FPT CLOUD

Whitepaper

High-Performance OpenStack:

Khi hạ tầng Cloud Việt Nam không còn là bottleneck

FPT Smart Cloud, FPT Corporation – May 2026

fptcloud.com

FPT CLOUD

Mục lục

Bài toán hiệu năng của hạ tầng Cloud

Trong nhiều năm, trong thế giới cloud tồn tại một "sự thật ngầm hiểu".

Hạ tầng ảo hóa luôn kém xa phần cứng vật lý về hiệu năng mạng, đặc biệt với các workload xử lý nhiều packet nhỏ hoặc lưu lượng không đều. Đây không chỉ là cảm giác chủ quan, mà là một giới hạn kỹ thuật đã được đo đạc và ghi nhận rộng rãi.

Khi hệ thống phải xử lý từ hàng trăm nghìn đến hàng chục triệu packet mỗi giây, các vấn đề bắt đầu lộ rõ:

- Throughput khó scale theo băng thông NIC, thậm chí hao hụt hơn một nửa so với khả năng lý thuyết của phần cứng
- Packet rate (PPS) chạm trần rất sớm vì guest overhead, context switching trên host ở mức cao, v.v.
- Latency tăng đột biến khi bật các tính năng security, hoặc khi nhiều guest "giành" steal CPU trên cùng một host.
- Packet drop xuất hiện ở ngưỡng tải cao và hệ thống không thể cam kết SLA 99.99% packet rate với near zero-drop như các "big cloud" thường công bố.

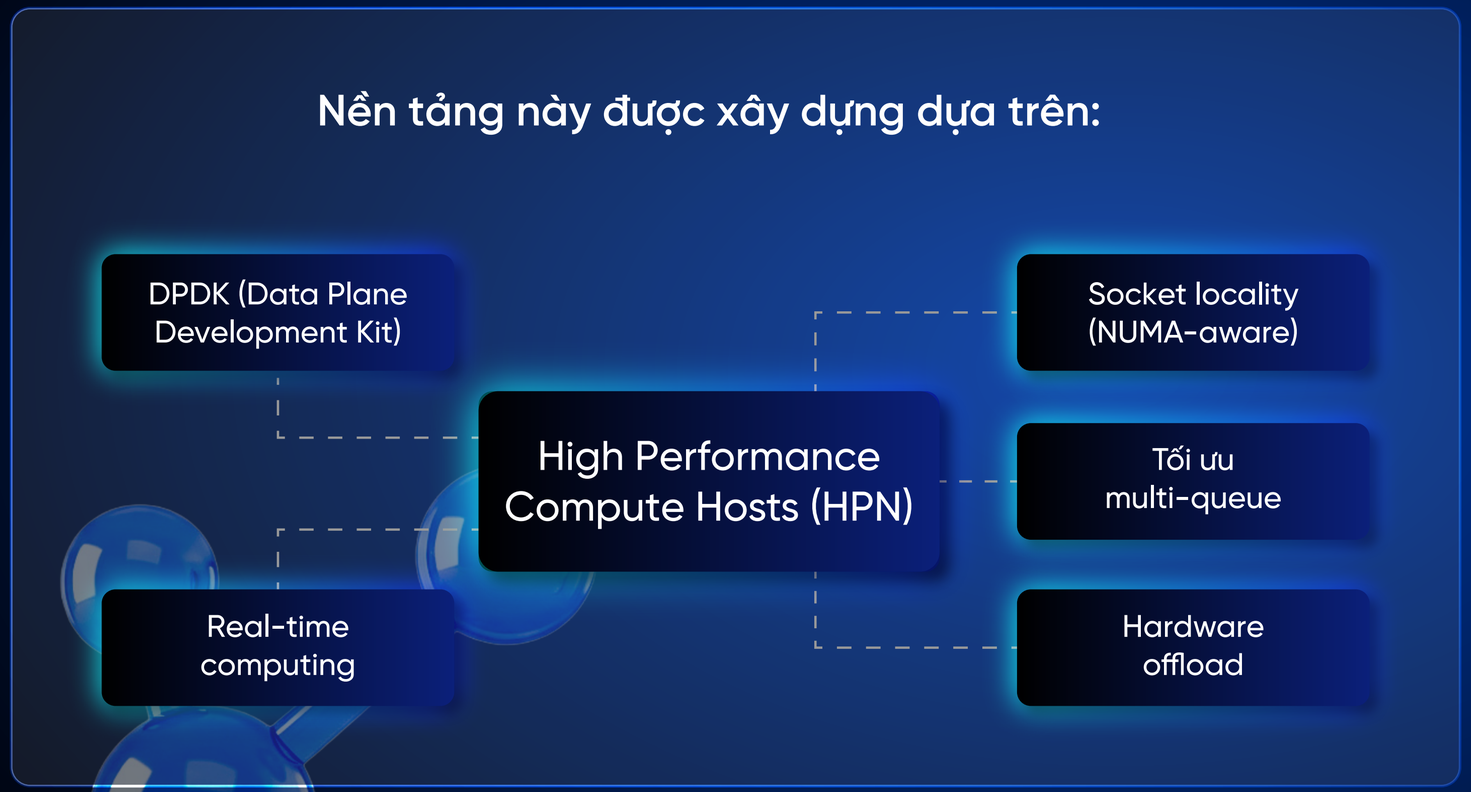
Vì vậy, các workload quan trọng như firewall, IDS/IPS hay các thành phần telco như 5G UPF, CGNAT, vRouter, load balancer, streaming application và những hệ thống edge realtime thường **không được tin tưởng chạy trên cloud nội địa.**

