

爱奇艺私有云IPv6网关的建设与实践

基础架构部 资深研发工程师 王庆



摘要

I

IPv6简介

国内IPv6现状、给云端带来的挑战

II

DPVS简介

DPVS的框架和实现

III

DPVS-IPv6的实现

DPVS对于IPv6的支持方法；IPv6->IPv4转发

IV

DPVS-IPv6的部署

高可用性、可扩展性的实用部署方式

V

后续工作&QA

IPv6简介

IPv4资源紧缺

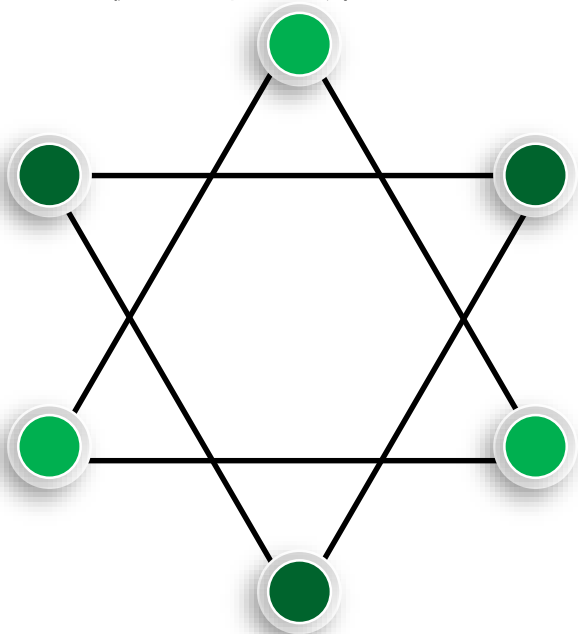
2011年2月，IANA分配完最后的一批地址库。

IPv6的内部环境

业务的平滑过渡、IDC的改造

IPv6的外部环境

爱奇艺可控IDC以外的网络环境对于IPv6的支持、可靠性



工信部

《推进互联网协议第六版（IPv6）规模部署行动计划》

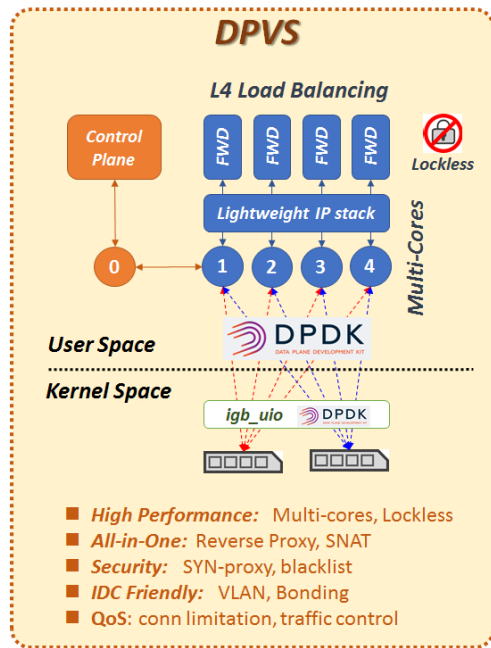
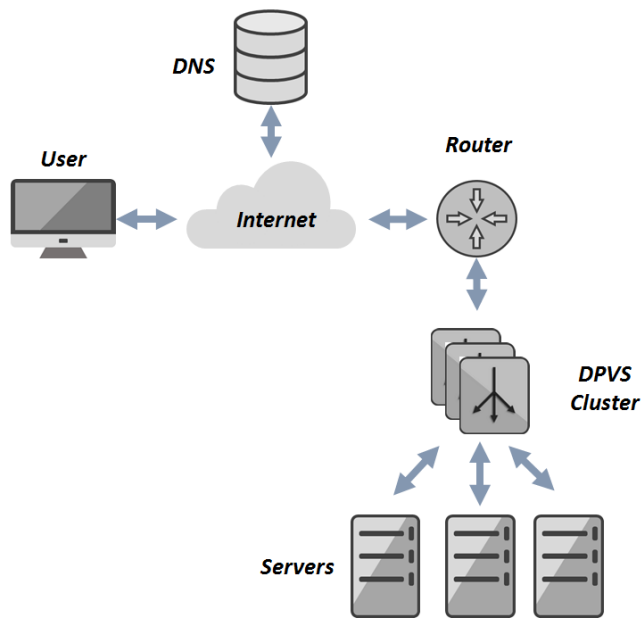
IPv6端的增加

