

Solarflare and OpenOnload

Solarflare Server Adapter Family



**Dual Port SFP+
SFN5122F & SFN5162F**

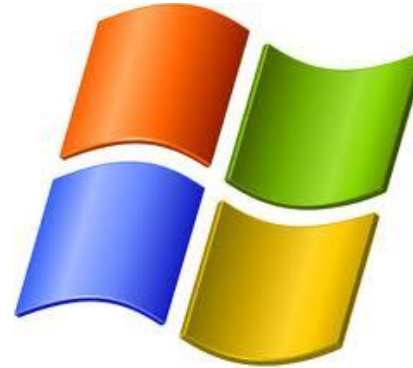
**Single Port SFP+
SFN5152F**

**Dual Port 10GBASE-T
SFN 5121T & SFN5161T**

**Single Port 10GBASE-T
SFN5151T**

- High performance
- Lowest power: < 2.5W per port SFP+
- Linux (< 7 μ sec), Windows, and Solaris
- Solarflare OpenOnload for ultra-low latency (< 4 μ sec)
- Hardware-assisted, scalable virtualization: up to 2048 vNICs

Platforms



Microsoft®
Hyper-V™



- High performance, low-latency
 - < 11uS ½ RTT TCP latency
 - Bidirectional line rate, both ports (40G) with < 20% CPU load
 - Stateless offloads (TSO, LRO, RSS)
 - Performance scales with multiple CPU cores
- Enterprise feature set
 - VLAN and teaming (including 802.3ad link aggregation)
 - Unattended installation
 - SNMP MIB
- Comprehensive management
 - Command line and GUI (MMC snap-in integrated into OS)
 - Local and remote management

- High performance, low-latency
 - $< 7\mu\text{S}$ $\frac{1}{2}$ RTT UDP latency
 - Bidirectional line rate, both ports (40G) with $< 20\%$ CPU load
 - Stateless offloads (TSO, LRO, RSS)
 - Performance scales with multiple CPU cores
- Full integration into Linux kernel
 - Driver “intree” (kernel.org)
 - Wide distribution and kernel support
 - Integration with ethtool, MTD, hwmon (sensors)
- Actively improving Linux networking
 - Solarflare is the maintainer of “ethtool”
 - Implemented Linux kernel RFS hardware offload

- Solaris 10
 - High performance low-latency Solaris GLDv3 driver
- Virtualization
 - VMware ESX 4.x high performance driver
 - SR-IOV support for highest performance
 - Citrix XenServer6, KVM, more coming
 - Windows Hyper-V VMQ acceleration
- BootROM
 - Full PXE and iSCSI boot support

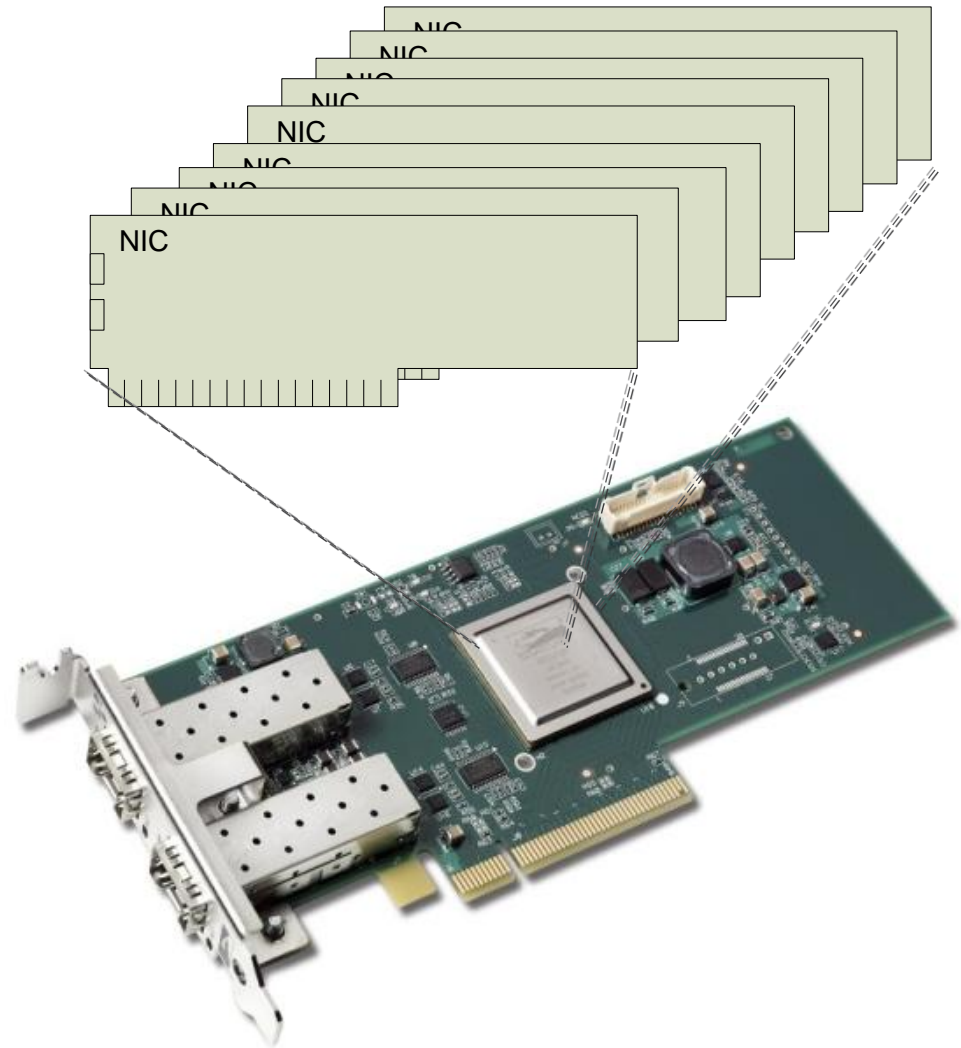
- Wide range of deployments
 - Financial
 - Many major exchanges worldwide
 - Cloud
 - Enterprise
 - Several Fortune 100
 - Scientific/HPC