› Câu hỏi đặt ra là: có phải chúng ta đang dùng phần cứng lỗi thời? Thực tế không phải.

Chúng ta dùng server và NIC đời mới, thậm chí có nhà cung cấp phải "cắn răng" đầu tư switch 100G đắt đỏ để bù lại phần suy hao. Vậy tại sao VM trên cloud ở Việt Nam (và nhiều nơi trong khu vực) vẫn gặp tình trạng này?

› Điều đáng buồn là: data plane của nhiều giải pháp cloud hiện nay không được thiết kế để packet processing ở quy mô lớn.

Cách chúng tôi tiếp cận



Thay vì cố "tối ưu thêm một chút" trên nền tảng cũ, chúng tôi chọn một hướng đi triệt để hơn:

> ***Tái kiến trúc toàn bộ dataplane của OpenStack để phục vụ packet processing theo đúng chuẩn telco***

Điều quan trọng cần hiểu là đây **không phải chỉ là "bật DPDK"**. Chúng tôi tạo ra một lớp compute hoàn toàn mới: High Performance Compute Hosts (HPN), được thiết kế cho high packet processing, latency nhạy cảm và hạn chế context switching đến mức tối thiểu.

Phương pháp kiểm chứng

Rất dễ để một nhà cung cấp đưa ra một bộ SLI về hạ tầng của họ. Những con số này đôi khi đáng tin, nhưng không chắc được thu thập từ các phép đo chuẩn tắc.

