Large-scale IPv6 deployment and Practice

@ Alibaba
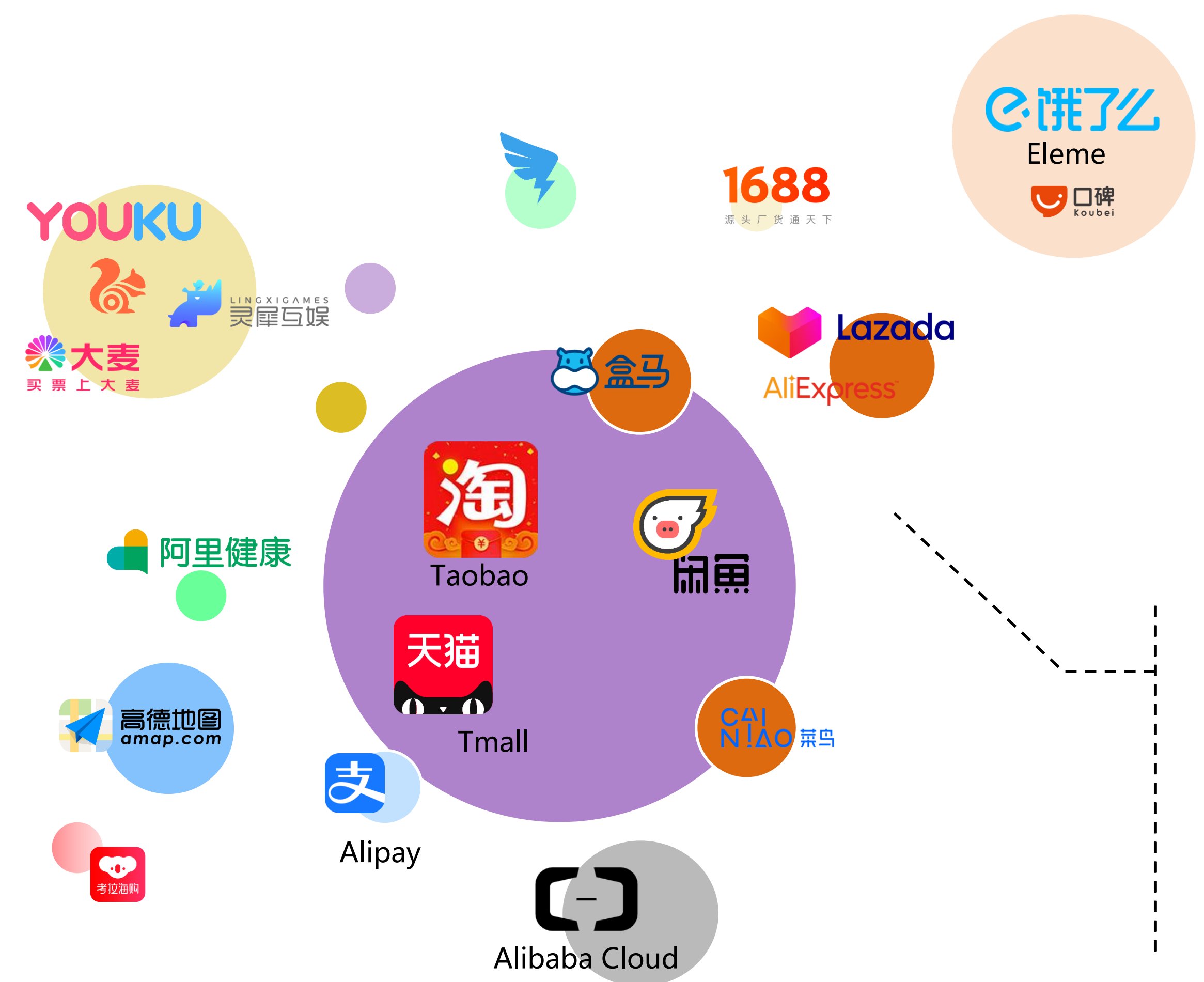
Linjian (Davey) Song
linjian.slj@alibaba-inc.com

