The bridge to possible

The New, Encrypted Protocol Stack

Andreas Enotiadis (MIG Mobility Sales CTO)
Bart Van de Velde (Sr. Director, Engineering, Networking CTO Office)
October 2023
OTE Group

Agenda

- The New Internet
- The New IP Protocol Implications
- What’s left?

In memory of
and based on the
brilliant work of
Mark Gallagher

14/09/1966-17/09/2021

Networking

QoE

Queueing

eSNI

DPI

DOH IP DNS

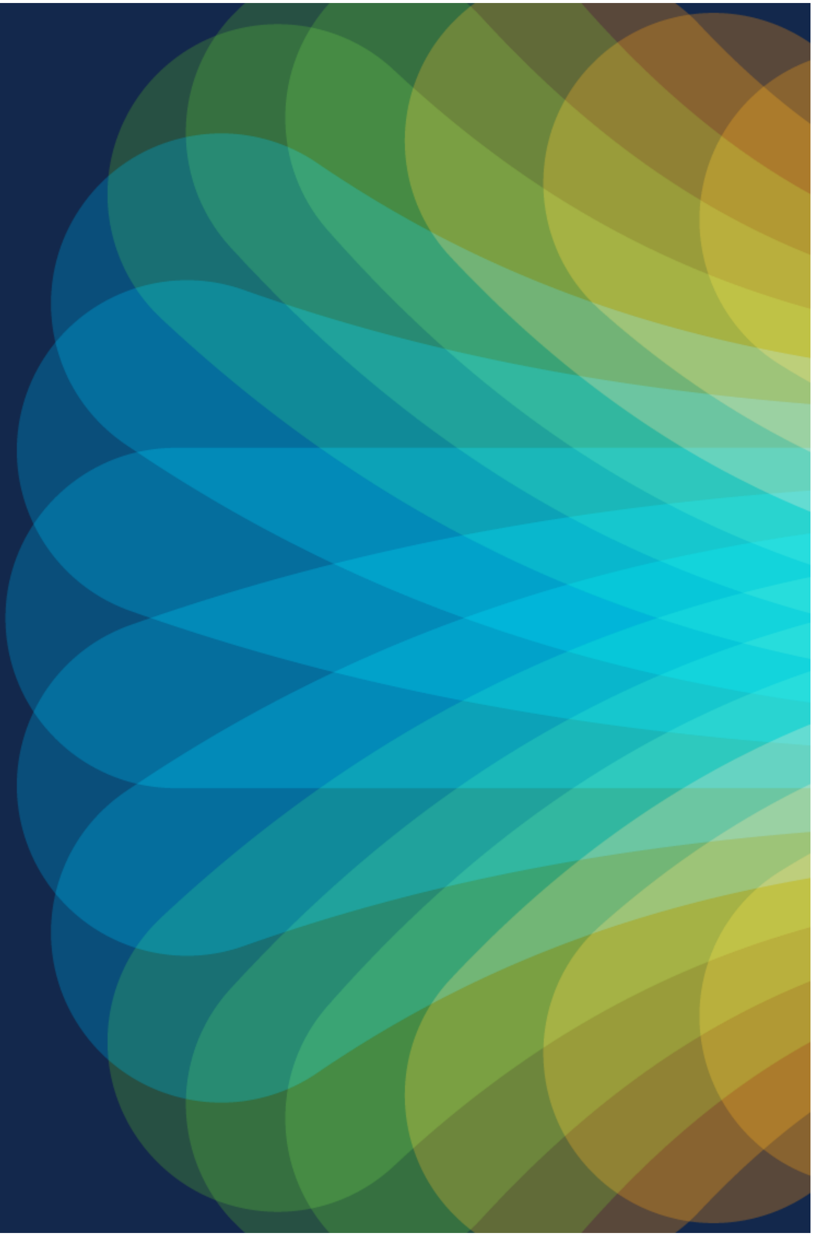
IETF TCP/IP

Google HTTP/3

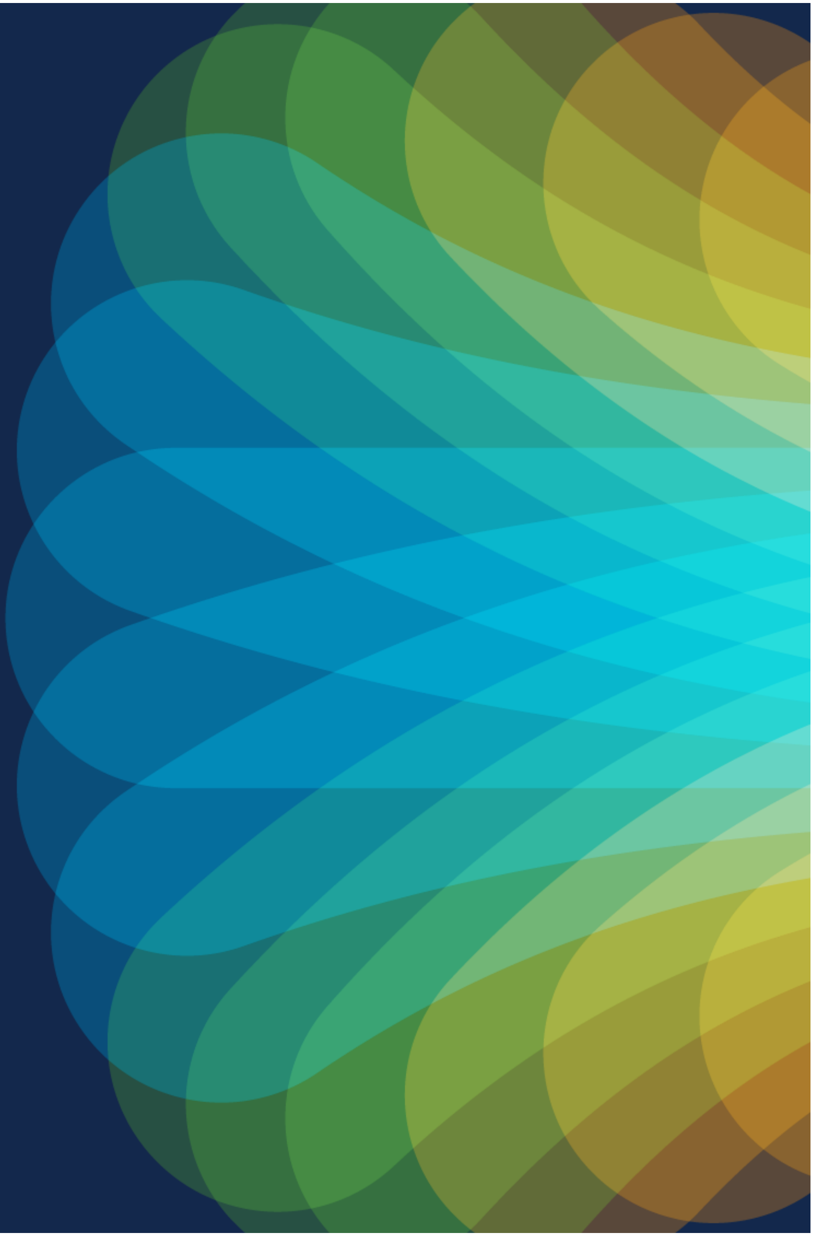
ECH TCP Traffic

Visibility UDP

QUIC

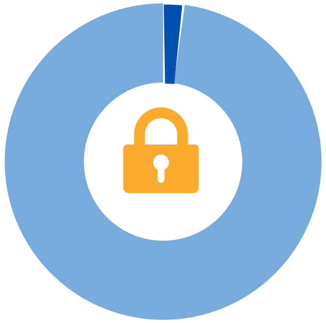


The New Internet

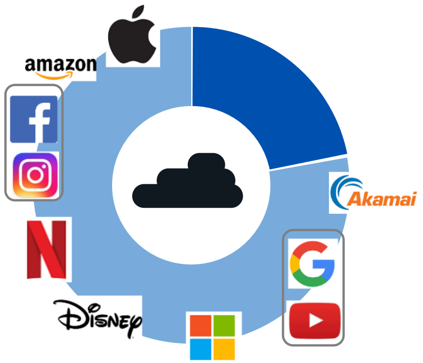


The Internet Reality – circa 2020 – Major US Carrier

>90% of
Volume: encrypted

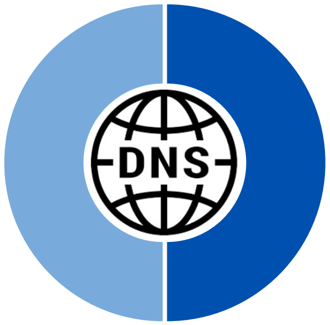


>70% of
Volume: to Cloud

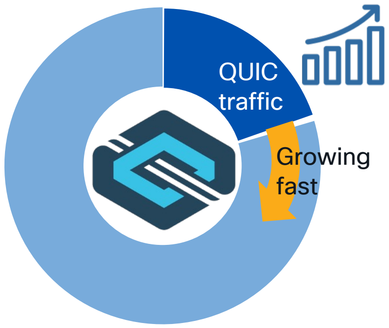


10 Cloud sites
“Elephant destinations”
not “Elephant flows”

~50% of
Flows: DNS



>20% of
Traffic: QUIC



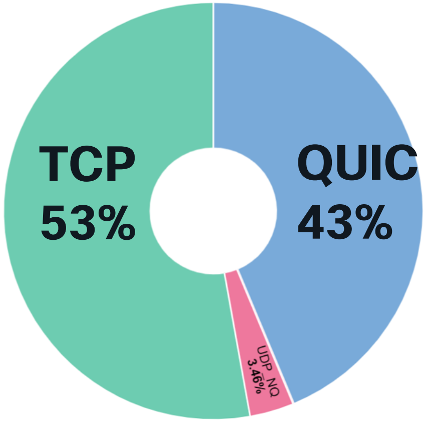
Many small flows
Micro-sessions

- Destination: all-encrypted world
- Cloud: concentrating the Internet

- Content: DNS is the load-balancer
- QUIC: Future Protocol of choice

Fast forward 18 months - Tier-1 EU Mobile Carrier

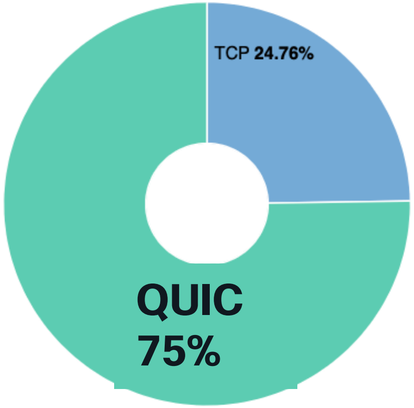
Overall Volume



QUIC has doubled in 18 months

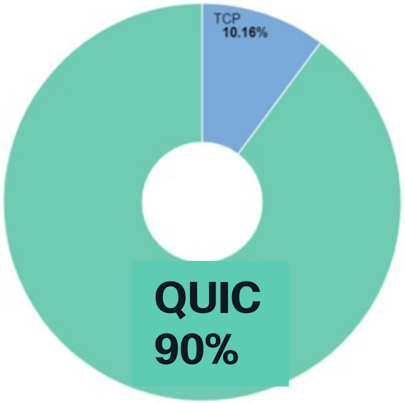
QUIC is 43% of total and rising

Volume



QUIC is “default”