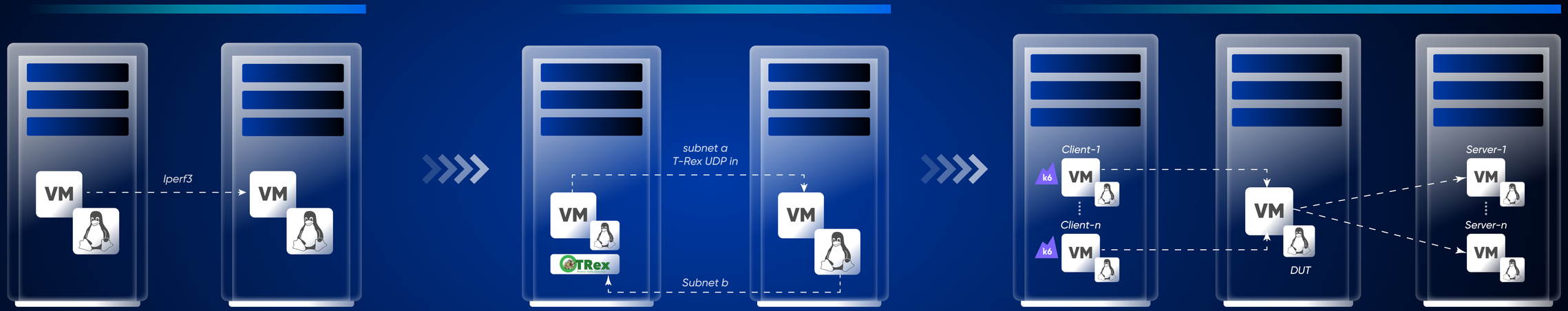
Vì vậy, chúng tôi đã dành nhiều thời gian nghiên cứu các công cụ mà telco và các cloud lớn thường dùng để đo đạc một cách lặp lại được, tiêu chuẩn hóa test case và giảm rủi ro "lệch pha" trong giai đoạn nghiệm thu các dự án. Trong bài viết này, chúng tôi dùng kết hợp T-Rex, iperf3, Grafana k6. Lý do phải dùng nhiều công cụ sẽ được tóm tắt như sau:

Một sai lầm phổ biến khi đánh giá hiệu năng network là dựa hoàn toàn vào **iperf3**. Công cụ này phụ thuộc nhiều vào TCP stack và hành vi của application. Trong khi đó, các hệ thống telco và realtime không quan tâm nhiều đến application layer. Họ quan tâm duy nhất: **Hạ tầng xử lý được bao nhiêu packet mỗi giây, độ trễ bao nhiêu, và drop có vượt ngưỡng hay không**.

Đó là lý do chúng tôi dùng **Cisco T-Rex** như một công cụ benchmark bổ sung, nhằm tiêu chuẩn hóa test case trong bài.

T-Rex cho phép sinh traffic ở L2/L3, giảm bias từ application, kiểm soát chính xác packet size, flow, PPS, và đo đúng năng lực forwarding thực tế của hệ thống.

Dưới đây là 3 mô hình benchmark chính được sử dụng. Mỗi khối ánh xạ tới một server compute vật lý, kết nối qua bond gồm 2 port uplink. Để tránh đi quá sâu, tôi sẽ không vẽ chi tiết carrier flow từ layer switching vào layer uplink.



Iperf Layer 2
(đường iperf chỉ đi trong một network namespace)

T-Rex Layer 3
(hai subnet A và B nằm trong hai network namespace riêng)

Grafana k6 Layer 3
(Ingress từ client và egress tới server nằm ở hai network namespace riêng)

Nói cách khác, đây là cách telco và các vendor lớn validate năng lực NFV của hạ tầng.
Chúng tôi sẽ đính kèm evidence từ iperf và Grafana k6, vốn có sẵn từ các ticket PoC thực tế của khách hàng.

Ba trụ cột hiệu năng cần đánh giá

Trước khi đi vào đo đạc, cần xác định rõ "đo cái gì". Thay vì chỉ nói "nhanh hơn", chúng tôi đánh giá dựa trên ba yếu tố cốt lõi:

Throughput

Packet Rate

Latency

Đây là ba yếu tố quyết định trực tiếp đến trải nghiệm và độ ổn định của workload. Ở góc nhìn của kỹ sư cloud, đôi khi cần trade-off giữa các yếu tố này, nhưng phần trade-off và tuning sẽ không nằm trong phạm vi bài viết.

