire for elasticity and flexibility among network functions will include tests where there is constant flux in the number of VM instances, the resources the VMs require, and the setup/teardown of network paths that support VM connectivity. Requests for and instantiation of new VMs, along with releases for VMs hosting VNFs that are no longer needed, would be a normal

operational condition. In other words, benchmarking should include scenarios with production life-cycle management of VMs and their VNFs and network connectivity in progress, including VNF scaling up/down operations, as well as static configurations.

5. All physical things can fail, and benchmarking efforts can also examine recovery aided by the virtual architecture with different approaches to resiliency.

6. The sheer number of test conditions and configuration combinations encourage increased efficiency, including automated testing arrangements, combination sub-sampling through an understanding of inter-relationships, and machine-readable test results.

3.4. Attention to Shared Resources

Since many components of the new NFV Infrastructure are virtual, test setup design must have prior knowledge of interactions/dependencies within the various resource domains in the SUT. For example, a virtual machine performing the role of a traditional tester function, such as generating and/or receiving traffic, should avoid sharing any SUT resources with the DUT. Otherwise, the results will have unexpected dependencies not encountered in physical device benchmarking.

Note that the term "tester" has traditionally referred to devices dedicated to testing in BMWG literature. In this new context, "tester" additionally refers to functions dedicated to testing, which may be either virtual or physical. "Tester" has never referred to the individuals performing the tests.

The possibility to use shared resources in test design while producing useful results remains one of the critical challenges to overcome. Benchmarking setups may designate isolated resources for the DUT and other critical support components (such as the host/kernel) as the first baseline step and add other loading processes. The added complexity of each setup leads to shared-resource testing scenarios, where the characteristics of the competing load (in terms of memory, storage, and CPU utilization) will directly affect the benchmarking results (and variability of the results), but the results should reconcile with the baseline.

The physical test device remains a solid foundation to compare with results using combinations of physical and virtual test functions or results using only virtual testers when necessary to assess virtual interfaces and other virtual functions.

4. Benchmarking Considerations

This section discusses considerations related to benchmarks applicable to VNFs and their associated technologies.

4.1. Comparison with Physical Network Functions

In order to compare the performance of VNFs and system implementations with their physical counterparts, identical benchmarks must be used. Since BMWG has already developed specifications for many network functions, there will be re-use of existing benchmarks through references, while allowing for the possibility of benchmark curation during development of new methodologies. Consideration should be given to quantifying the number of parallel VNFs required to achieve comparable scale/capacity with a given physical device or whether some limit of scale was reached before the VNFs could achieve the comparable level. Again, implementation based on different hypervisors or other virtual function hosting remain as critical factors in performance assessment.

4.2. Continued Emphasis on Black-Box Benchmarks

When the network functions under test are based on open-source code, there may be a tendency to rely on internal measurements to some extent, especially when the externally observable phenomena only support an inference of internal events (such as routing protocol convergence observed in the data plane). Examples include CPU/Core utilization, network utilization, storage utilization, and memory committed/used. These "white-box" metrics provide one view of the resource footprint of a VNF. Note that the resource utilization metrics do not easily match the 3x4 Matrix, described in Section 4.4.

However, external observations remain essential as the basis for benchmarks. Internal observations with fixed specification and interpretation may be provided in parallel (as auxiliary metrics), to assist the development of operations procedures when the technology is deployed, for example. Internal metrics and measurements from open-source implementations may be the only direct source of performance results in a desired dimension, but corroborating external observations are still required to assure the integrity of measurement discipline was maintained for all reported results.

A related aspect of benchmark development is where the scope includes multiple approaches to a common function under the same benchmark. For example, there are many ways to arrange for activation of a network path between interface points, and the activation times can be compared if the start-to-stop activation interval has a generic

and unambiguous definition. Thus, generic benchmark definitions are preferred over technology/protocol-specific definitions where possible.

4.3. New Benchmarks and Related Metrics

There will be new classes of benchmarks needed for network design and assistance when developing operational practices (possibly automated management and orchestration of deployment scale). Examples follow in the paragraphs below, many of which are prompted by the goals of increased elasticity and flexibility of the network functions, along with reduced deployment times.

