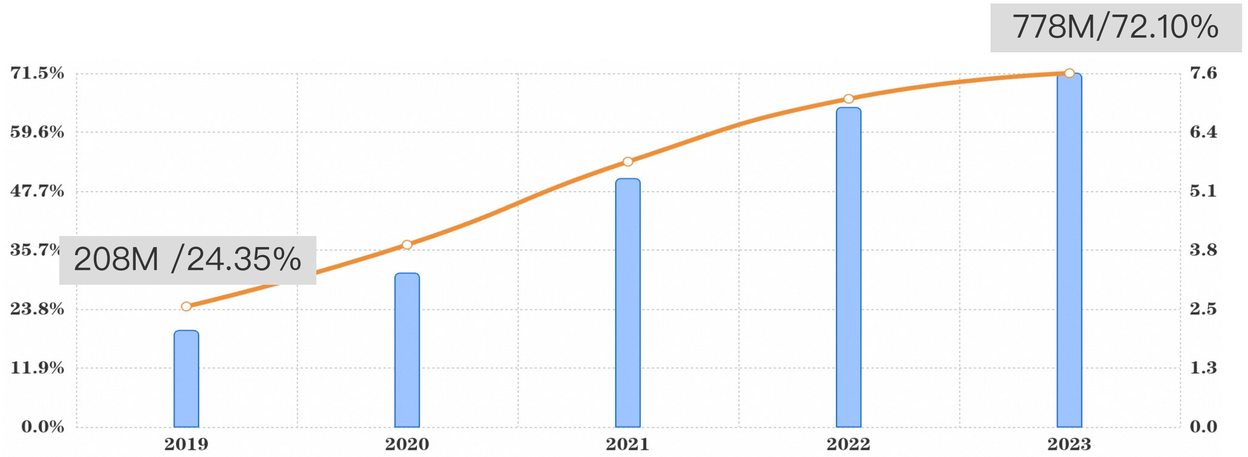
Large-scale IPv6 deployment and Practice @ Alibaba

Linjian (Davey) Song, linjian.slj@alibaba-inc.com
Network Infrastructure, Alibaba Cloud

