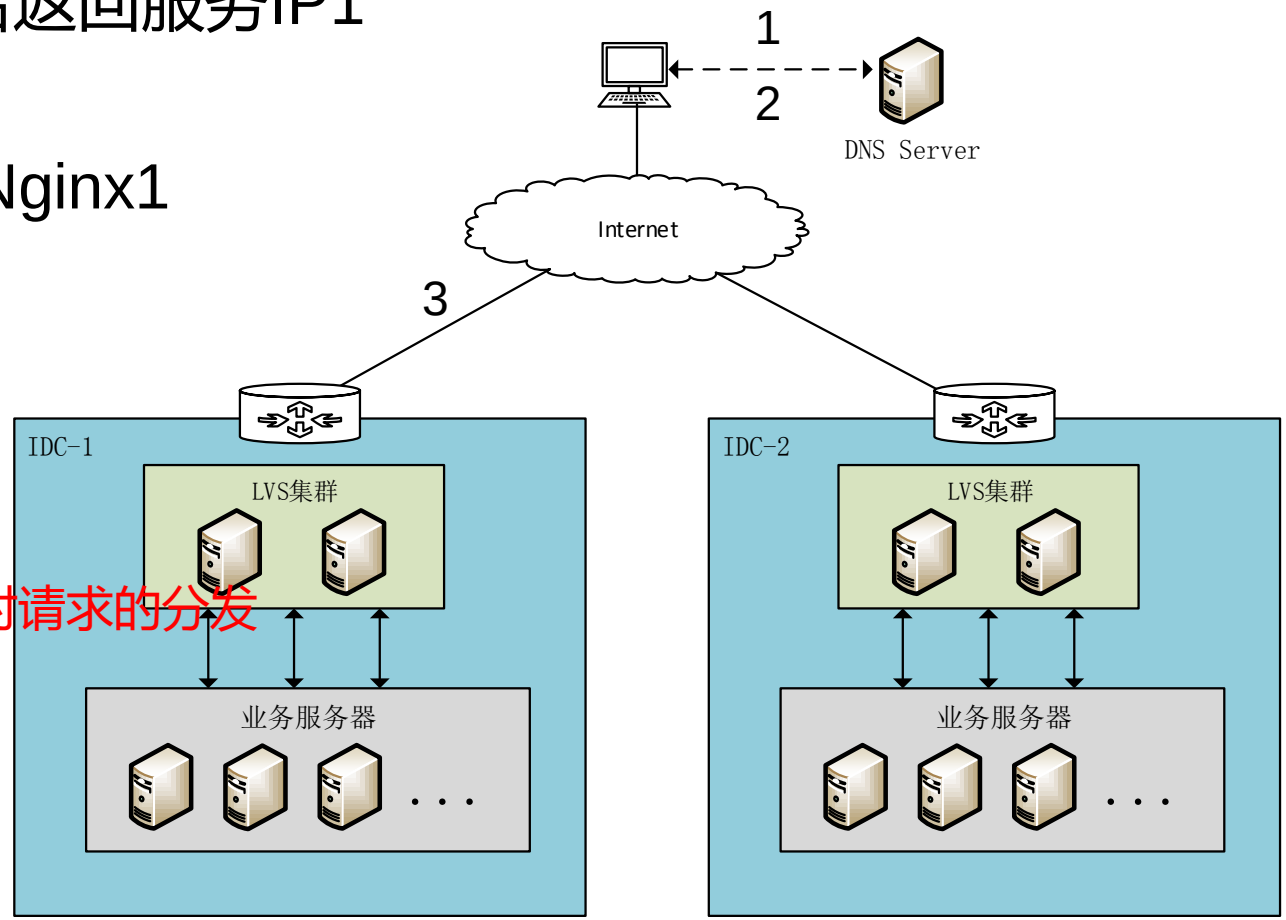
LVS应用介绍

张勇

目录

LVS应用场景与基本功能介绍

- 用户访问 www.360.cn
- 请求DNS Server，DNS递归查找后返回服务IP1
- 用户请求服务IP1
- 数据包到达LVS，LVS决定调度给Nginx1
- Nginx1 接收请求，处理



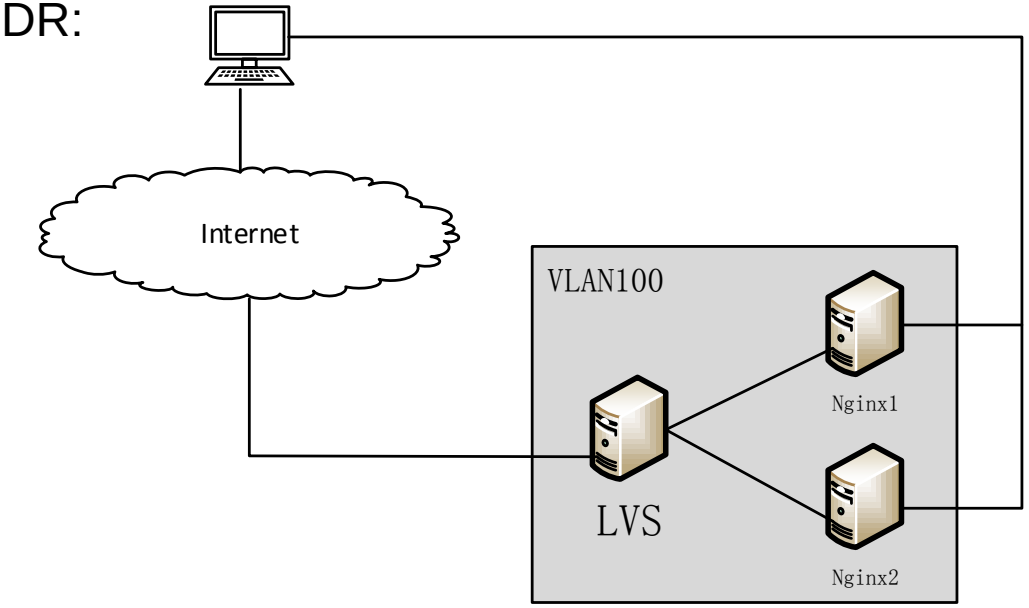
4: IPVS负责对请求的分发

- LVS在360内部的应用
  - 2014
    - 万兆服务器200+台
    - 业务数目（VIP:VPort）：3000+
  - 2015
    - 万兆服务器350+台
    - 业务数目（VIP:VPort）：10000+
    - 单集群挂载的RS数目（Max）：8000+
    - 单集群的流量峰值（正常业务）：24Gb+

LVS应用场景与基本功能介绍——转发模式

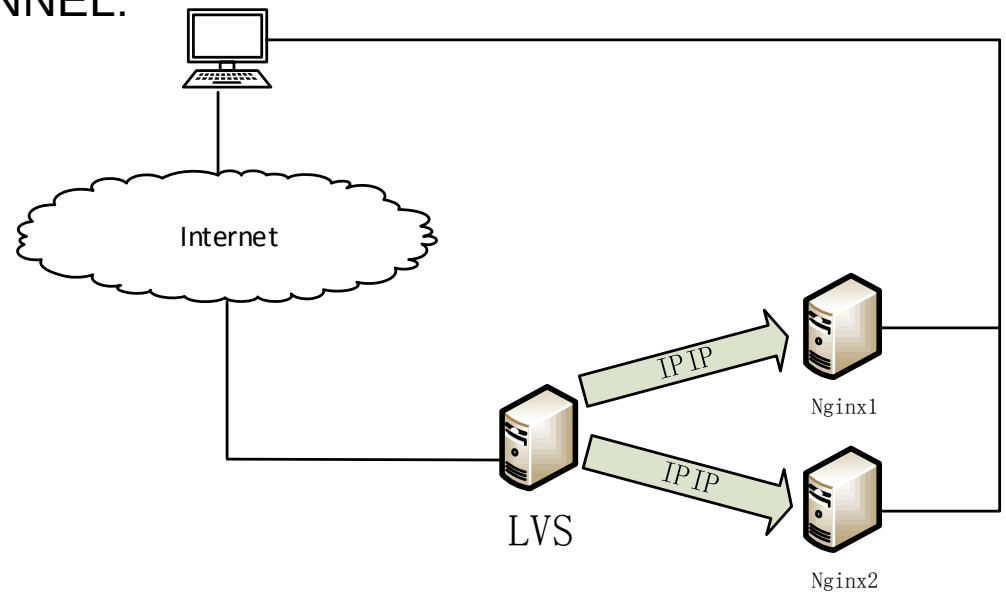
- IPVS支持DR、 NAT、 TUNNEL三种转发模式

DR:



1. Lvs与nginx服务器需要在同一个vlan
2. Nginx服务器上需要绑定vip

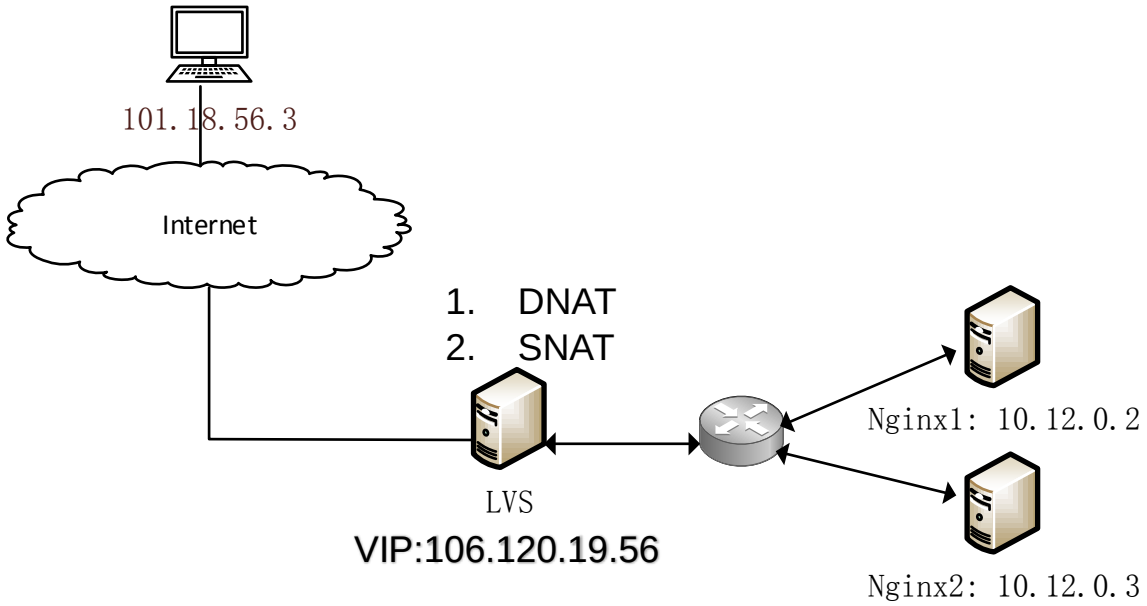
TUNNEL:



1. Nginx服务器上需要配置ipip隧道
2. 隧道头增加开销
3. Nginx服务器上需要绑定vip

LVS应用场景与基本功能介绍——转发模式

- IPVS支持DR、 NAT、 TUNNEL三种转发模式

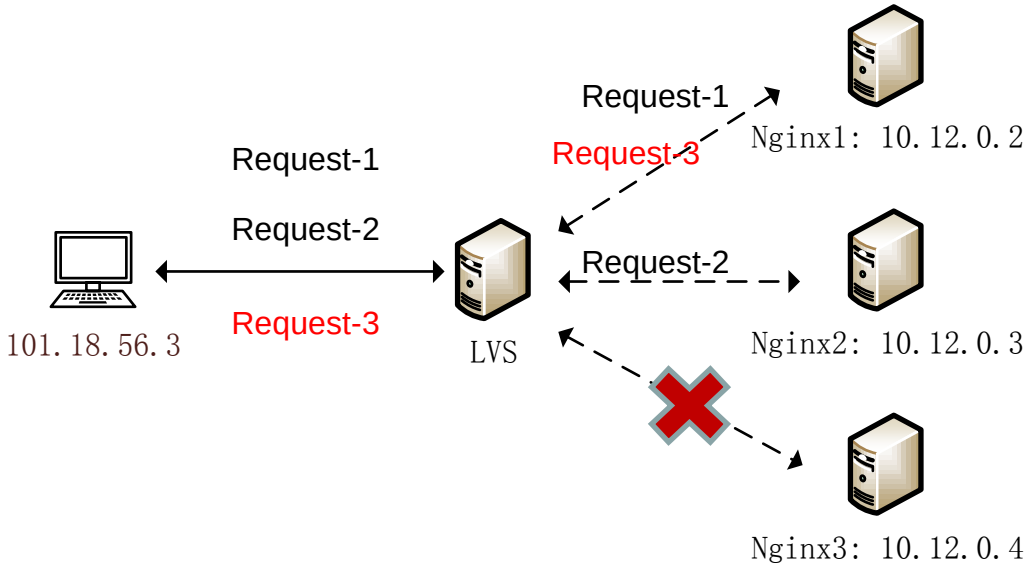


	SRC IP	DST IP
	101.18.56.3	106.120.19.56
1. DNAT	101.18.56.3	10.12.0.2
	10.12.0.2	101.18.56.3
2. SNAT	106.120.19.56	101.18.56.3

1. Nginx服务器上需要配置路由策略，默认下一跳指给LVS

LVS应用场景与基本功能介绍——keepalived

- 问题
  - 配置管理
  - RS宕机？
  - LVS宕机？
- Keepalived
  - 提供配置文件的管理方式
  - 负责健康检查，保证服务高可用
  - 支持vrrp协议，实现lvs的主备冗余