Volume



Meta has gone full QUIC

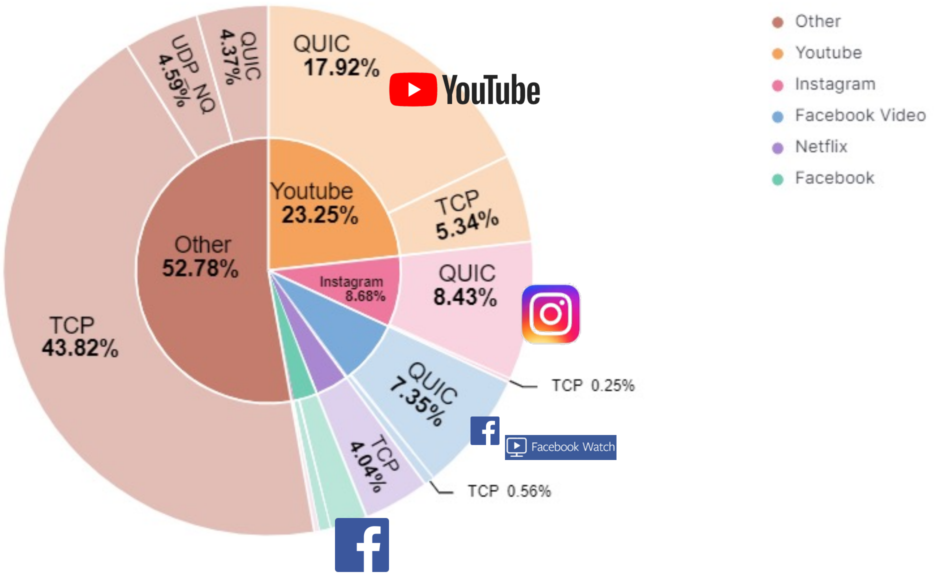
(snapshot 11/2/2022)

Network Traffic by Volume and Flows

The big flows that matter are predominantly QUIC

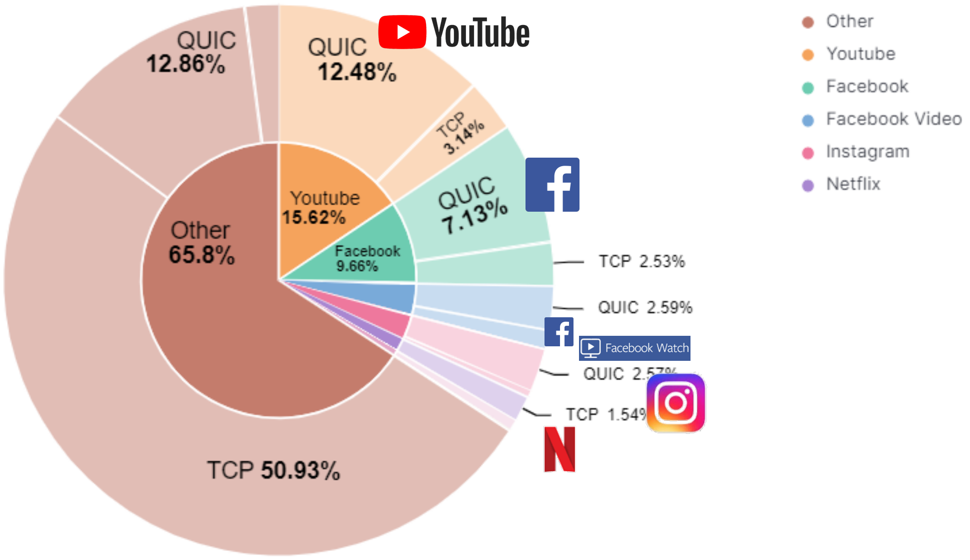
Overall Volume by Apps

Big 5 is 48% of traffic
QUIC is 40% of traffic
“other traffic” still largely TCP, QUIC now visible (4.3%).



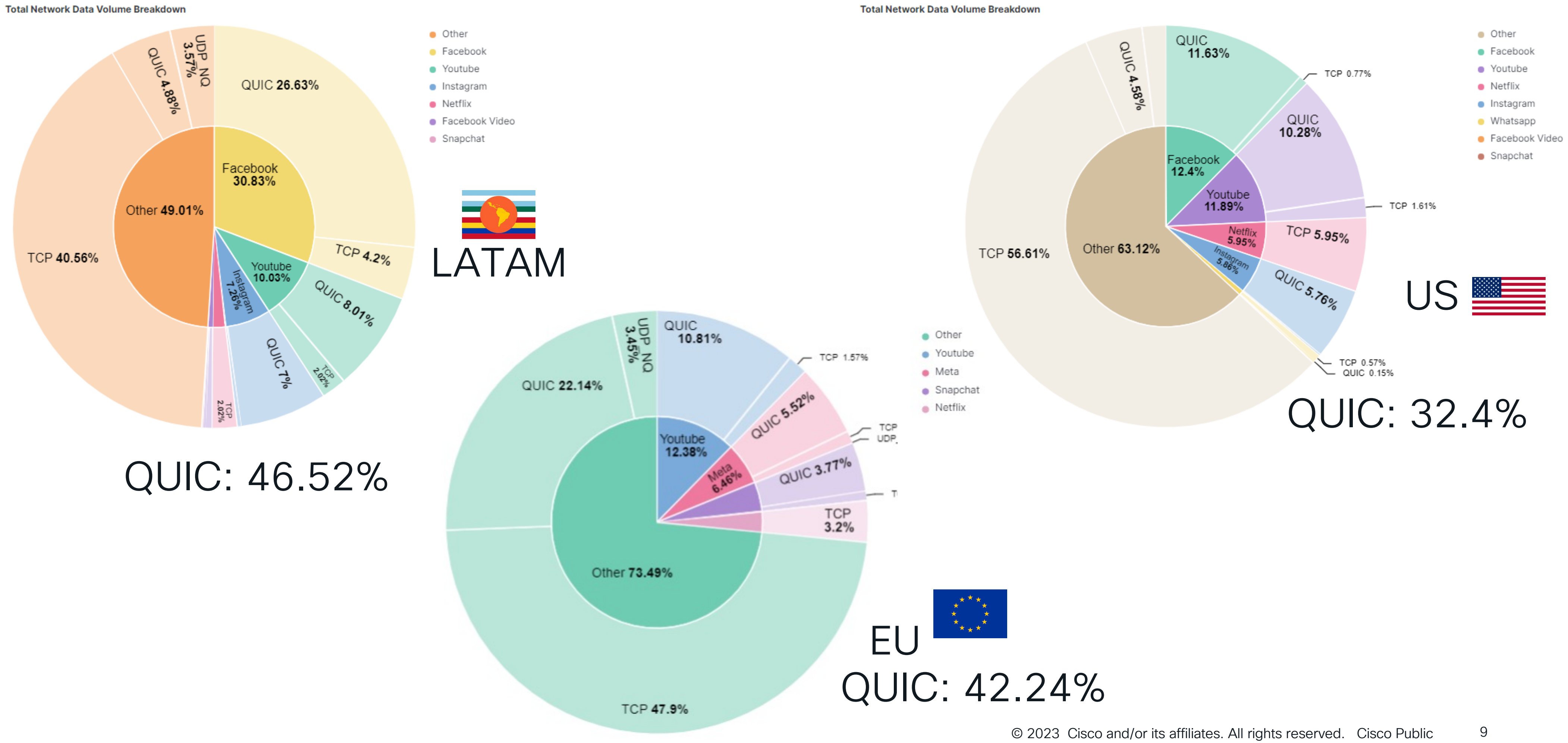
Total Flows by Apps

Lots of TCP sessions (likely IOT related, transactional related)
Big 5 APPs QUIC sessions are very targeted and high efficienc
(video related behaviour); fewer but higher in volume



(snapshot 11/2/2022)