爱奇艺终端设备有1.3亿IPv6用户

iOS APP Store

2016年5月需要IPv6-Only验证通过

DPVS简介



流量入口/负载均衡

流量出口/代理上网

安全网关

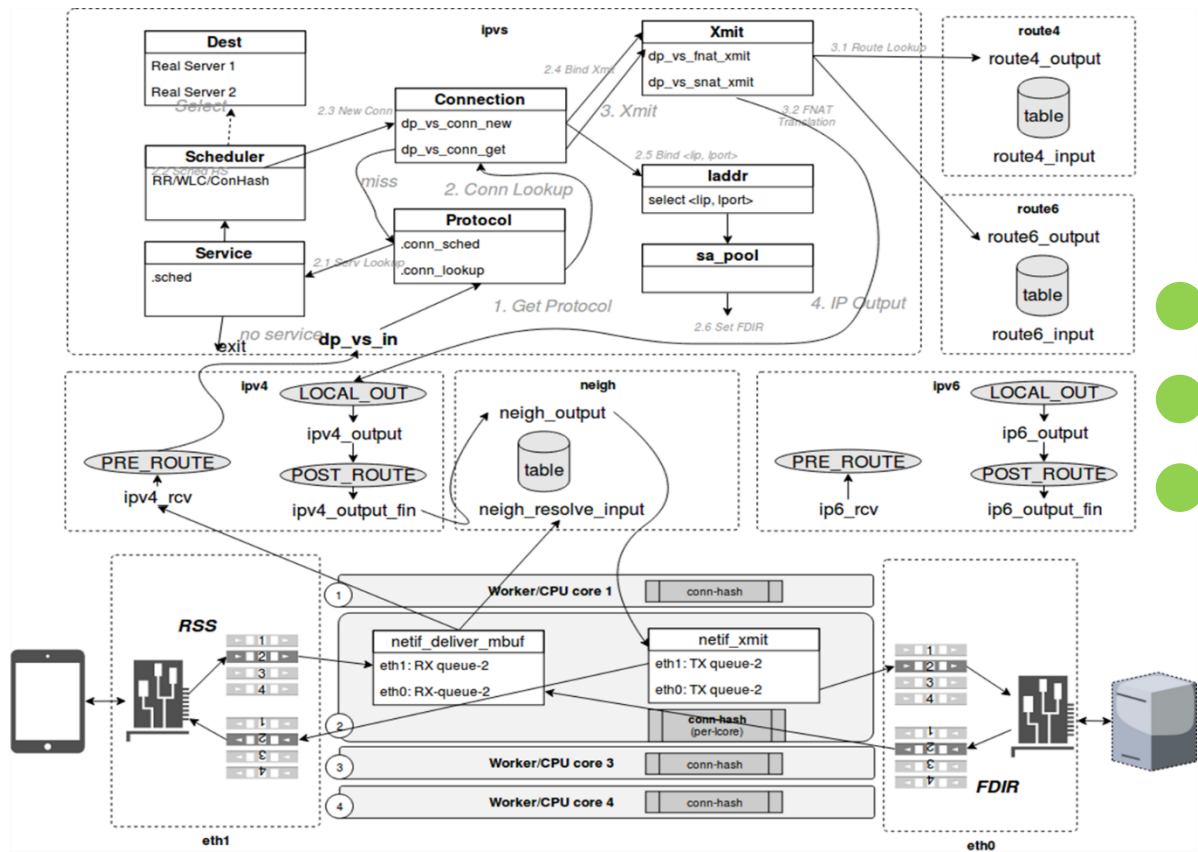
用户态实现

网卡队列/CPU
