Date: July 02, 2025

Planet Network International

OTT Analysis

Enhancing Over-The-Top (OTTs) service quality and connectivity assessment in mobile networks

ITU-T Study Group 12 Regional Group for Africa (SG12RG-AFR) and Workshop on Telecommunication Service Quality

Freetwon, Sierra Leone, 1-4 July 2025

Gorgui NIANG
IT Engineer

SUMMARY

Principles and features

Data sources and Pre-requisites

Deliverables

Advantages of this approach

Challenge of Measuring OTT QoS

Principles and Features

The objective of this module is to analyse more precisely the way that subscribers consume their data, and which type of services they are being provided to, by Over-the-Top Providers (also known as “OTT Providers”), such as :

- Social Networks: Facebook, Instagram, Twitter …
- Video: YouTube, Netflix, BeIN Sports …
- Communications: WhatsApp, Skype, WeChat, Teams …
- Webmail: Gmail, Yahoo …
- News: NYT, Local news …
- Security: Kaspersky, Avast …
- Games: King …
- Market Places: Amazon …
- Cloud Services: Azure, Amazon AWS
- Cloud Storage: Dropbox …
- …

Principles and Features

The module then reconstructs the active subscriber base of each OTT service provider, which can then be used to :

- estimate the activity and the revenues of the associated entities
- Analyze the evolution of the international usage, with subscribers switching from “legacy/Telecom” calls to “VoIP Services” such as WhatsApp, Skype or Viber

Principles and Features

To perform such analysis, three methodologies can be considered :

➔ Deep Packet Inspection (also called « DPI ») – IPDR

1. This approach requires physically placing a Probe on each router of the operator, making the link between the operator's network and the internet (and eventually the cache servers that some Content Delivery Network Companies, such as Akamai, and major Web players, such as Google, install within the operator's networks to reduce the use of the « external bandwidth »)

2. It analyses each IP packet, and determines
   a. Session Start date
   b. Session End date
   c. Which IP + port it comes from
   d. Which IP + port it goes to
   e. Which Protocol is used
   f. The number of bytes used, both for envelope and content

3. This data then needs to be correlated to two data sources:
   a. The link between target public IP and OTT provider
   b. The link between origin public IP and Subscriber Party

Principles and Features

International VoIP CDR reconstruction

4. This approach requires to place probe on the Internet access of the country

5. The packets are then analyzed and used to reconstruct VoIP sessions from VoIP OTT providers, such as Skype, WhatsApp Calls, and determines

a. Start date

b. Which IP + port it comes from

c. Which IP + port it connects to

d. Call type : MO or MT

e. The duration of the call

6. This data then needs to be correlated to ;

a. The link between the « local » public IP and Subscriber Party

b. The link between the « international » public IP, and the associated country

Principles and Features

➔ DNS Logs Analysis

7. **This approach is not intrusive, and is based on the logs of the DNS servers**
8. **Every time a subscribers tries to contact an OTT provider, he knows its name / URL, but then needs to find the associated target IP of the server that hosts the service. The DNS server performs this « Domain Name Resolution, and, given the name, provides the IP address**
9. **Each DNS log contains**
   a. Request Date
   b. Which IP + port requested it
   c. Which URL was requested
   d. The associated IP address

10. This data then needs to be correlated to two data sources:

   a. The link between URL template and OTT provider
   b. The link between origin public IP and Subscriber Party

Principles and Features

The advantages and draw backs of each approach are the following:

➔Deep Packet Inspection – IPDR

11. **Advantages :**
    a. Provides the actual number of bytes used
12. **Draw backs:**
    a. Intrusive in the network
    b. Calculation intensive, hence heavy investment
    c. Requires additional hardware every time a new gateway is installed
    d. Requires the maintenance of extensive and up to date lists of all IP addresses used by the Service providers, which are very numerous, and change very often
    e. Potential risk of access to subscribers' personal information, if not properly encrypted
    f. Cannot analyse encrypted traffic
    g. Generates a huge amount of data

Principles and Features

International VoIP CDR reconstruction

13. **Advantages:**

    a. Identifies the international calls entering and exiting the country using VoIP OTT

    b. Based on the law, it can be used to apply international interconnection taxes

    c. Does not require to analyze all IP traffic, just what is exchanged with other countries, hence less data to analyze, and fewer points to tap (at PoP, so sometimes not even within operator's networks)

    d. Reasonable volume of data generated : corresponds to « classical » International Voice CDR, hence its volume is less than the volume generated by MSC/Softswitch

