



IT大咖说  
知识共享平台

# 基于DPDK实现高性能L4技术实践

## 1

## 产生背景

K8s + haproxy ingress

✓ Haproxy性能不能满足需求

K8s + LVS

✓ DR, TUN, NAT模式需要修改容器配置

✓ FULLNAT不支持高版本内核

业务日益增长的高并发、低延时的需求

## 2

## 目标

K8s + ha接近网卡线速的报文转发性能

支持常见的4层负载均衡功能，配置灵活

充分发挥双路服务器性能，避免计算资源闲置

集群可实现高可用功能

部署简单，避免改动客户端和后端服务

## 3

## 功能特性

### 1) 协议支持

负载均衡支持包含TCP、UDP协议的四层负载均衡，配备健全的链路跟踪机制，以及多种调度策略，用户可以根据服务的需要创建合适自己的负载均衡。

### 2) 高可用性

支持容器的健康检查，保证应用可用性：负载均衡会定时探测后端服务是否正常运行，健康检查频率可自定义；一旦探测到异常，则不会将流量再分配到这些异常实例，保证应用可用性。

### 3) 集群部署，多层次容错机制

负载均衡采用集群部署，支持热升级，机器故障和集群维护对用户完全透明，结合DNS使用还可支持全局负载均衡。

## 4

# 功能特性

## 4) 灵活性

支持多种流量调度算法，使得流量分配更均匀：负载均衡支持加权轮询和最小连接数这两种调度算法，可根据自身需求选择相应的算法来分配用户访问流量，并支持设置后端容器权重，使得流量调度更均匀，提升负载均衡能力。

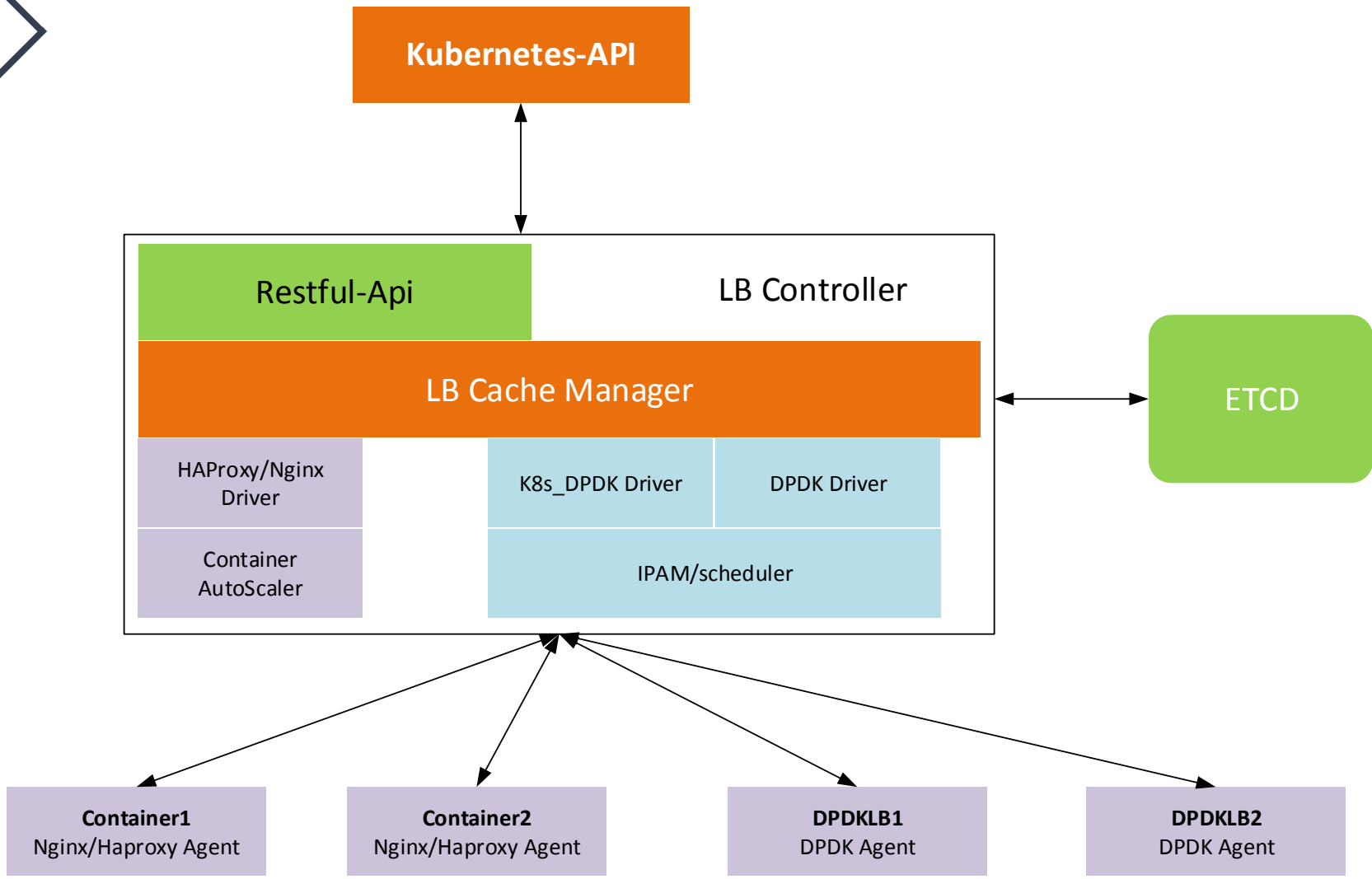
支持会话保持，满足用户个性化需求：负载均衡通过IP地址实现会话保持，可将一定时间内来自同一用户的访问请求，转发到同一个后端容器上进行处理，从而实现用户访问的连续性。

