

VNET/P: Bridging the Cloud and High Performance Computing Through Fast Overlay Networking

¹Lei Xia, ²Zheng Cui, ³John Lange, ⁴Yuan Tang, ¹Peter Dinda, ²Patrick Bridges



1. Northwestern University
2. University of New Mexico
3. University of Pittsburgh
4. University of Electronic Science and Technology of China



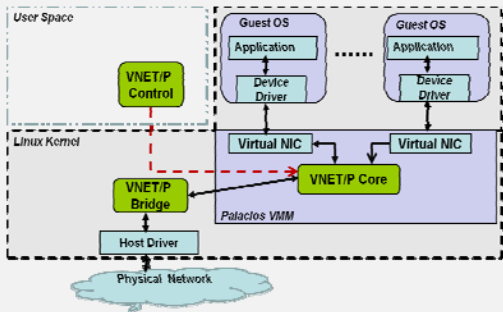
VNET Model

- VNET: A layer 2 virtual overlay network for the user's virtual machines
 - Provide location independence to VMs
 - Carry VMs' traffic via configurable overlay network
- Virtual machines on virtual networks as the abstraction for computing
- Virtual network as a fundamental layer for measurement and adaptation

Bridge the Cloud and HPC

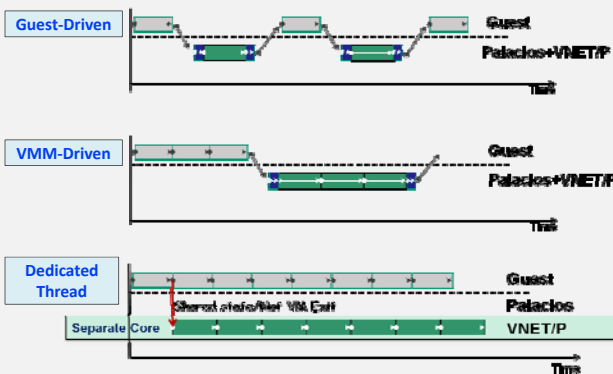
- Adaptive cloud/HPC-spanning model:** Seamlessly bridge distributed cloud and tightly-coupled resources
 - Applications can dynamically span to both environments
 - To provide effective on-demand resources for HPC
- Adaptive cloud/HPC-spanning model is currently limited by performance of virtual networking**
 - How can we provide high performance inter-VM traffic while VMs are located on the **same** data center/cluster?

Architecture of VNET/P



- Move virtual networking directly into VMM
- Enable optimizations that can only happen inside VMM

Optimizations

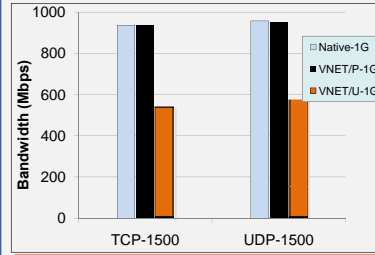


High throughput mode avoids most VM exits

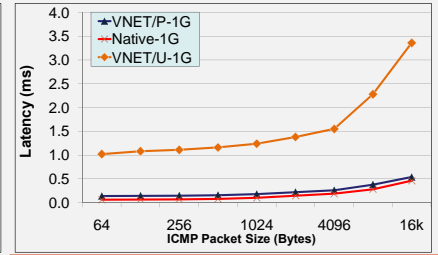
Future Work

- Further performance improvement
 - Further performance optimizations (*in submission*)
 - Optimistic interrupts, Cut-through forwarding, Noise isolation
 - Move VNET up to guest through guest code injection (*to appear in ICAC'12*)
- An Ethernet abstraction for VMs on non-Ethernet interconnects (InfiniBand, SeaStar, etc)