LVS应用场景与基本功能介绍

- 负载均衡系统
  - 请求分发模块：IPVS
  - 配置管理与HA：keepalived
  - 管理系统
  - 监控（基础监控、业务监控）

目录

- LVS应用场景与基本功能介绍
- **LVS在大规模网络环境下存在的问题**
- LVS的攻击防御
- LVS在360云中的基本应用
- 将来的工作

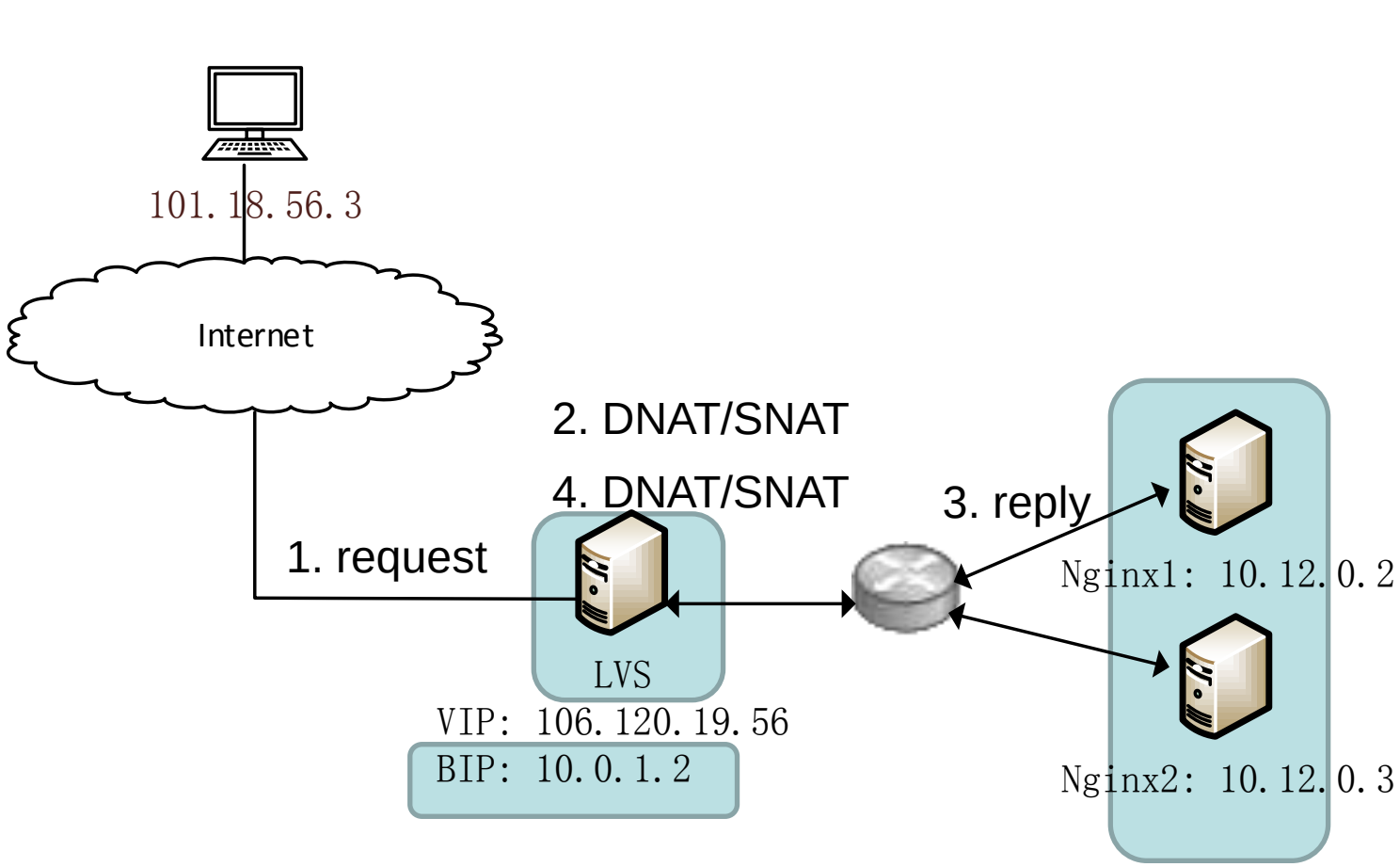
LVS在大规模网络环境下存在的问题

- 运维的成本较高
  - DR：配置VIP、二层规模
  - NAT： 配置路由
  - TUNNEL：配置IPIP、配置VIP
- 基于VRRP的主备模式存在单点瓶颈
  - 单个VIP的流量过高

- 使用跨网段的fullnat转发+集群部署

LVS在大规模网络环境下存在的问题——fullnat

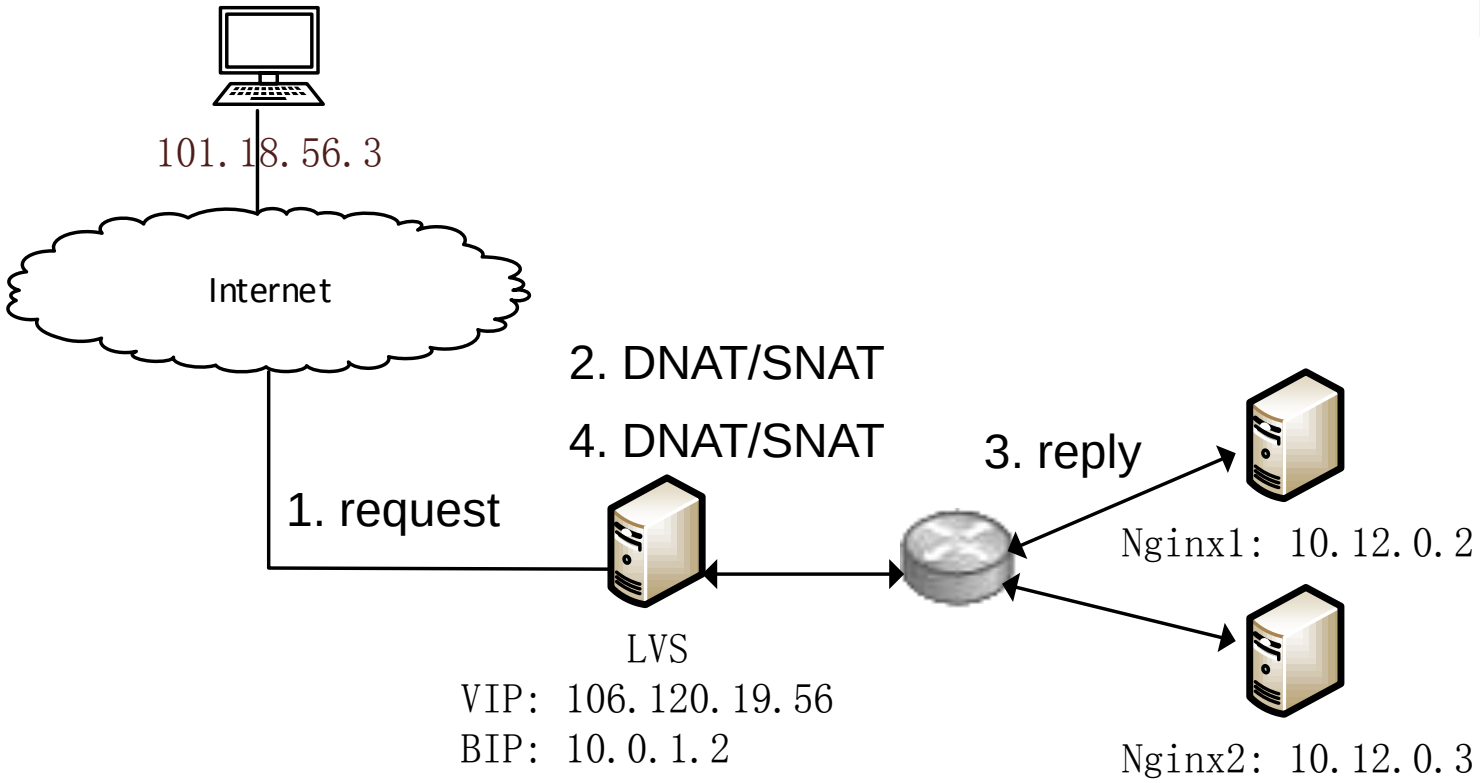
- 跨网段的fullnat转发

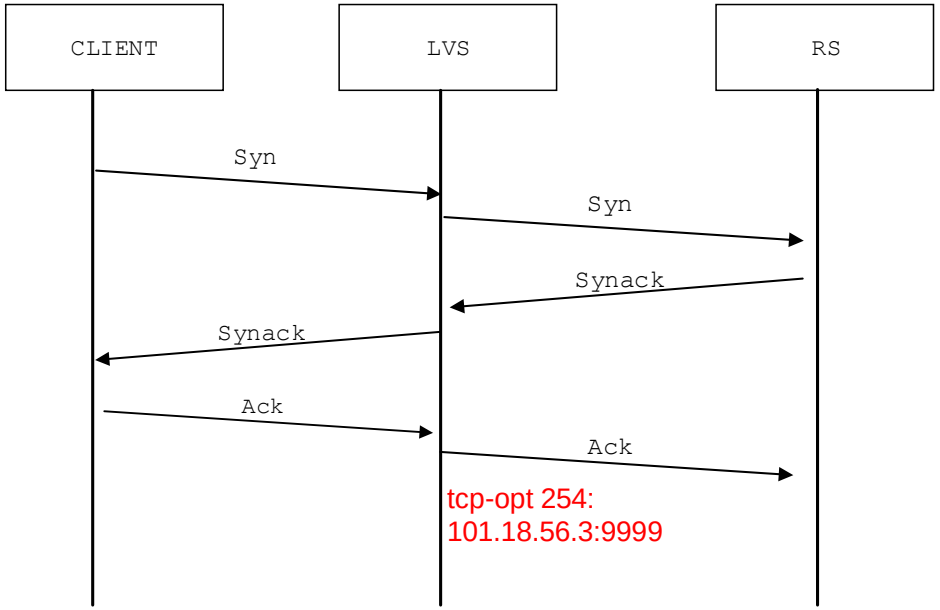


	SRC IP	DST IP
1:	101.18.56.3	106.120.19.56
2:	10.0.1.2	10.12.0.2
3:	10.12.0.2	10.0.1.2
4:	106.120.19.56	101.18.56.3

LVS在大规模网络环境下存在的问题——fullnat

- 跨网段的fullnat转发(用户真实IP)