- Time to deploy VNFs: In cases where the general-purpose hardware is already deployed and ready for service, it is valuable to know the response time when a management system is tasked with "standing up" 100s of virtual machines and the VNFs they will host.

- Time to migrate VNFs: In cases where a rack or shelf of hardware must be removed from active service, it is valuable to know the response time when a management system is tasked with "migrating" some number of virtual machines and the VNFs they currently host to alternate hardware that will remain in service.

- Time to create a virtual network in the general-purpose Infrastructure: This is a somewhat simplified version of existing benchmarks for convergence time, in that the process is initiated by a request from (centralized or distributed) control, rather than inferred from network events (link failure). The successful response time would remain dependent on data-plane observations to confirm that the network is ready to perform.

- Effect of verification measurements on performance: A complete VNF, or something as simple as a new policy to implement in a VNF, is implemented. The action to verify instantiation of the VNF or policy could affect performance during normal operation.

Also, it appears to be valuable to measure traditional packet transfer performance metrics during the assessment of traditional and new benchmarks, including metrics that may be used to support service engineering such as the spatial composition metrics found in [RFC6049]. Examples include mean one-way delay in Section 4.1 of [RFC6049], Packet Delay Variation (PDV) in [RFC5481], and Packet Reordering [RFC4737] [RFC4689].

4.4. Assessment of Benchmark Coverage

It can be useful to organize benchmarks according to their applicable life-cycle stage and the performance criteria they were designed to assess. The table below (derived from [X3.102]) provides a way to organize benchmarks such that there is a clear indication of coverage for the intersection of life-cycle stages and performance criteria.

	SPEED	ACCURACY	RELIABILITY
Activation			
Operation			
De-activation			

For example, the "Time to deploy VNFs" benchmark described above would be placed in the intersection of Activation and Speed, making it clear that there are other potential performance criteria to benchmark, such as the "percentage of unsuccessful VM/VNF stand-ups" in a set of 100 attempts. This example emphasizes that the Activation and De-activation life-cycle stages are key areas for NFV and related Infrastructure and encourages expansion beyond traditional benchmarks for normal operation. Thus, reviewing the benchmark coverage using this table (sometimes called the 3x3 Matrix) can be a worthwhile exercise in BMWG.

In one of the first applications of the 3x3 Matrix in BMWG [SDN-BENCHMARK], we discovered that metrics on measured size, capacity, or scale do not easily match one of the three columns above. Following discussion, this was resolved in two ways:

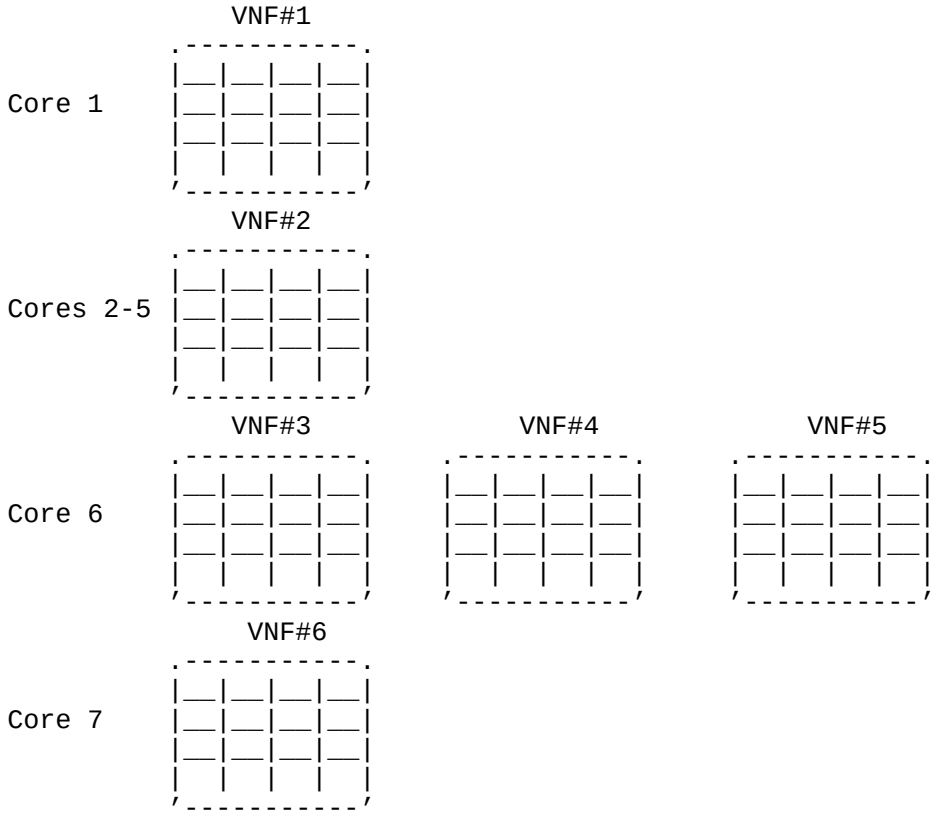
- Add a column, Scale, for use when categorizing and assessing the coverage of benchmarks (without measured results). An example of this use is found in [OPNFV-BENCHMARK] (and a variation may be found in [SDN-BENCHMARK]). This is the 3x4 Matrix.