绑定

连接表无锁化

高性能负载均衡

DPVS框图

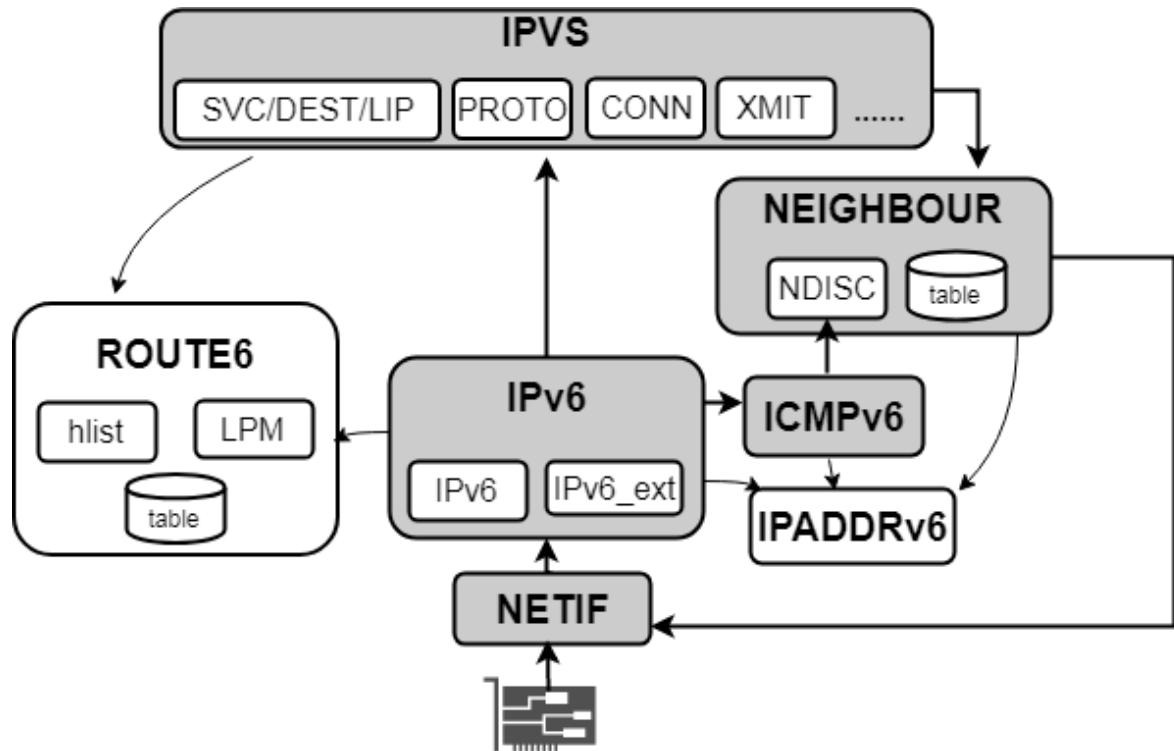


● DR/NAT/TUNNEL/FULLNAT/SNAT

● Syn_proxy/Blacklist/Limit conn

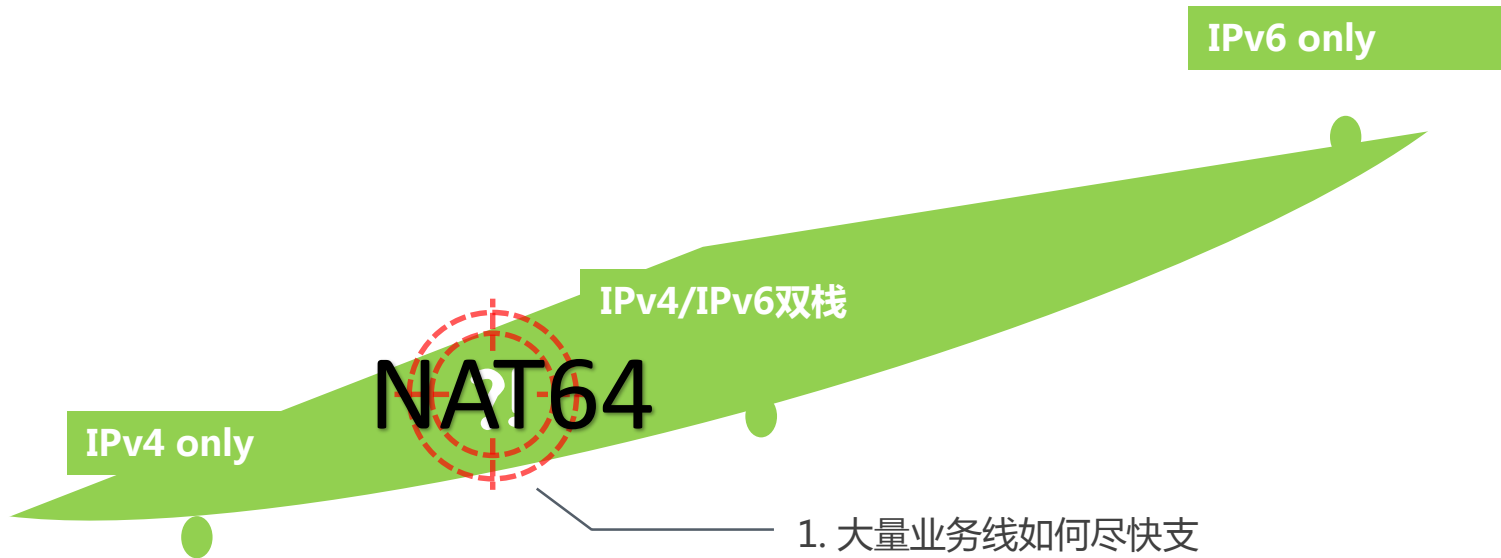