Connections 2024

IPv6 development in China in one slides

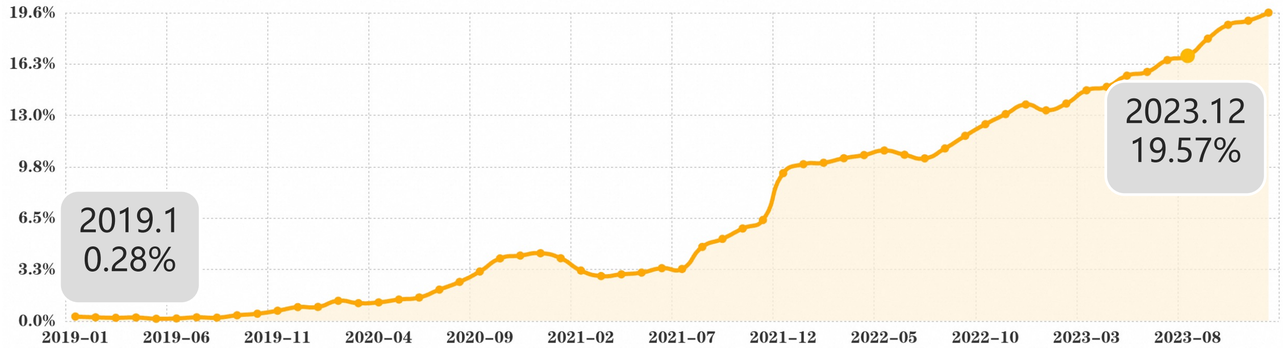


The number of active IPv6 internet users

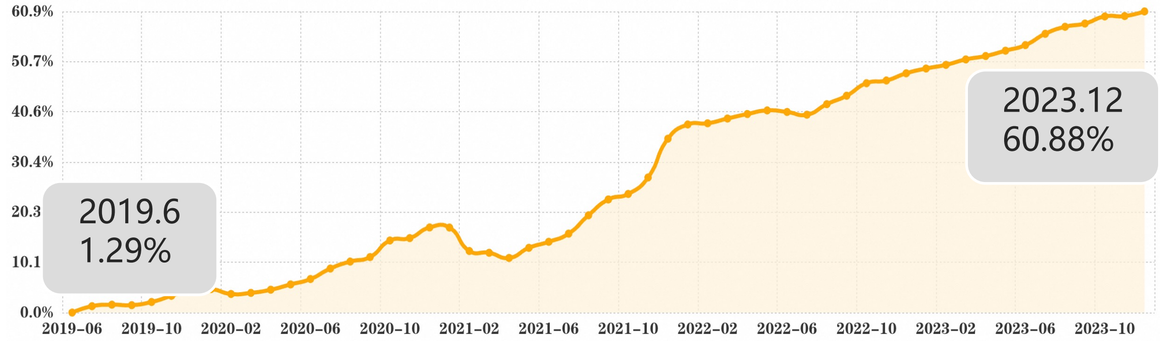
Top100 Apps in China

Govt Website in China

Financial Institution in China



Fixed network IPv6 traffic proportion



Wireless/Mobile network IPv6 traffic proportion

* Data Source : https://m.china-ipv6.cn/complete/#/

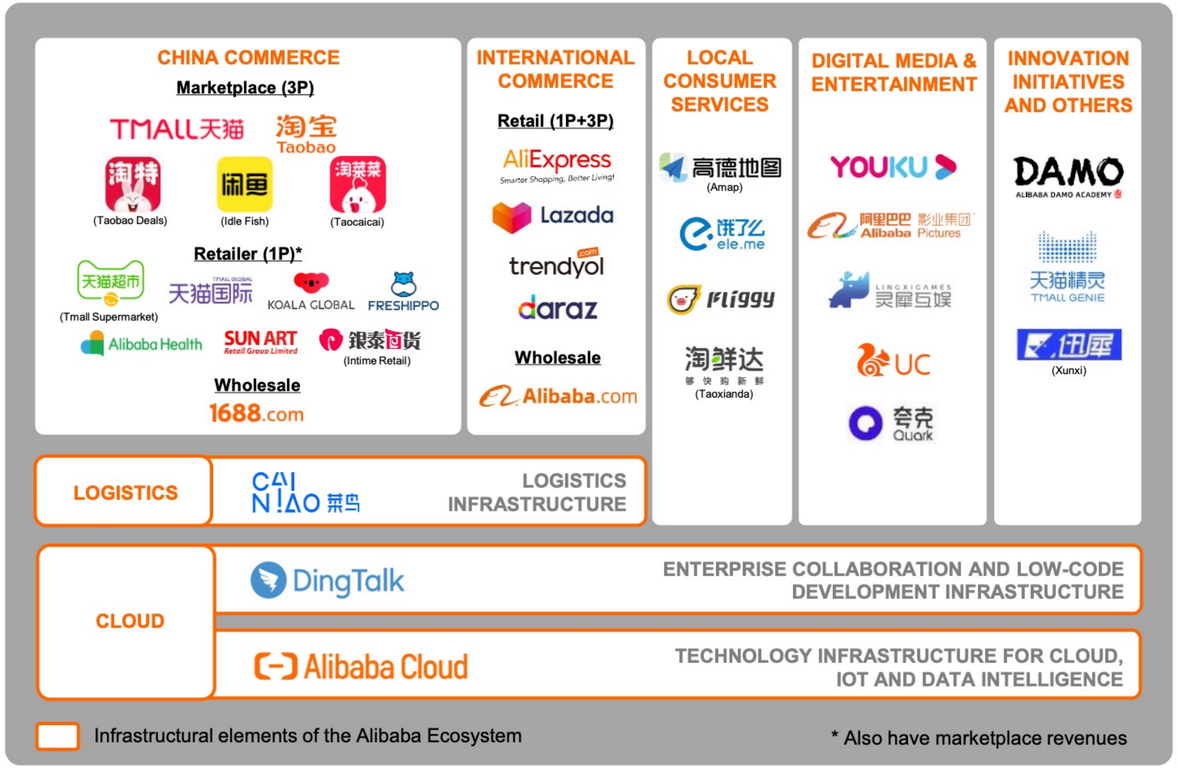
Brief Introduction on Alibaba Group

Alibaba Cloud

Alibaba Group's MISSION IS TO MAKE IT EASY TO DO BUSINESS ANYWHERE

Alibaba's Vision for Fiscal Year 2036

The Alibaba Ecosystem



Challenges in IPv6 transition for large-scale online service

Motivation and Challenges in IPv6 Deployment

Alibaba Cloud

Motivation

- IPv4 exhaustion and increasing prices
- Government IPv6 mandates
- IPv6 innovation and new capabilities

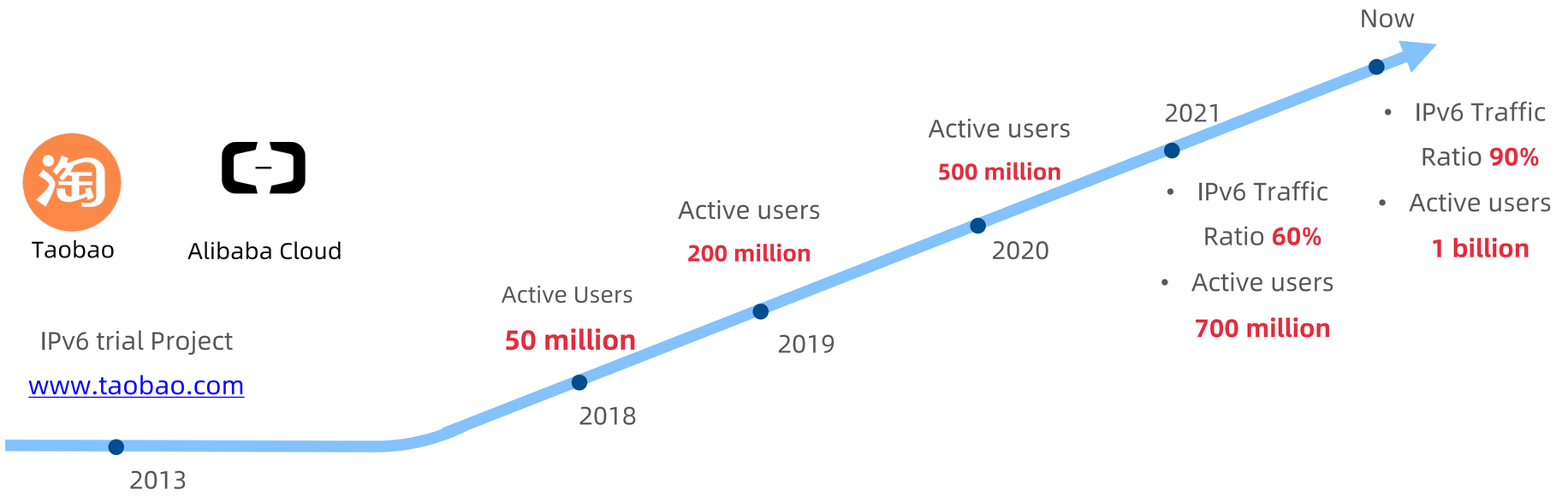
Challenges

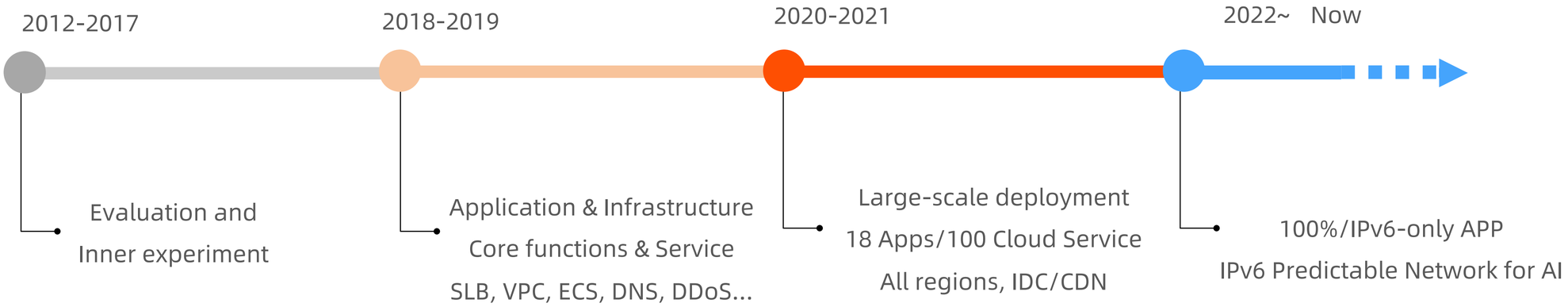
- Stability and risk issues during IPv6 transition for large-scale online service.
- Legacy infrastructure must support new services
- Application transition when Network is not fully ready (the early 2 years).
- Huge work on network and application upgrading

Technical Consideration

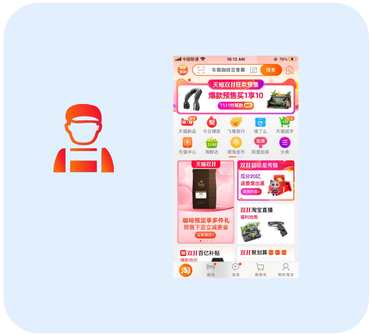
1. How to migrate online services smoothly?
2. How to decouple the upgrade progress on both network and application?
3. The IPv6/SRv6 network towards The Era of AI