Principles and Features

➔ International VoIP CDR reconstruction

14. **Draw backs:**

    a. Intrusive in the PoP

    b. Requires additional hardware every time a new PoP is installed (which is not that often)

    c. Algorithms used to identify the call and calculate its duration depend on the way each OTT has implemented his VoIP service. This implementation can evolve over time and there is no formal guarantee that it will be possible to reconstruct the same information in the future, although it might take years

    d. Usually, local VoIP calls do not exit the country, as OTT providers optimize the bandwidth and take the shortest route to connect the two parties. However, there is no formal guarantee that local calls, in some countries, do not pass through a proxy located outside of the country, hence making a « local » call appear as « international ».

Principles and Features

➔ DNS Logging

15. **Advantages:**
    a. Nothing to install in the operator's network
    b. Robust with respect to network evolutions (no additional hardware, no integration)
    c. DNS logs are already generated and can be transferred as is to the satellite servers, or the satellite servers themselves can act as DNS log servers
    d. Resilient to SSL encryption, which represents 70% of the traffic
    e. Generates less volume of data to analyze than DPI (still it is a lot !!!)
    f. Easier Maintenance, as URL contain the « name » of the OTT service provider
    g. Operators already use the DNS logs to monitor potential security breaches, for their internal network as well as for their subscribers

Principles and Features

DNS Logging

16. **Draw backs:**

a. Does not provide the number of bytes nor the duration of the session

b. Some purely technical private applications might not require the use of the DNS server and hence are not seen. But this does not apply to OTT providers.

c. In some cases, subscribers have the possibility to reconfigure the DNS server they want to use. But this is rare, and usually only possible on the PC, with the mobile serving as router. Hence it can be considered neglectable for the purpose of OTT analysis

d. In some even rarer cases, end users can encrypt their DNS requests, in which case the information is not available. However, this represents nowadays a neglectable proportion of the traffic, and Governments can decide by law that DNS request encryption is not authorized (cf current actions in the USA)

Principles and Features

This module adapts to your environment and collects any of these data sources:

- ➔IPDR from the DPI system
- ➔VoIP CDR from the Reconstruction systems
- ➔DNS logs from the DNS

And then it can perform consistency controls:

- ➔Consistence between DNS logs and actual data sessions
- ➔Consistence between DNS logs and VoIP calls

Principles and Features

In terms of integration, for each scenario:

➔ For DPI

17. If probes have already been installed, then they generate IPDR files
18. These files just need to be sent to the satellite servers
19. If probes need to be installed, we can provide such probes and perform their integration, for a specific quotation

➔ For VOIP CDR reconstruction

20. If probes have already been installed, then they generate VoIP CDR files
21. These files just need to be sent to the satellite servers
22. If probes need to be installed, we can provide such probes and perform their integration, for a specific quotation

Principles and Features

For DNS logs collection

23. DNS logs are usually already generated by the operators: in this scenario, Operators will keep the existing file transfer mechanism and add this data source to the flows already sent to the Satellite servers
24. If DNS logs are not already generated, the operator will just have to install a log server and push the files: no need to keep a consequent historical information (few days, just in case) and hence reasonable storage
25. In some cases, the satellite server can even be used to log the DNS requests directly from the DNS server (syslog server).

NB: One key element to consider when dealing with IPDR and DNS logs is the volume of Data which needs to be collected and pushed / pulled onto the central server, hence putting a huge load on the network connections.

We can expect the volume of data generated by DNS logs to be over 10 times more than “classical” CDR, hence accounting to dozens of billions of “tickets” every day, for which a specific archive strategy will need to be defined.