- Wide range of partners



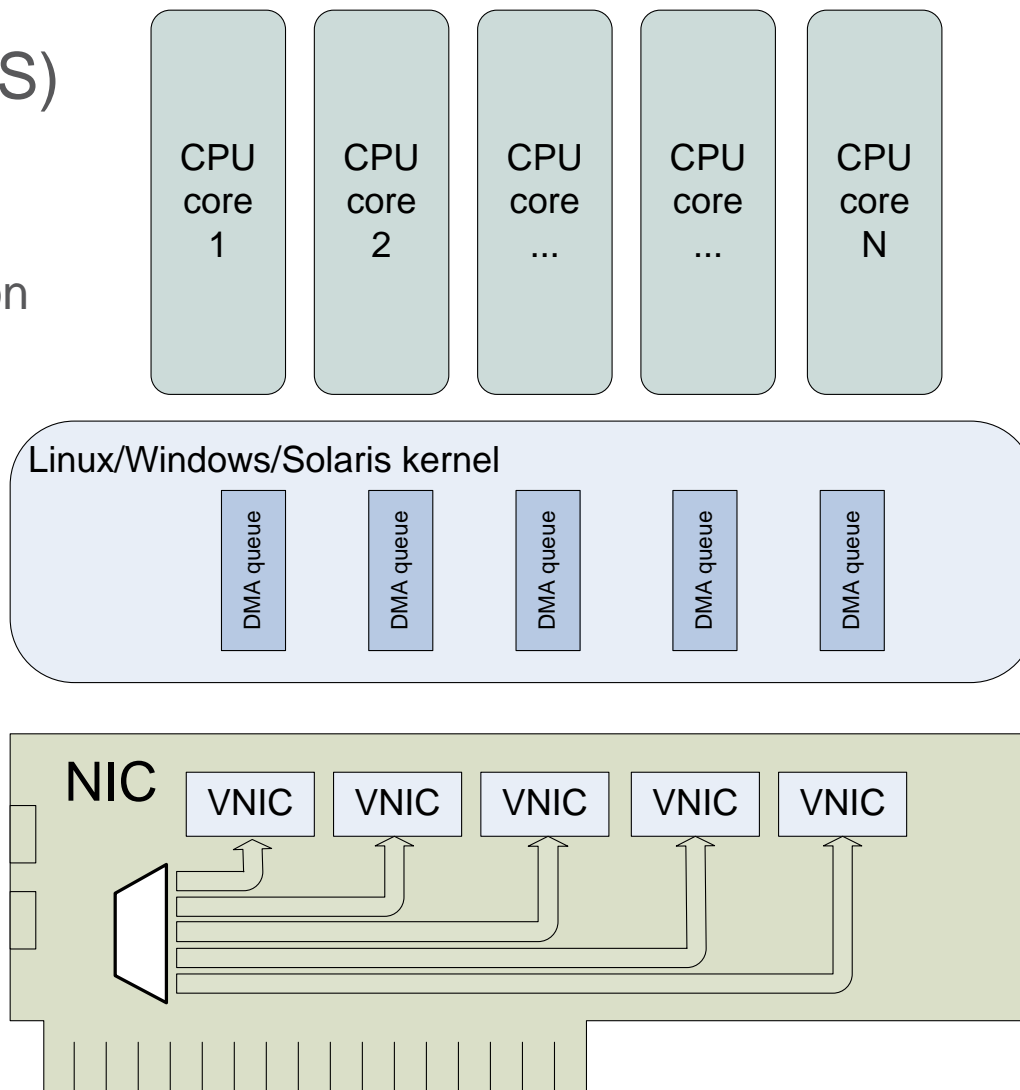
- Solarflare server adapters are “virtualisable”
 - 2048 vNICs
 - High performance in many domains



