

知 DL380 Gen10+服务器搭配QL41000系列网卡时，k8s集群只能通过ip+端口的方式访问主机，无法通过ip+端口的方式访问pod容器

驱动安装升级 徐鑫 2024-01-12 发表

组网及说明

DL380 Gen10+服务器

Marvell QL41000系列网卡

问题描述

k8s集群内节点端口只能通过机器的ip+端口访问，无法通过节点ip+端口进行访问
换句话说就是只能通过ip+端口的方式访问主机，但是无法通过ip+端口的方式访问pod容器

过程分析

k8s内部使用IPIP隧道数据包的方式进行内部DNS解析，这种特定的数据包不支持硬件tx校验和协助，所以QLogical网卡的驱动对数据包进行校验会导致数据包损坏，进而导致影响k8s内部DNS解析的访问，而宿主机使用的DNS是外部，常规的DNS不使用IPIP数据包类型。

只针对Marvell QLogic的这个网卡。

https://support.hpe.com/hpesc/public/docDisplay?docLocale=en_US&docId=a00114668en_us

CA文档解决方案需要禁用 tx-checksum offload (ethtool -K <interface> tx off)，以便 DNS 解析正常运行，现场操作也确实如此。但用户并不想做这种动作。

根据CA文档提供的bug网站 (<https://bugs.launchpad.net/ubuntu/+source/linux/+bug/1909062>)，表示Marvell has developed a fix for the qede driver, which checks the packet type, and if it is IPPROTO_IPIP, then csum offloads are disabled for socket buffers of type IPIP. 厂商其实提供了新的qede驱动，可以解决问题，封装到了新的内核驱动中。(补丁网站: [qede: fix offload for IPIP tunnel packets · torvalds/linux@5d5647d · GitHub](https://github.com/torvalds/linux/pull/5d5647d))

https://support.hpe.com/connect/s/software/details?language=en_US&softwareId=MTX_db0757dac9da41c2b256760033&tab=revisionHistory

最新版本驱动解决了DNS解析的问题:

This product addresses an issue where the 10GbE Adapters having an DNS resolution issue on Kubernetes container-orchestration system

524SFP+的底层是QL41132HLCU网卡 (<https://www.marvell.com/products/hpe/products.html>)，使用同一个驱动模块，升级后解决。

解决方法

临时解决: 使用ethtool禁用硬件的tx数据包解析 // ethtool -K <interface> tx off

永久解决: 升级驱动至8.55.14.0-2 及以上