- If using the matrix to report results in an organized way, keep size, capacity, and scale metrics separate from the 3x3 Matrix and incorporate them in the report with other qualifications of the results.

Note that the resource utilization (e.g., CPU) metrics do not fit in the matrix. They are not benchmarks, and omitting them confirms their status as auxiliary metrics. Resource assignments are configuration parameters, and these are reported separately.

This approach encourages use of the 3x3 Matrix to organize reports of results, where the capacity at which the various metrics were measured could be included in the title of the matrix (and results for multiple capacities would result in separate 3x3 Matrices, if there were sufficient measurements/results to organize in that way).

For example, results for each VM and VNF could appear in the 3x3 Matrix, organized to illustrate resource occupation (CPU Cores) in a particular physical computing system, as shown below.



The combination of tables above could be built incrementally, beginning with VNF#1 and one Core, then adding VNFs according to their supporting Core assignments. X-Y plots of critical benchmarks would also provide insight to the effect of increased hardware utilization. All VNFs might be of the same type, or to match a production environment, there could be VNFs of multiple types and categories. In this figure, VNFs #3-#5 are assumed to require small CPU resources, while VNF#2 requires four Cores to perform its function.

4.5. Power Consumption

Although there is incomplete work to benchmark physical network function power consumption in a meaningful way, the desire to measure the physical Infrastructure supporting the virtual functions only adds to the need. Both maximum power consumption and dynamic power consumption (with varying load) would be useful. The Intelligent Platform Management Interface (IPMI) standard [IPMI2.0] has been implemented by many manufacturers and supports measurement of instantaneous energy consumption.

To assess the instantaneous energy consumption of virtual resources, it may be possible to estimate the value using an overall metric based on utilization readings, according to [NFVIaas-FRAMEWORK].

5. Security Considerations

Benchmarking activities as described in this memo are limited to technology characterization of a DUT/SUT using controlled stimuli in a laboratory environment, with dedicated address space and the constraints specified in the sections above.

The benchmarking network topology will be an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network or misroute traffic to the test management network.

Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the DUT/SUT.

Special capabilities SHOULD NOT exist in the DUT/SUT specifically for benchmarking purposes. Any implications for network security arising from the DUT/SUT SHOULD be identical in the lab and in production networks.

6. IANA Considerations

This document does not require any IANA actions.

7. References

7.1. Normative References

[NFV.PER001] ETSI, "Network Function Virtualization: Performance & Portability Best Practices", ETSI GS NFV-PER 001, V1.1.2, December 2014.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, DOI 10.17487/RFC2544, March 1999, <http://www.rfc-editor.org/info/rfc2544>.

[RFC4689] Poretsky, S., Perser, J., Erramilli, S., and S. Khurana, "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", RFC 4689, DOI 10.17487/RFC4689, October 2006, <http://www.rfc-editor.org/info/rfc4689>.