Principles and Features

When the objective of the module is to reconstruct subscriber bases, or analyze the usage on specific sites / URL (and not provide detailed investigation of all sessions of all subscribers), we recommend the following approach:

➔ **Pre-process the data on the satellite servers**

26. **Perform a first level filtering directly on the satellite servers, to reduce the number of «tickets »**
    a. Filtering based on the country associated with the IP
    b. Filtering based on a specific list of IPS
    c. Filtering based on specific sites / OTT platforms
    d. NB: tickets that do not correspond to any filter are regrouped, for each subscriber, as « other », to keep the representativity of the traffic
27. **Perform a first level cleansing, to reduce the size of each « ticket »**
    a. Remove all un-necessary information
    b. Clean the URL

Principles and Features

➔ Pre-process the data on the satellite server

28. **Compress the tickets**
    a. If possible, regroup tickets, for a given subscriber and a given site on a given period
    b. Encode the information
    c. Apply compression algorithm

➔ Keep the information collected on the central server for a limited time

29. **Keep collected raw files for a limited period (for example a couple of weeks)**
30. **Keep cleansed, structured detailed data, for a longer period (for example a couple of months)**
31. **Keep aggregated data for a very long period (for example a couple of years)**

Principles and Features

That way, only the relevant data is uploaded to the central server, with controlled impact on the required bandwidth and the storage capacity

NB: The conservation period can be adapted to the customer's requirements and can evolve over time. This aspect mainly depends on the hardware and storage investment, following a linear rule.

NB: During the integration phase, an analysis of the main visited sites / OTT providers is performed, to adapt to the local environment.

Such an analysis is also performed on a regular basis to detect significant newcomers.

New Service providers can be simply added through PARAM 'Center.

Once on the central server, the IP address + Port will be linked to the SGSN/SGW logs, or the DHCP logs, depending on where the DNS server is installed inside the operator's network, to get the subscriber identifier (PARTY_ID / MSISDN).

Principles and Features

The solution also includes:

➔ A configuration interface, in PARAM 'Center, to manage the list of OTT service providers to monitor, and the URL template / IP addresses they use

➔ Multiple dashboards on active subscriber base, with different levels of usage:

32. Very active user
33. Normal user
34. Simple visitor

Finally, the solution also includes cross-platform consistency controls, to make sure that

- All IPDR / VoIP CDR / DNS logs are indeed associated with a SGSN/SGW or DHCP session
- All GGSN/PGW data sessions are indeed associated with at least one IPDR/DNS log (unless some very exceptional behavior that should remain neglectable)

Data sources and Pre-requisites

The Solution requires the collection of the following data sources:

➔ Based on the decided approach

35. DNS logs
36. IPDR, and the installation of the probes
37. Reconstructed VoIP CDR, and the installation of the probes

S-GW/SGSN CDR or alternatively DHCP logs, depending on DNS server's location within the operator's network.

NB: Based on the strong and explicit requirement specified by Regulators:

- Use DNS logs

- Use CDR generated by already existing probes inside the MNO network if they have done the investment

Consistency checks between both data sources to make sure the MNO sent all the information

Data sources and Pre-requisites

➔ Functionally, DNS logs contain the following information:

- CDR_DATE: Request Date and Time
- IP_ID: IP address
- PORT_ID: port used by the subscriber to perform that request
- PROTOCOL_ID: Protocol (UDP/TCP)
- DOMAIN_ID: Domain name
- STATUS_ID: Status

Data sources and Pre-requisites

➔ Similarly, IPDR contains the following information:

- CDR_ID
- CDR_DATE: Session start date
- IP_ID: IP address
- PORT_ID: port used by the subscriber to perform that request
- PROTOCOL_ID: Protocol (UDP/TCP)
- OTHER_IP_ID
- OTHER_PORT_ID
- DOMAIN_ID: Domain name
- UP_VOL : uploaded number of bytes
- DOWN_VOL : downloaded number of bytes
- CDR_DUR : duration

Data sources and Pre-requisites

➔ And Reconstructed VoIP CDR have the following format:

- CDR_ID
- CDR_DATE: Start Date
- CALLING_IP_ID: IP address
- CALLING_PORT_ID: port used by the subscriber to perform that request
- CALLED_IP_ID
- CALLED_PORT_ID
- SERVICE_PROVIDER_ID : Skype, Whatsapp …
- CDR_DUR : duration

Data sources and Pre-requisites

➔For mobile users, to identify the subscriber behind the IP+port, the solution correlates that information with the SGSN / S-Gateway CDR, which provide:

- SESSION_DATE: Session start Date
- SESSION_DUR: Session duration
- PARTY_ID: Subscriber MSISDN
- IMSI_ID: Subscriber IMSI
- IP_ID: Allocated Public IP
- MIN_PORT_ ID: min allocated port within the chunk
- MAX_PORT_ID: max allocated port

The integration of such CDR is performed within the Unrated Usage Module

Data sources and Pre-requisites

➔ Similarly, the DHCP logs, if required, contain the following information

- SESSION_DATE: Session start Date
- SESSION_DUR: Session duration
- PARTY_ID: Subscriber MSISDN
- IMSI_ID: Subscriber IMSI
- IP_ID: Allocated Private IP

Deliverables

The solution integrates, within PARAM 'Center, the configuration of the OTT service providers, and the URL templates that correspond.

