NFV PM
Thought Experiments

Yaakov (J) Stein

QoE and QoS

Customers are willing to pay a service provider
for communications services with QoE guarantees

Unfortunately, QoE is subjective
and thus difficult to measure / guarantee

So we use **Q**uality **o**f **S**ervice **K**ey **P**erformance **I**ndicators as *proxies* for QoE

A QoS KPI is an *objective* communications parameter
that can be *easily measured*
and that can be *related* to QoE

Note that the *only* reason to measure a QoS parameter
is its relationship to QoE

If we show that the value of some parameter

- does not influence QoE, or
- influences QoE in nonconsistent manner

then there is no reason to measure it

PM parameters and QoE

One can show that QoE indeed depends on PM parameters
for *conventional communications services*

Unfortunately, the precise relationship depends on the service type

$$\textbf{QoE = f (application; QoS}_1\textbf{, QoS}_2\textbf{, … QoS}_n\textbf{)}$$

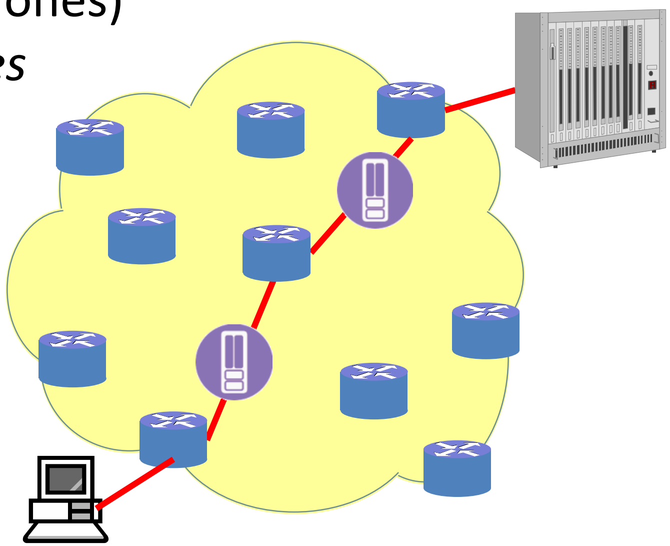
For example:

- non-real-time TCP traffic suffers data-rate reduction under packet loss, but is relatively immune to increase in delay
- real-time interactive traffic employing **P**acket **L**oss **C**oncealment is relatively immune to low rates of packet loss, but suffers from delay
- timing distribution is immune to delay
  but sensitive to low frequency packet delay variation

Researchers have found the formula relating QoE to QoS KPIs
for many service types (e.g., work of Markus Fiedler from BTH, Sweden)

Rich communications services

In today's networks (especially NFV-enabled ones)
we can define *rich communications services*



By that term we mean services with both
- transport (traditional L2/L3 networking)
- computational (L4/L7 middleboxes)

components

For example, a service may be defined
- from site X to site Y
- with data-rate at least R and latency no more than L
- with network functions A, B, C and D performed along the way

We have seen that for traditional (transport-only) services
- end users are willing to pay for service QoE guarantees
- service QoE can be related to network QoS KPIs
- hence, end users are willing to pay for network QoS guarantees

Finding QoS KPIs for rich services

For rich services the end user still experiences some end-to-end QoE

The questions we ask are:

- **can we find easily measurable parameters that determine this QoE ?**
- **can we find the precise relationship between QoE and these KPIs ?**

It makes sense to study 3 types of potential parameters :

1. transport QoS KPIs (identical to conventional QoS parameters) $\mathbf{T_i}$
2. NFVI KPIs (e.g., CPU/memory/storage usage as collected by VIM) $\mathbf{N_j}$
3. KPIs of the individual VNFs deployed (as collected by VNFM) $\mathbf{V_k}$

and to look for a relationship of the form

$$\mathbf{QoE = f\ (application;\ T_i\ ,\ N_j\ ,\ V_k)}$$

However, there are no guarantees that
our well-known QoS KPIs are meaningful in this new setting

The effect of SDN on PM

Before embarking on our search for QoS parameters
it is worth noting the effect of SDN on Performance Monitoring

- SDN makes end-to-end paths more deterministic
  with conventional networks
  - Carrier Ethernet and TE MPLS are path-pinned
  - IP and LDP-based MPLS can arbitrarily change routing

  with SDN all services are path-pinned

- SDN networks emphasize consistency
  before changing flow behavior, old packets are allowed to drain from the system

- SDN networks support application level flow granularity
  in conventional networks multiple application types are mixed in a single FEC

Thus PM parameters

- do not abruptly change in SDN networks
- can be matched to service type

So we can conclude that SDN actually makes determining QoE easier!