V6ops@IETF114

Brief Introduction on Alibaba Application and service

YOUKU

LINGXI GAMES 灵犀互娱

大麦
买票上大麦

1688
源头厂货通天下

e 饿了么
Eleme

口碑
Koubei

盒马

Lazada
AliExpress

淘
Taobao

闲鱼

阿里健康

天猫
Tmall

CAINIAO 菜鸟

高德地图
amap.com

支
Alipay

考拉海购

Alibaba Cloud

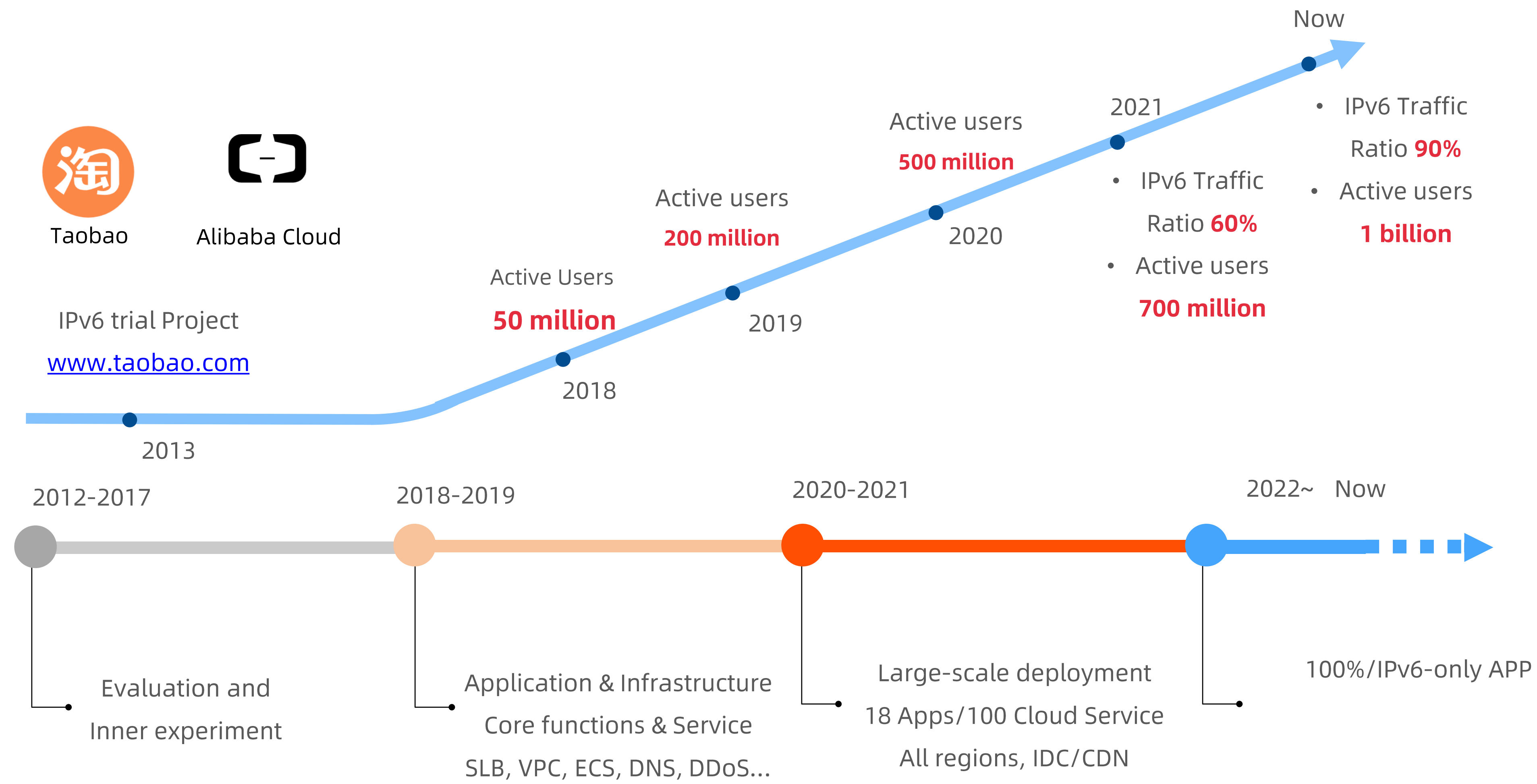China Active users：**1 B**

International Active users：**305 M**

- E-commerce, Entertainment, Travel, Social, Smart Logistics, Local life service, etc.
- Cloud, IDC, CDN, DNS, etc.

Alibaba Application and Service

Challenges in IPv6 transition for large-scale online service

IPv6 Deployment Phases in Alibaba

Taobao

Alibaba Cloud

IPv6 trial Project

www.taobao.com

2013

Active Users
50 million

2018

Active users
200 million

2019

Active users
500 million

2020

2021

- IPv6 Traffic
  Ratio **60%**