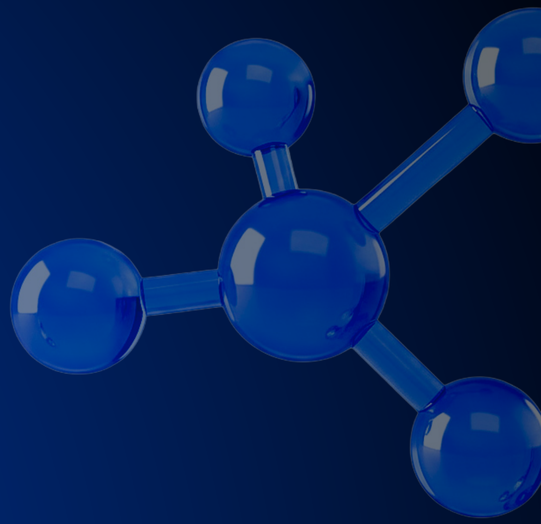
1. Throughput — Khi virtual network đạt line-rate thực sự

Trong môi trường thực tế, throughput cao không chỉ là "đạt peak", **mà còn là duy trì ổn định dưới tải.**
Tại một AZ ở Hà Nội, chúng tôi dùng NIC/port 10G và DUT là Cisco Catalyst 8000V (16 vCPU / 32 GB RAM) chạy trên KVM. Hệ thống đạt:

~6.4 Gbps sustained throughput (Layer 3) *khi dồn tải với packet nhỏ 64B.*

Chúng tôi cũng benchmark các nhà cung cấp khác tại Việt Nam và ghi nhận con số khoảng

~2.3 Gbps Layer 3 *trên VM KVM thông thường trong một cụm production đã có tải, ở điều kiện tương đương (mạng ảo, không tính VLAN hay public network).*



› *Điểm đáng chú ý là 6.4 Gbps Layer 3 đã tiệm cận giới hạn lý thuyết của vhost/virtio trong môi trường 10G.*

Evidence iperf3 Layer 2

```
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[  5]   0.00-1.00   sec  1.09 GBytes  9.35 Gbits/sec    0   3.16 MBytes
[  5]   1.00-2.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   2.00-3.00   sec  1.09 GBytes  9.36 Gbits/sec    0   3.16 MBytes
[  5]   3.00-4.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   4.00-5.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   5.00-6.00   sec  1.09 GBytes  9.37 Gbits/sec    0   3.16 MBytes
[  5]   6.00-7.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   7.00-8.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   8.00-9.00   sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
[  5]   9.00-10.00  sec  1.09 GBytes  9.38 Gbits/sec    0   3.16 MBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[  5]   0.00-10.00  sec  10.9 GBytes  9.38 Gbits/sec    0             sender
[  5]   0.00-10.04  sec  10.9 GBytes  9.34 Gbits/sec                  receiver

iperf Done.
```

Băng thông tầng network ảo hóa hao hụt không quá 10% là một yếu tố ít được nhìn thấy trên thị trường cloud nước ta mặc dù nó đã được đề cập trong chuẩn thông tư 1145 của Chính phủ.

Evidence Grafana k6 Layer 3

Đường ingress và egress gần như trùng nhau, hầu như không có dropped. Có thể tính throughput theo công thức:

Throughput = packet size × (pps subnet 1 + pps subnet 2)

Tại một AZ khác ở TP.HCM, hạ tầng dùng NIC/port 25G. Khi chạy workload thực tế với iperf3 multi-queue từ nhiều VM nhỏ **(mỗi VM chỉ 4 vCPU và 4 GiB RAM)**, đặt trên các host vật lý khác nhau, throughput đạt: **~25 Gbps sustained (Layer 2)**

Cần nhấn mạnh việc VM nằm trên các host khác nhau, vì các loại mạng ảo như VXLAN, Geneve, v.v. thường có overhead rõ rệt khi traffic đi liên-host.

Tôi có thể giải thích thêm nguyên nhân, nhưng có thể tóm gọn: encapsulate/decapsulate gói tin rất "nặng" và đôi lúc mất kiểm soát khi cụm cloud đã vào production.

Trong test này, VM của chúng tôi không xuất hiện hiện tượng "collapse" khi tăng tải. Traffic được giữ ổn định và không thấy suy hao nghiêm trọng.