The pattern persists worldwide into 2023



LATAM

QUIC: 46.52%

US

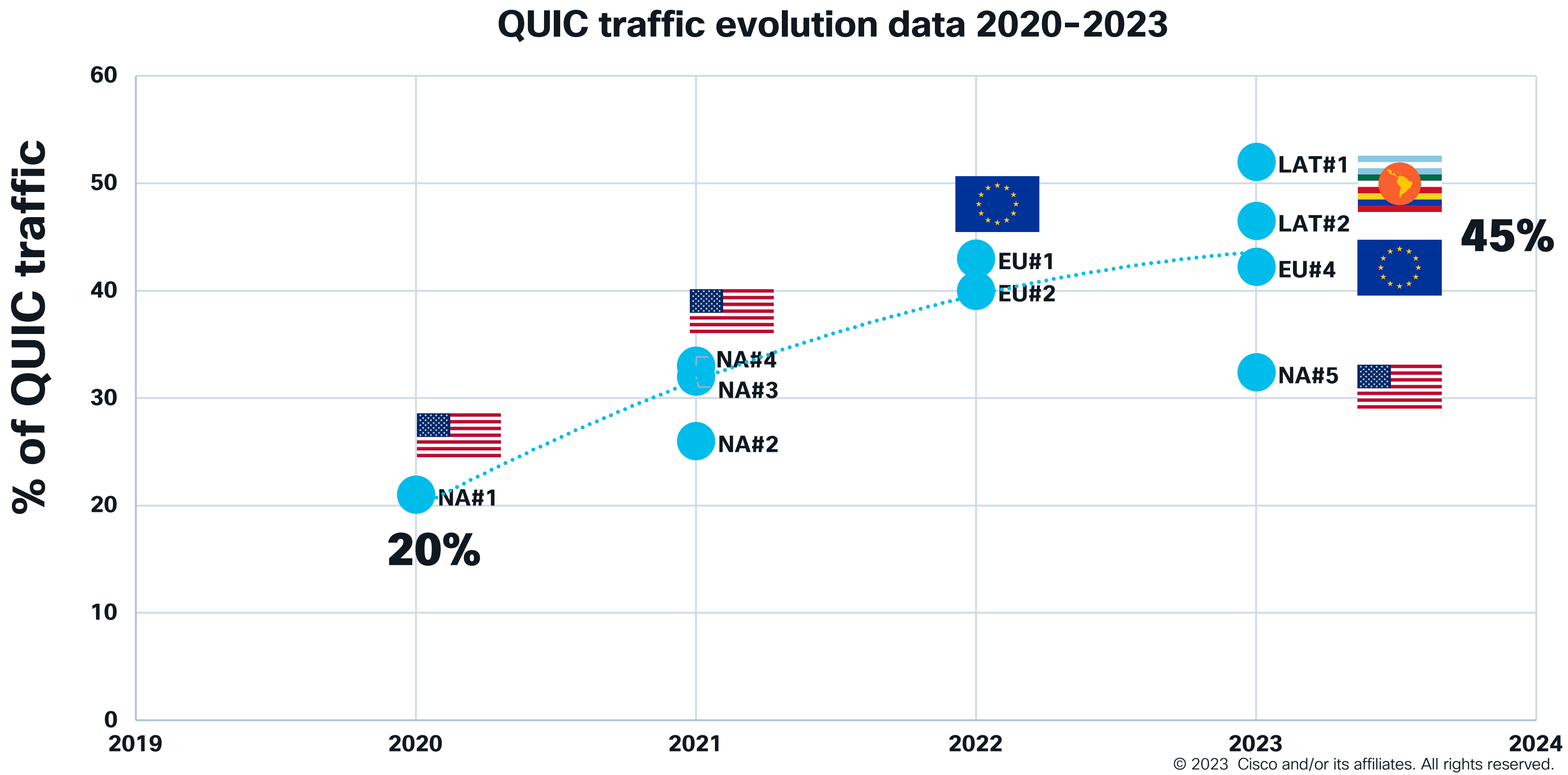
QUIC: 32.4%

EU

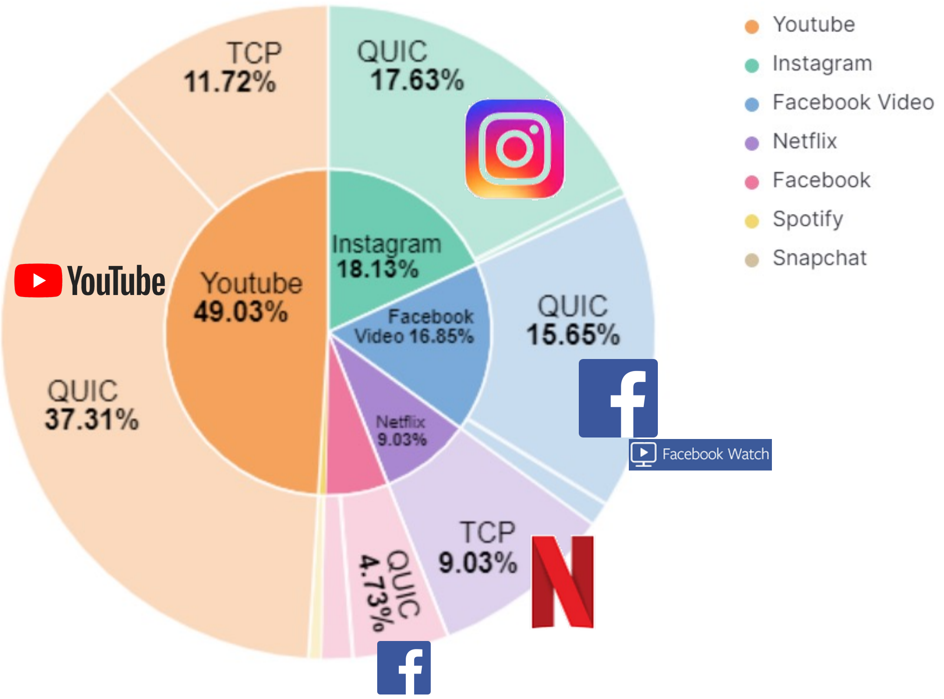
QUIC: 42.24%

QUIC is growing across the world

various snapshots



Top 5 Apps – QUIC is dominant 80/20 rule now

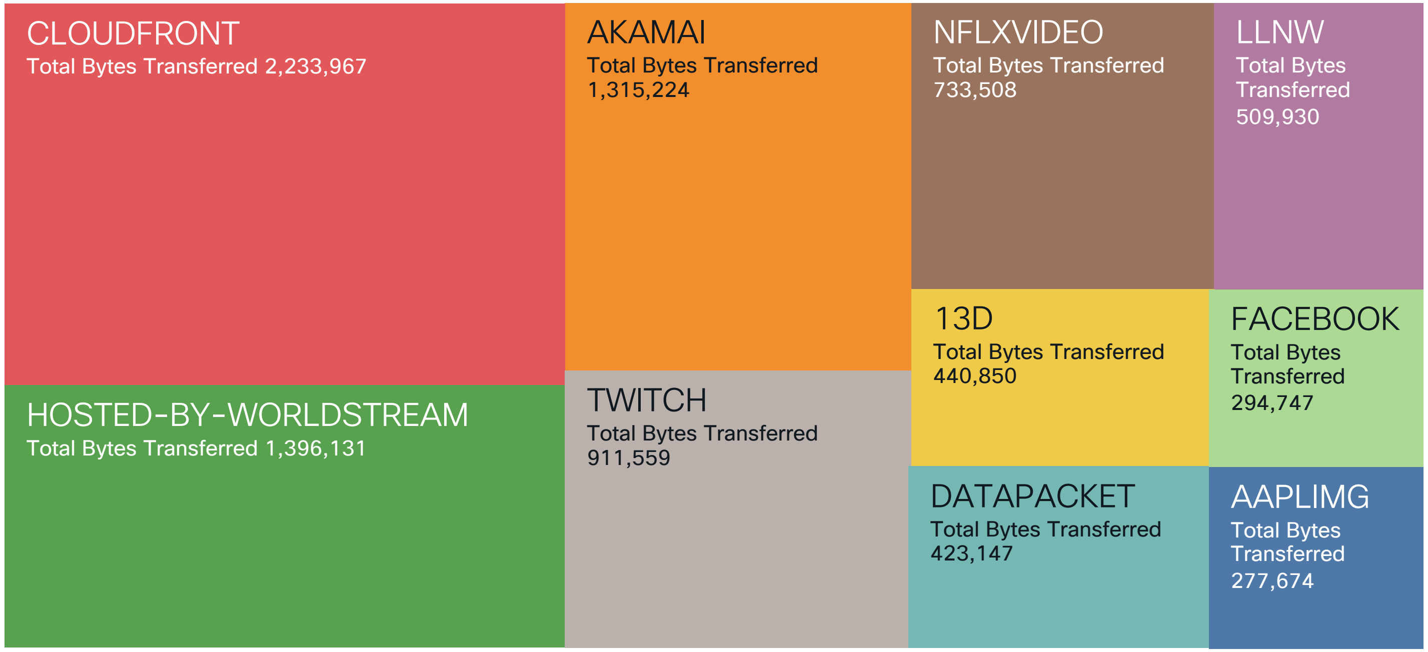


April 10 2022

Fixed Broadband: It’s not that different – May 2022

if different sources

Data Volume Distribution by Hostname



CDN

Hosting

Gaming

Video Streaming

Profile aligned with Fixed Broadband traffic (browser driven traffic)

QUIC : 41%   TCP: 53%   UDP (other): 6%

*source Tier 1 EU SP

The New IP stack

New Stack, New Behaviour

An application driven global transition

HTTP/3 Stack = UDP+QUIC+TLS

Old App Stack

- HTTP/1.1/2
- TLS
- TCP
- IP

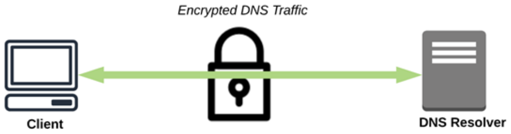
New App Stack

QUIC – RFC 9000
HTTP/3 – RFC9114

- HTTP/3
- QUIC + TLS1.3
- UDP
- IP

DoH

DoT – RFC7858
DoH – RFC8484



DNS communication over HTTPS/TLS

eSNI / ECH

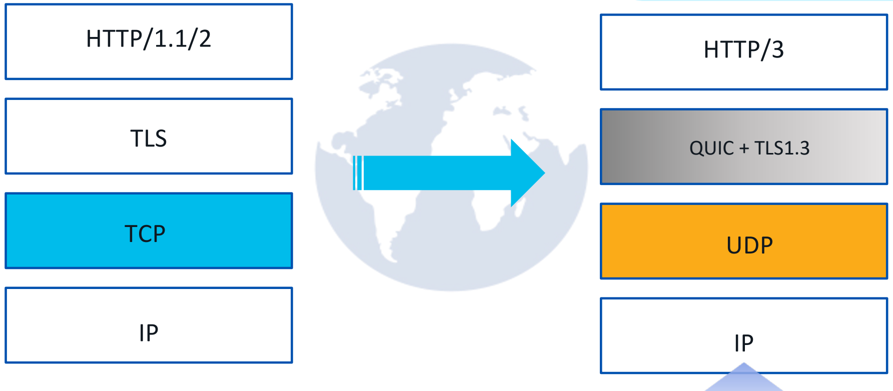
RFC8744

Large Scale Adoption

DPI is gone

HTTP/3 Stack = UDP+QUIC+TLS+H3+DoH+eSNI/ECH

Old App Stack



HTTP/1.1/2

TLS

TCP

IP

New App Stack

QUIC – RFC 9000
HTTP/3 – RFC9114

HTTP/3

QUIC + TLS1.3

UDP

IP

- *Improved Security*
- *Multi-session*
- *Improved QoE*
- *APP friendly design*

Large Scale Adoption

✚ DoH

DoT – RFC7858
DoH – RFC8484

Application Controlled DNS
DNS Traffic not observable

Google & CloudFlare serve 50% of global DNS requests
Both support DoH
All major OSs & Browsers support DoH (Firefox Defaults for US to CloudFlare)