The Solution also includes the following dashboards:

➔ User active base - Overview

38. Per operator
39. Per period (Day / Month / Quarter / Year)
40. Per OTT service provider
41. Number of subscribers
42. Number of requests / sessions and associated volumes and duration (depending on source availability)

Deliverables

➔ User active base – Per OTT

43. Per OTT service provider
44. Per operator
45. Per period (Day / Month / Quarter / Year)
46. Number of subscribers
47. Number of requests / sessions and associated volumes and duration (depending on source availability)

➔ User active base – Per OTT - geo-location information

48. Per OTT service provider
49. Per operator
50. Per period (Day / Month / Quarter / Year)
51. Per Location
52. Number of subscriber
53. Number of requests / sessions and associated volumes and duration (depending on source availability)

Deliverables

➔ International VoIP traffic – Incoming – Overview (depending on source availability)

54. Per operator
55. Per period (Day / Month / Quarter / Year)
56. Per OTT service provider
57. # Calls and Duration

➔ International VoIP traffic – Incoming – Detail (depending on source availability)

58. Per operator
59. Per period (Day / Month / Quarter / Year)
60. Per OTT service provider
61. Per Country of Origin
62. # Calls and Duration

Deliverables

➔ International VoIP traffic – Outgoing – Overview (depending on source availability)

63. Per operator
64. Per period (Day / Month / Quarter / Year)
65. Per OTT service provider
66. # Calls and Duration

➔ International VoIP traffic – Outgoing – Detail (depending on source availability)

67. Per operator
68. Per period (Day / Month / Quarter / Year)
69. Per OTT service provider
70. Per Country of Destination
71. # Calls and Duration

Deliverables

➔ Consistency controls (depending on source availability)

72. DNS versus IPDR
73. DNS versus reconstructed VoIP CDR
74. DNS logs versus active SGSN / S-GW session
75. IPDR versus active SGSN / S-GW session
76. Reconstructed VoIP CDR versus active SGSN / S-GW session

Advantages of this approach

The advantages of this approach are the following

➔ No Impact on operator's operations

- DNS approach is not intrusive on the operator's network
- Each platform is already generating and stores these files
- We use raw files: no need to prepare specific decoded feeds
- We assist the operator to push the data onto our platform

➔ Robust Approach

- Optimize the volume of data collected and processed
- Optimize the bandwidth requirements between satellites and central server
- Analyzes almost all traffic, including encrypted
- Capacity to analyze over 80 billion logs every day

Advantages of this approach

The advantages of this approach are the following

➔ Data Quality remains guaranteed

- Consistency controls secure the proper reception of all traffic
- Internal platform consistency controls are applied every day for each subscriber

➔ Respect user privacy

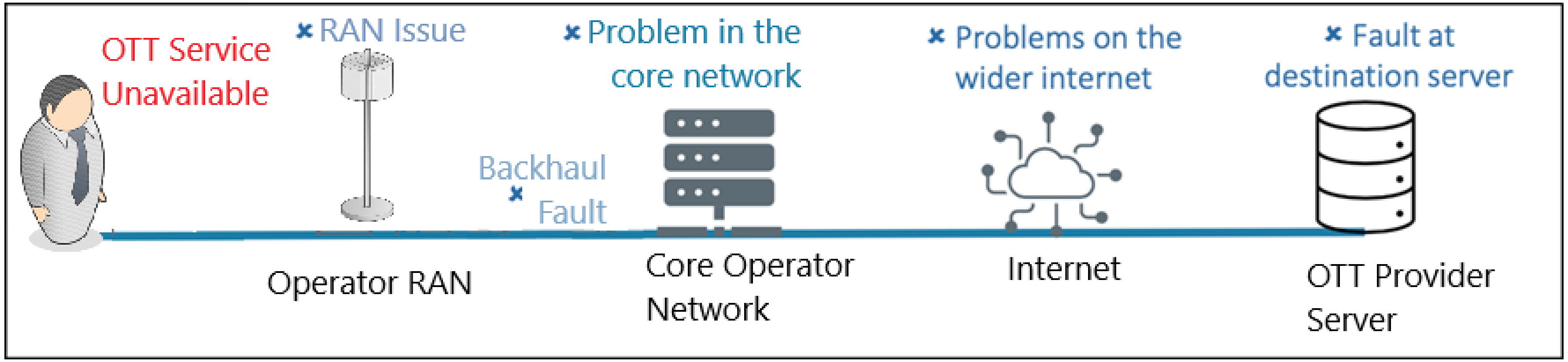
- No possibility to access unauthorized information (SMS content, instant location …)
- No information about subscriber identity