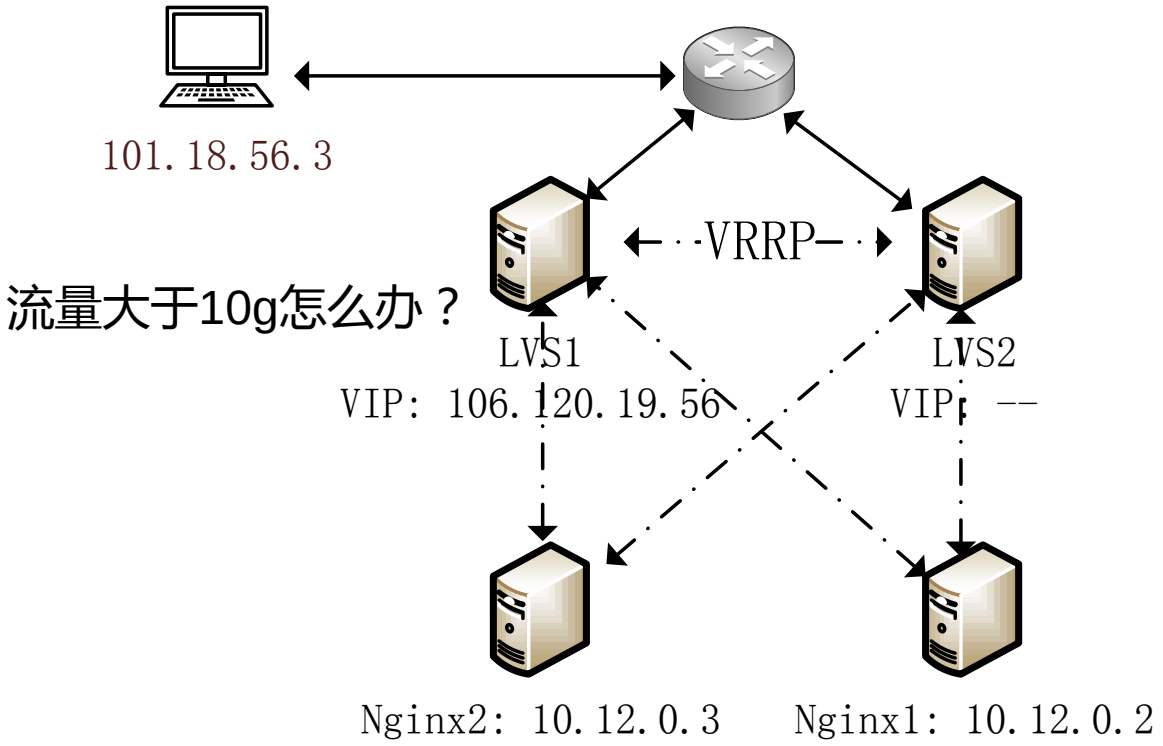


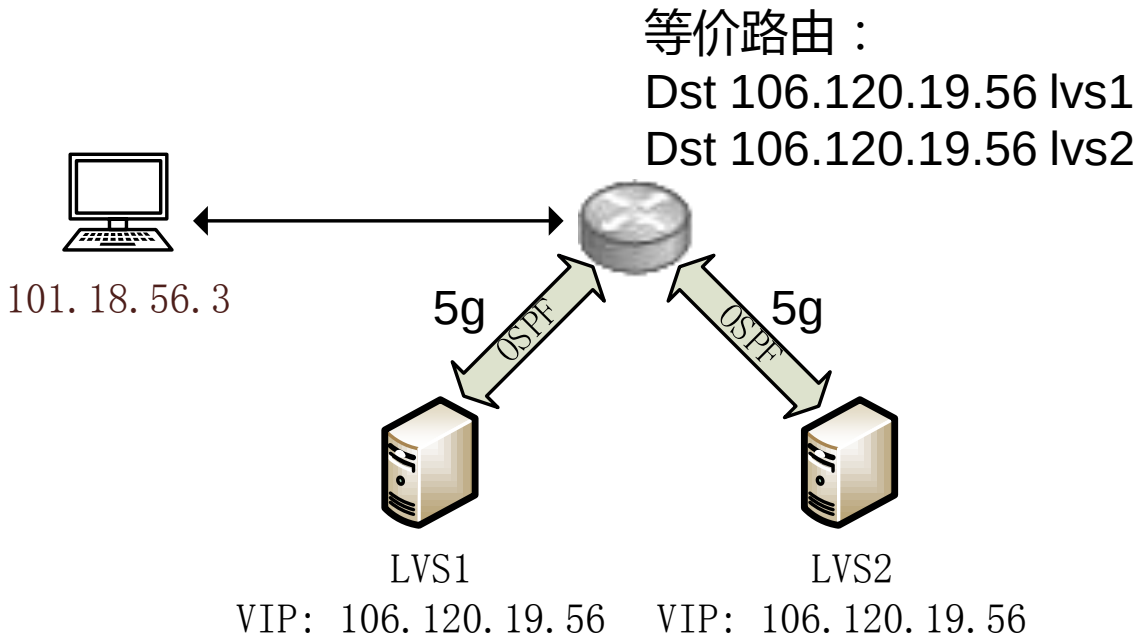
RS Kernel：

1. 在创建sock的时候保存tcp-opt数据
2. 在上层调用inet_getname时，使用上面保存的数据替换掉IP

LVS在大规模网络环境下存在的问题——集群部署

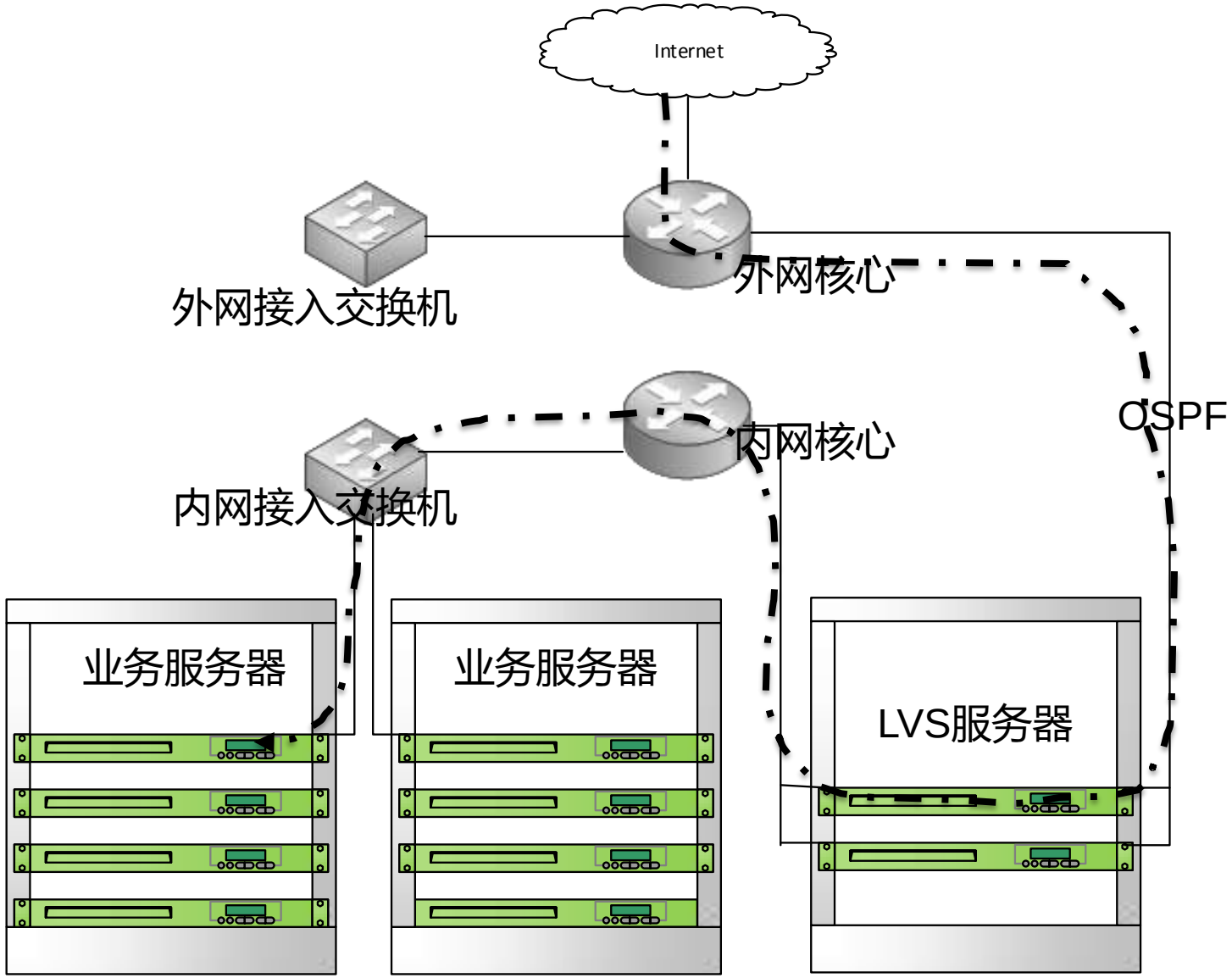
- 基于vrrp探活的单点瓶颈





集群模式下，vip不响应arp，会导致同网段不通