✚ eSNI / ECH

RFC8744

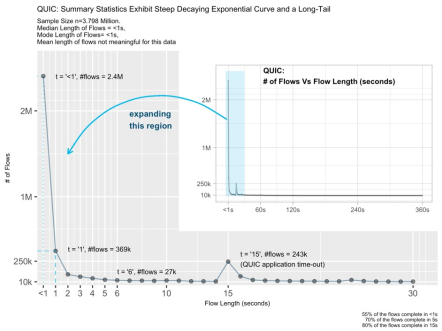
Target Domain is opaque / unobservable

DPI Ineffective

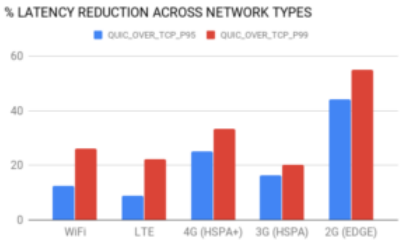
including alternative hints e.g. DNS or SNI analysis

QUIC Moves Control of the User Experience to the App

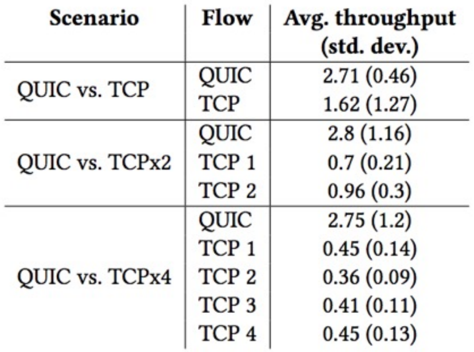
Apps do not play nice – they will deliver over everyone else



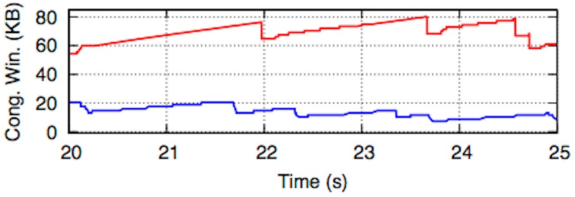
70% of interactions complete in <5s**



The poorer the network, the better the improvement*



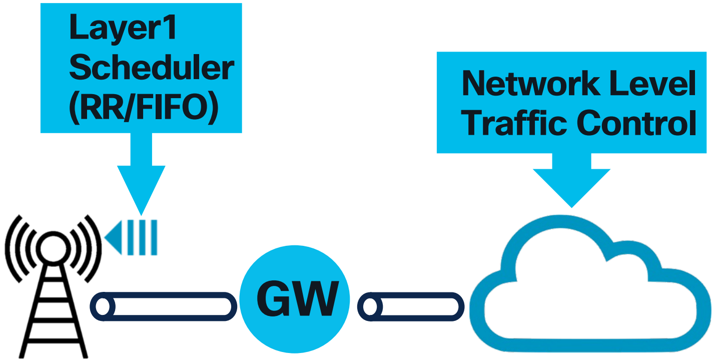
Scenario	Flow	Avg. throughput (std. dev.)
QUIC vs. TCP	QUIC	2.71 (0.46)
	TCP	1.62 (1.27)
QUIC vs. TCPx2	QUIC	2.8 (1.16)
	TCP 1	0.7 (0.21)
	TCP 2	0.96 (0.3)
QUIC vs. TCPx4	QUIC	2.75 (1.2)
	TCP 1	0.45 (0.14)
	TCP 2	0.36 (0.09)
	TCP 3	0.41 (0.11)
	TCP 4	0.45 (0.13)



QUIC is “Unfair”***

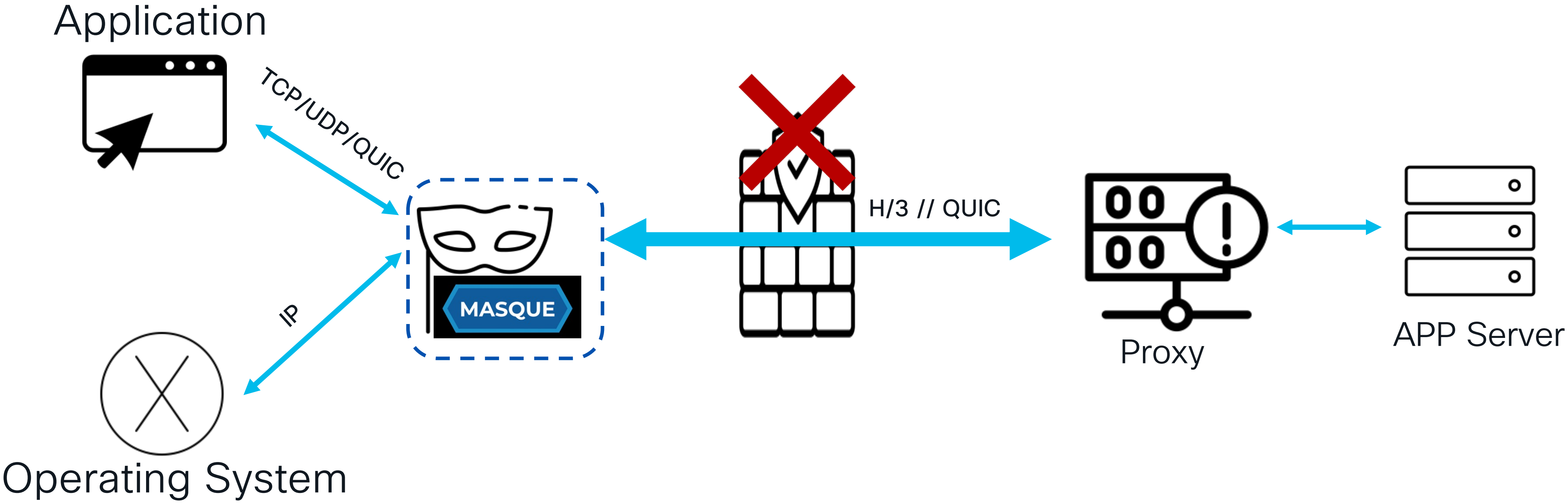
Impacted Areas

(e.g. wireless access)



*uber engineering;**Cisco Analysis, cust.data;***APNIC study

Tunneling is a new threat vector (exfiltration tool?)

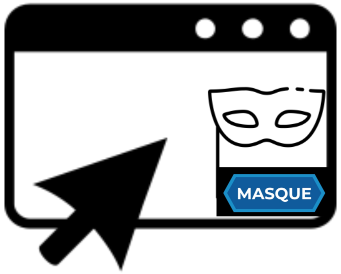


MASQUE

Multiplexed Application Substrate over QUIC Encryption

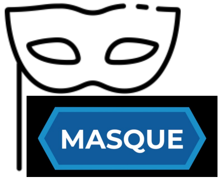
Goal is to develop mechanism(s) that allow configuring and concurrently running multiple proxied stream- and datagram-based flows inside an HTTP connection.

