



Hamilton Institute

Very High Speed Networking in VMs and Bare Metal

Professor Luigi Rizzo
University of Pisa

Friday, July 5th, 2013

Abstract: In this talk I will give a survey of solutions and tools that we have developed in recent years to achieve extremely high packet processing rates in commodity operating systems, running on bare metal and on virtual machines.

Our NETMAP framework supports processing of minimum-size frames from user space at 10 Gbit/s (14.88 Mpps) with very small CPU usage. Netmap is hardware-independent, supports multiple NIC types, and it does not require IOMMU or expose critical resources (e.g. device registers) to userspace. A libpcap library running on top of netmap gives instant acceleration to pcap clients without even the need to recompile applications.

VALE is a software switch using the netmap API, which delivers over 20~Mpps per port, or 70 Gbit/s with 1500-byte packets. Originally designed to interconnect virtual machines, VALE is actually very convenient also as a testing tool and as a high speed IPC mechanism.

More recently we have extended QEMU, and with a few small changes (using VALE as a switch, paravirtualizing the e1000 emulator, and with small device driver enhancements), we reached guest-to-guest communication speeds of over 1 Mpps (with socket-based clients) and 5 Mpps (with netmap-based clients).

Netmap and VALE are small kernel modules, part of standard FreeBSD and also available as add-on for Linux. QEMU extensions are also available from the author and are being submitted to the qemu-devel list for inclusion in the standard distributions.

Links: <http://info.iet.unipi.it/~luigi/netmap/>, <http://info.iet.unipi.it/~luigi/vale/>

Biography: Luigi Rizzo is an Associate Professor of Computer Engineering at the University of Pisa, Italy. His research focuses on computer networks and operating systems. In particular, he has done some highly cited work on multicast congestion control, FEC-based reliable multicast, network emulation, and more recently on packet scheduling, fast network I/O, virtualization. Much of his work has been implemented and deployed in popular operating systems and applications, and widely used by the research community. His contributions include the popular dummynet network emulator (a standard component of FreeBSD and OSX, and now also available for linux and windows); one of the first publicly available erasure code for reliable multicast; the qfq packet scheduler; and the netmap framework for fastpacket I/O.

Venue: Seminar Room, Hamilton Institute, Rye Hall, NUI Maynooth

Time: 2.00pm - 3.00pm

Travel directions are available at www.hamilton.ie