Evidence iperf Layer 2

```
[ ID] Interval           Transfer     Bitrate         Retr
[  5]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  38324             sender
[  5]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[  7]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  7889              sender
[  7]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[  9]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  10184             sender
[  9]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[ 11]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  71524             sender
[ 11]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[ 13]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  13970             sender
[ 13]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[ 15]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  6545              sender
[ 15]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[ 17]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  30971             sender
[ 17]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[ 19]   0.00-60.00  sec  21.6 GBytes  3.09 Gbits/sec  6467              sender
[ 19]   0.00-60.04  sec  21.6 GBytes  3.09 Gbits/sec                  receiver
[SUM]   0.00-60.00  sec   173 GBytes  24.7 Gbits/sec  185874              sender
[SUM]   0.00-60.04  sec   173 GBytes  24.7 Gbits/sec                  receiver

iperf Done.
```

Một điểm thú vị là iperf vẫn có retransmission. Tuy nhiên, nếu retransmit khá cao mà packet drop vẫn được khống chế ở khoảng 0.001% tổng PPS, thì overhead nhiều khả năng nằm ở network stack của VM.

Nói cách khác, dataplane chuyển mạch của hạ tầng vẫn chưa bị quá tải. Hệ thống còn dư địa để tăng sức chịu tải, và khách hàng có thể giảm retransmission bằng cách tăng tài nguyên VM hoặc tối ưu application.

Chốt lại hai ý quan trọng

Dataplane không còn là bottleneck.

Bottleneck đã được đẩy lên application layer, đúng chỗ nó nên nằm.

Trên cloud truyền thống, điều ngược lại thường xảy ra. Và mỗi lần khách hàng "claim" về hạ tầng, nhà cung cấp lại phải tốn thêm OPEX để scale đội ngũ hỗ trợ.

2. Packet Rate — Yếu tố quyết định thực sự

Throughput cao không đồng nghĩa hệ thống "khỏe". Thứ quyết định là **packet mỗi giây** – nhất là với packet nhỏ hoặc mixed traffic.

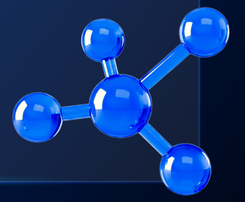
Với benchmark T-Rex và DUT là Cisco C8000V, chúng tôi thu được kết quả:

So với hệ thống không dùng DPDK (**~240-260 Kpps**), đây là mức tăng **6-8 lần**.

Hơn nữa, chúng tôi đã tùy chỉnh dự án open-source **VyOS** để hướng tới một sản phẩm NFV độc quyền trong thời gian tới.

Các test T-Rex UDP thực tế tới DUT VyOS custom (32 core, 32 GiB RAM) đạt kết quả mà chúng tôi chưa từng thấy ở bất kỳ nhà cung cấp NFV nào tại Việt Nam:

- *Hệ thống đạt hơn 12 triệu packet/giây (12+ Mpps, giới hạn bởi hệ thống phát tải)*
- *Không có packet drop tại peak*
- *CPU utilization: ~92% (polling ổn định, không jitter)*