Options for Masque

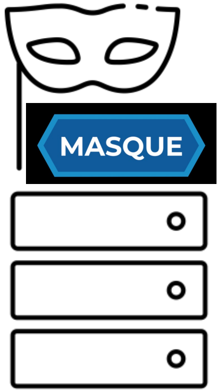


Inside the App

Inside the O/S

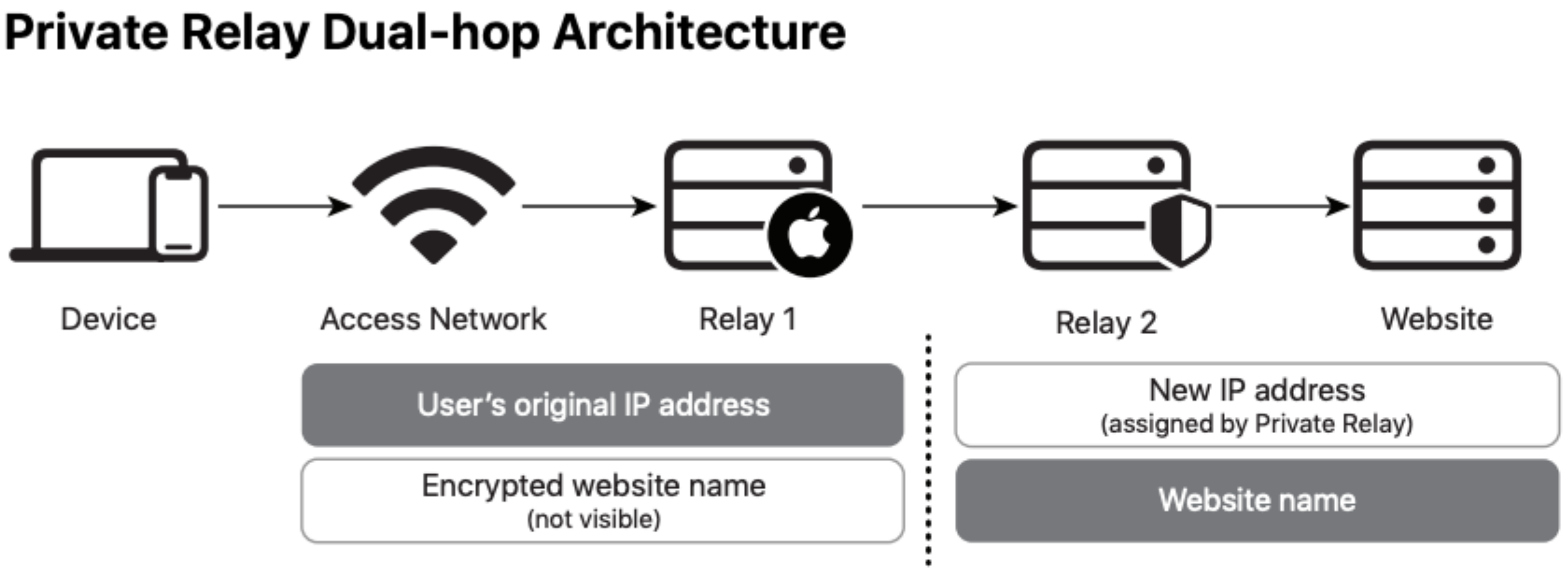


Client to O/S



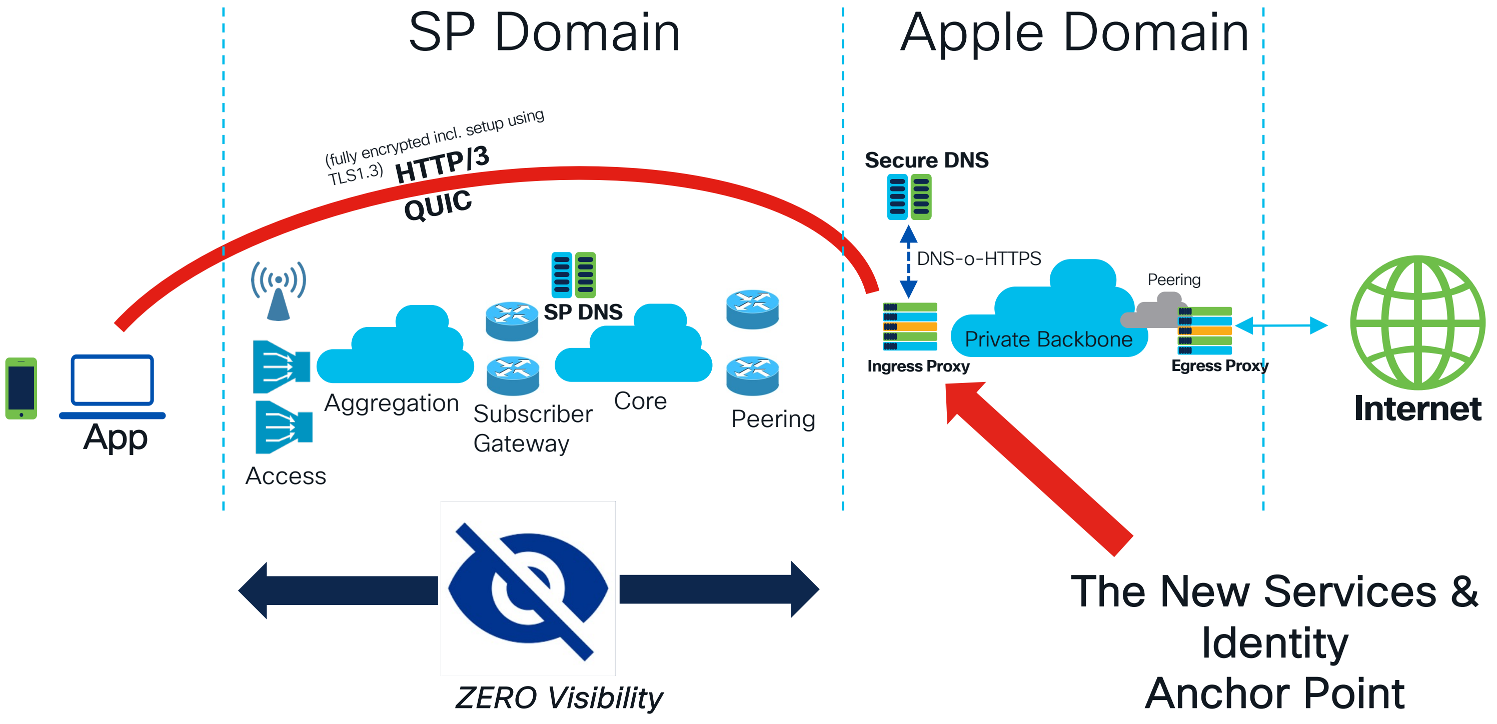
Network Appliance
(tunnel IP)

Apple Private Relay: Dual Hop Masque



Decoupling users from content

SP Domain has less insights on traffic

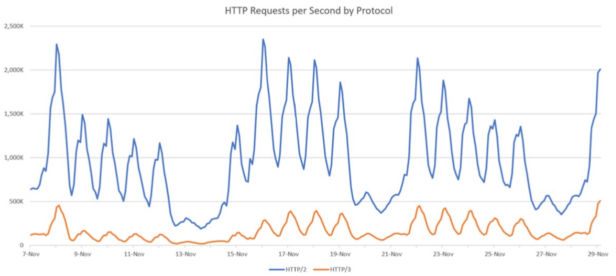


The New Services & Identity Anchor Point

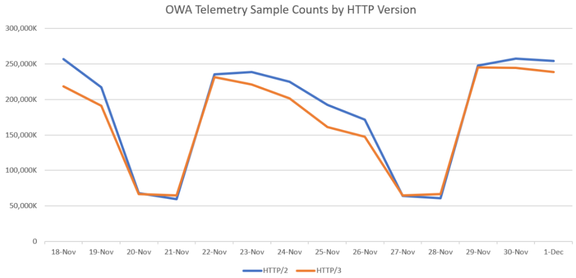
QUIC at MSFT*

Pervasive across Products

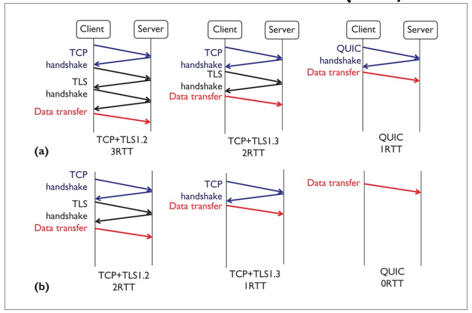
- 70% of worldwide front-end servers deployed latest Windows Server with HTTP/3 support
- Chart below shows **all** EXO H2/H3 usage; including browser, mobile and desktop clients



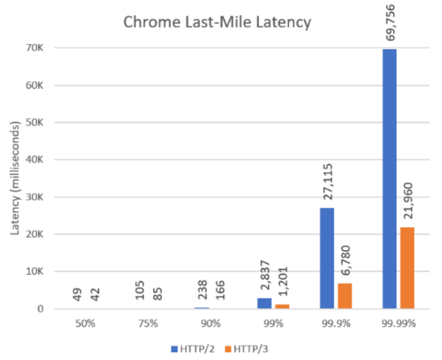
Easy to adopt

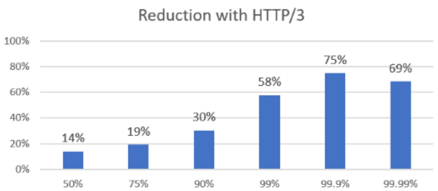


Outlook runs on Quic/H3



SMBoQUIC – No VPN





Outlook web access *actually* runs better using H/3

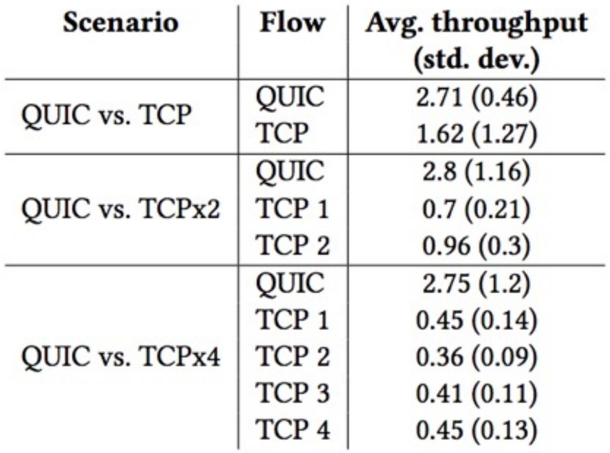
* Source: EPIQ 20212, Nick banks, MSFT

QUIC/H3/DoH stack is in business

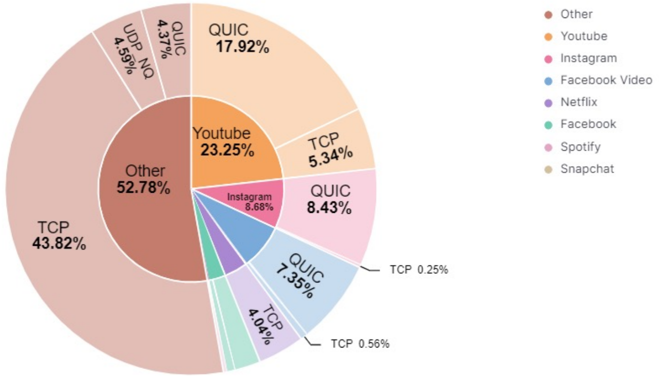
fastly
CLOUDFLARE
Akamai
android
Google
Microsoft
aws
Uber
YouTube

Content Delivery   Security   Privacy   Loadbalancing   App Infrastructure   App Experience

Net Neutrality has effectively been subverted



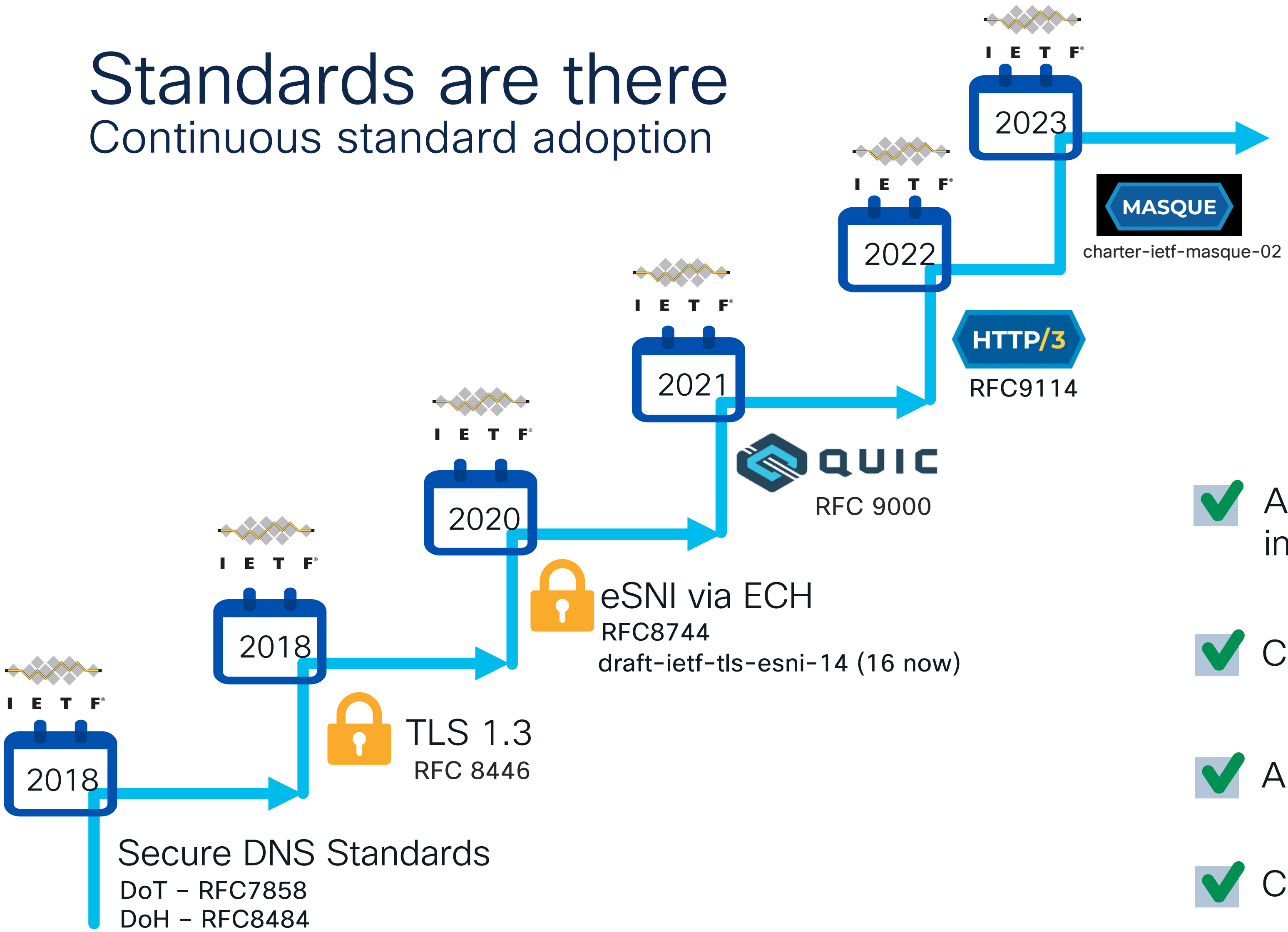
Scenario	Flow	Avg. throughput (std. dev.)
QUIC vs. TCP	QUIC	2.71 (0.46)
	TCP	1.62 (1.27)
QUIC vs. TCPx2	QUIC	2.8 (1.16)
	TCP 1	0.7 (0.21)
	TCP 2	0.96 (0.3)
QUIC vs. TCPx4	QUIC	2.75 (1.2)
	TCP 1	0.45 (0.14)
	TCP 2	0.36 (0.09)
	TCP 3	0.41 (0.11)
	TCP 4	0.45 (0.13)