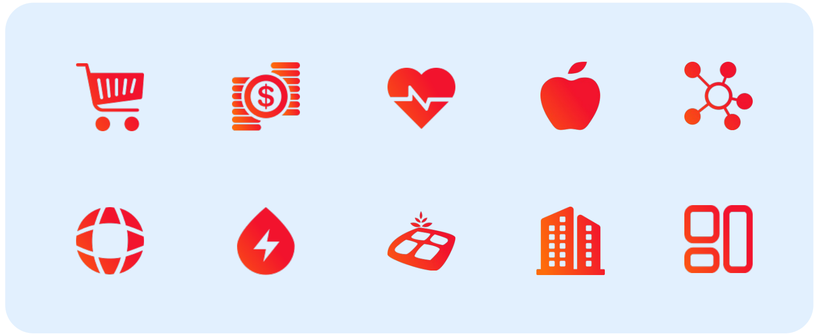
IPv6 Deployment Phases in Alibaba





High Level Online Services Architecture - Application





1. APP dual-stack adaption
2. IPv6 security
3. IPv6 transition: SLB64 ->SLB66
4. Application backend upgrade
5. APP monitoring & measurement

APP per device
dual-stack policy

TCP option (253)
carrying IPv6 address

App Client

SDK

Users scheduling

Internet
(ISP network)

CDN

DNS

Public BGP Networking

Security Defense
(DDoS/ Risk control)

Access Point

SLB64

IPv4

L4-LB

L7-LB

IPv6

SLB66

Network infrastructure (Alibaba IDC)

App Servers

Business Capability

Middleware

Data Base

AI

BIG Data

Container Platform

High Level Online Services Architecture – Cloud Platform

Alibaba Cloud

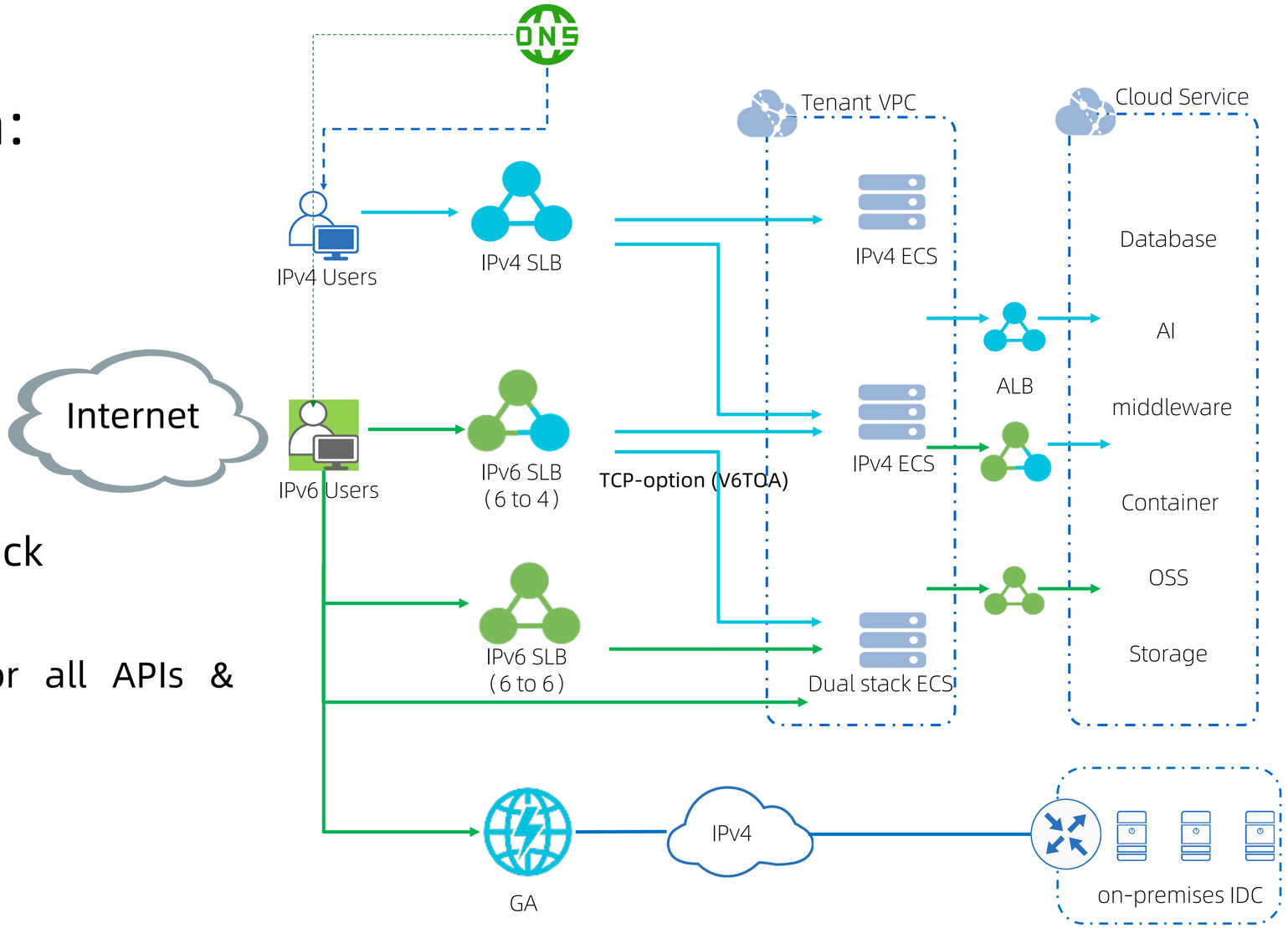
IPv6 transition solution:

Solution 1: Proxy

Solution 2: SLB64

Solution 3: SLB66/Dual stack

Towards IPv6-only: IPv6 for all APIs & system calls in a transaction



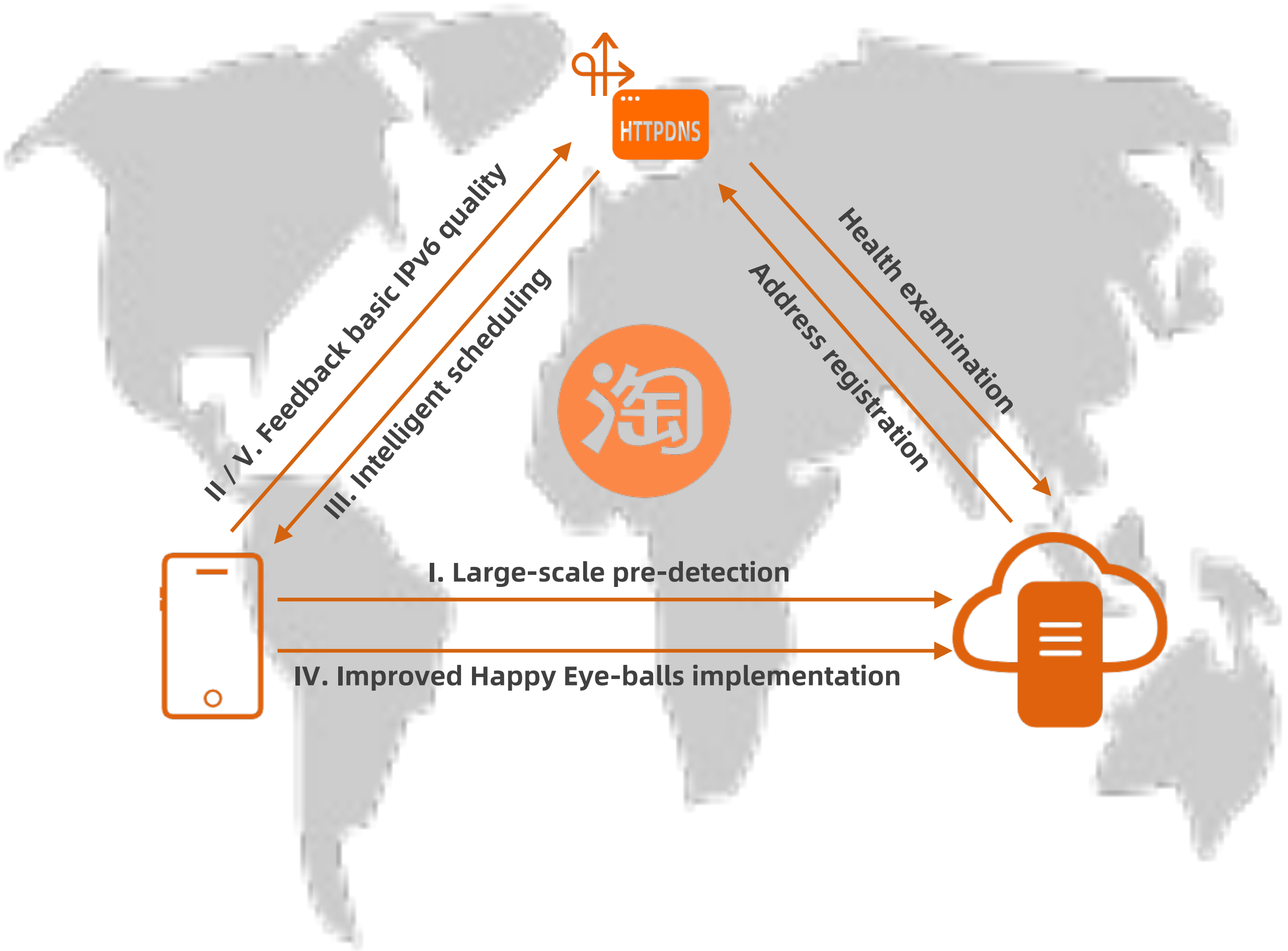
IPv6 Measurement

Alibaba Cloud

Alibaba application performance measurement Platform

- In mobile networks, the TCP connection success rate for IPv6 stands at 98%, closely matching that of IPv4. Conversely, this success rate drops to approximately 70-80% in wired Wi-Fi networks.
- The latency of TCP connections under both IPv6 and IPv4 protocols is approximately 50~60 milliseconds, with no significant differences observed
- The Coverage of dual-stack network is around 90% in 4G/5G network in 2023, and only around 28% in wired WIFI network

Dual-Stack adaption for Apps



Intelligent decision on IPv6 adoption on specific user groups

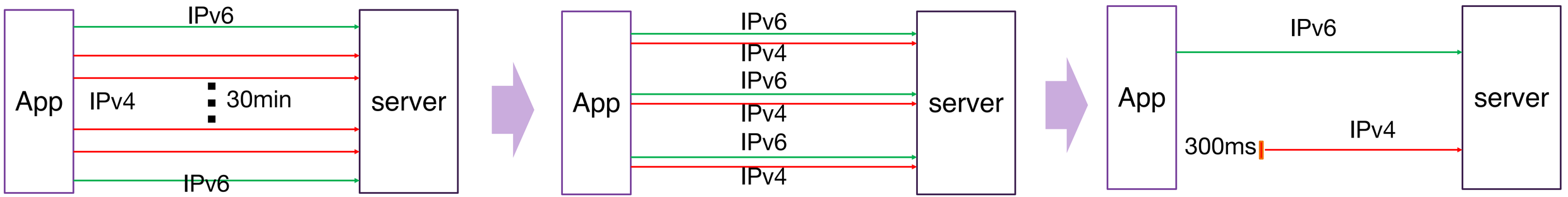
- **Large-scale pre-detection on IPv6 network quality with variety protocols:**
  - ICMP Ping
  - TCPing
  - HTTP
- **Refined intelligent scheduling base on variety factors:**

FACTORS	scheduling	Content
• Geo Location • Android/IOS • ISP network • 4G/5G/WIFI • Etc.	→	• IPs • Ports • HTTP2/3 • IP Priority • Etc.

- **Improved Happy Eye-balls implementation:**
  - **Compound DNS**: get AAAA and A response at same time in one request;
  - **Priority Address List**: the client received initial address list has been SORTED by intelligent scheduling system already;
  - **State 3-tuple**: the history of addresses of client attempting recording is IP-Port-Protocol tuple rather than just IP.

Increasing IPv6 traffic ratio

- 500+ Domains for Taobao APP, Set priority to enable AAAA records and third-party Domain
- Pre Pull for IPv6 address of domains based on intelligent decision
- **Dual-stack adaption and Happy eyeballs**



30 min Penalty
for IPv6 failure

IPv6 first
fallback to IPv4

Race IPv4 and IPv6 with
a connection attempt delay

IPv6 innovation in Alibaba Predictable Network

Alibaba Cloud

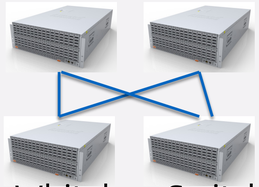
SONiC P4 Single-chip whitebox + IPv6 only(SRv6) + Programmability

Global Service Center

NetO Network Orchestrator

Global WAN

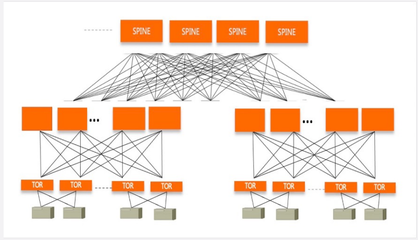
ePF（elastic peering fabric）



Fabric Controller

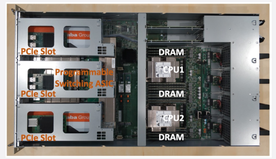
Whitebox Switch

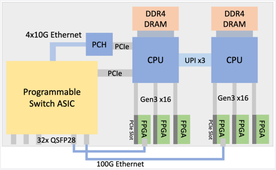
SRv6



Mega-scale
DC Network

High Performance
Large-scale AI Cluster





Hyper-converged Edge Platform

IPv6 simplicity and Fine-grained traffic engineering for one network evolution

Where are we now?