Virtual NICs for performance scaling

VNIC per CPU core (RSS)

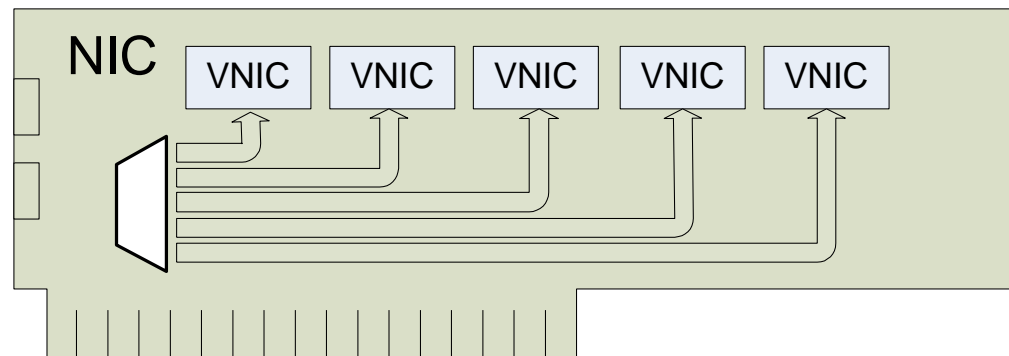
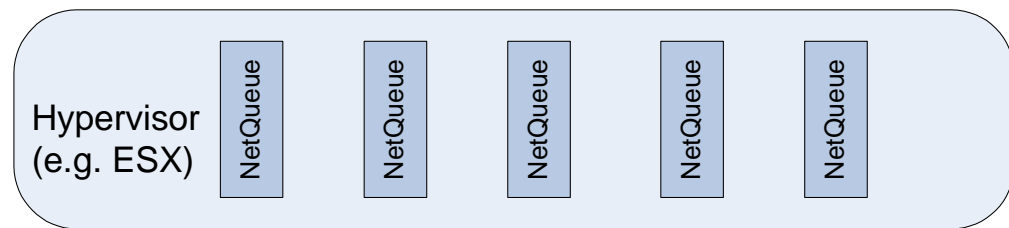
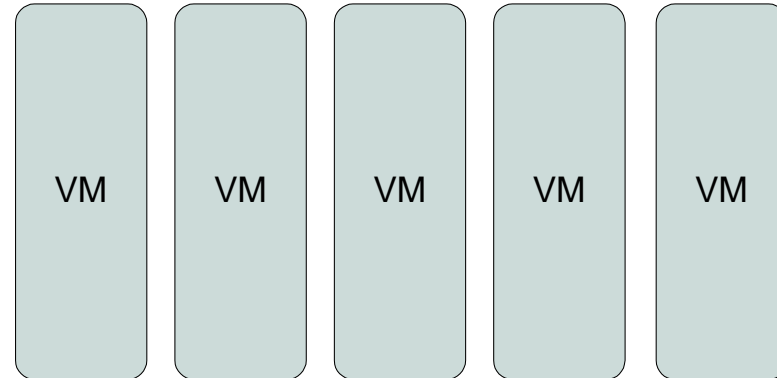
- RX queue per CPU core
- TX queue per CPU core
- Complete CPU core separation
- Performance scales across CPUs