- Net Neutrality implicit assumption is that during network congestion the network will impartially impact all flows – **and that all flows will respond in the same way (TCP assumption)**
- App owned flow control breaks this assumption conclusively
- Therefore ~50% of the traffic in the internet is no longer conformant to neutrality principles

Standards are there

Continuous standard adoption



✓ At scale, in production

✓ Client — android, chromium, Microsoft Edge, Mozilla

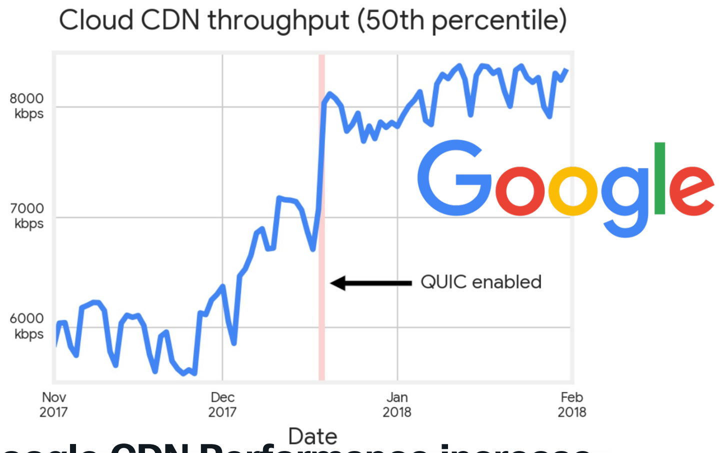
✓ Application

✓ Cloud

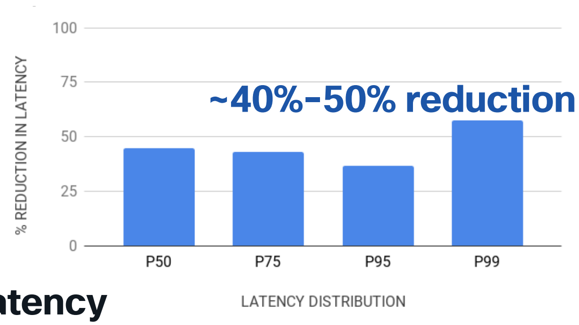
The consumers are observing benefits

QoE Drives QUIC Adoption

1.8B Daily Active Users – 3B Monthly QUIC and H/3 are protocols of choice*

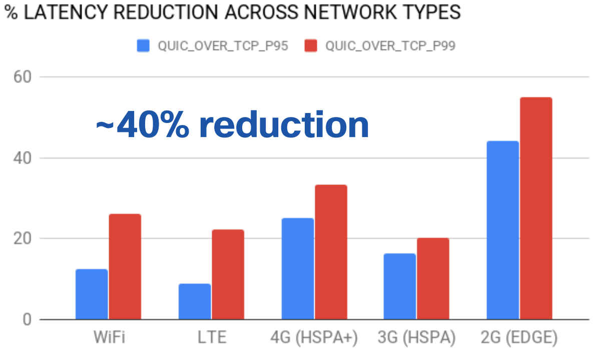


Google CDN Performance increase



~40%-50% reduction

Latency reduced significantly**



~40% reduction

The more fragile the network, the more QUIC excels**

*source Facebook engineering

** source Uber engineering

SP Services Portfolio needs assessment
(non-exhaustive list)

Differentiated Billing

- *Zero rated Apps*
- *App aware service*

Regulated Services

- *Site blocking*
- *Traffic intercept*

Traffic Management

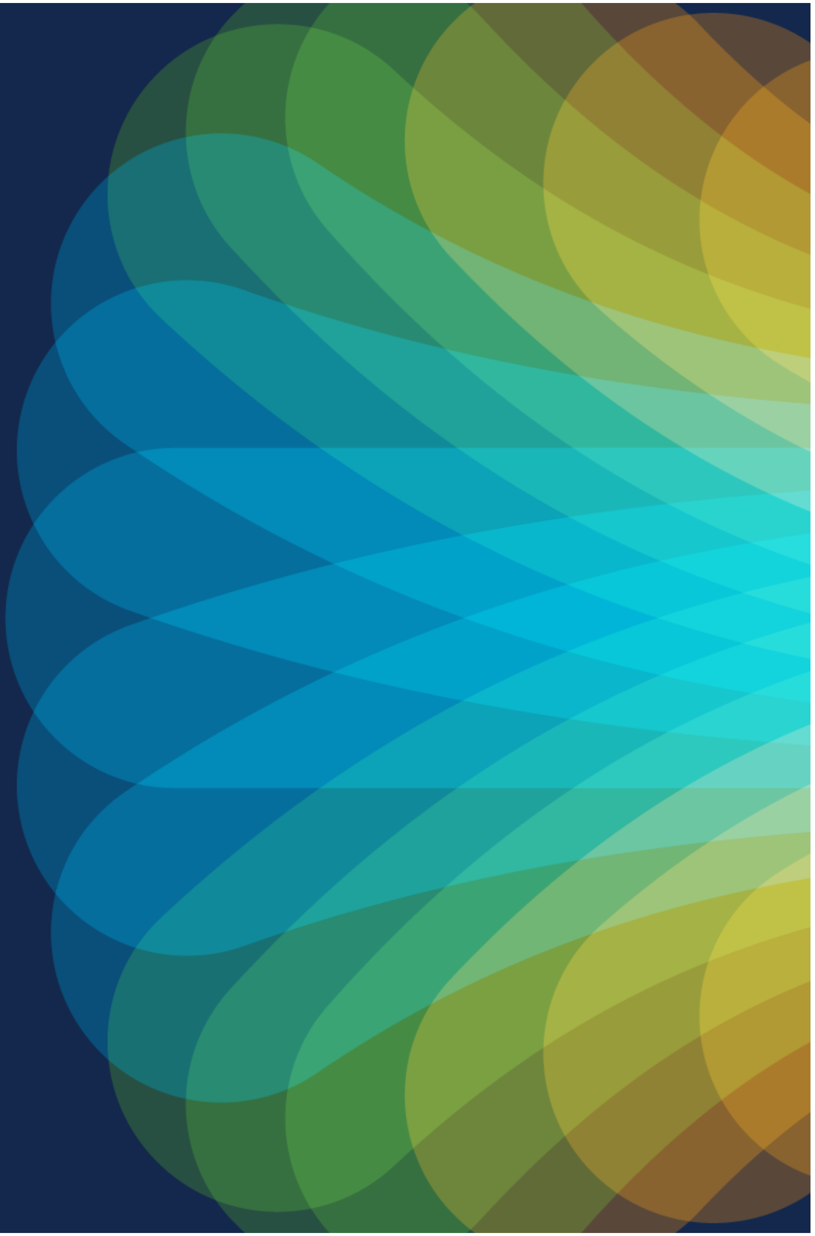
- *Peering*
- *Optimal interconnect*

Business Services

- *VPN*
- *Security*

non-exhaustive list

So what can we do?



Customers are looking for solutions

Example Use Cases Asked

Manage video downloads vs video streaming, downloads being the priority

DPI won't work anymore in QUIC

Recognise type of flow and act accordingly

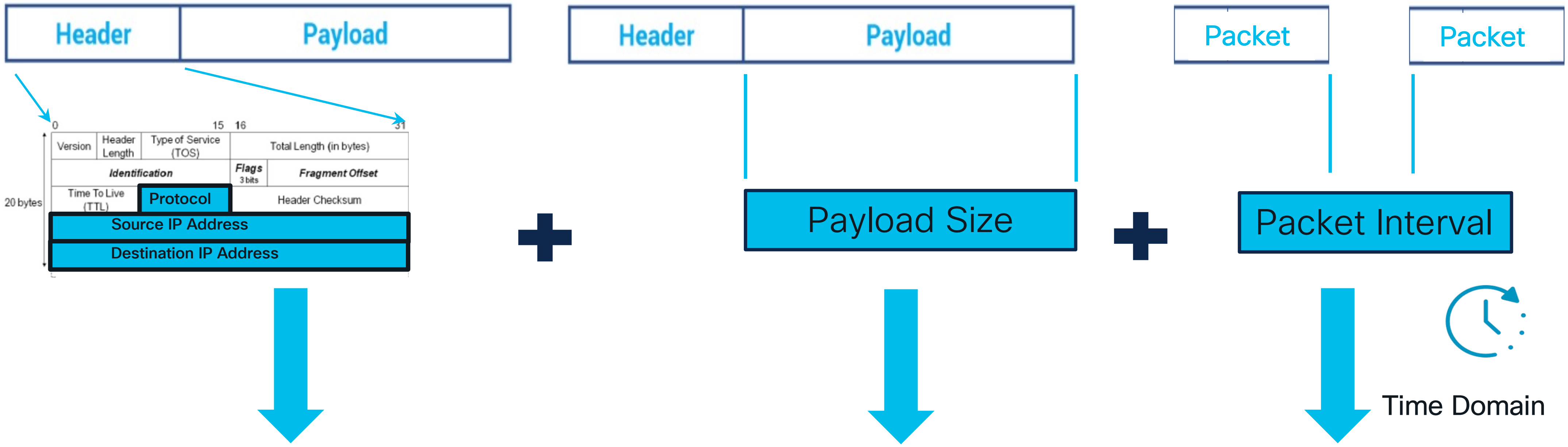
Manage Snap video vs Snap apps

Same problem

Account for encrypted traffic in terms of source/destination

More generically: Identify and manage QUIC flows; mitigate impact on Radio; optimise against industry metrics; future-proof network smarts