Evidence T-Rex UDP Layer 3

```
Global Statistics

connection   : localhost, Port 4501                       total_tx_L2  : 6.31 Gbps
version      : STL @ v3.08                                total_tx_L1  : 8.28 Gbps
cpu_util.    : 92.49% @ 13 cores (13 per dual port)       total_rx     : 6.2 Gbps
rx_cpu_util. : 0.0% / 0 pps                               total_pps    : 12.32 Mpps
async_util.  : 0% / 120.81 bps                            drop_rate    : 0 bps
total_cps.   : 0 cps                                      queue_full   : 1,170,127,002 pkts

Port Statistics

   port    |         0         |         1         |       total
-----------+-------------------+-------------------+------------------
owner      |              root |              root |
link       |                UP |                UP |
state      |      TRANSMITTING |      TRANSMITTING |
speed      |          200 Gb/s |          200 Gb/s |
CPU util.  |             92.49% |             92.49% |
--         |                   |                   |
Tx bps L2  |         3.16 Gbps |         3.16 Gbps |         6.31 Gbps
Tx bps L1  |         4.14 Gbps |         4.14 Gbps |         8.28 Gbps
Tx pps     |         6.16 Mpps |         6.16 Mpps |        12.32 Mpps
Line Util. |            2.07 % |            2.07 % |
---        |                   |                   |
Rx bps     |         3.11 Gbps |         3.09 Gbps |          6.2 Gbps
Rx pps     |         6.07 Mpps |         6.04 Mpps |        12.11 Mpps
----       |                   |                   |
opackets   |       5295100722 |       5295147488 |       10590248210
ipackets   |       5201502958 |       5184552840 |       10386055798
obytes     |     493126347648 |     493137809012 |      986264156660
ibytes     |     468750367552 |     464905962240 |      933656329792
tx-pkts    |          5.3 Gpkts |          5.3 Gpkts |        10.59 Gpkts
rx-pkts    |          5.2 Gpkts |         5.18 Gpkts |        10.39 Gpkts
tx-bytes   |         493.13 GB |         493.14 GB |         986.26 GB
rx-bytes   |         468.75 GB |         464.91 GB |         933.66 GB
-----      |                   |                   |
oerrors    |                 0 |                 0 |                 0
ierrors    |                 0 |                 0 |                 0

status:  |

Press 'ESC' for navigation panel...
status: [OK]

tui>_
```

Đặt con số này vào bối cảnh: nhiều hệ thống KVM datapath kernel, dù zero-load, đôi khi đã choke ở mức vài trăm Kpps.

Trong điều kiện production-loaded, các VM chạy trên host HPN của chúng tôi đạt trung bình **~6–7 Mpps**. Có thể đánh giá hạ tầng đã đạt **multi-Mpps trong môi trường carrier-grade**.

Điều này cực kỳ quan trọng cho realtime workload, telco và streaming use case, nơi mỗi packet đều cần được xử lý "đúng nhịp" và nhất quán.

3. Latency & Stability — Điều khách hàng thực sự cảm nhận

Nếu throughput là “tốc độ tối đa”, thì latency là “cảm giác thực tế”. Với benchmark trên một NFV khác là Checkpoint R81 (16 core, 32 GiB RAM)

Latency trung bình:
~0.16–0.20 ms

Jitter:
~0.015–0.017 ms

Ở mức này, latency gần như không đáng kể đối với đa số workload. Ngược lại, khi dùng datapath kernel trên cloud trong cùng điều kiện:

- Latency có thể tăng tới **165–270 ms**
- Packet drop bắt đầu xuất hiện

Sự khác biệt không còn là “tốt hơn một chút”, mà là khác biệt về bản chất.

Trong môi trường production, NFV firewall Checkpoint được chúng tôi đo qua Zabbix:

- *Latency duy trì ~0.3–0.5 ms*
- *CPU steal: ~1e-7 (gần như bằng 0)*
- *Không xuất hiện queue buildup*
- *Không có burst drop*

Evidence Zabbix agent

Kết luận về latency của hạ tầng:

- Workload chạy deterministic
- Không có jitter ngẫu nhiên
- Không có "spike" gây lỗi khó debug

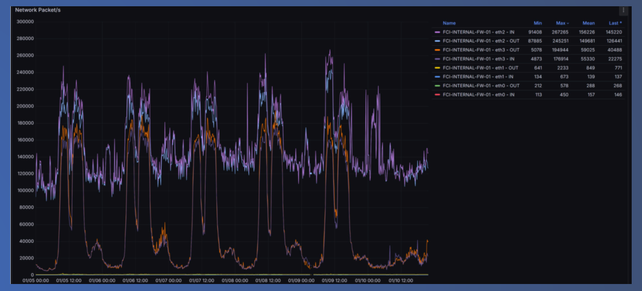
Đây là yếu tố then chốt với các hệ thống realtime.