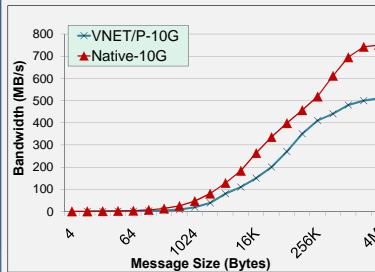
Performance of VNET/P



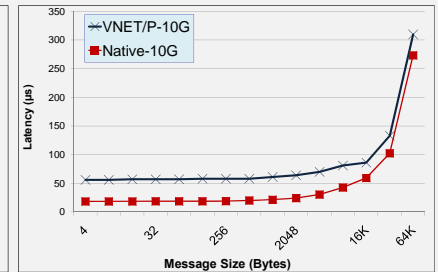
VNET/P achieves **native** bandwidth on 1Gbps network



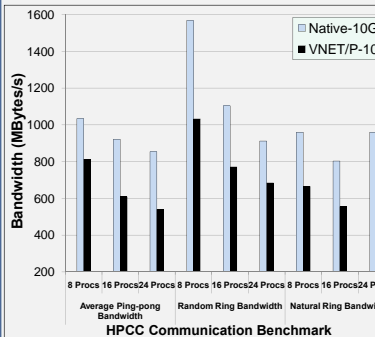
VNET/P achieves **2 times** of native latency on 1Gbps network



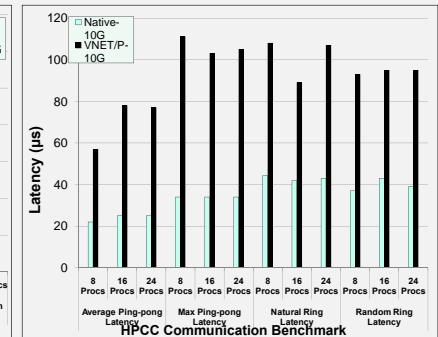
VNET/P achieves around **65%-70%** of native bandwidth on 10Gb network



VNET/P achieves **3 times** of native latency on 10Gb network



VNET/P achieves **scalable** bandwidth and latency on multiple nodes



VNET/P achieves **scalable** bandwidth and latency on multiple nodes

App	Native-1G	VNET/P-1G	VNET/P-1G (%)	Native-10G	VNET/P-10G	VNET/P-10G (%)
ep.B.8	103.15	101.94	98.8%	102.18	102.12	99.9%
ep.B.16	204.88	203.9	99.5%	208	206.52	99.3%
ep.C.8	103.12	102.1	99.0%	103.13	102.14	99.0%
ep.C.16	206.24	204.14	99.0%	206.22	203.98	98.9%
mg.B.8	4400.52	3840.47	87.3%	5110.29	3796.03	74.3%
mg.B.16	1506.77	1498.65	99.5%	9137.26	7405	81.0%
cg.B.8	1542.79	1319.43	85.5%	2096.64	1806.57	86.2%
cg.B.16	160.64	159.69	99.4%	592.08	554.81	93.7%
fb.B.16	1575.83	1290.78	81.9%	1433.3	1228.39	85.8%
ia.B.8	38.88	34.61	89.0%	50.15	50.04	99.8%
ia.B.16	35.99	35.78	99.4%	23.09	23	99.6%
ia.C.8	89.54	82.15	91.7%	132.08	131.87	99.8%
ia.C.16	84.36	82.22	97.6%	77.77	76.84	98.8%
lu.B.8	6818.52	5495.23	80.6%	7173.65	6021.78	83.9%
lu.B.16	7847.99	6694.12	85.3%	12981.86	9643.21	74.3%
sp.B.8	1361.38	1215.85	89.3%	2634.23	2421.98	91.9%
sp.B.16	1489.32	1399.6	94.0%	3010.71	2916.81	96.9%
tu.B.9	3423.52	3297.04	96.3%	5229.01	4076.52	78.0%
tu.B.16	4599.38	4348.99	94.6%	6315.11	6105.11	96.7%

VNET/P achieves **95%** of native performance on most of NAS application benchmarks

Conclusion

- Extend virtual networking for VMs down to clusters and supercomputers
 - **Such model is limited currently by virtual networking performance**
- VNET/P: high performance virtual overlay networking for tightly-coupled parallel systems
 - Overlay networking directly implemented into VMM
 - High performance on 1Gb/10Gb networks
- Software-based overlay network can be extended into tightly-coupled environments**