LVS在IDC中的实际部署



交换机对packet做HASH的元素有：
Ingress端口，源目的IP，源目的端口

问题：交换机一般不支持一致性hash
LVS集群内的session同步

LVS在大规模网络环境下存在的问题——小结

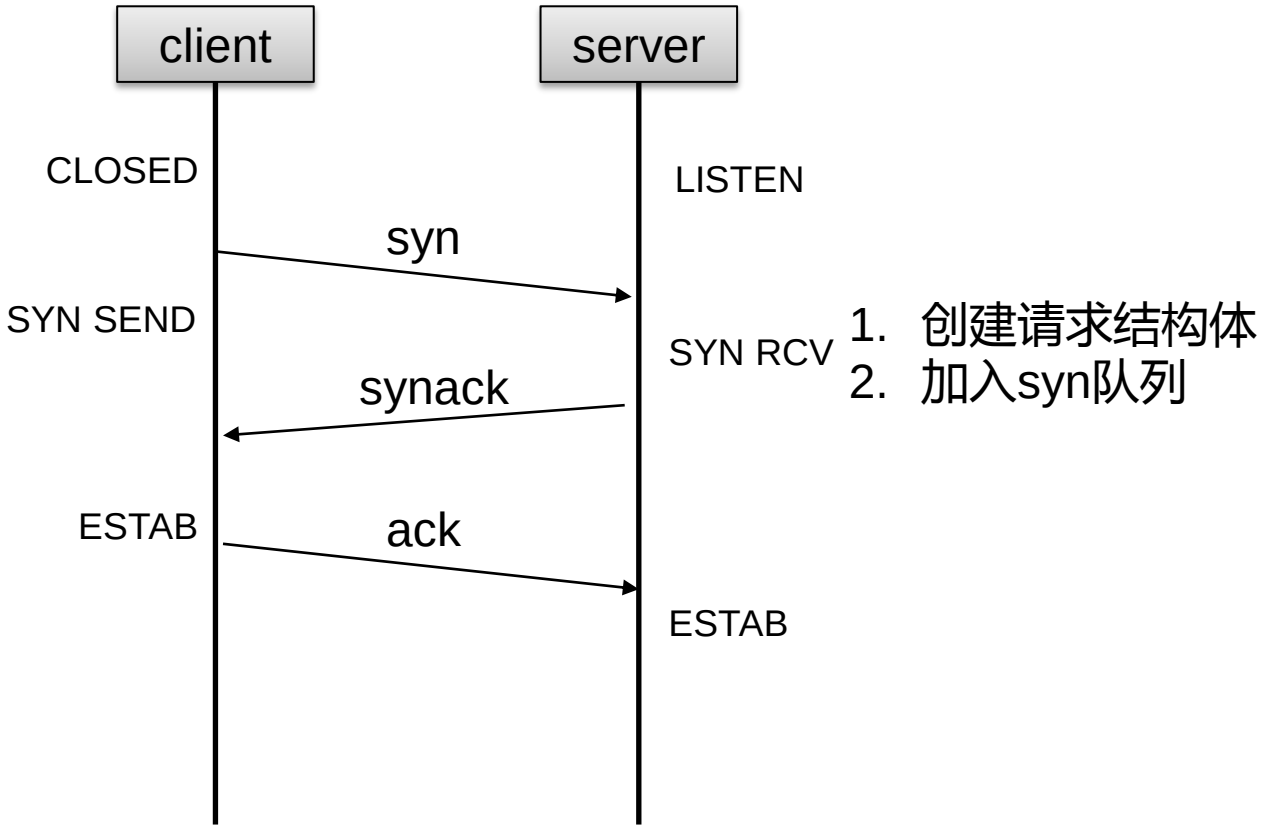
- 二层限制/配置复杂
  - 跨网段的fullnat转发
  - 引入真实IP问题
  - 为后端服务器定制内核模块
- LVS主备模式单点瓶颈
  - 利用等价路由实现AA模式的集群部署