PM for rich services

SDN provides a different path computation mechanism
but is still limited to pure transport networking

On the other hand, NFV changes the paradigm
by enabling nontrivial data processing

As a first step to understand this new world
we will embark on a set of *gedanken experiments* (AKA *thought* experiments)

We will show that standard transport QoS KPIs (type 1. above)
whether viewed end-to-end or hop-by-hop
can be meaningless for NFV-enabled networks

Our proofs will rely on demonstrating that with specific but genuine VNFs
radical degradation of the value of standard QoS parameters
may not change (or even improve) the user's experience!

Packet loss can be *problematic*

Experiment 1 IPS

An IPS function discards packets that it deems to be malicious

This leads to an increase in measured PLR

Discarding these packets are in the user's best interest (except for false-alarms)
So, one would wish the QoE to increase

However, QoE(PLR) can not universally increase
as most other packet loss degrade QoE

Note that these discards are not comparable to those of a policing function

Policing discards do *not* degrade measured PLR
as it is only computed on in-profile packets

This example may seem contrived, and its effects rare

But it gets much worse!

Packet loss can be *meaningless*

Experiment 2   TCP proxy

A TCP proxy terminates the end-to-end TCP session
creating two segmental sessions

While the transmitted byte-stream is maintained, its segmentation is not

For example,
3 packets may enter the proxy, and either 2 or 4 exit it !

Thus, counting packets becomes meaningless in the presence of a proxy

And thus PLR becomes undefined!

The obvious remedy is to completely abandon counting packets
and to measure the more meaningful entity – *volume* !
By volume we mean the number of bytes received
irrespective of the number of encapsulating packets

So, we posit that the correct KPI is loss of volume, not loss of packets

Let's see if that helps ....

Volume loss can be *meaningless*

Experiment 3 WAN optimization – compression

By compression we mean

- lossless data compression and data deduplication
- audio or video compression, transcoding or transrating

All of these mechanisms input some number of bytes
and output some (often much smaller) other number of bytes

The compressions are designed to be either truly information lossless
or at least to introduce no noticeable QoE degradation

Thus, QoE should remain unchanged although volume changes significantly!

Obviously QoE can not be completely independent of volume
a reductio ad absurdum argument would show QoE>0 with zero volume

So, we once again reach the conclusion that
sometimes QoE must decrease with reduced volume, and sometimes not

The obvious remedy is to completely abandon measuring byte volume
and to measure the more meaningful entity – Shannon information !

Information loss can be *meaningless*

Experiment 4 WAN optimization – caching server (CDN server)

A caching server is a mechanism that conserves network resources
by storing information that may be consumed multiple times

Observing a particular flow
frequently the output from a caching server
is precisely the input (pass-through)

But when the flow consists of information that has been previously cached
then packets with non-zero information are output
although no packets were input

At a flow level this seems to show information creation by the server
in paradoxical contradiction to principles of information theory

So, even measured loss of Shannon information content loss
can not be used as an end-to-end QoS parameter!

Note, the information integrated over all time and aggregated over all flows
is conserved, but is not related to the QoE of any particular flow

Synthetic OAM packets

Network engineers will immediately object to our line of reasoning

Certainly PLR is well-defined
 and the fault lies totally with our measurement methodology!

The proper way to measure PLR in such cases
 is to introduce synthetic OAM packets
 designed to bypass the computational functionality
and thus measure true end-to-end transport PLR!

That argument is completely true, and completely irrelevant !

We aren't interested in measuring QoS parameters as an academic exercise
The purpose of measuring them is to predict QoE on user traffic

Traffic that does not traverse all the elements of user packets
 i.e., that is not *fate sharing* with true user traffic
can not be expected to assist in the prediction of the QoE of such user traffic!

Delay may be *problematic*

Perhaps the problem is only with PLR
and other traditional KPIs remain useful ?

The second most useful QoS parameter is end-to-end propagation delay

Of course, many of out previous examples
already cast doubt on the meaningfulness of delay

If packets are combined and resegmented as in experiment 2 (TCP proxy)
then we need to track individual bytes, not packets
resulting in very erratic statistics

If packet contents changes as in experiment 3 (compression)
then how do track bytes ?

If packets are not even sent as in experiment 4 (caching)
what does propagation delay even mean ?

But, there is an even stronger argument !