[RFC4737] Morton, A., Ciavattone, L., Ramachandran, G., Shalunov, S., and J. Perser, "Packet Reordering Metrics", RFC 4737, DOI 10.17487/RFC4737, November 2006, <http://www.rfc-editor.org/info/rfc4737>.

[RFC7498] Quinn, P., Ed. and T. Nadeau, Ed., "Problem Statement for Service Function Chaining", RFC 7498, DOI 10.17487/RFC7498, April 2015, <http://www.rfc-editor.org/info/rfc7498>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <http://www.rfc-editor.org/info/rfc8174>.

7.2. Informative References

[Approved_ETSI_NFV]
ETSI, Network Functions Virtualisation Technical Committee, "ETSI NFV", <http://www.etsi.org/standards-search>.

[BareMetal]
Popek, G. and R. Goldberg, "Formal requirements for virtualizable third generation architectures", Communications of the ACM, Volume 17, Issue 7, Pages 412-421, DOI 10.1145/361011.361073, July 1974.

[Draft_ETSI_NFV]
ETSI, "Network Functions Virtualisation: Specifications", <http://www.etsi.org/technologies-clusters/technologies/nfv>.

[IPMI2.0] Intel Corporation, Hewlett-Packard Company, NEC Corporation, and Dell Inc., "Intelligent Platform Management Interface Specification v2.0 (with latest errata)", April 2015, <http://www.intel.com/content/dam/www/public/us/en/documents/specification-updates/ipmi-intelligent-platform-mgt-interface-spec-2nd-gen-v2-0-spec-update.pdf>.

[NFVIaas-FRAMEWORK]
Krishnan, R., Figueira, N., Krishnaswamy, D., Lopez, D., Wright, S., Hinrichs, T., Krishnaswamy, R., and A. Yerra, "NFVIaaS Architectural Framework for Policy Based Resource Placement and Scheduling", Work in Progress, draft-krishnan-nfvrg-policy-based-rm-nfviaas-06, March 2016.

[OPNFV-BENCHMARK]
Tahhan, M., O'Mahony, B., and A. Morton, "Benchmarking Virtual Switches in OPNFV", Work in Progress, draft-ietf-bmwg-vswitch-opnfv-04, June 2017.

[RFC1242] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, DOI 10.17487/RFC1242, July 1991, <http://www.rfc-editor.org/info/rfc1242>.

[RFC5481] Morton, A. and B. Claise, "Packet Delay Variation Applicability Statement", RFC 5481, DOI 10.17487/RFC5481, March 2009, <http://www.rfc-editor.org/info/rfc5481>.

[RFC6049] Morton, A. and E. Stephan, "Spatial Composition of Metrics", RFC 6049, DOI 10.17487/RFC6049, January 2011, <http://www.rfc-editor.org/info/rfc6049>.

[SDN-BENCHMARK] Vengainathan, B., Basil, A., Tassinari, M., Manral, V., and S. Banks, "Terminology for Benchmarking SDN Controller Performance", Work in Progress, draft-ietf-bmwg-sdn-controller-benchmark-term-04, June 2017.

[X3.102] ANSI, "Information Systems - Data Communication Systems and Services - User-Oriented Performance Parameters Communications Framework", ANSI X3.102, 1983.

Acknowledgements

The author acknowledges an encouraging conversation on this topic with Mukhtiar Shaikh and Ramki Krishnan in November 2013. Bhavani Parise and Ilya Varlashkin have provided useful suggestions to expand these considerations. Bhuvaneswaran Vengainathan has already tried the 3x3 Matrix with the SDN controller document and contributed to many discussions. Scott Bradner quickly pointed out shared resource dependencies in an early vSwitch measurement proposal, and the topic was included here as a key consideration. Further development was encouraged by Barry Constantine’s comments following the BMWG session at IETF 92: the session itself was an affirmation for this memo. There have been many interesting contributions from Maryam Tahhan, Marius Georgescu, Jacob Rapp, Saurabh Chattopadhyay, and others.

Author’s Address

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ 07748
United States of America

Phone: +1 732 420 1571
Fax: +1 732 368 1192
Email: acmorton@att.com