- Active users
  700 million

Now

- IPv6 Traffic
  Ratio **90%**
- Active users
  1 billion

2012-2017

Evaluation and
Inner experiment

2018-2019

Application & Infrastructure
Core functions & Service
SLB, VPC, ECS, DNS, DDoS...

2020-2021

Large-scale deployment
18 Apps/100 Cloud Service
All regions, IDC/CDN

2022~   Now

100%/IPv6-only APP

Motivation and Challenges in IPv6 Deployment

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 (in 2 years period).
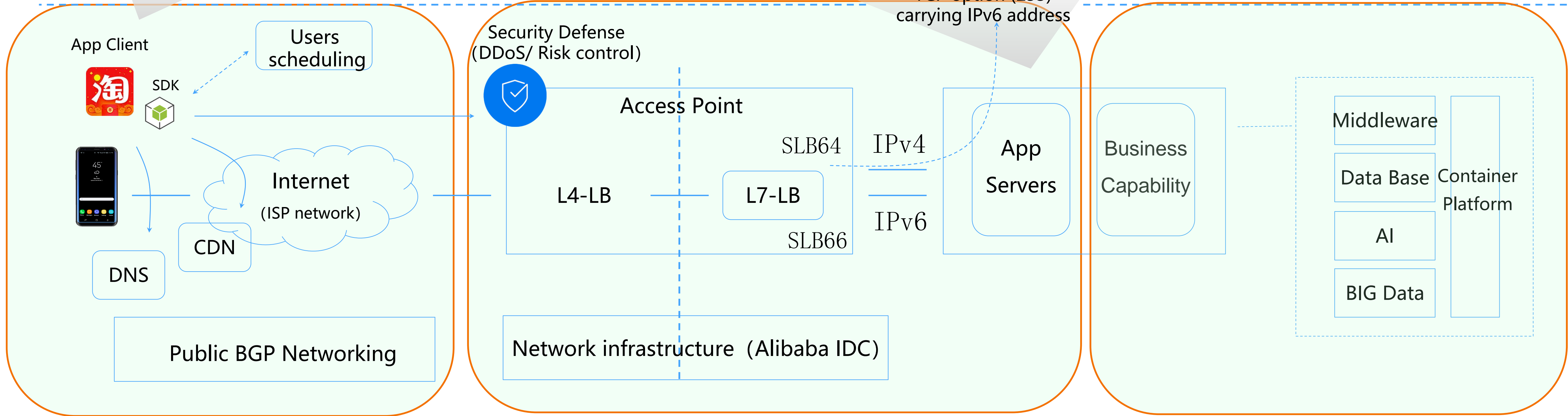- Coordinating 20+ Business unites and 500+ people

Technical Consideration

1. How to migrate online services smoothly?
2. How to measure IPv6 network and adapting dual-stack?
3. IPv6 MTU issues

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

Users scheduling

Security Defense
(DDoS/ Risk control)