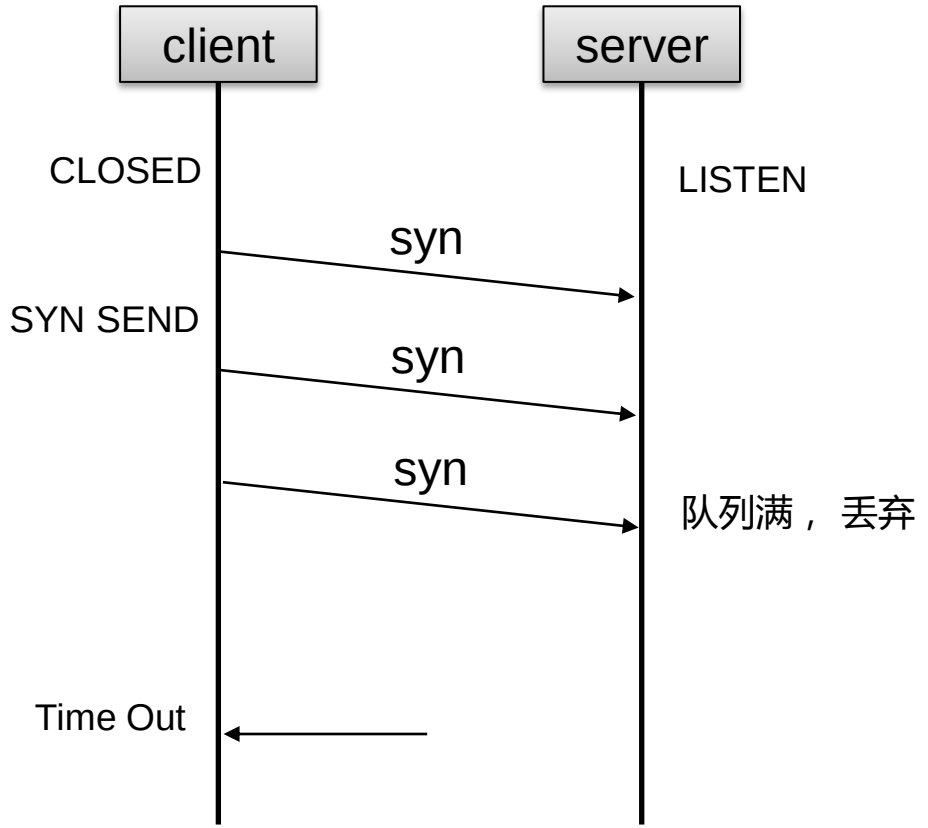
目录

- LVS应用场景与基本功能介绍
- LVS在大规模网络环境下存在的问题
- **LVS的攻击防御**
- LVS在360云中的基本应用
- 将来的工作

LVS的攻击防御- SYN Flood

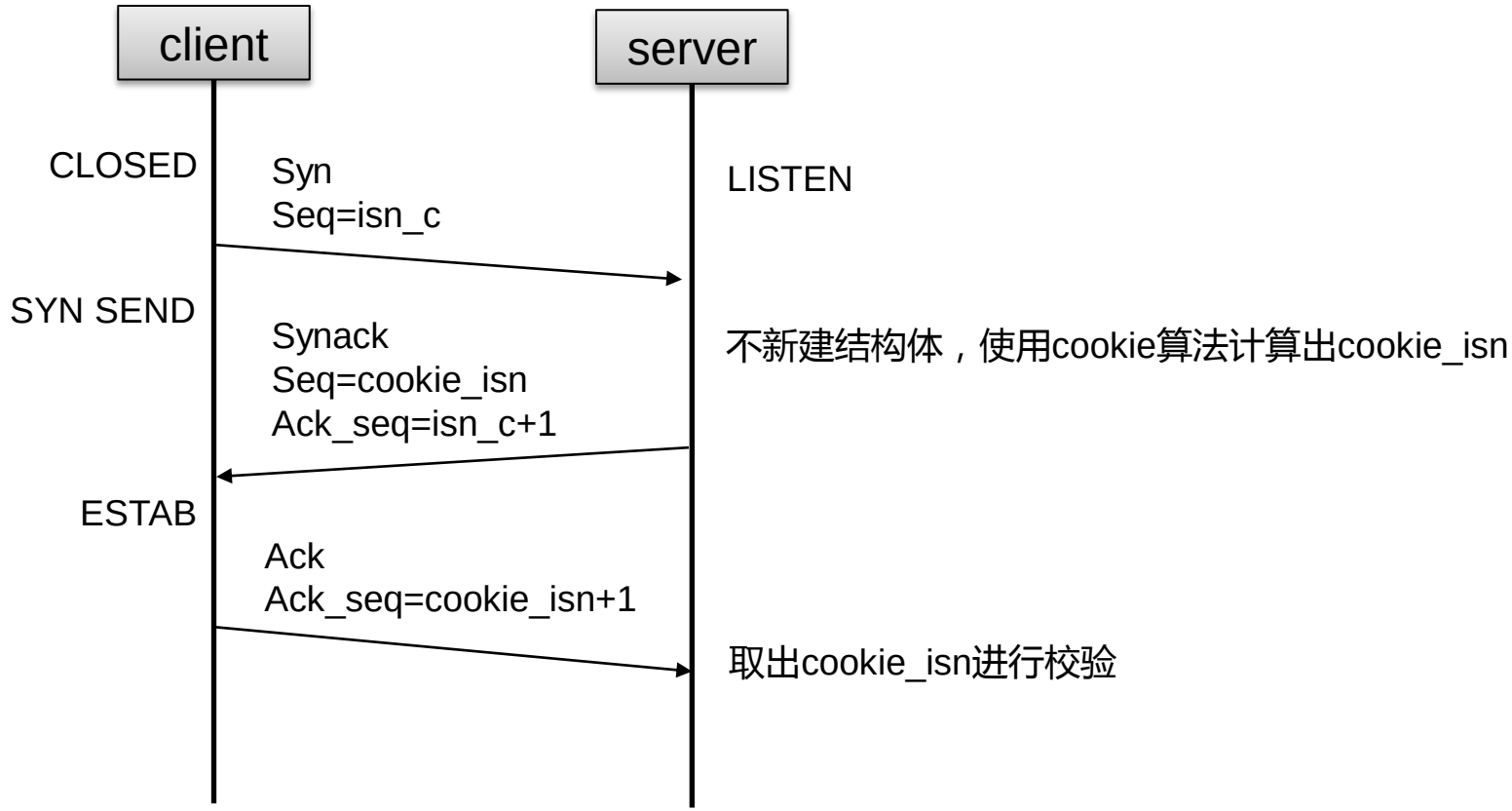
- 攻击原理





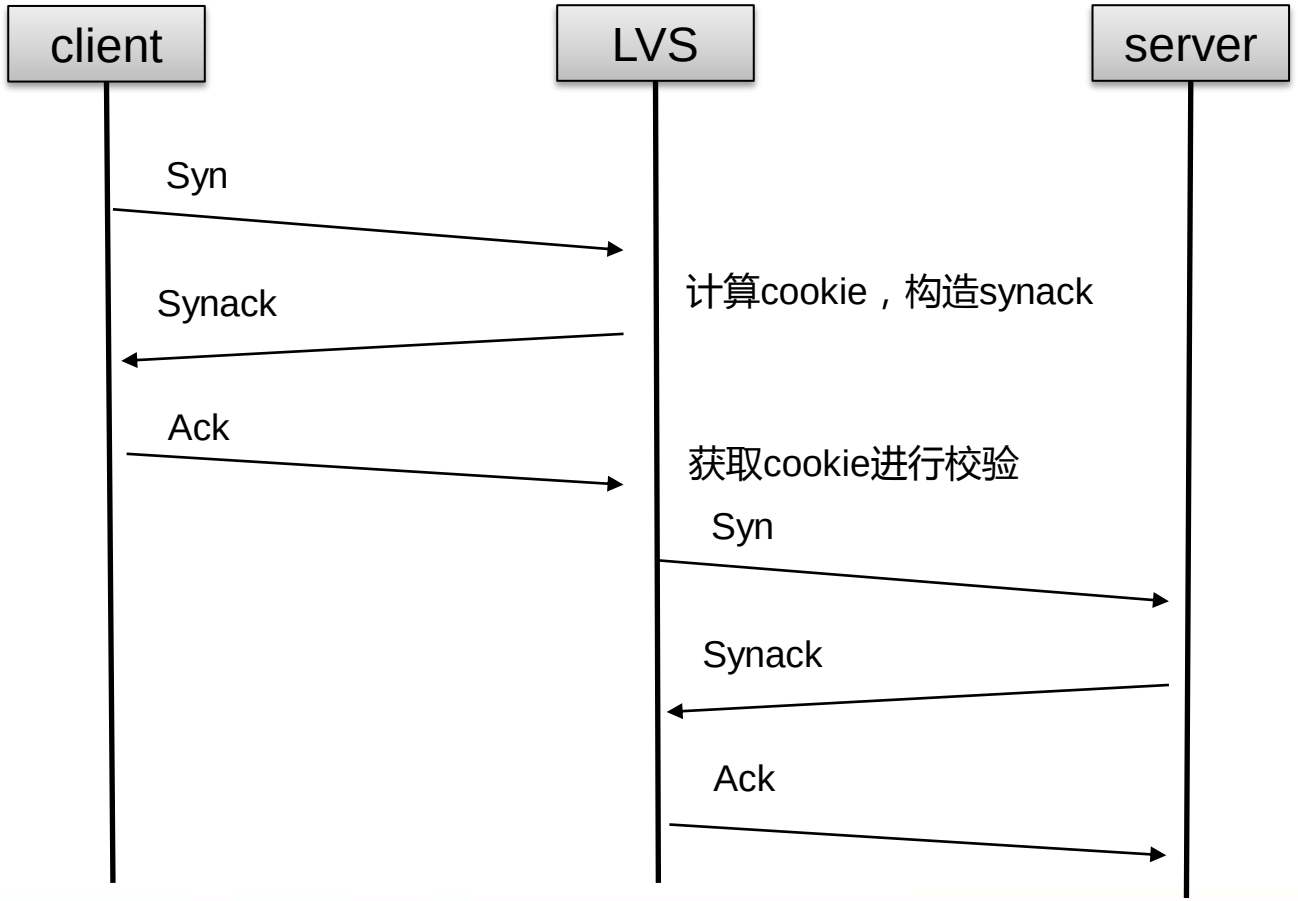
LVS的攻击防御- SYN Flood

- Kernel的防护方案



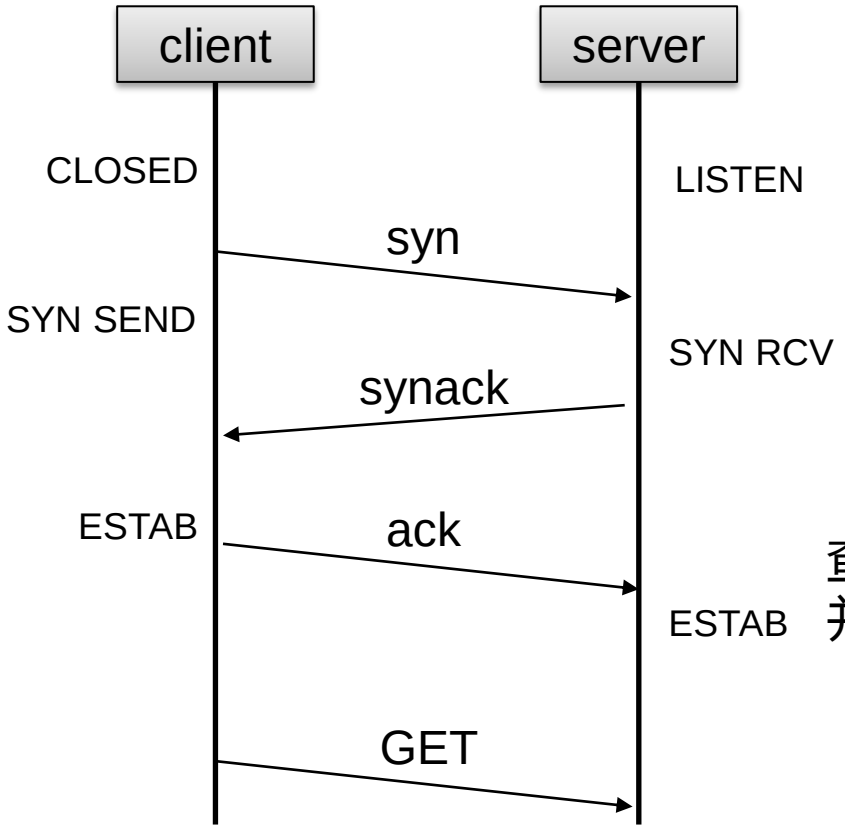
LVS的攻击防御- SYN Flood

- LVS的解决方案

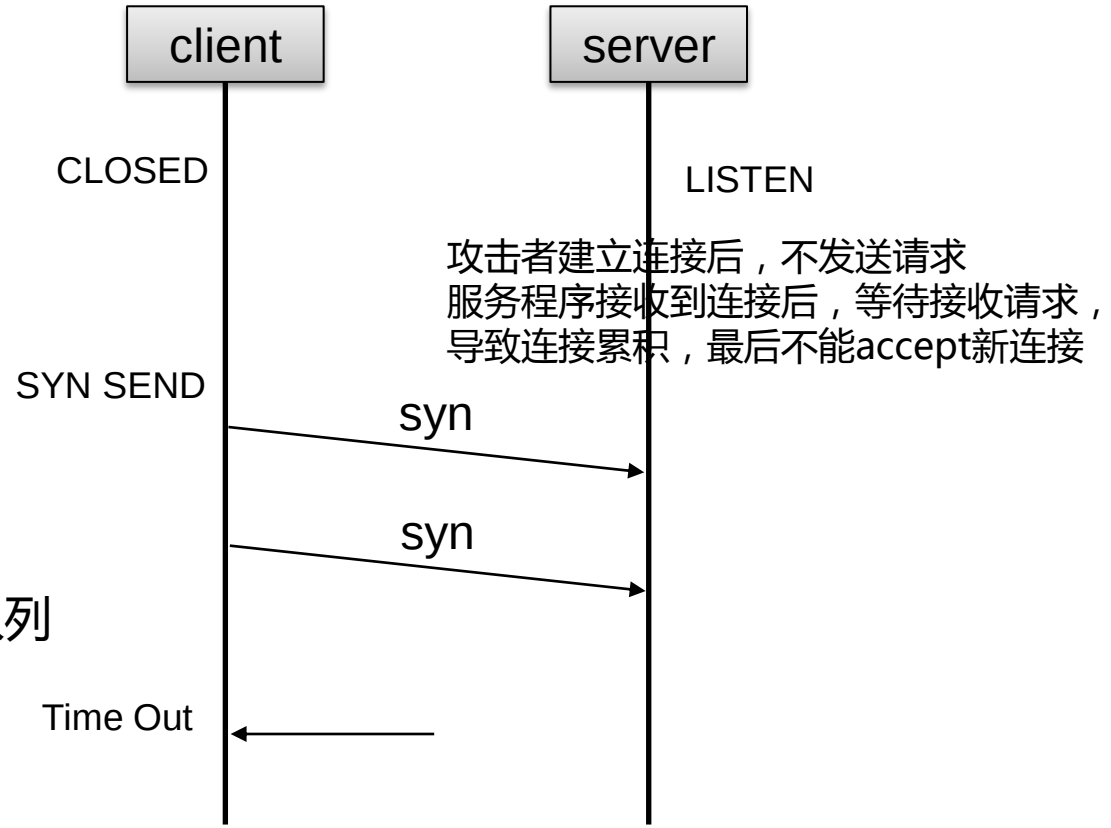


LVS的攻击防御- 慢连接