Alibaba Cloud

Positive Developments

- Established a fully connected IPv6 production network.
- IPv6 performance rivals or exceeds IPv4 in Alibaba tests.
- IPv6 traffic has surged to over 60% on wireless and 19% on fixed networks.
- Over 90% of Alibaba App users access via IPv6 on wireless networks.
- Dual-stack adoption has been confirmed as an effective strategy
- Widespread deployment of IPv6/SRv6 in data center networks.

Challenges

- Last-mile connectivity issues persist, with poor IPv6 wired network performance.
- Although practical, dual-stack hasn‘t reduced IPv6 demand or the financial burden of IPv4 trade.
- With long-term legacy network/system reliance, there's no clear plan or pathway to retire IPv4.

Conclusion

Alibaba Cloud

- IPv6 is a fully mature technology. Alibaba’s infrastructure and applications support IPv6 and enable 1 Billion users globally using IPv6.
- Stability and risk control are crucial considerations on IPv6 upgrade for large-scale online service
- Upgrading legacy infrastructure is much more complicated than building a new one. Keep IPv6 into consideration in the first place
- New momentum is observed. IPv6/SRv6 can simplify the network stack and provide efficiency in DC network

Alibaba Cloud

Thank You

Happy Eyeballs, Version 3: *Better Connectivity Using Concurrency*

Tommy Pauly
February 2024

Happy eyeballs is an algorithm that helps clients use IPv6 more

Why do **clients** want to use IPv6?

IPv6 is the **future**!

It's **new**! *(Only 26 years old...)*

Performance

IPv6 connections to dual-stack hosts on dual-stack networks:

120ms at **50th** percentile

400ms at **90th** percentile

IPv4 connections to dual-stack hosts on dual-stack networks:

170ms at **50th** percentile

670ms at **90th** percentile

IPv4 connections to IPv4-only hosts on dual-stack networks:

126ms at **50th** percentile

530ms at **90th** percentile

Fewer NATs

More optimized servers

More server locations

Better routing

Clients are a key element of the IPv6 transition.

IPv6 connectivity is a **prerequisite**,
but not enough

Clients need to **reliably** and **consistently** use IPv6

Why not **always** use IPv6 when it is available?

IPv6 connectivity is sometimes **broken**

Servers might be broken or slow on IPv6

Networks might be broken or slow on IPv6

Waiting on a single address is a **bad idea**

Network brokenness

Misconfigured firewalls

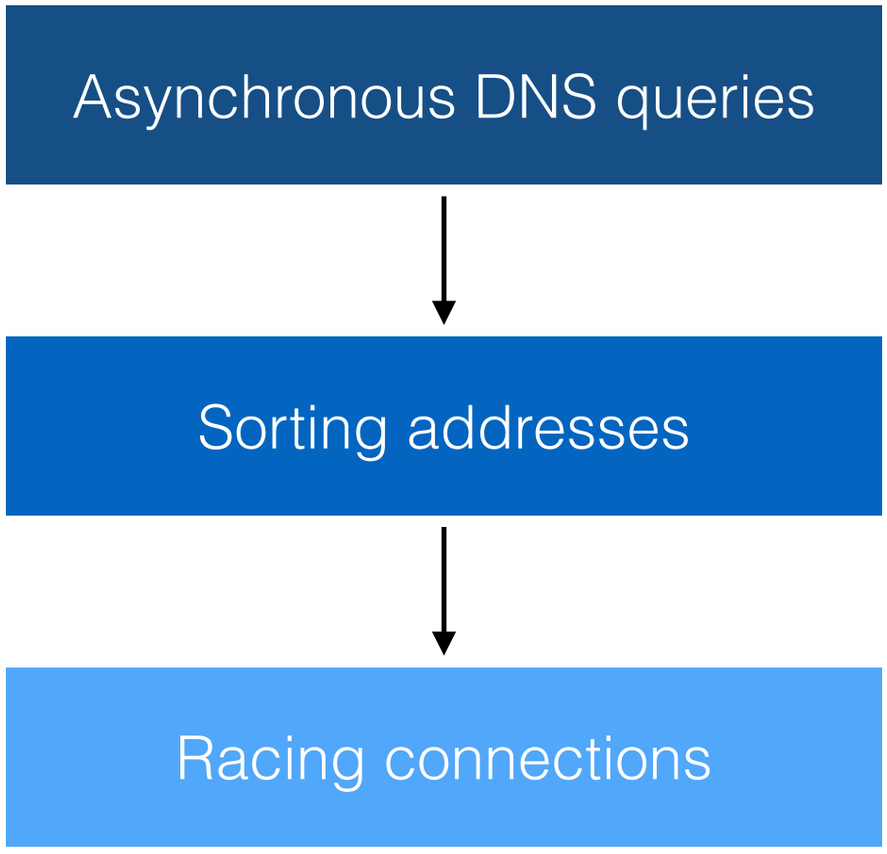
(Buggy) TCP terminating proxies

Issues are getting solved but not totally fixed

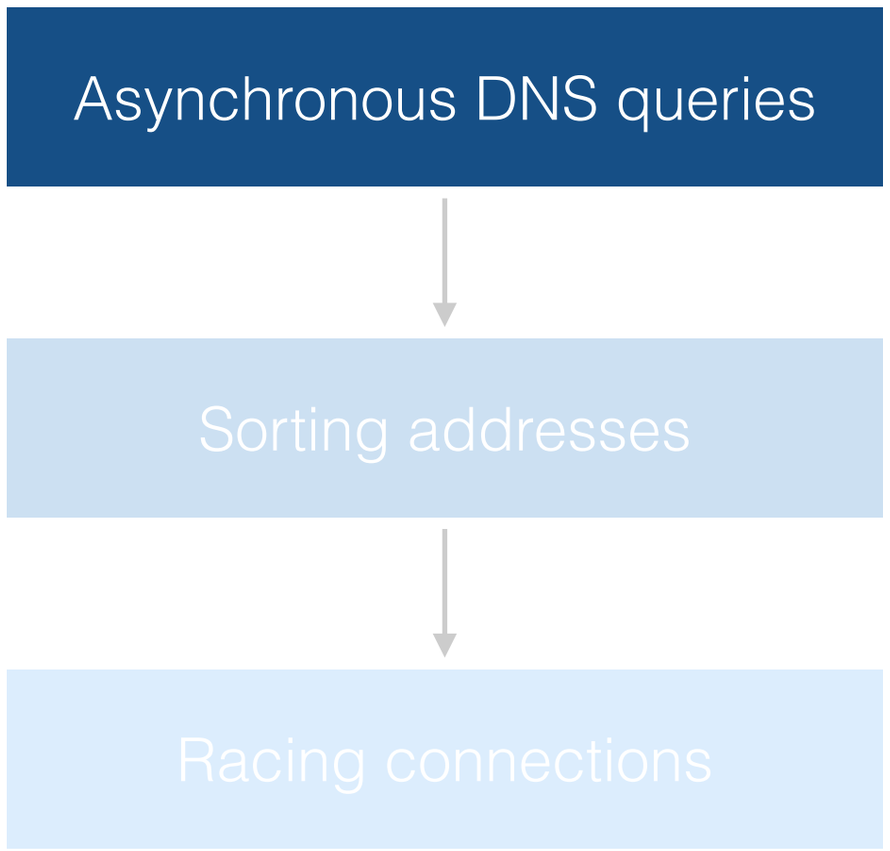
So what does Happy Eyeballs do?