Production Proof: FPT Software

Benchmark là cần thiết, nhưng production mới là câu trả lời cuối cùng.
Trong môi trường thực tế của đối tác FPT Software, chạy trên hạ tầng NIC/port 10G: Sau khi chuyển firewall sang các host hiệu năng cao HPN của chúng tôi, lấy một instance đại diện (16 core, 32 GiB RAM):

Packet rate tăng từ

~445K lên > 611K pps



Packet drop giảm về

~0.001% tổng packets

Mấu chốt: số vCPU cấp cho NFU

giảm từ 48 xuống 16,

nhưng hai metric trên không hề kém hơn và latency cải thiện rõ rệt

Throughput tăng từ ~1.9 Gbps lên ~3.4 Gbps (đã bật các filtering/validating features trên NFV Checkpoint)

Evidence Zabbix agent

Kết luận từ phía đội ngũ kỹ thuật khách hàng:

- Hiệu năng tăng gần gấp đôi
- Latency thấp và ổn định như baremetal tại on-premise
- Chi phí compute giảm đáng kể

Đây không chỉ là cải thiện kỹ thuật, mà là **tối ưu chi phí vận hành thực tế.**

Sự khác biệt và vị thế trong khu vực

Một lo ngại phổ biến khi dùng tăng tốc phần cứng như DPDK là sẽ mất đi các tính năng cloud. Vì vậy, chúng tôi đã kiểm chứng đầy đủ:

- *Live migration*
- *Resize / cold migrate VM*
- *Volume operations*
- *Security Group CRUD*
- *Floating IP SNAT/DNAT*
- *Trunking ports*
- *East-West / North-South traffic*
- *Metadata, DHCP*

Và cam kết tất cả đều hoạt động ổn định.

Bạn không cần phải chọn giữa **hiệu năng** và **tính năng cloud** – bạn có thể có cả hai. Nhiều hệ thống "có DPDK" nhưng không đạt được kết quả tương tự.

Sự khác biệt nằm ở cách thực thi và mức độ tùy chỉnh:

01 CPU pinning và cấu hình theo chuẩn telco

02 NUMA-aware và locality tối ưu

03 Hardware offloading theo workload

04 Multi-queue scaling đúng cách

05 Tư duy datapath khác biệt, và chỉnh sửa các dự án open-source theo nhu cầu thực tế

Những thứ này không thể đạt được chỉ bằng việc "bật một option".

Dựa trên benchmark microservices nội bộ, so sánh với các nghiên cứu phổ biến, và production validation với khách hàng lớn, chúng tôi tin rằng mình đang làm chủ một trong những nền tảng Cloud Open Infra hiệu năng cao nhất tại Đông Nam Á. Quan trọng hơn, trong giai đoạn phát triển tiếp theo, nền tảng này có tiềm năng cạnh tranh trực tiếp với hyperscale quốc tế trong các workload network-intensive.

Kết luận

Cloud không còn là bottleneck.

Với kiến trúc đúng và sự đầu tư tỉ mỉ vào nghiên cứu các công nghệ tăng tốc phần cứng, chúng tôi tin rằng có thể đem lại:

- Throughput tiệm cận line-rate
- Packet processing ở mức multi-Mpps
- Latency sub-millisecond ổn định
- Near-zero packet loss trong production

Tất cả trên hạ tầng ảo hóa.

What's next?

Nếu bạn đang vận hành

- Firewall / Security appliance
- Telco VNF
- Streaming system
- Finance, realtime commercial
- Governance system yêu cầu latency thấp và PPS cao

Chúng tôi cung cấp

- PoC benchmark theo workload thực tế
- Đánh giá bottleneck hiện tại
- Tư vấn và migration sang hạ tầng hiệu năng cao
- Luôn luôn phát triển thêm các giải pháp cạnh tranh về hiệu năng cao cloud computing

FPT CLOUD

Liên hệ chúng tôi

fptcloud.com
support@fptcloud.com
1900 638 399

Hà Nội: Số 10 Phạm Văn Bạch, Phường Cầu Giấy
TP. Hồ Chí Minh: Tòa nhà PJICO, Số 186 Điện Biên Phủ, Phường Xuân Hòa
Tokyo: 33F, Sumitomo Fudosan Tokyo Mita Garden Tower, 3-5-19 Mita, Minato-ku