● Vlan/Bond

DPVS-IPv6



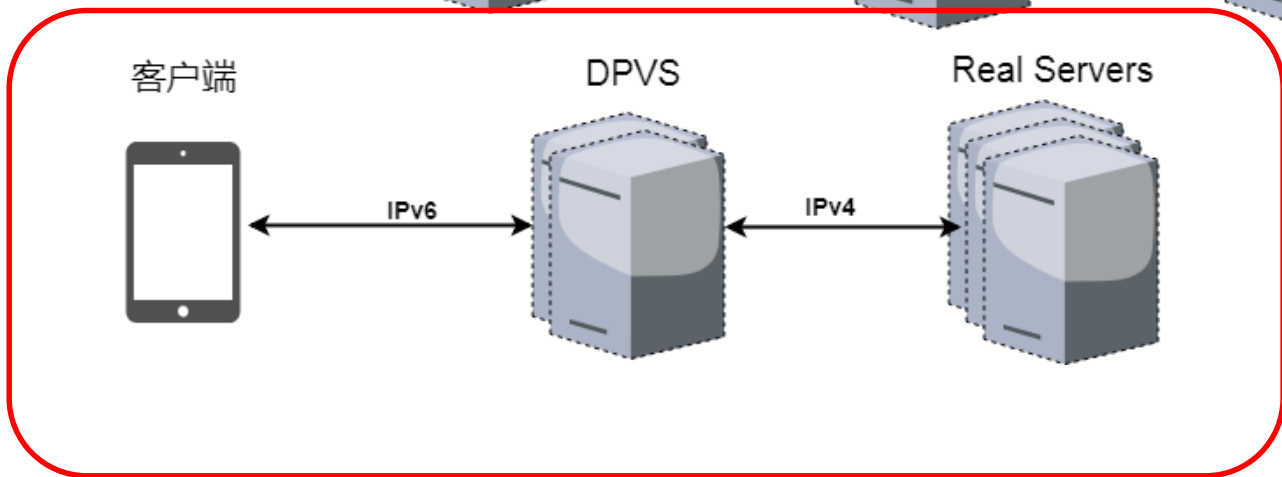
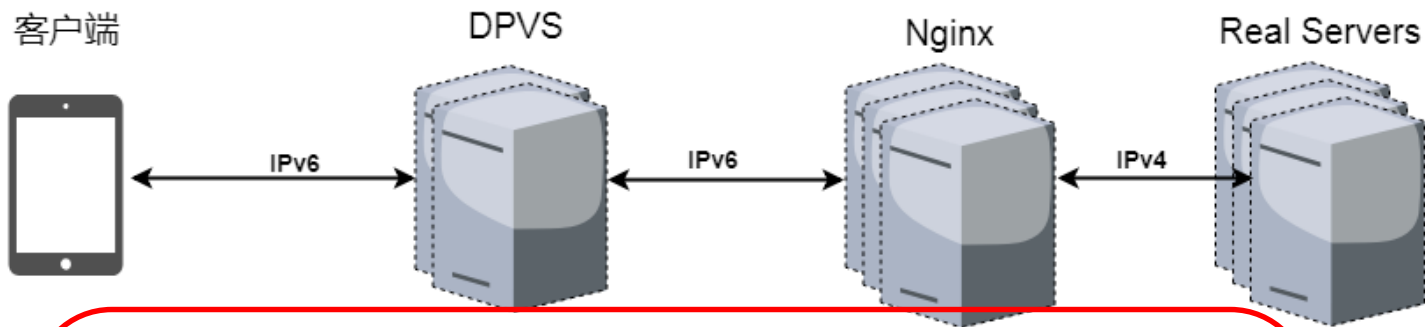
- 多播的处理
- IPv6扩展头
- TOA改造