Make the users' eyeballs "happy" by making connections **quickly** to servers that **work**

Algorithm (RFC 8305)



Algorithm (RFC 8305)



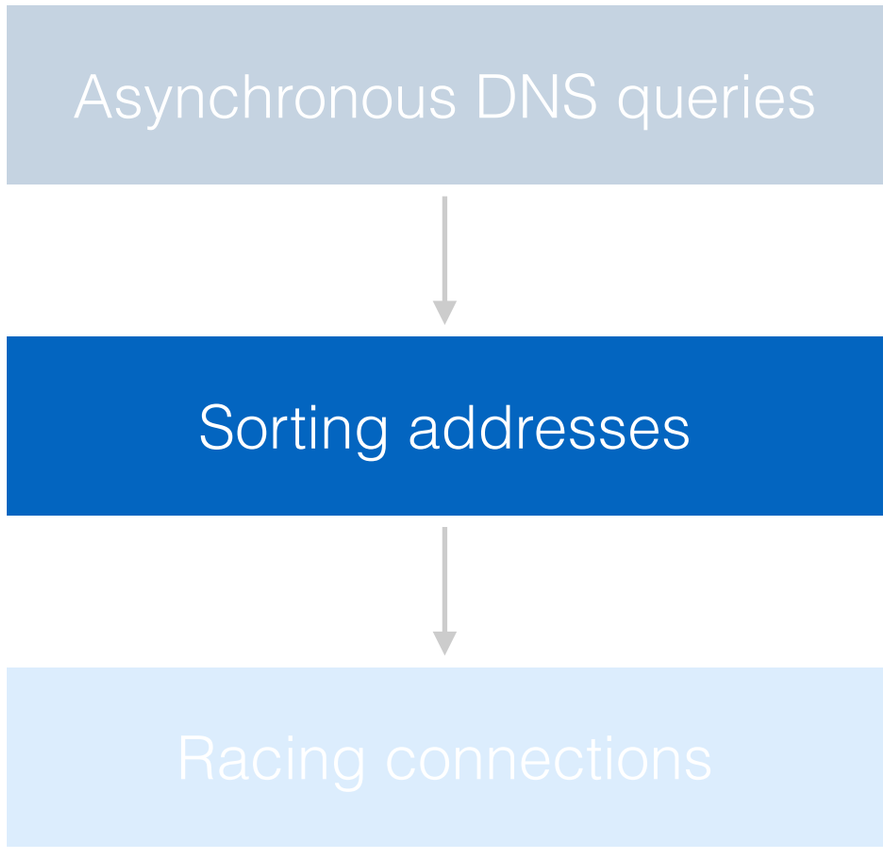
Prefer IPv6 DNS resolver addresses

Issue AAAA / A queries in parallel (AAAA first)

Act on AAAA responses immediately

Wait up to 50ms for AAAA if A comes back first

Algorithm (RFC 8305)



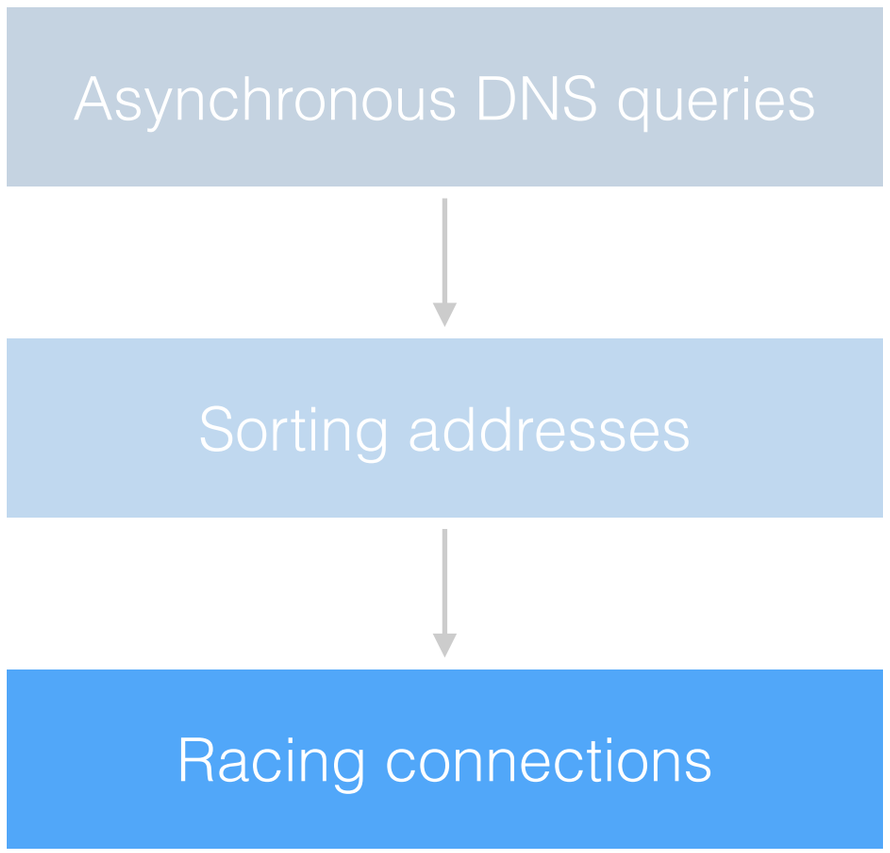
RFC 6724 address sorting

Historical round trip times

Place IPv6 at the start; two IPv6 addresses first if available

Use IPv4 first if historical data shows IPv6 brokenness

Algorithm (RFC 8305)