Virtual NICs for Virtual Machines

VNIC per VM

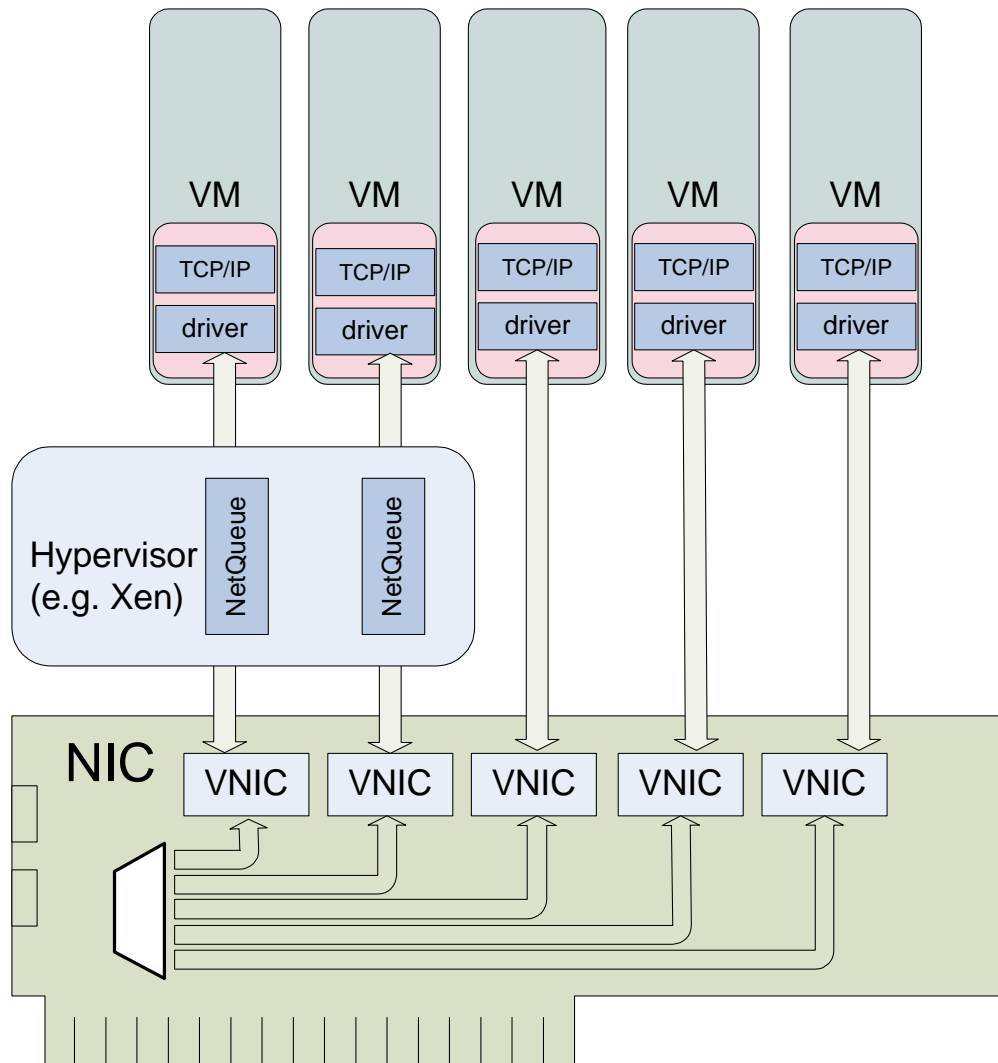
- RX queue per VM
- TX queue per VM
- VM separation
- Used for VMware NetQueue and Windows HyperV VMQ



Virtual NICs for Virtual Machines (2)

VNIC per VM

- Same model used for SR-IOV
- In this case VM has direct access to VNIC(s) via SR-IOV VF