SDK

Access Point

SLB64   IPv4

App Servers

Middleware

Internet
(ISP network)

L4-LB   L7-LB   IPv6

Business Capability

Data Base   Container Platform

SLB66

CDN

AI

DNS

BIG Data

Public BGP Networking

Network infrastructure (Alibaba IDC)

High Level Online Services Architecture – Cloud Platform
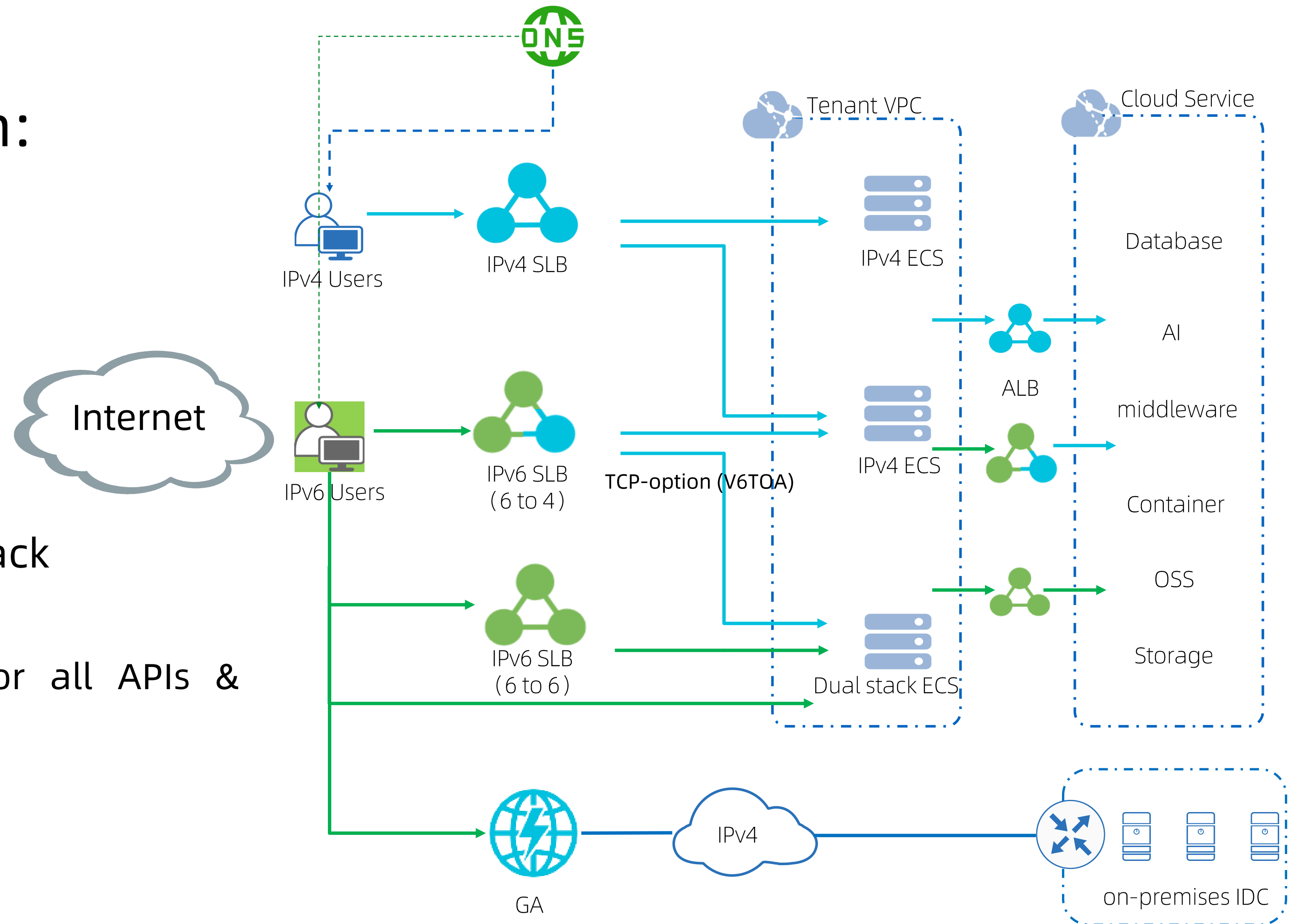
IPv6 transition solution:

Solution 1: Proxy

Solution 2: SLB64

Solution 3: SLB66/Dual stack

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

ONS

Tenant VPC

Cloud Service

IPv4 Users

IPv4 SLB

IPv4 ECS

Database

AI

ALB

Internet

IPv6 Users

IPv6 SLB
（6 to 4）

TCP-option (V6TOA)

IPv4 ECS

middleware

Container

IPv6 SLB
（6 to 6）

Dual stack ECS

OSS

Storage

GA

IPv4

on-premises IDC

IPv6 Measurement

An application performance measurement platform for IPv6:

TCP Success Rete, RTT , IPv6 Coverage and Broken dual-stack WIFI

- The TCP success rate was around 75% in 2018, 90% in early 2019 and 99.8% this year

- The RTT of the first packet was around 250ms, less than 200ms in early 2019 and now around 150ms this year

- The Coverage of dual-stack network is around 90% in 4G/5G network in early 2022, and around 27% in WIFI access network
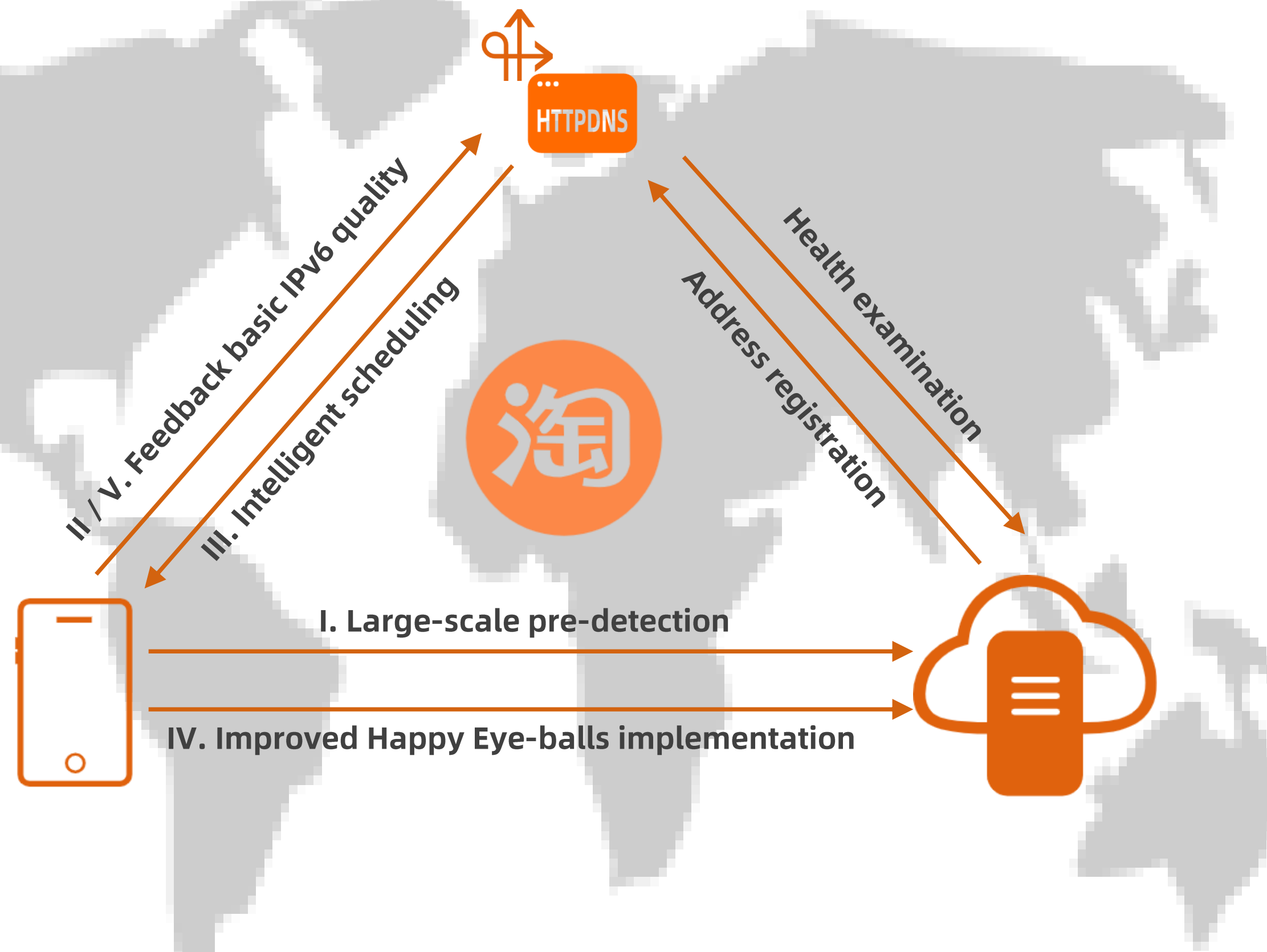
Broken dual-stack WIFI Situation (Last-mile issue)