➔ Financial opportunity

- Provides a powerful tool to control TAX and VAT collection on OTT Providers

Challenge of Measuring OTT QoS

There are many points of failure as this basic diagram highlights:



The performance of OTT on mobile data services is dependent on numerous factors including :

a) Local RAN conditions;

b) The operators network including load balanced / geographical components;

c) Upstream internet connection or peered link; and

d) OTT providers equipment, operational processes and systems.

The solution of measuring OTT performance:

➔ includes focused testing for (b) to remove areas outside of operator control (factors c and d).

➔ includes end-to-end assurance of all componts (factors a-d) to measure real OTT subscriber service levels.

Deliverables

➔ OTT Template QoS KPI Breakdowns

- Per operator
- Per period (Day / Month / Quarter / Year)
- Per OTT service template type
- Per Network Infrastructure Components
- Per Technology Type (fixed line, 2G, 3G, 4G,5G)

➔ OTT Template QoS KPI Area Examples

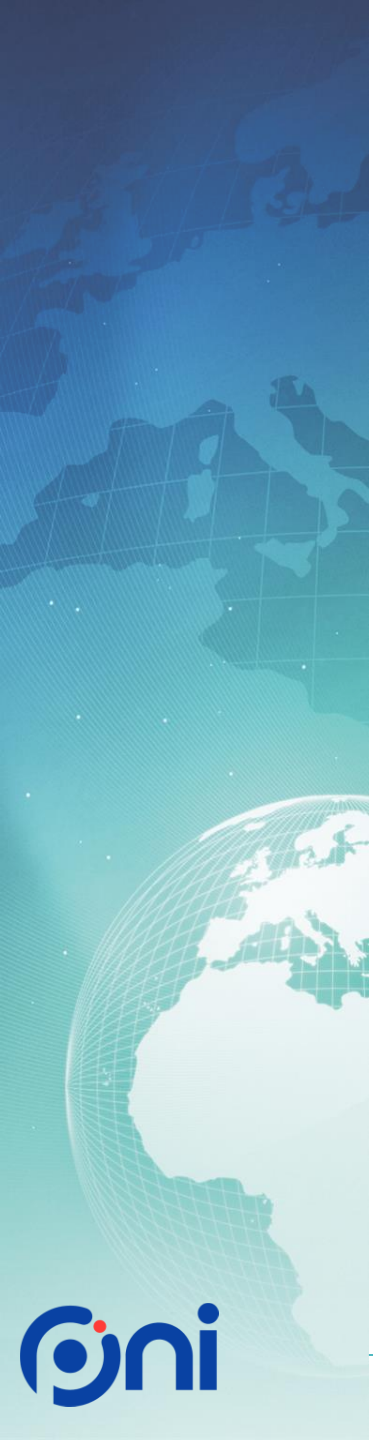
- Network Behaviour (Bandwidth, Jitter, Latency)
- OTT Service Availability
- Application Response Times
- Content Quality

➔ OTT Application QoS KPI Breakdowns

- Per operator
- Per period (Day / Month / Quarter / Year)
- Per OTT Application
- Per Network Infrastructure Components
- Per Technology Type (fixed line, 2G, 3G, 4G,5G)

➔ OTT Application QoS KPI Area Examples

- Network Behaviour (Bandwidth, Jitter, Latency)
- OTT Service Availability
- Application Response Times
- Content Quality



THANK YOU