Delay can be *meaningless*

Experiment 5 Web browsing

Studies show that users

- are universally satisfied if web pages stabilize in less than 2 seconds
- are universally frustrated if web pages don't stabilize within 8 seconds

This is already hard to track for web pages with embedded components
where an embedded image may load with significant delay
and cause the page to graphically reassemble

But, let's look at the simplest point-to-point case

The browser itself is a software function that is part of the service
and runs software (e.g., javascript) downloaded as part of the data

This software may an unbounded amount of time to run
before finally stabilizing the graphical representation of the web page

Thus, delay from request to page stabilization
is not uniquely determined by network propagation delay!

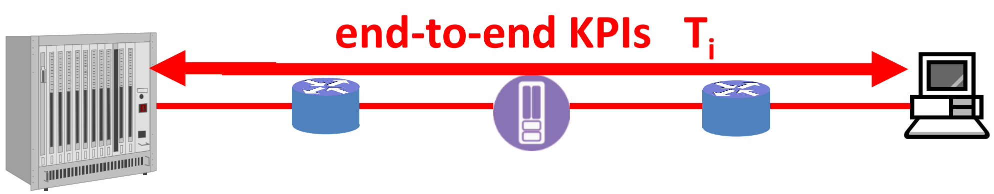
(only a *greater than or equal to* relationship exists)

What about hop-by-hop KPIs ?

What we have shown is that for NFV-enabled networks

QoE is not a function of flow-level **end-to-end** QoS parameters such as Packet Loss Ratio and network delay

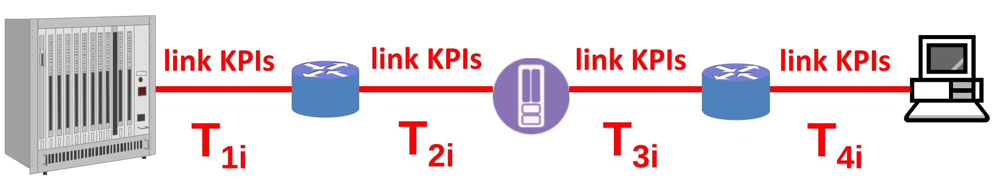
$$\mathbf{QoE \neq f\ (application;\ T_i\ )}$$



Of course, there *may* still be meaning to **hop-by-hop** QoS parameters

$$\mathbf{QoE = f\ (application;\ T_{\lambda i}\ )}$$

where λ runs over all the links between different VNFs of the flow in question



What about hop-by-hop KPIs ?

What we have shown is that for NFV-enabled networks

QoE is not a function of flow-level **end-to-end** QoS parameters such as Packet Loss Ratio and network delay

$$\mathbf{QoE \neq f\ (application;\ T_i\ )}$$

Of course, there *may* still be meaning to **hop-by-hop** QoS parameters

$$\mathbf{QoE = f\ (application;\ T_{\lambda i}\ )}$$

where λ runs over all the links between different VNFs of the flow in question

Hop-by-hop parameters are

- always easy and consistent to measure
- frequently hard to combine into meaningful end-to-end parameters

The latter is particularly the case in networks with routing where we can't be sure which links and nodes are allocated to a service and the allocations may change over time

However, there are other problems as well!

HbH KPIs can be *hard to track*

In order to relate HbH QoS parameters to a service and its QoE
we need to know which NEs and links a service traverses

A prevalent way of mapping physical entities to services
is to track packets on their way through the network

The alternative is:

- using synthetic OAM traffic to measure HbH KPIs
- dictate or receive list of NEs and links the service traverses

Experiment 6 NAT

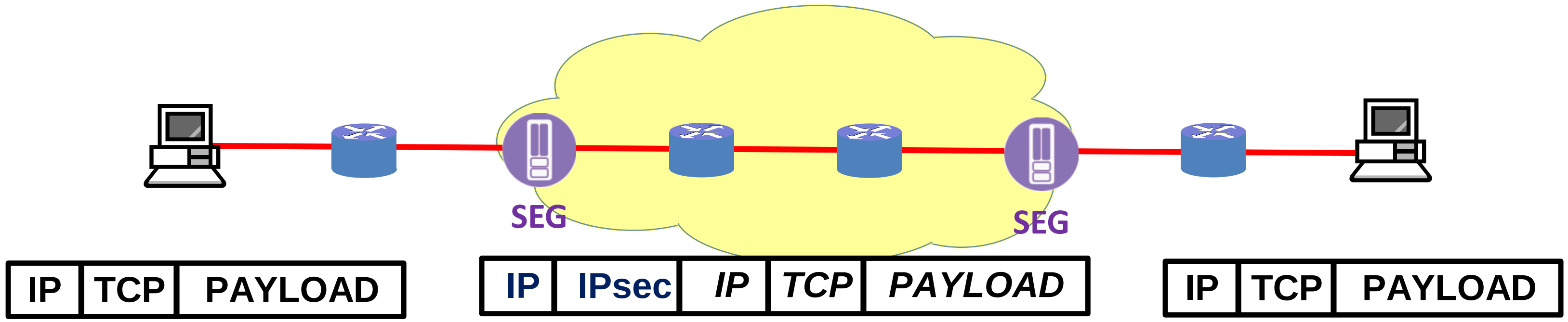
Consider a VNF implementing NAT

NAT changes the packet headers
making identifying packets based on their headers difficult