NAT64的需求

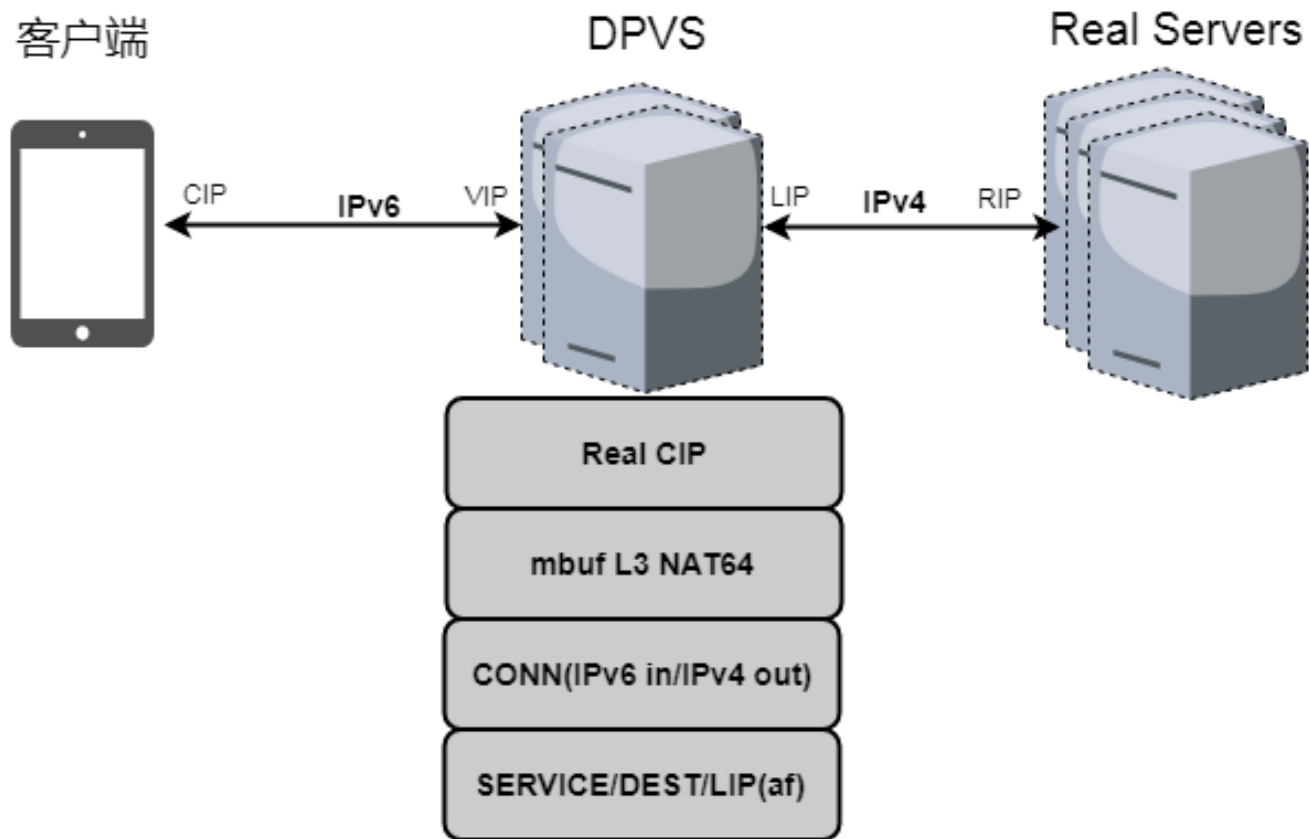


1. 大量业务线如何尽快支持IPv6 ?
2. 云端大规模存量IPv4服务器如何改造 ?
3. 内网交换机升级带来的大范围影响 ?