## 5) 易用性

提供多种管理途径，轻松操纵负载均衡：用户可通过控制台轻松实现负载均衡器的配置、释放等功能。后续会开放标准的API或SDK提供给用户，从而自己开发对负载均衡的控制管理。

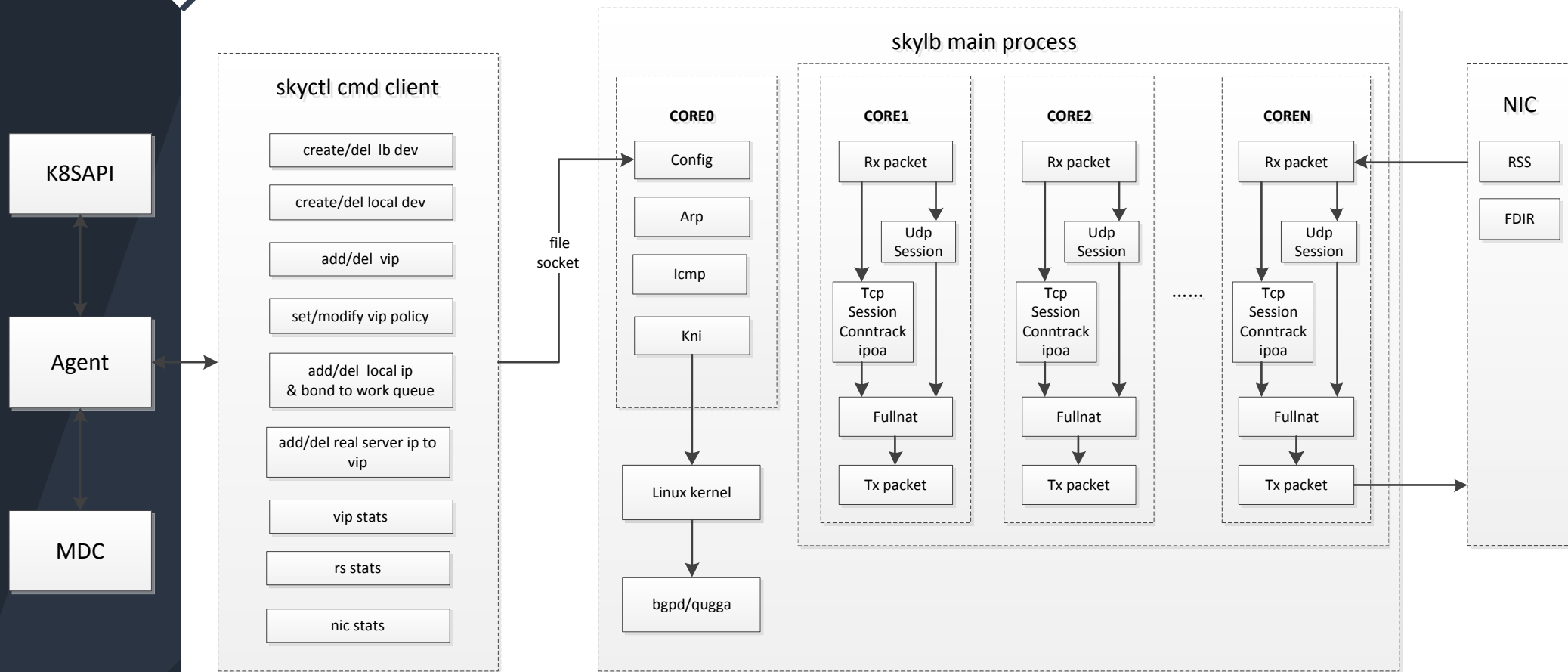
# 集群架构

5



# 6

## DBDKLB架构



## 7

## DBDKLB 组网架构

支持传统二层vlan 负载均衡以及三层bgp组网。

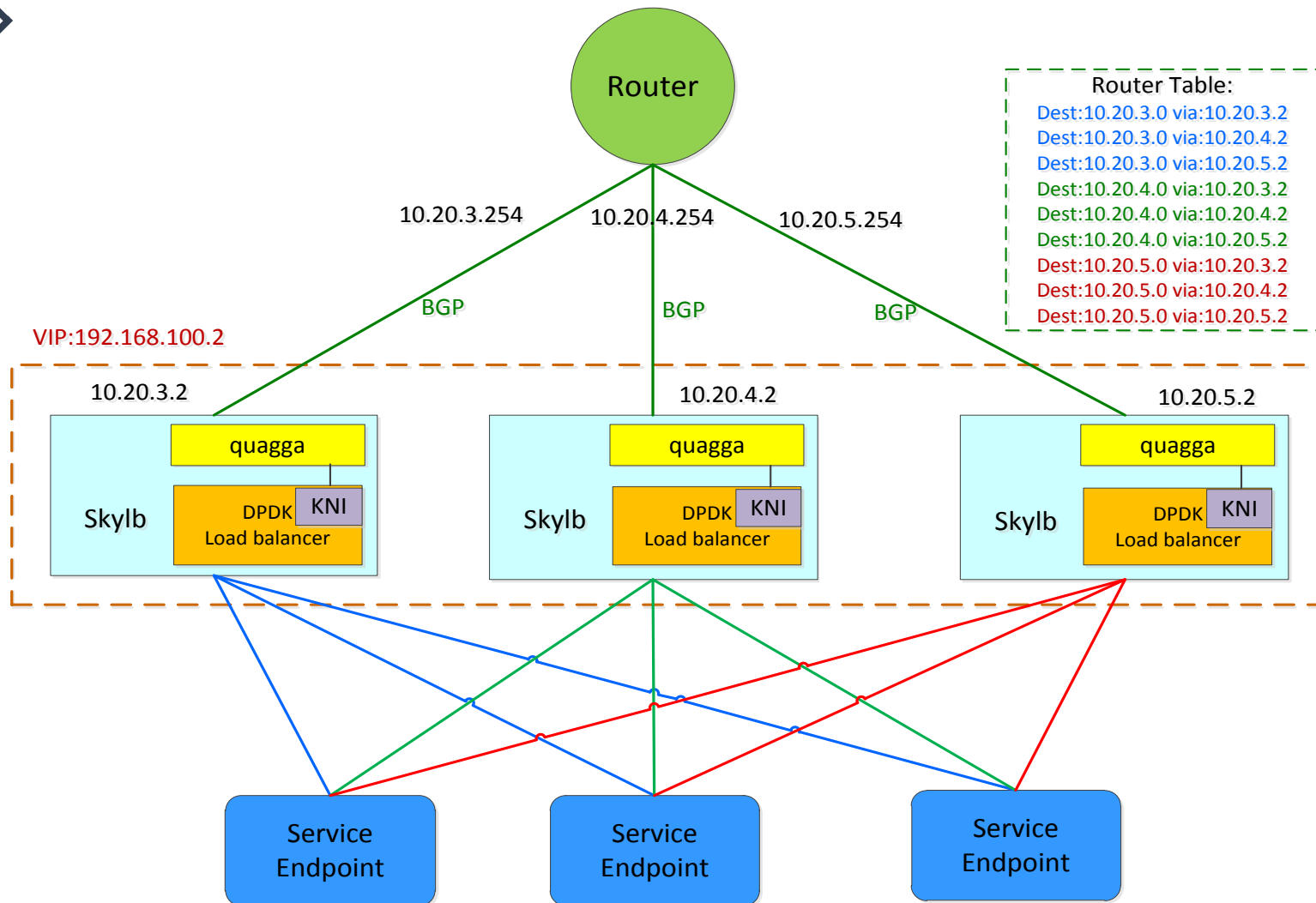
配合switch ecmp 实现负载均衡的高可用

下一页基于bgp 介绍负载均衡的基本组网情况



# 8

## DBDKLB 组网架构



## 9

## 性能

protocol	schedule	TPS	ERROR	jupiter rx-pps	jupiter rx-bps
TCP	ipport	2064789.7	0	4.78M	8.41G
TCP	rr	2060283.5	0	4.78M	8.41G
TCP	lc	777078.5	0	-	-
UDP	ipport	4212952	0	8.28M	7.75G
UDP	rr	4272837.6	0	8.28M	7.75G
UDP	lc	812356.2	0	-	-

Github 地址：<https://github.com/tiglabs/jupiter>

## 10

## 问题及解决

**问题：**

virtio-user代替KNI与内核网络交互，不支持master/slave进程模式

**原因：**

Virtio-user创建vdev设备的内存空间非进程间共享，slave进程无法访问，出现core dump

**解决方案：**

在该BUG未解决前，使用KNI与内核网络交互。

## 11

## 问题及解决

**问题：**

dpdk线程占用CPU 100%，无法了解业务处理的负载情况

**分析：**

无

**解决方案：**

分别统计每秒线程收发包占用的时间和业务处理占用的时间，计算各自百分比。缺点是业务负载情况不够精确，时间统计浪费CPU。



IT大咖说  
知识共享平台

谢谢!