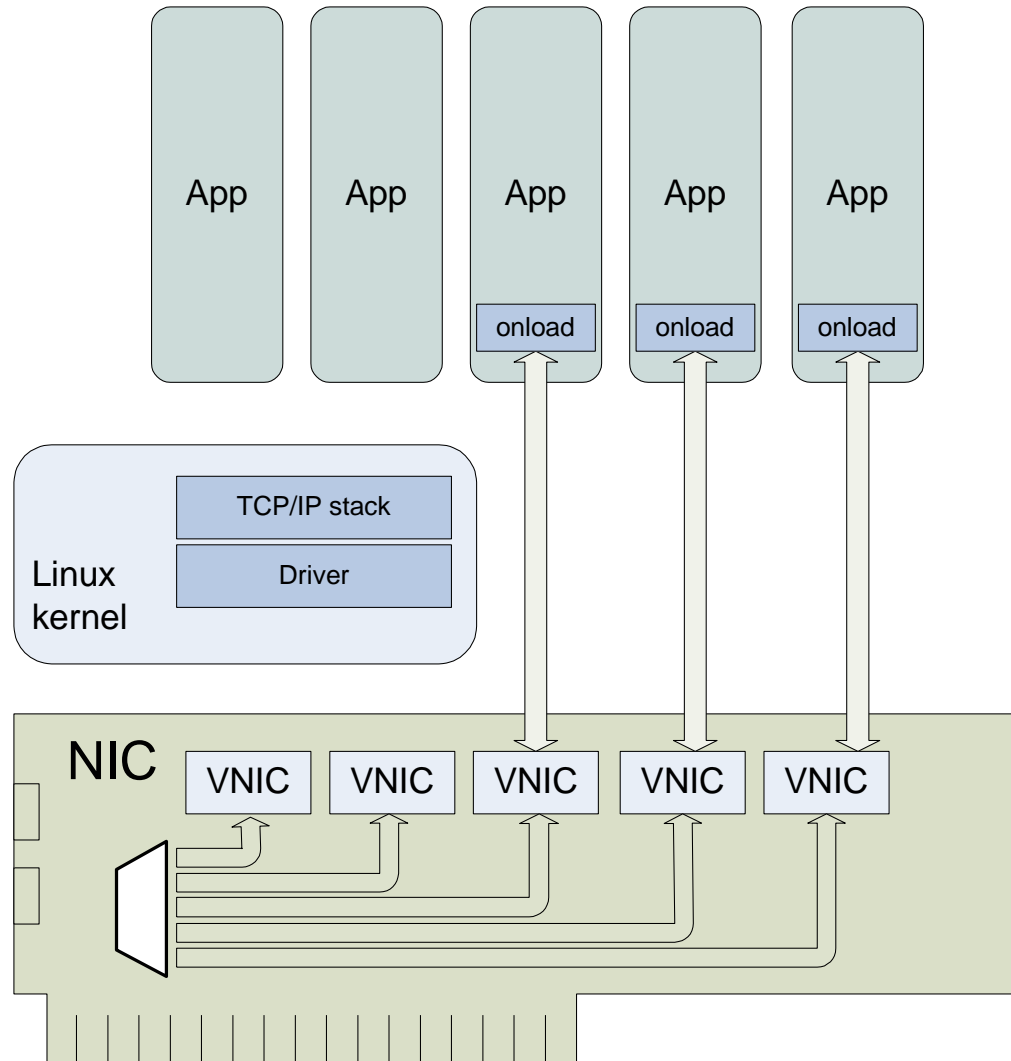


Virtual NICs for application acceleration



VNIC per application

- Used by OpenOnload

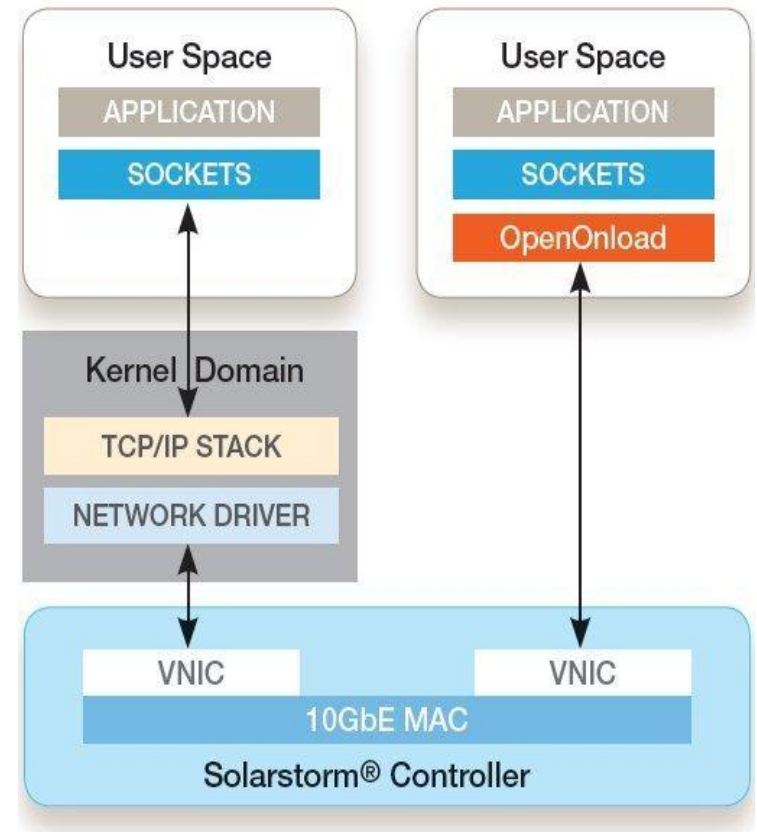


Solarflare OpenOnload® Application Acceleration Middleware



- TCP and UDP Acceleration
 - Kernel bypass
 - App gets direct access to hardware
 - Fewer context switches, copies
 - Benchmarks
 - Reduces latency by 50%
 - Increases message rates 2x to 3x
 - “Real” applications even more benefit
- Compatibility
 - No recompile/application mods
 - Regular Ethernet/IP network
 - Unicast and multicast
 - “Just works”