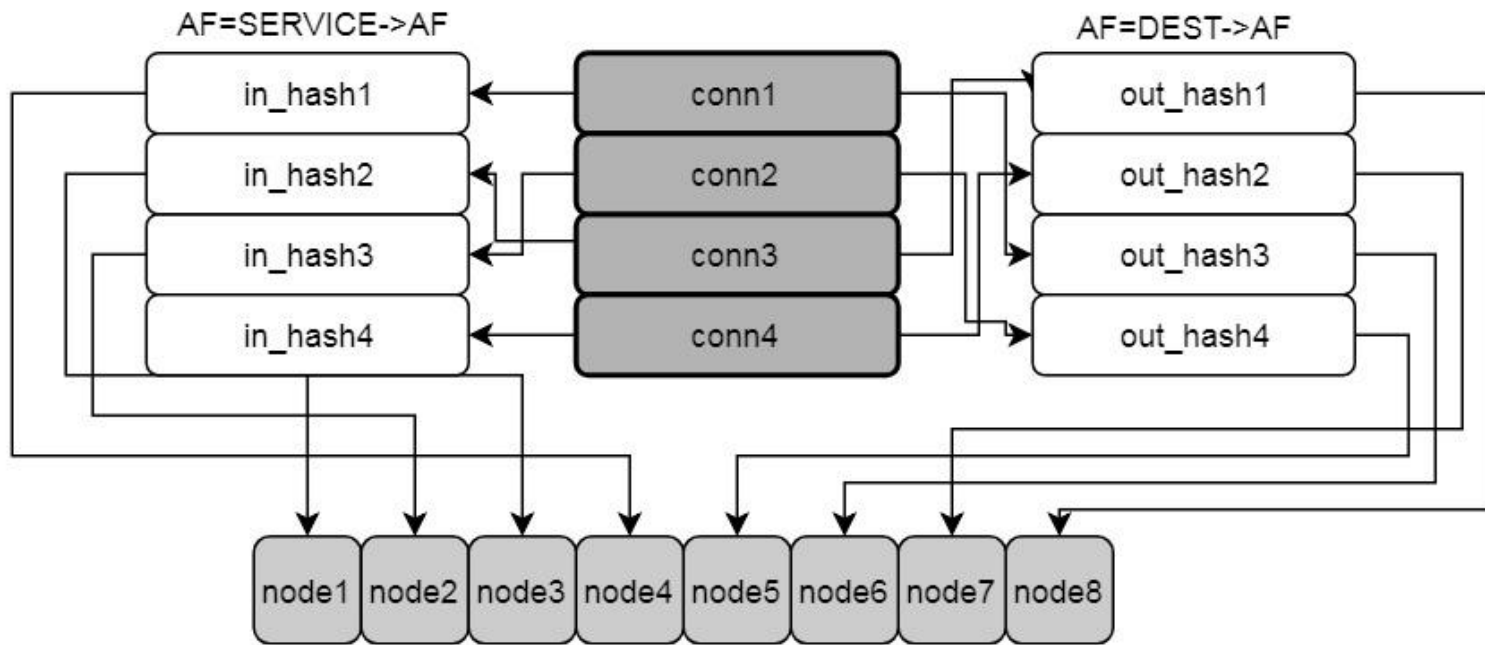
NAT64的方案



DPVS-NAT64



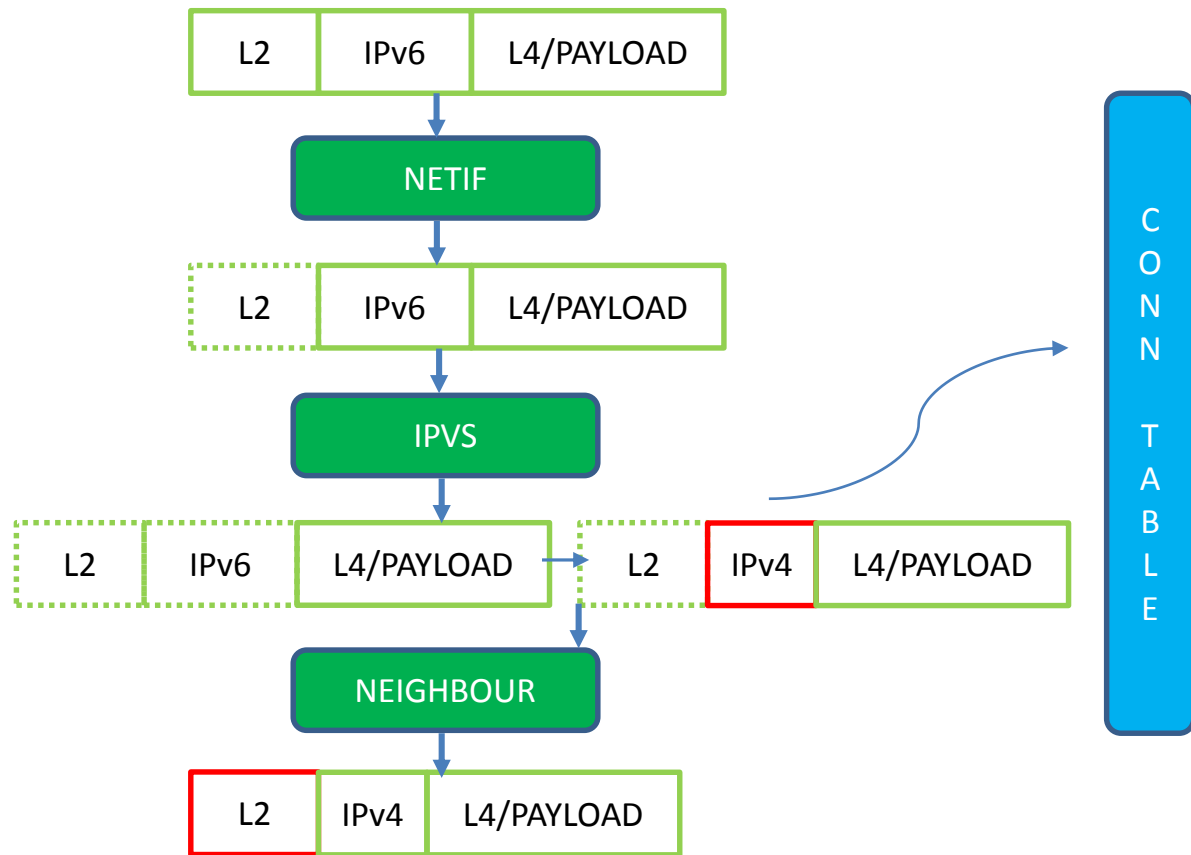
NAT64 CONN



```
struct conn_tuple_hash {
    struct list_head list;
    int direct; /* inbound/outbound */

    /* tuple info */
    int af;
    uint16_t proto;
    union inet_addr saddr; /* pkt's source addr */
    union inet_addr daddr; /* pkt's dest addr */
    uint16_t sport;
    uint16_t dport;
} __rte_cache_aligned;
```

NAT64 mbuf (in bound)