170 thousand cases (10% sample rate)

Dual-Stack adaption for Apps

II / V. Feedback basic IPv6 quality

III. Intelligent scheduling

Health examination

Address registration

HTTPDNS

淘

I. Large-scale pre-detection

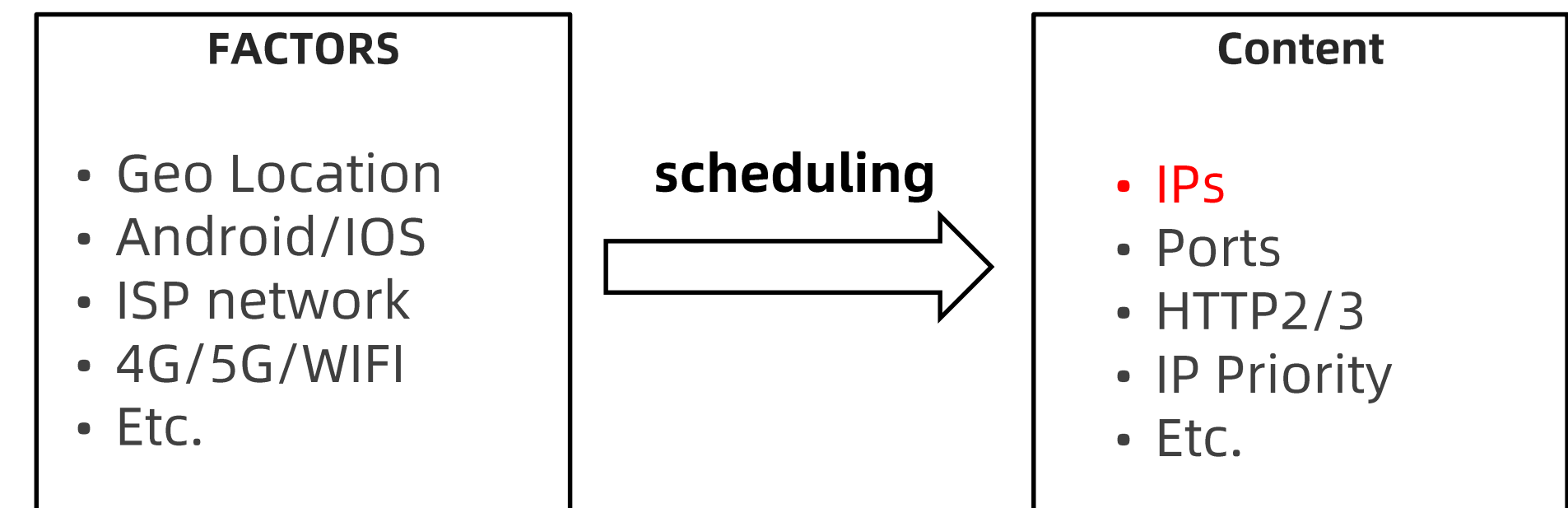IV. Improved Happy Eye-balls implementation