- 攻击原理

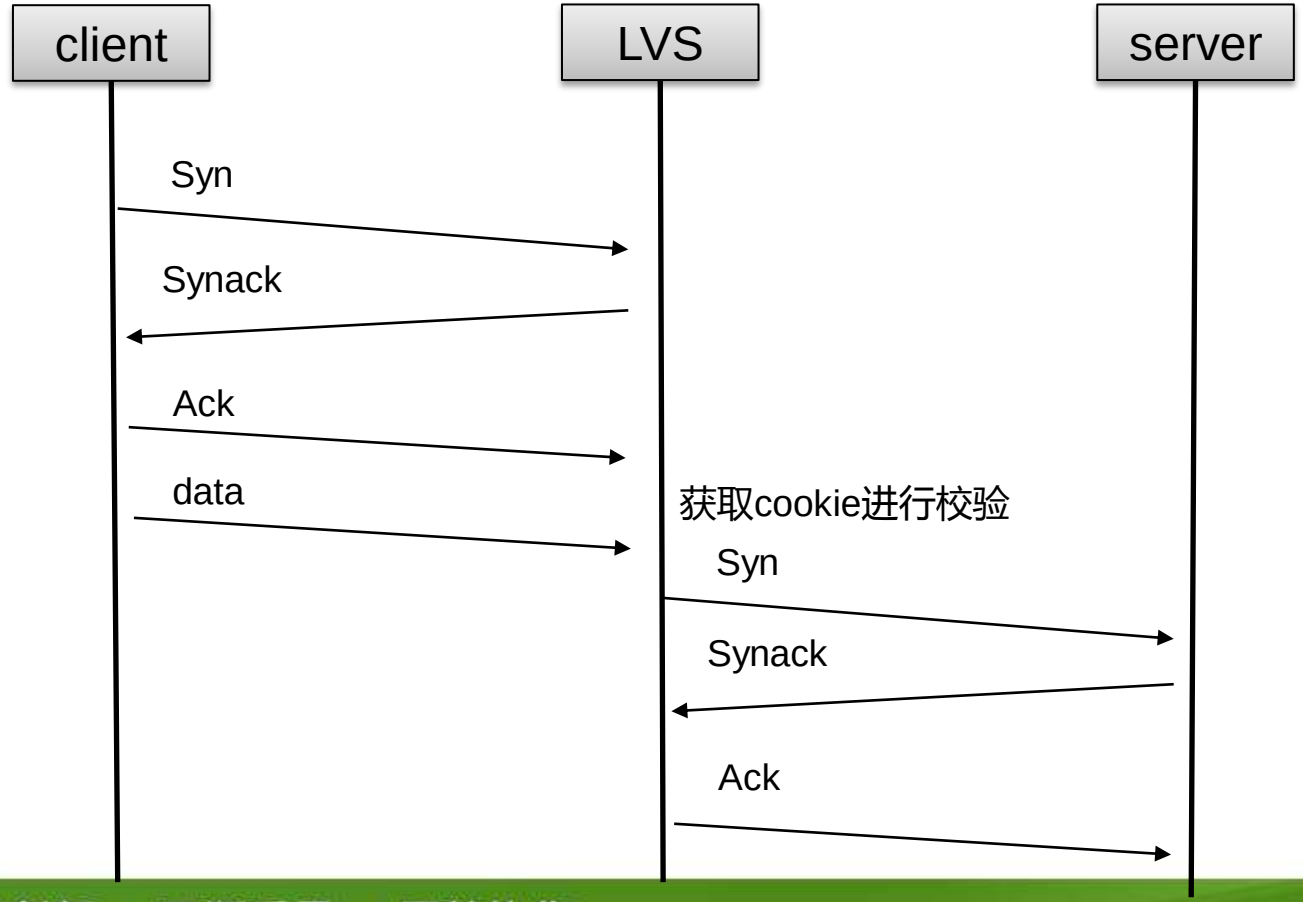


查询syn队列，
并将结构体加入accept队列



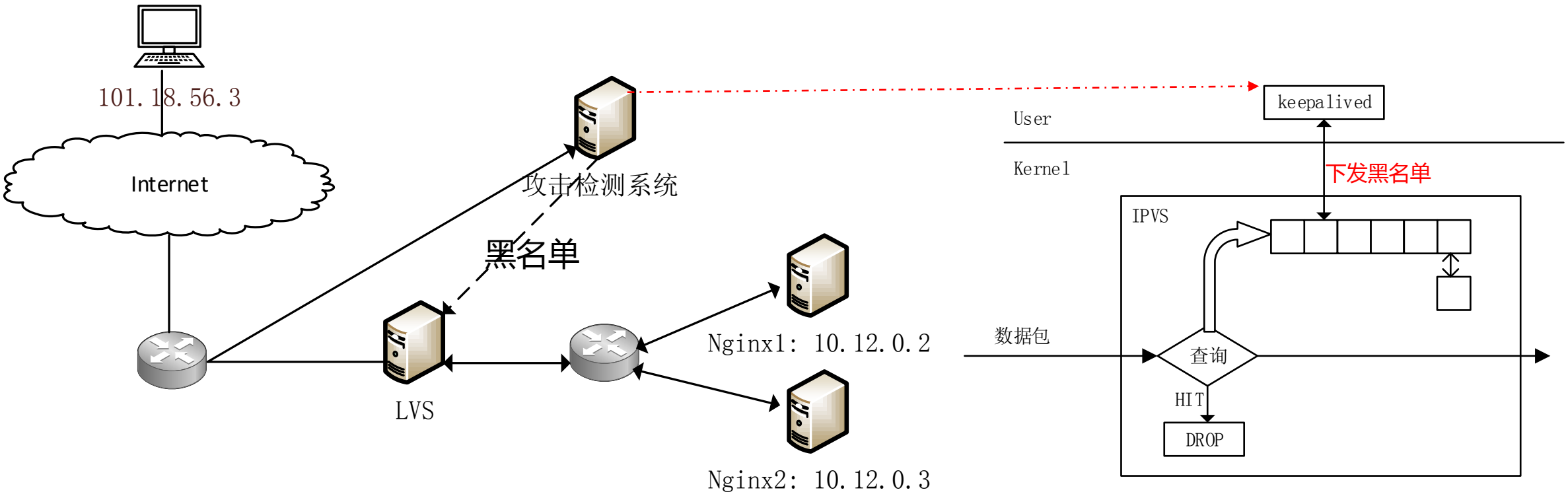
LVS的攻击防御- 慢连接

- LVS的解决方案
- 注： 不适用于Server先发包的业务场景



LVS的攻击防御- 其他复杂攻击

- 黑名单功能
  - 为攻击识别检测系统提供封禁接口



LVS的攻击防御——小结

- SYNFLOOD
  - Synproxy
- 慢连接
  - Synproxy + waitdata
- 其他复杂攻击
  - 提供黑名单封禁接口

目录

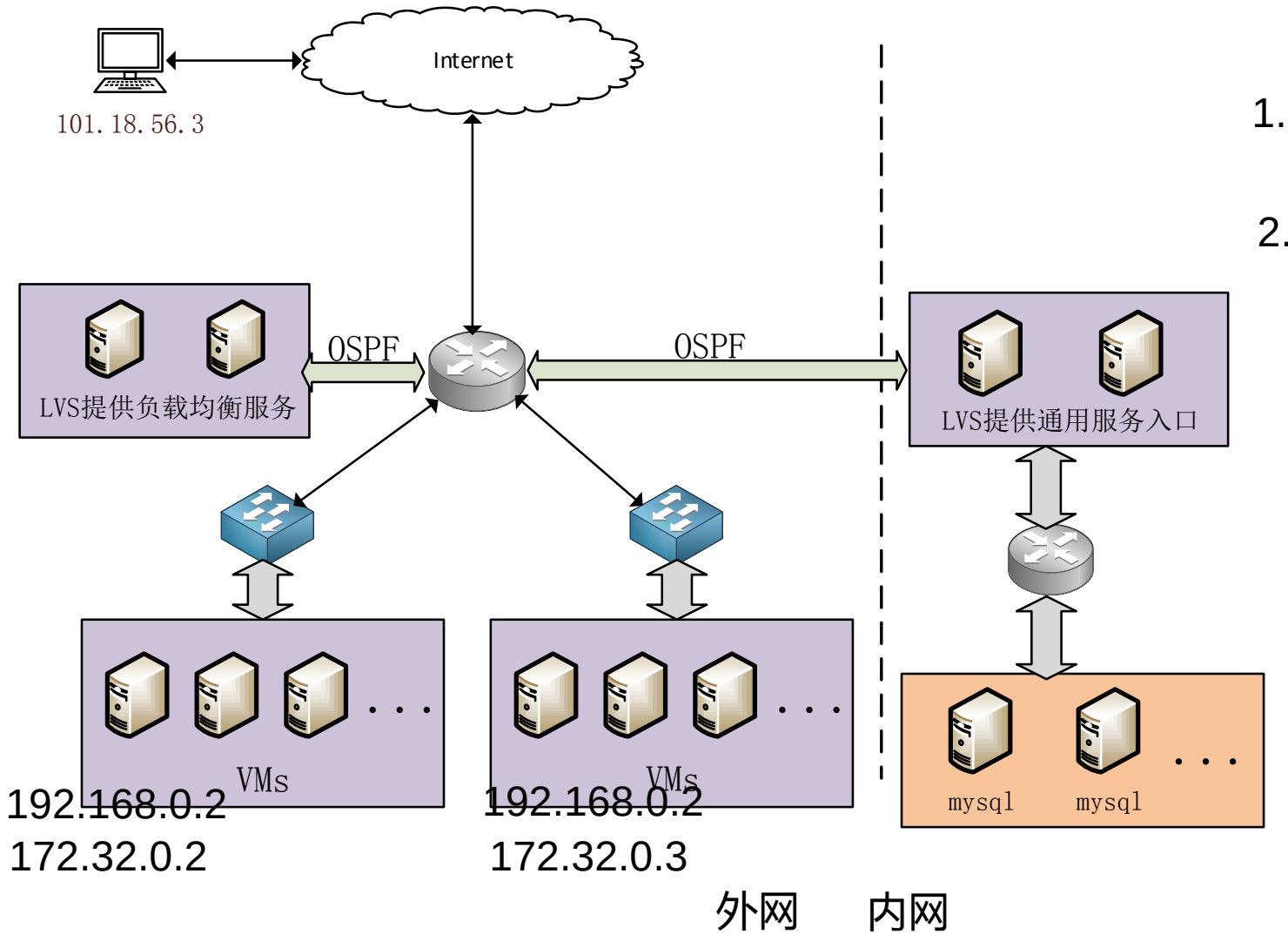
- LVS应用场景与基本功能介绍
- LVS在大规模网络环境下存在的问题
- LVS的攻击防御
- **LVS在360云中的基本应用**
- 将来的工作

LVS在360云中的基本应用

- 为云主机提供负载均衡服务
  - 简化在云主机内的配置
  - 为不同“区”下的云主机提供负载均衡
  - 采用fullnat转发模式的集群部署

- 内部通用服务的快速接入 —— mysql/memcache/redis…

LVS在360云中的基本应用



1. 自定义网络功能的引入，云主机IP地址可重叠

2. 获取用户真实IP，需要适配多种操作系统（windows、ubuntu、centos…）

目录

- LVS应用场景与基本功能介绍
- LVS在大规模网络环境下存在的问题
- LVS的攻击防御
- LVS在360云中的基本应用
- **将来的工作**

将来的工作

- IPVS：rwlock过渡到RCU
- 40g网卡
- fullnat的session同步
- 云平台中用户真实IP问题

Q/A