There is some information that will not go away

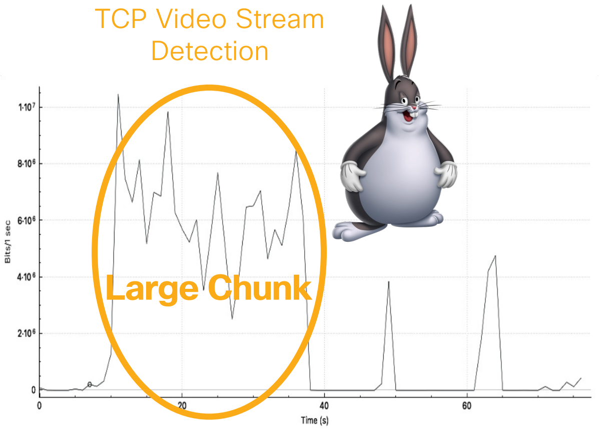


CDN Information

Traffic Volume in Time Information

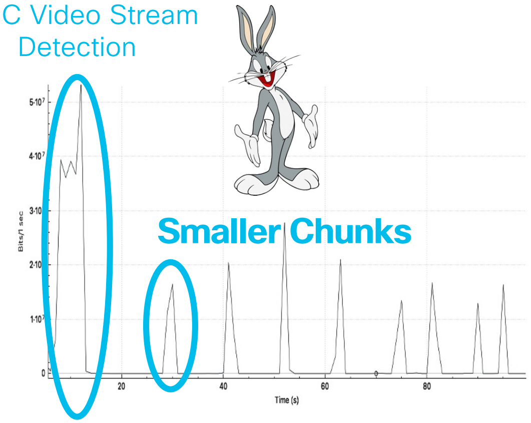
App (e.g. Video) Behavior varies by protocol and use case

TCP Video Stream Detection



TCP based ABR video players prefer larger, sustained downloads due to high cost of establishing the TCP session and reducing time spent in TCP slow start.
Often use HTTP/2 connection. (DASH/HLS) to fix HOL.

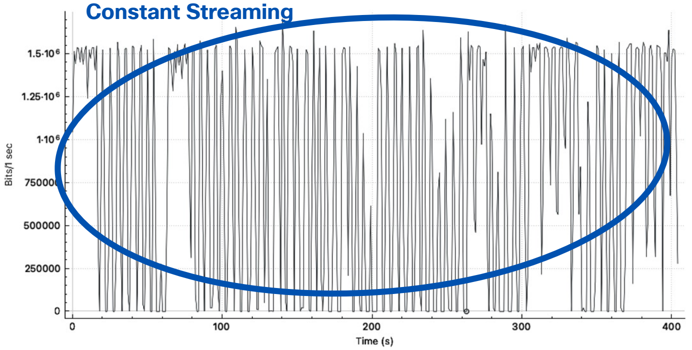
QUIC Video Stream Detection



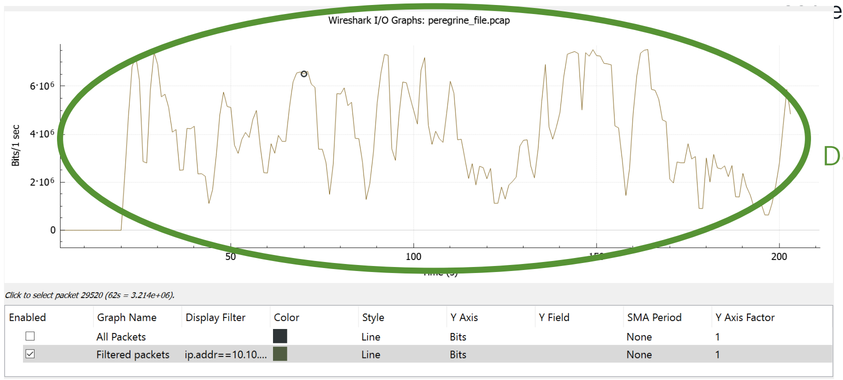
QUIC based ABR video players prefer requesting video in smaller chunks.

Multiple QUIC Streams in many cases to (different) ers

UDP Video Live Stream Detection



UDP based video players are extremely reliant on consistent network performance. Small buffer, sustained T'put
Applications: YouTube Live, WebEx, Microsoft Teams, Zoom



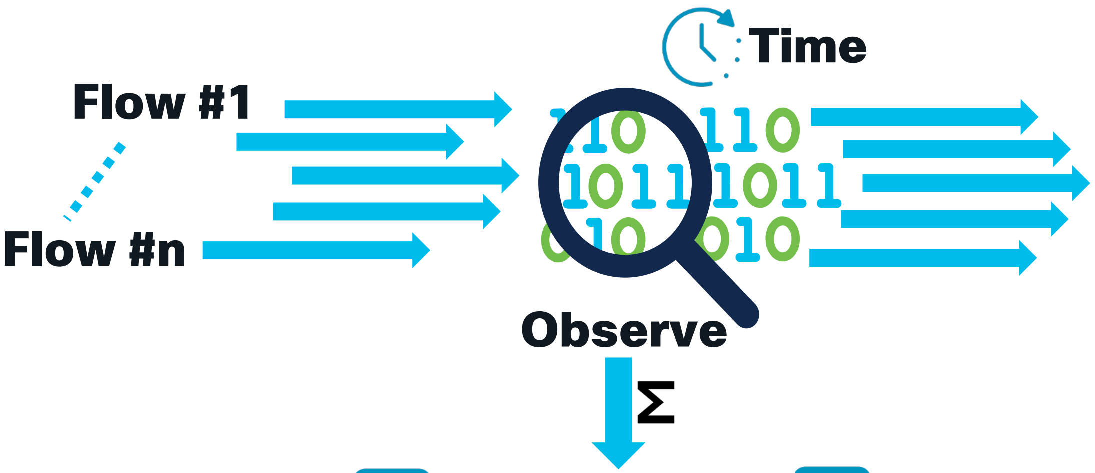
Click to select packet 29520 (62s = 3.214e+06).

Enabled	Graph Name	Display Filter	Color	Style	Y Axis	Y Field	SMA Period	Y Axis Factor
☐	All Packets		■	Line	Bits		None	1
☑	Filtered packets	ip.addr==10.10....	■	Line	Bits		None	1

Download Stream Detection

Time Domain Flow recognition

- Observe all flows
- Profile per flow (Time domain matched)
- The resulting profile will allow to distinguish the nature of the flow
  - Content Download
  - (x-Form) Streaming content
  - Real time 2 way communication
  - Video/non-video
  - Short lived flows



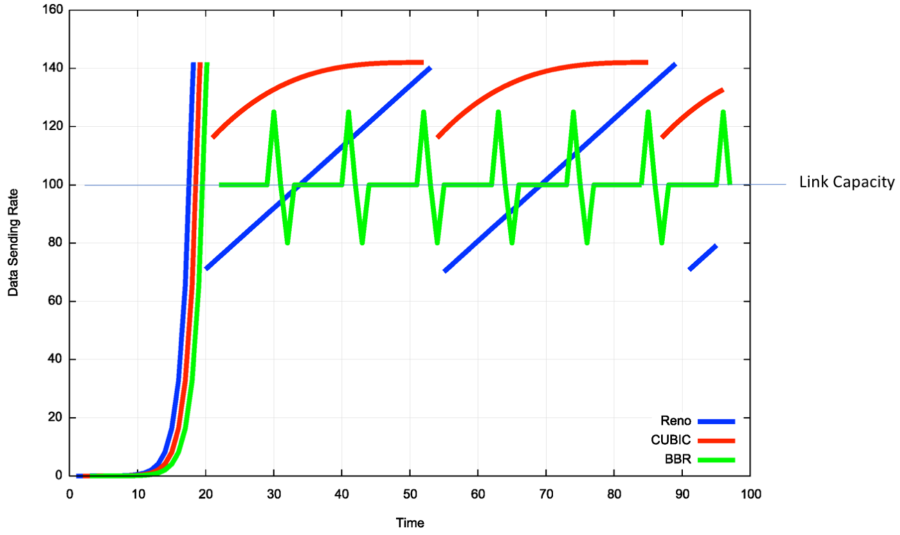
Category

- ● Download
- ● Real time 2-way
- ● Short-Mid-Long Form Video
- ● Real time one way
- ● Short form interactions
- ● ...

Inferring congestion

- Different congestion algo's have different behaviour
- Time-domain observation + anomaly detection -> congestion inference

Reno vs CUBIC vs BBR behaviour*

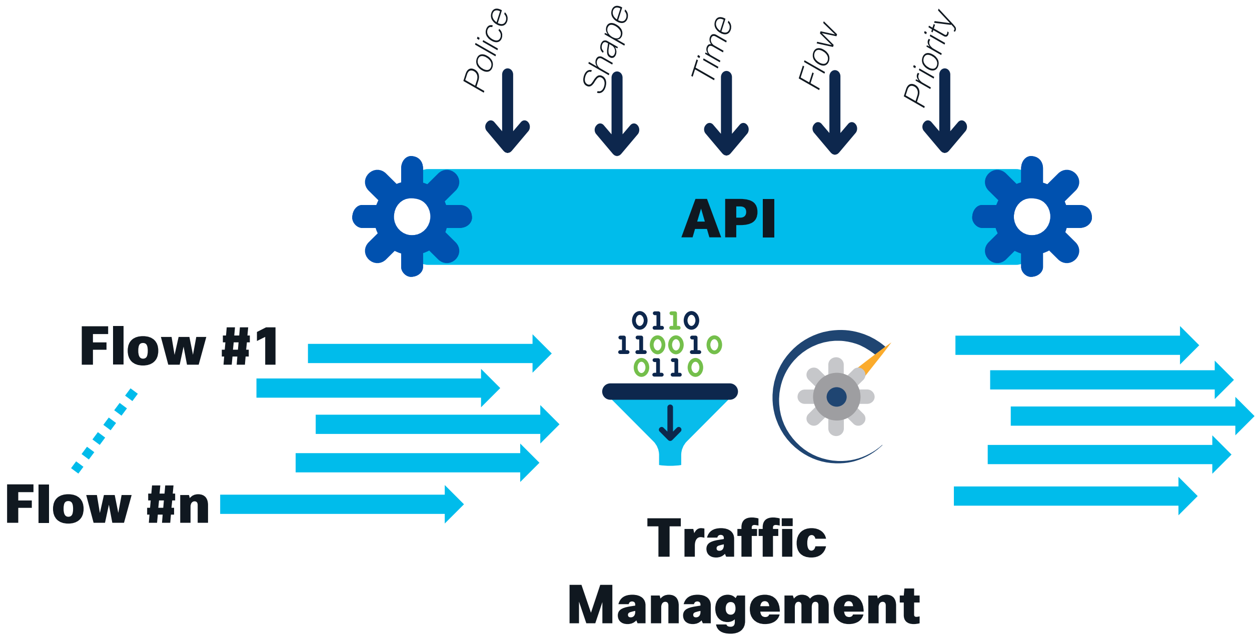


- Assessment of various flows in parallel
- Understand Protocol behaviour: congested or not
- This serves as input for Policy Application