Intelligent decision on IPv6 adoption
on specific user groups

- **Large-scale pre-detection to research basic network IPv6 quality with variety protocols:**
  - ICMP Ping
  - TCPing
  - HTTP

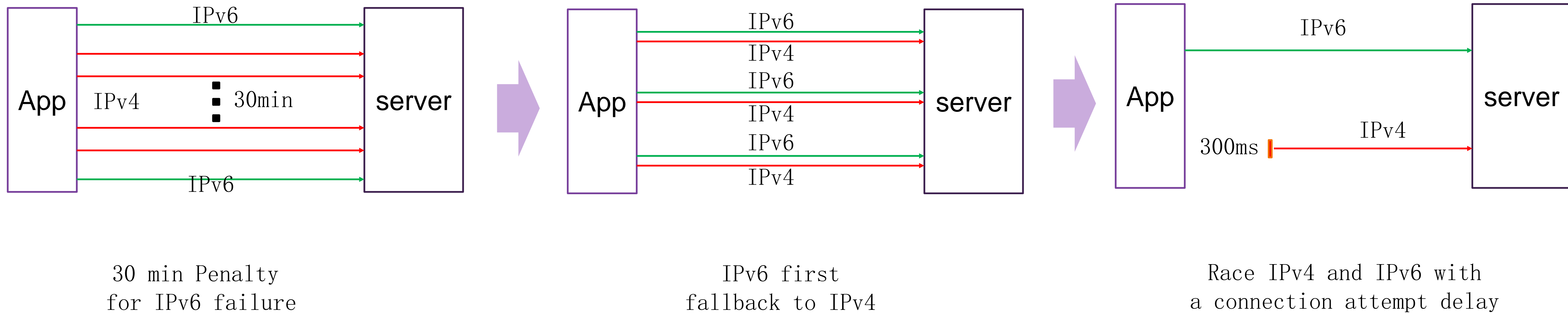- **Refined intelligent scheduling base on variety factors:**

FACTORS		Content
• Geo Location • Android/IOS • ISP network • 4G/5G/WIFI • Etc.	**scheduling** ⟹	• 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 MTU issue

Failures due to large IPv6 packets drop due to NAT, Proxy, FW etc..

For all connections

- **IPv6 MTU 1240 bytes**

- **TCP MSS 1180 bytes** (IPv4 MSS 1440)

- 1280(IPv6 MTU) - 40(IPv6 header) -20(TCP header) -
  40(maximum TCP option)

- TCP option carrying IPv6 address and port

For CDN connections

- **IPv6 MTU 1450 bytes**

- **TCP MSS 1390 bytes**

5 Years later, where are we now?

- Good aspects

  - A fully connect IPv6 Production network has been built
  - IPv6 user penetration (active users) is over 90% from the view of Alibaba Apps via LTE network.
  - IPv6 Traffic ratio : more than 40% in LTE network and 10% in metropolitan area network (https://m.china-ipv6.cn/complete/#/)
  - IPv6 performance is almost close to IPv4 and better than IPv4 in some cases (test inside Alibaba)
  - Dual-stack is proved to be the best approach adopting IPv6

- Bad aspects

  - The last-mile issue still exists, long-tail user side network
  - IPv4 address price still hurts, more and more
  - Dual-stack may last for long time due to legacy network/system. No plans and no path to sunset IPv4

Conclusion and take away

- 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 a new one. Keep IPv6 into consideration in the first place
- Dual stack is the only approach adopted in large scale in Alibaba. However, the IPv6 community should consider a countdown timer for IPv4

Thank You