OpenOnload



- Trivial to deploy and to use
 - \$ myapp arg1 arg2
 - \$ onload myapp arg1 arg2
- Open source
 - GPL
 - Very wide range of kernels/distros supported
 - EnterpriseOnload
 - Bundled subscription and 24/7 support

- Just a library and a kernel module
 - **No** application changes
 - **No** recompile
 - **No** kernel patches
 - **No** protocol changes
- Picks up existing Linux network configuration
 - IP addresses and route table
 - Bonding (aka teaming)
 - VLANs
 - Multicast (IGMP)
 - Kernel settings, e.g. socket buffer sizes

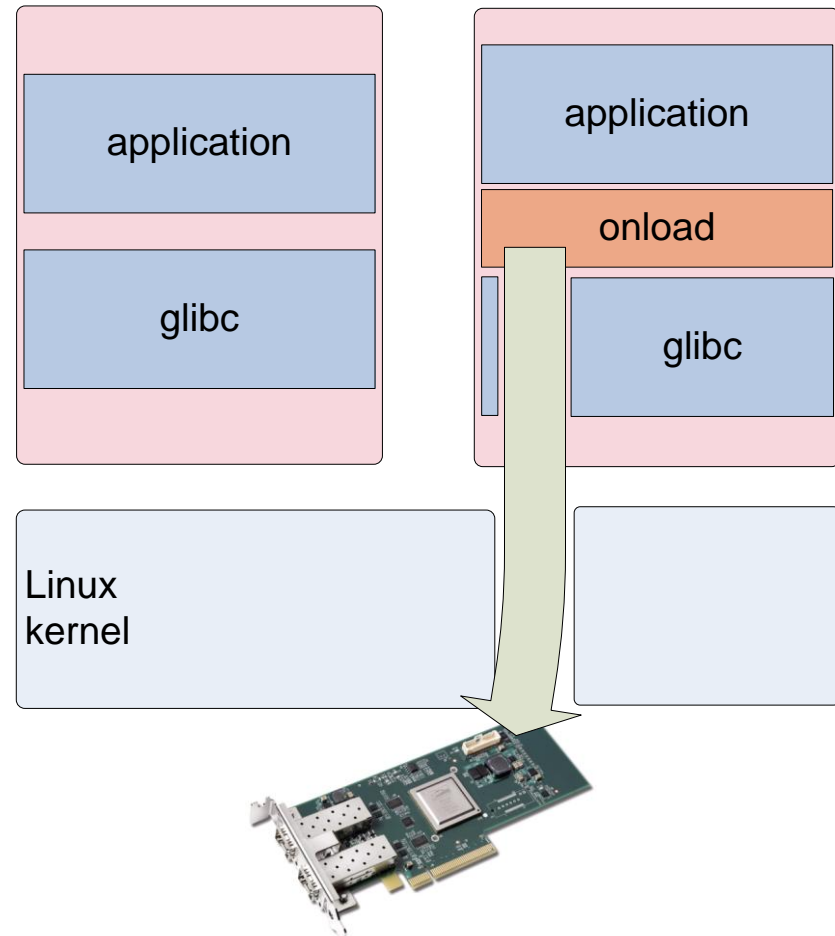
Solarflare OpenOnload® Application Acceleration Middleware



