

我们引用 Soong-Sil University 发布的第三方测试报告:Considerations for Benchmarking Network Performance in Containerized Infrastructure。该报告包含多个测试场景，我们关注 BMP2BMP（Baremetal POD to Baremetal）场景，即测试运行在物理服务器上 pod 的网络通信性能。

在 BMP2BMP 中测试如下几类插件：

- OpenStack-Kuryr (OVS bridge)
- Flannel-CNI (docker bridge-Flannel bridge)
- MACVLAN, IPVLAN / Data acceleration(SR-IOV)

BMP2BMP 场景环境架构如下图 2 所示：有两个 Bare metal 环境，第一个环境有两个 pod，第二个环境有一个 pod。

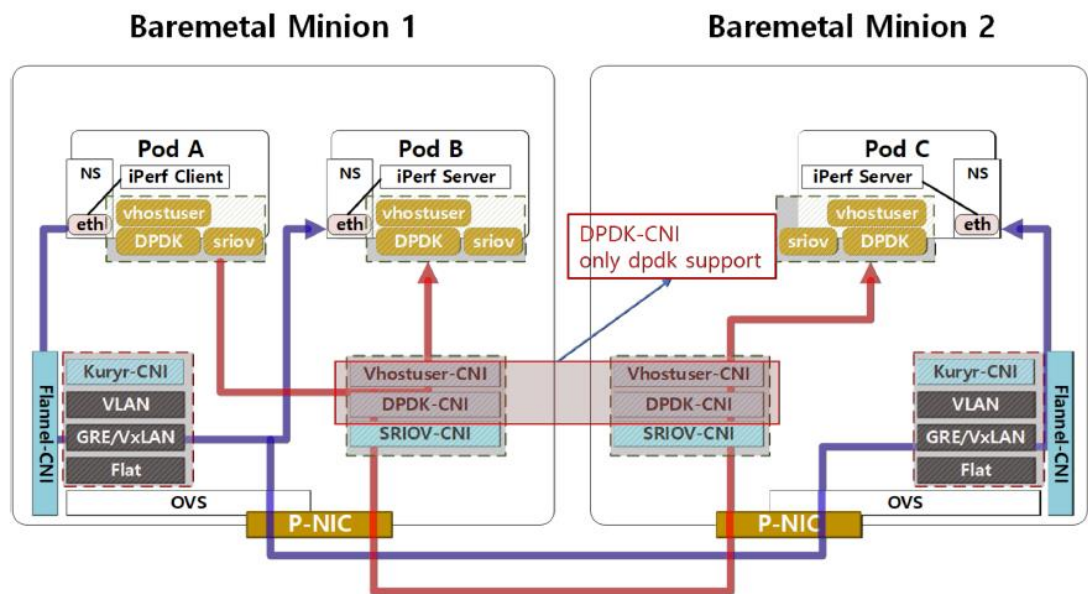


图 2. Pod 网络性能测试环境架构图

BMP2BMP 场景两个服务器的配置如下所示：

节点	部件	配置
Minion1	CPU	Intel(R) Xeon(R) Gold 6148 2.40GHz * 2
Minion2	MEMORY	DDR4 2400 MHz 32GB * 6

	SR-IOV NIC	Mellanox ConnectX-5 (40G SFP+)
--	------------	--------------------------------

BMP2BMP 场景软件配置如下所示：

软件配置	OS	Ubuntu 16.04 Server LTS
	Cloud OS	Openstack queens by Devstack
	COE	kubernetes v1.9.0 and docker 18.06
	CNI	default cni plugin driver and kuryr, flannel, sr-iov, vshot-user, multus