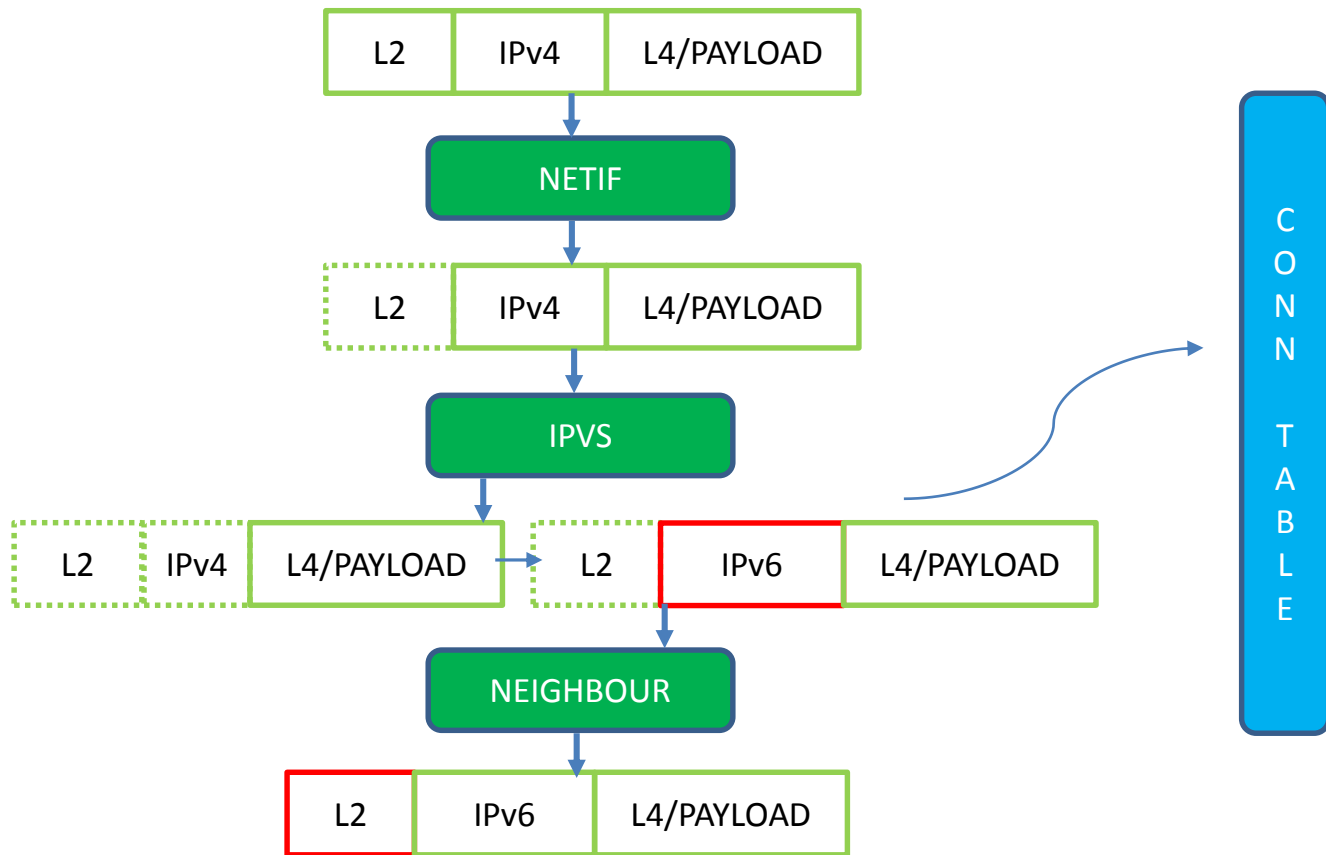
NAT64 mbuf (in bound)

Version	Traffic class	Flow lable	
Payload len		Next_header	HOPS
CIP			
VIP			

Verson_ihl	TOS	Total_len
Packet_id		Frag_offset
TTL	PROT	CKSUM
LIP		
RSIP		

OPTION

NAT64 mbuf (out bound)