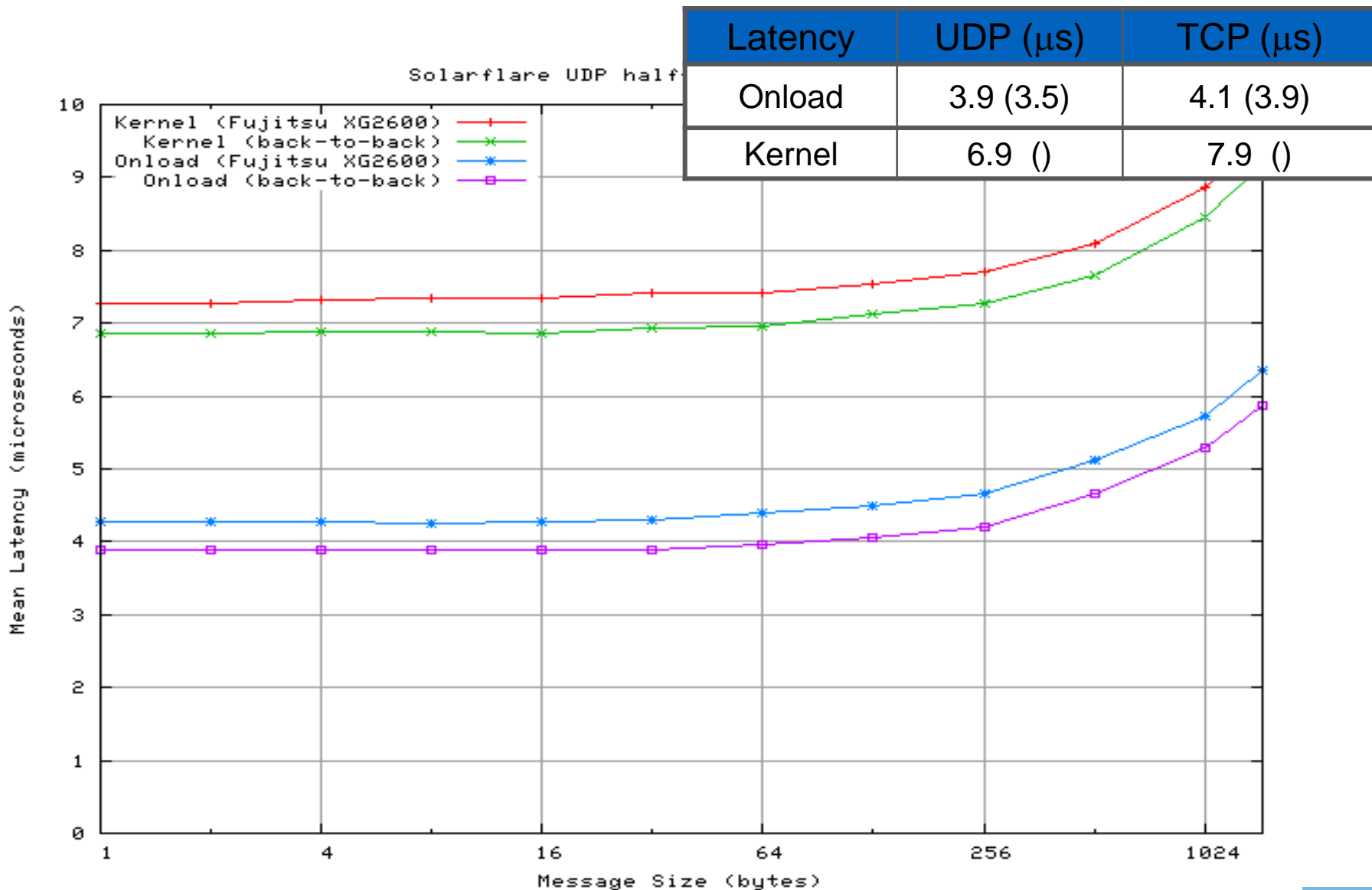
- Mature technology
 - 8+ years of development
- Goal is to make any application “just work”
 - Occasionally we find apps that do not; this is considered a “bug” in Onload and we fix it.
- ~ 100% coverage of API’s
 - epoll, poll, select
 - socket options
 - fork + exec
 - interop with non-Solarflare interfaces
(via “hand-over” to kernel stack)