在一个物理服务器上两个 pod 的通讯的测试结果：

吞吐量：macvlan > ovs-vlan (20% lower) > SR-IOV > ipvlan，详情如下图 3 所示：

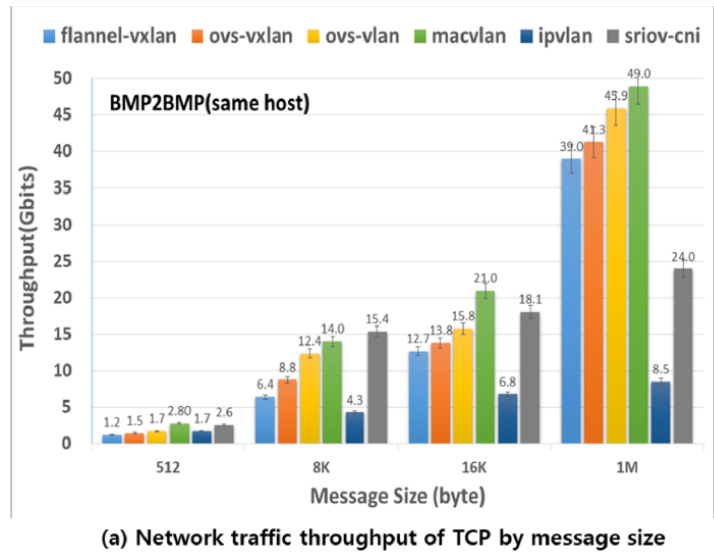


图 3. 同一个物理机上两个 pod 网络通讯的吞吐量

网络延迟：SRIOV(up to 16K) > ovs-vlan > ipvlan > macvlan，如下图 4 所示：

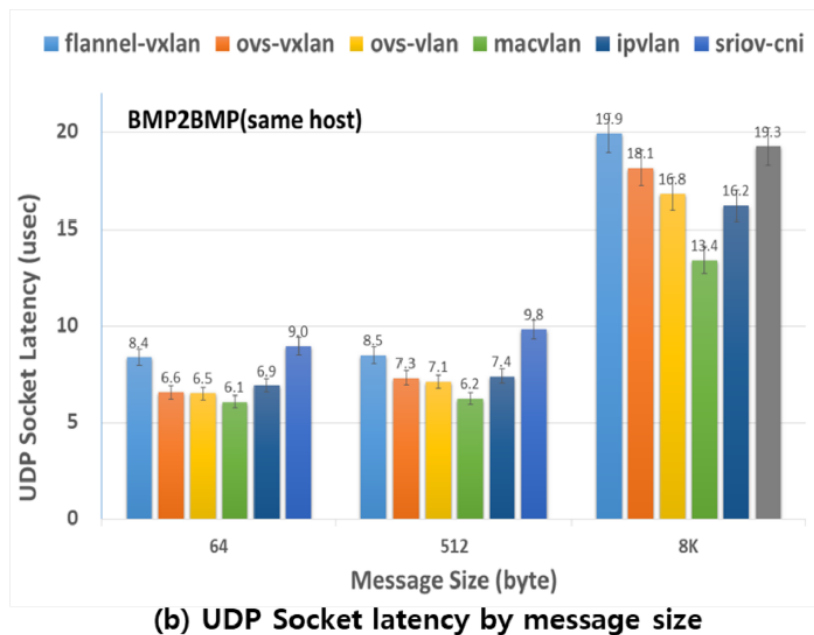


图 4. 同一个物理机上两个 pod 网络通讯的网络延迟

两个物理服务器上 pod 的通信性能测试中，结合吞吐量和网络延迟，整体性能排名：

MACVLAN > SR-IOV > ovs-vlan > ipvlan

具体数据如下图 5 所示：

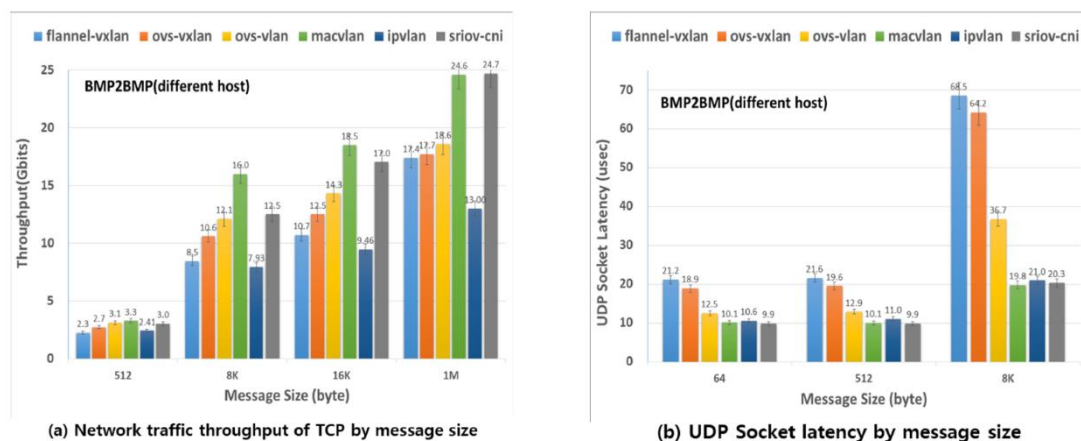


图 5. 两个物理机上 pod 网络通讯的网络性能

我们再关注另外一个测试报告：Performance of Container Networking Technologies，在该报告中吞吐量测试时：容器所用的 macvlan 网络的极限值和线性与物理网卡非常接近（线性好于 ovs 和 Linux Bridge），如下图 6 所示：

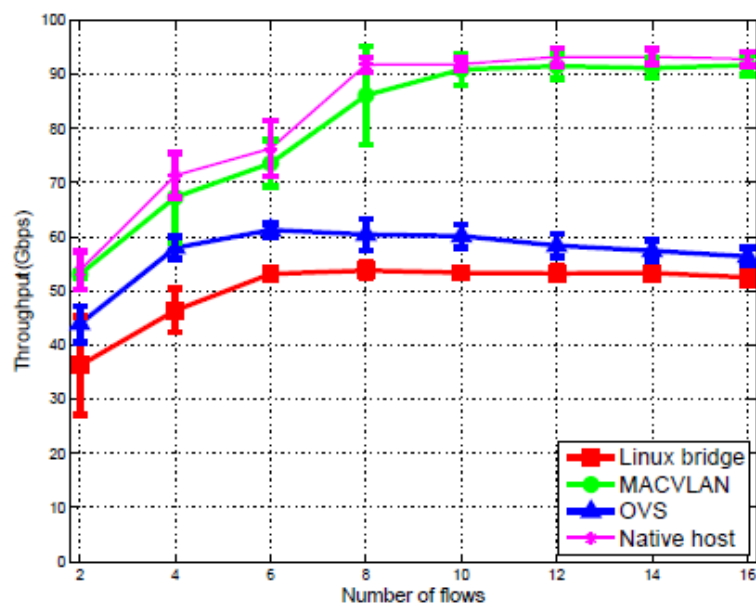


Figure 7: CHHC multi flows throughput with networking modes.

图 6. 吞吐量极限测试

从以上两个测试报告中，我们可以得出结论：如果客户在生产环境想提升 OpenShift 中的容器网络性能，不必考虑替换默认 SDN（OVS）。使用 macvlan 等多网络平面即可，macvlan 的性能接近物理网卡。这样既能保证 OpenShift 原生架构不会破坏（带来 PaaS 方案的多厂商支持的问题），也有助于提升容器的业务网络的性能。