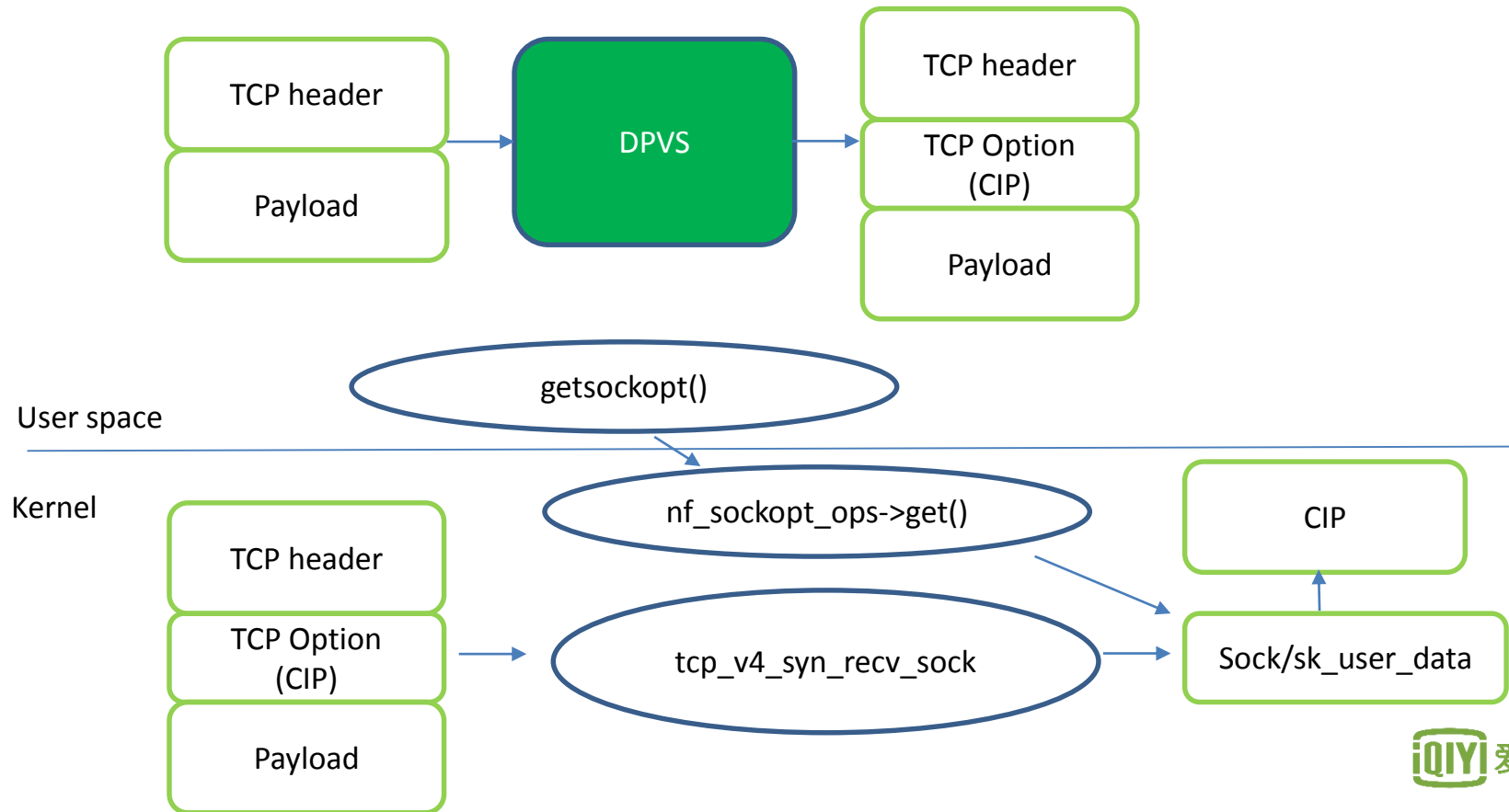


NAT64 mbuf (out bound)

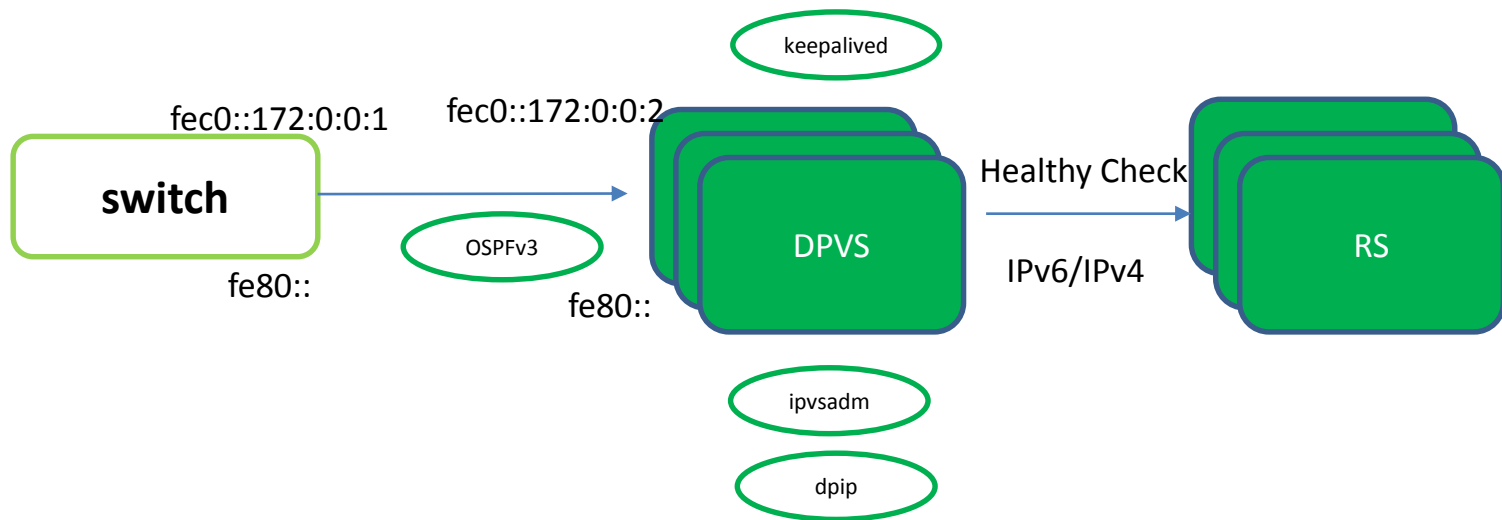
Verson_ihl	TOS	Total_len
Packet_id		Frag_offset
TTL	PROT	CKSUM
RSIP		
LIP		

Version	Traffic class	Flow lable	
Payload len		Next_header	HOPS
VIP			
CIP			

NAT64 TOA



DPVS-IPv6的部署



后续工作

突破10G网卡的瓶颈

UDP支持的完善

25G网卡的支持

SNAT NAT64

UOA支持IPv6

IPv4局域网可以与IPv6
外网通信

<https://github.com/iqiyi/dpvs>

Q&A