Start attempts to addresses in parallel, staggered by a delay

Delay based on the TCP retransmission timeout

Race ends when one connection completes

Racing through TLS

Performance impact

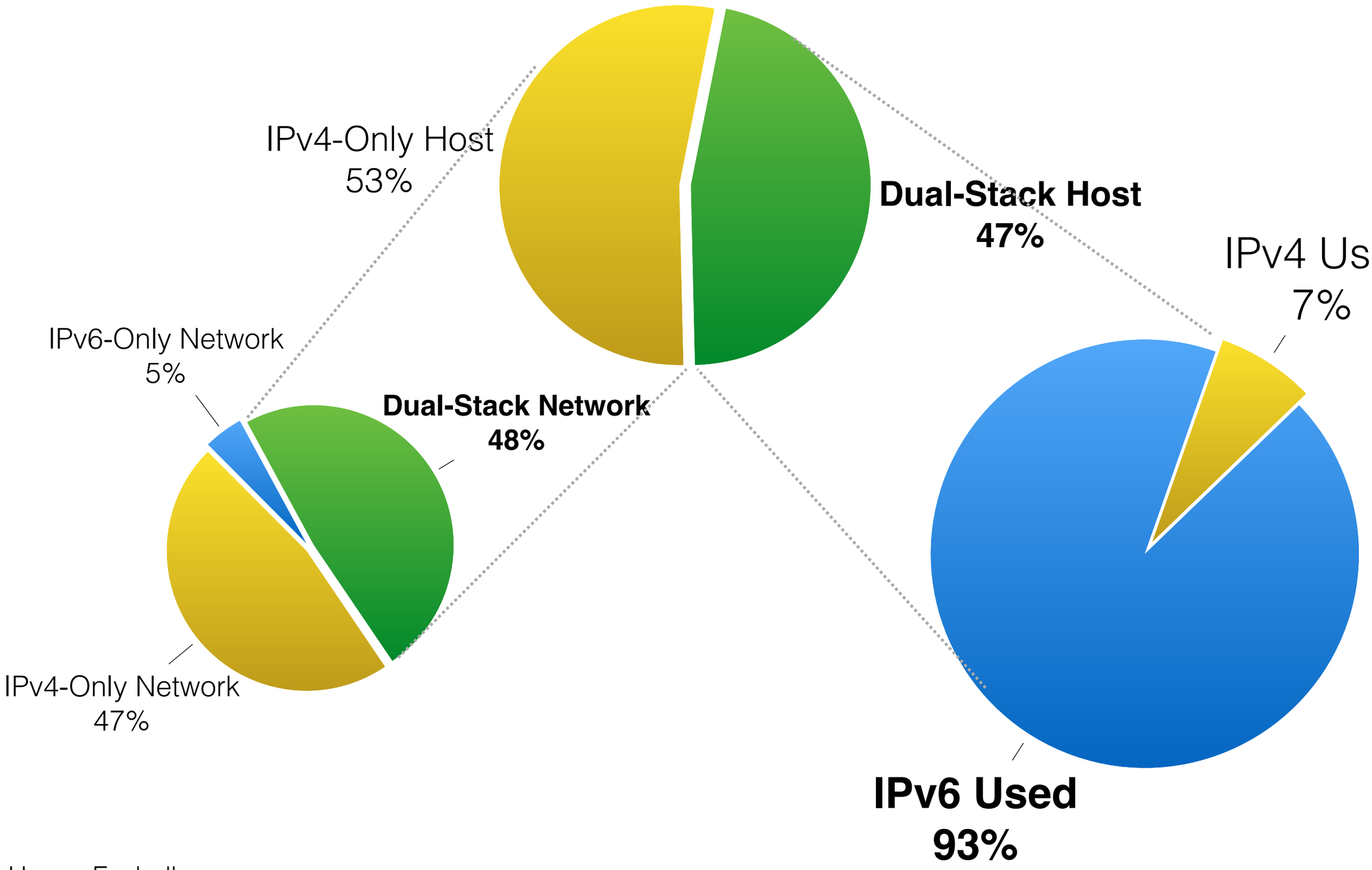
Measurements

- IPv6 availability on networks
- IPv6 availability on hosts

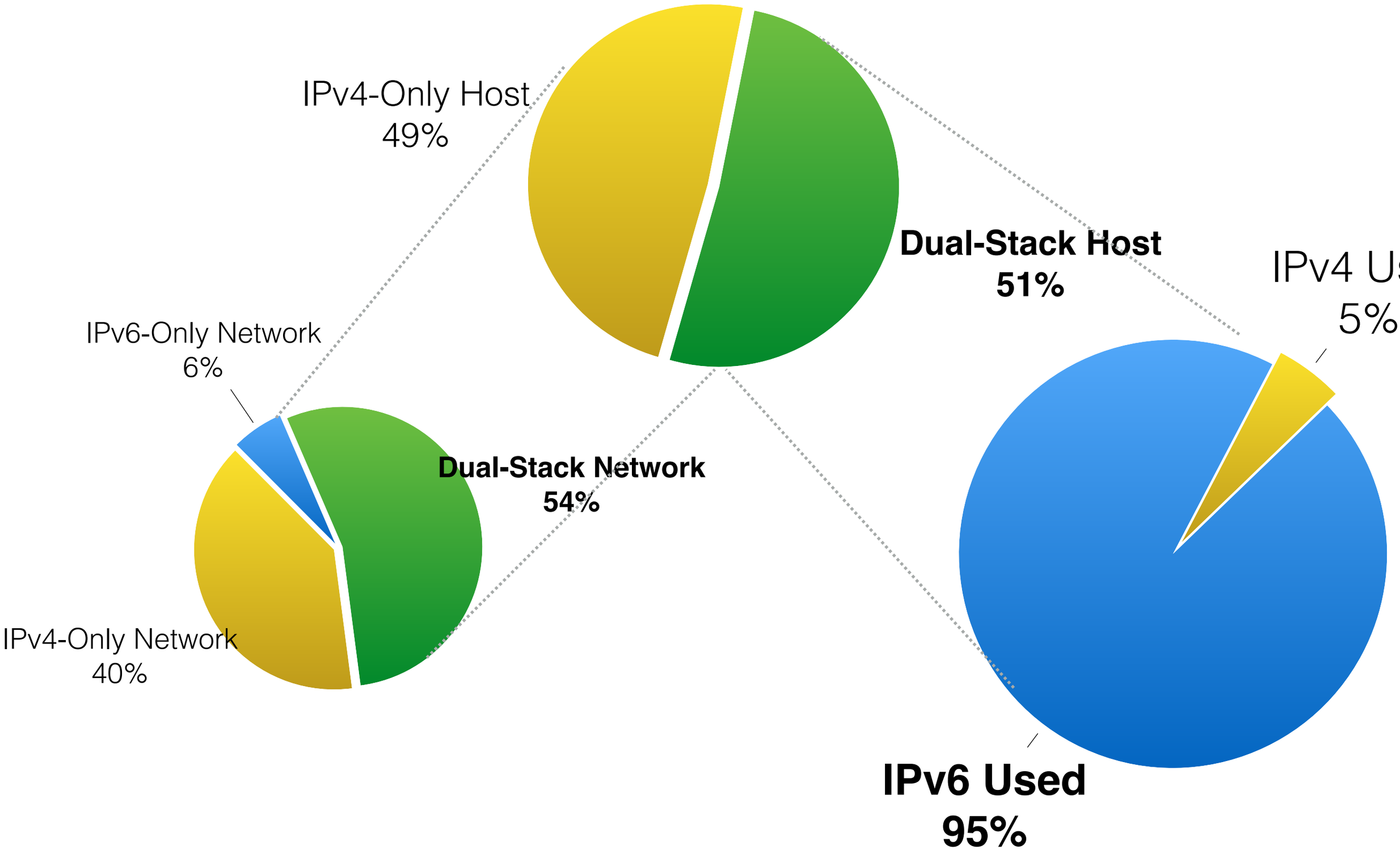
IPv6 success via Happy Eyeballs

Measured on a per-client connection basis

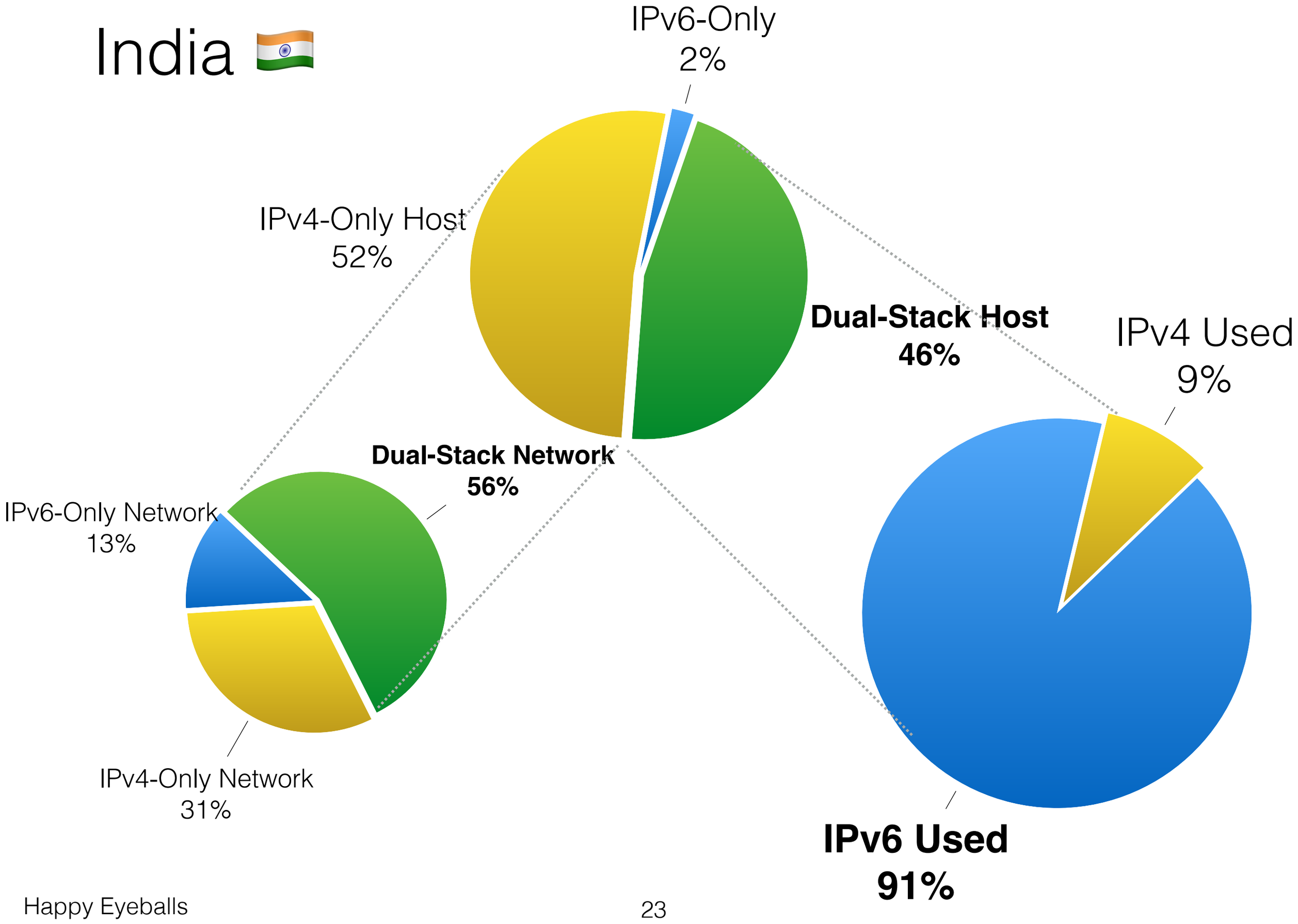
Worldwide 🌐


IPv4-Only Host
53%
Dual-Stack Host
47%
IPv6-Only Network
5%
Dual-Stack Network
48%
IPv4-Only Network
47%
IPv4 Used
7%
IPv6 Used
93%

USA 🇺🇸


IPv4-Only Host
49%
Dual-Stack Host
51%
IPv4 Used
5%
IPv6-Only Network
6%
Dual-Stack Network
54%
IPv4-Only Network
40%
IPv6 Used
95%

India 🇮🇳


IPv6-Only
2%
IPv4-Only Host
52%
Dual-Stack Host
46%
Dual-Stack Network
56%
IPv6-Only Network
13%
IPv4-Only Network
31%
IPv4 Used
9%
IPv6 Used
91%

Happy Eyeballs allows using IPv6
as **much** as possible,
as **soon** as possible,
even if there is still brokenness

Hiding brokenness

Techniques to expose brokenness

Measurement of networks to see how often fallback to IPv4 needs to occur

Should clients automatically report issues? Via what mechanism?

Still an area of open discussion and investigation

Happy Eyeballs, **Version 3**

What has changed?

New transport protocol (QUIC)

More IPv6-only networks and mechanisms

Encrypted DNS

New DNS record types (SVCB/HTTPS)

New application protocol and security protocol implications

RFC 9460: SVCB/HTTPS records

(*Service binding*)

Algorithm with SVCB

Query SVCB/AAAA/A RRs in parallel

Use SVCB priority to sort associated A/AAAA answers and address hints

- A/AAAA answers not associated with SVCB are prioritized last
- Prefer ECH keys if present

Sort endpoints (address, protocol, ECH keys) and stagger connection attempts

Race until full handshakes complete (TCP / TLS / QUIC, etc)


QUIC

Generalizing for QUIC

Prefer services with QUIC-capable ALPNs when sorting endpoints, after ECH keys and SvcPriority

- QUIC provides improved delivery and congestion control, connection migration, etc.

Adjust connection establishment logic to not just mention TCP

- Race until QUIC completes
- Also allow racing until TLS above TCP completes

How to engage

ALLDISPATCH at IETF 119

GitHub issues and pull requests

IETF v6ops mailing list discussion

Questions?