* https://blog.apnic.net/2017/05/09/bbr-new-kid-tcp-block/

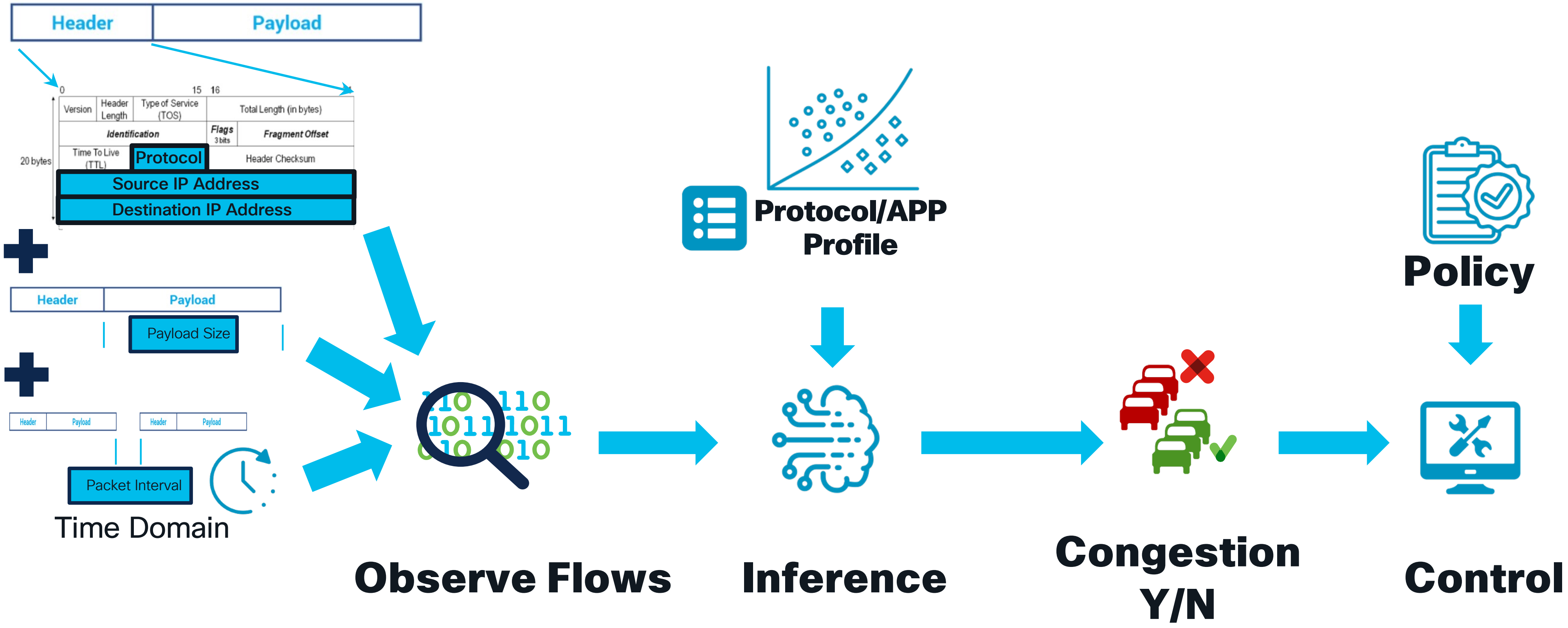
Programmable Traffic Management

- Traffic can be controlled in various ways.
  - Buffer
  - Discard
  - Flow control
  - …
- e.g. CUTO is a pre-compiled example where the parameters are implicitly configured



Overall System Logic

Basis for building use cases



Header | Payload

0 15 16

Version | Header Length | Type of Service (TOS) | Total Length (in bytes)

Identification | Flags 3 bits | Fragment Offset

20 bytes

Time To Live (TTL) | Protocol | Header Checksum

Source IP Address

Destination IP Address

Header | Payload

Payload Size

Header | Payload | Header | Payload

Packet Interval

Time Domain

Protocol/APP Profile

Policy

Observe Flows

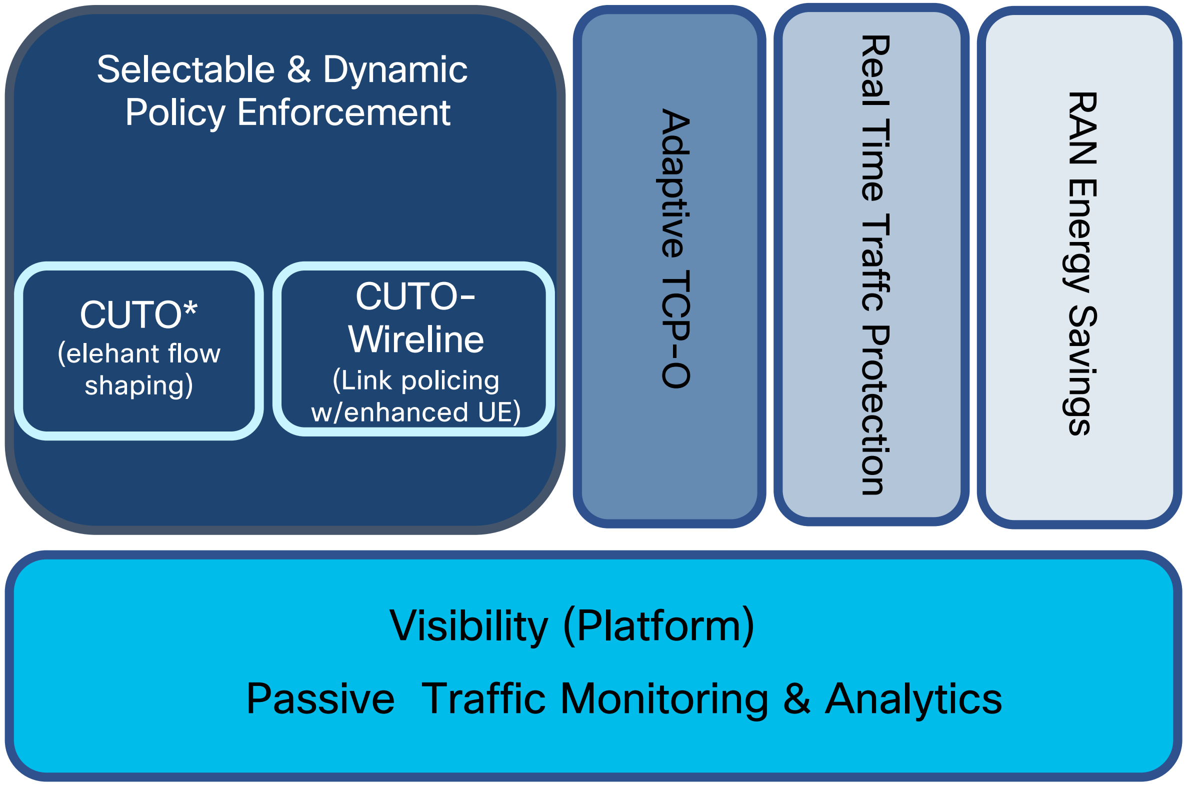
Inference

Congestion Y/N

Control

Use Cases Summary

Non-exhaustive list



Selectable & Dynamic Policy Enforcement

CUTO* (elehant flow shaping)

CUTO-Wireline (Link policing w/enhanced UE)

Adaptive TCP-O

Real Time Traffic Protection

RAN Energy Savings

Visibility (Platform)

Passive Traffic Monitoring & Analytics

*Cisco Ultra Traffic Optimization

Visibility (Platform)

(Passive) Traffic Monitoring & Analytics

Policy Enforcement Engine

Dynamic Policy Enforcement per (APN|MSISDN|Link|Base Station|...)

CUTO (Dynamic Congestion Alleviation by Elephant Flow Shaping)
CUTO-Wireline (Hard interconnect link policing while maintaing an enhanced User Experience)

Protection for Real-Time Traffic

Manage overall link congestion dynamically to protect RTP traffic (videoconf, collaboration, etc)

Adaptive TCP-O

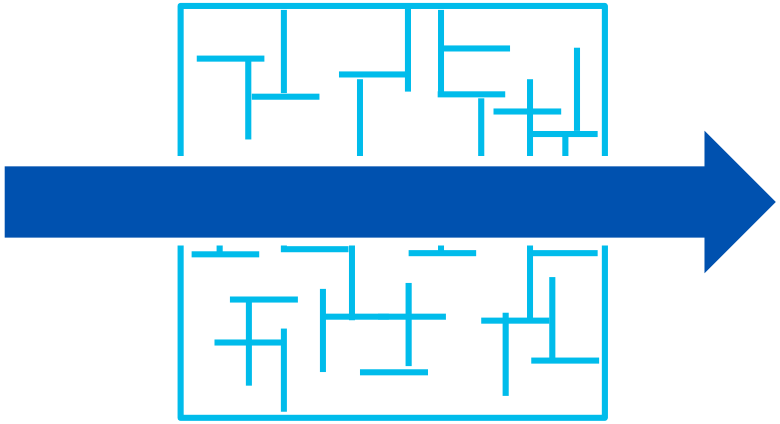
Based on current observed network state across all traffic across all protocols (including UDP & QUIC)

RAN Energy Savings / Sustainability

Dynamically switching bands on/off at a cell site to match IP based real-time traffic demand & QoE from customers.

Why does this scale

Simple



- I only use state on the important/interesting stuff
  - 20% of the flows generate 80% of the volume

Smart

- I only use state if I need it
  - when there is a reason e.g. congestion

Summary

- Traffic is encrypted, application controlled, and obfuscated
- Traditional DPI approaches (w)(d)on’t work
- This evolution will affect Service Provider consumer offering policy
- An IP centric approach is feasible and addresses several use cases

The bridge to possible

Thank you