Of course, the measuring system could observe the NAT's input and output
and deduce the needed header mappings

HbH KPIs can be *impossible to track*

Experiment 7 IPsec tunnel security gateway (SEG)



The security gateway VNF prepends a new header to the packets and additionally encrypts the original headers and payload

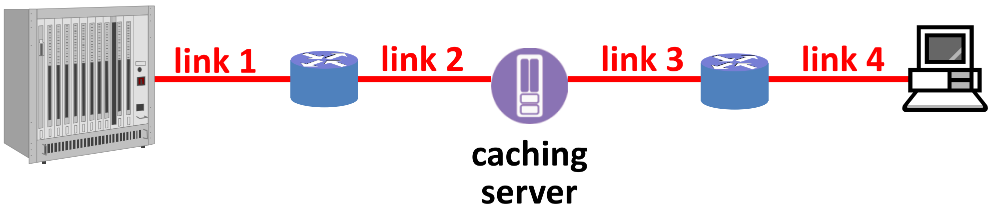
Thus making it impossible to map packets entering the network to packets inside the network

For both experiment 6 (NAT) and 7 (SEG) we *can* observe the packets end-to-end

Thus, end-to-end measurements are just as easy as before !

Hop-by-hop KPIs can be *meaningless*

Experiment 8   traffic caching (again)



When the caching server is providing packets to the end user
the KPIs of links 1 and 2 do not influence the QoE at all
since the packets received by the end-user does not traverse them!

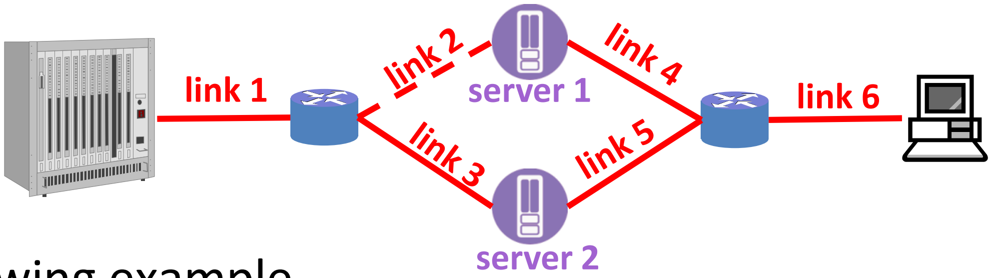
This may seem to be similar to rerouting
where links that once were on-path are not longer on-path

However, for rerouting, at least in principle
the network knows which links are on-path

For caching, this information is fundamentally unknowable to the network

Hop-by-hop KPIs can be *counterintuitive*

Experiment 9   Influence of rerouting or protection switching



Consider the following example

A rich communications service

- initially traverses links 1, 2, 4, and 6 and utilizes server 1

Due to severe degradation of QoS KPIs on link 2, the service is rerouted

- to traverse links 1, 3, 5, and 6 and to utilize server 2

Server 2 happens to perform the desired functionality better, due to

- upgraded software
- more available CPU power and/or memory and/or storage

causing the QoE to improve!

Thus, QoS *degradation* may lead to QoE *improvement* !

Are these examples too contrived?

One may claim that the examples above
were specifically chosen to invalidate traditional QoS parameters
and that in more typical cases we could still reasonably employ QoS KPIs

Beyond the fact that all of our gedanken experiements
utilized actually existing functionalities
such an argument is misleading or even meaningless

The whole idea of rich communications services is that

- new functionalities will be appearing at an accelerated rate
- such new functionalities may be inserted anywhere along the service path
- and that we do not know a priori

We can not make assumptions on what VNFs do

We can not make assumptions as to which VNFs are where

We must assume the worst case
that all of our aforementioned cases may occur

Summary

Starting from the fundamental assumption that QoS parameters
are only useful in so far as they relate to service QoE

We used a series of gedanken experiments
to test the usefulness of traditional QoS KPIs in NFV-enabled networks

We discovered that in general

- end-to-end QoS parameters such as packet loss rate and delay are useless
- hop-by-hop QoS parameters are useless and even counterintuitive

This leads us to conclusion that new research is required
to determine if there are *any* KPIs that quantitatively relate to QoE