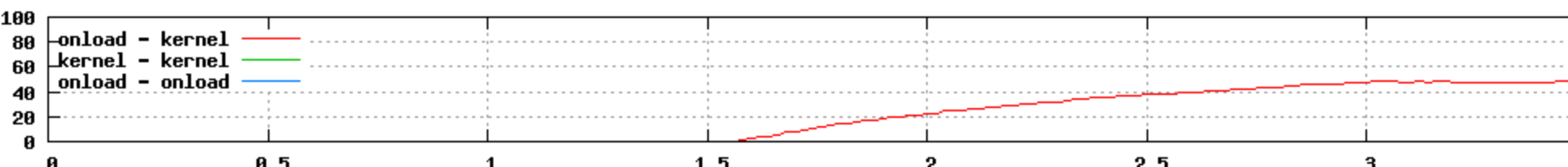
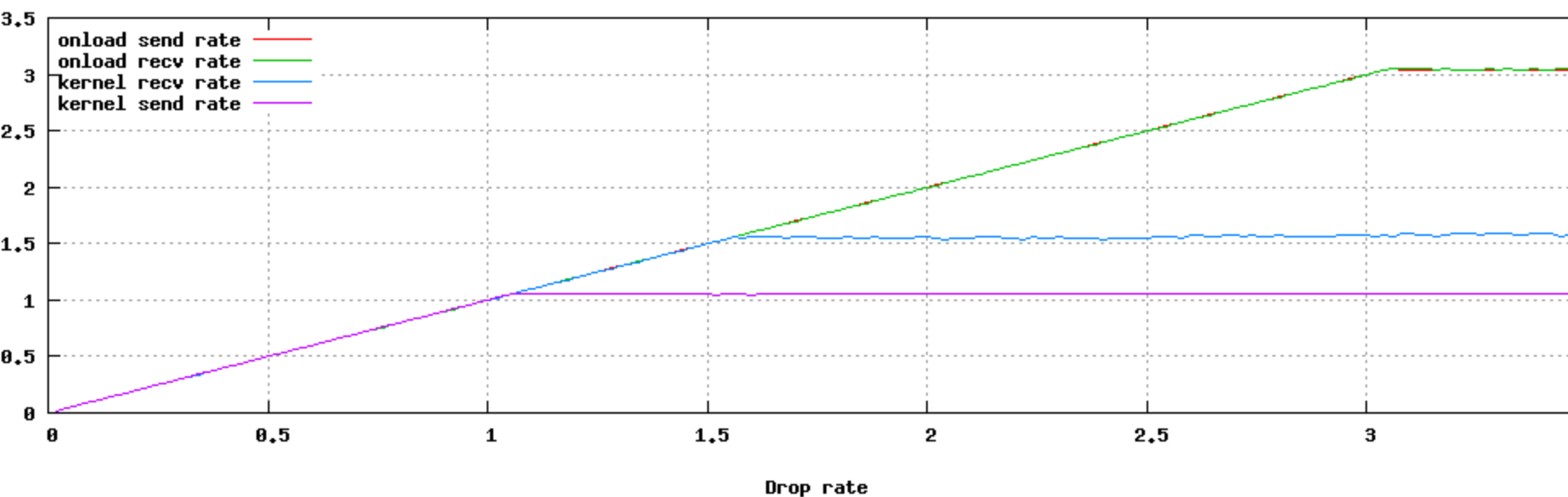
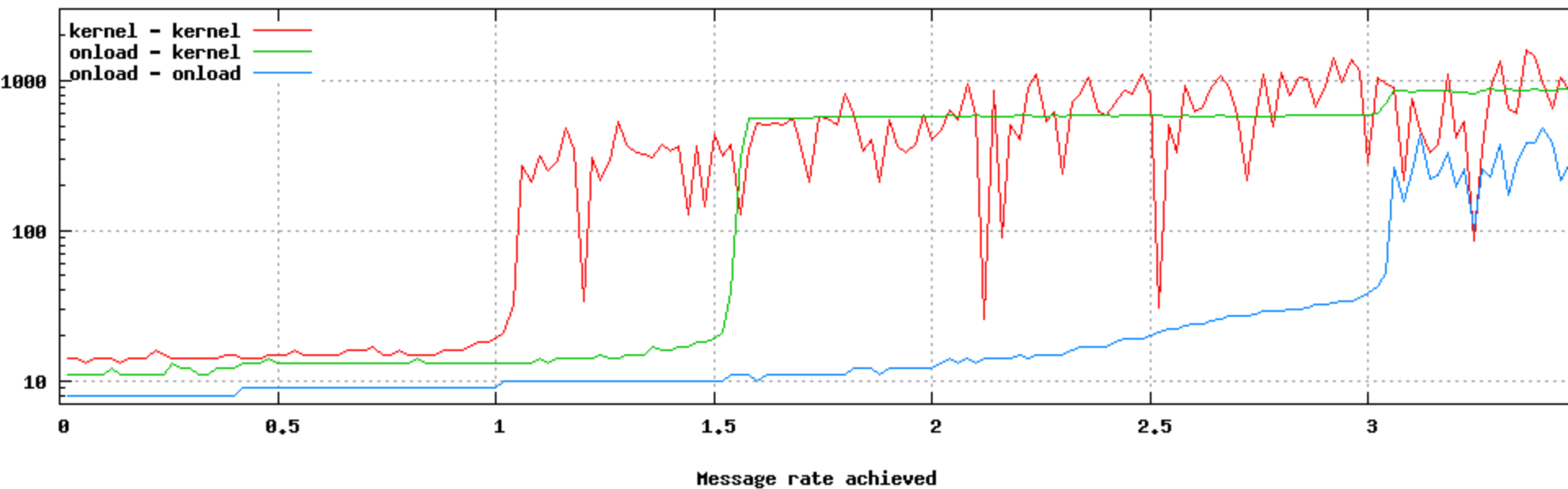
Onload implementation

- Uses `LD_PRELOAD` to intercept libc calls
 - `socket()`
 - `bind()`
 - `accept()`
 - `recvmsg()`
 - `setsockopt()`
 - `poll()`, `select()`, `epoll()`
 - many others...

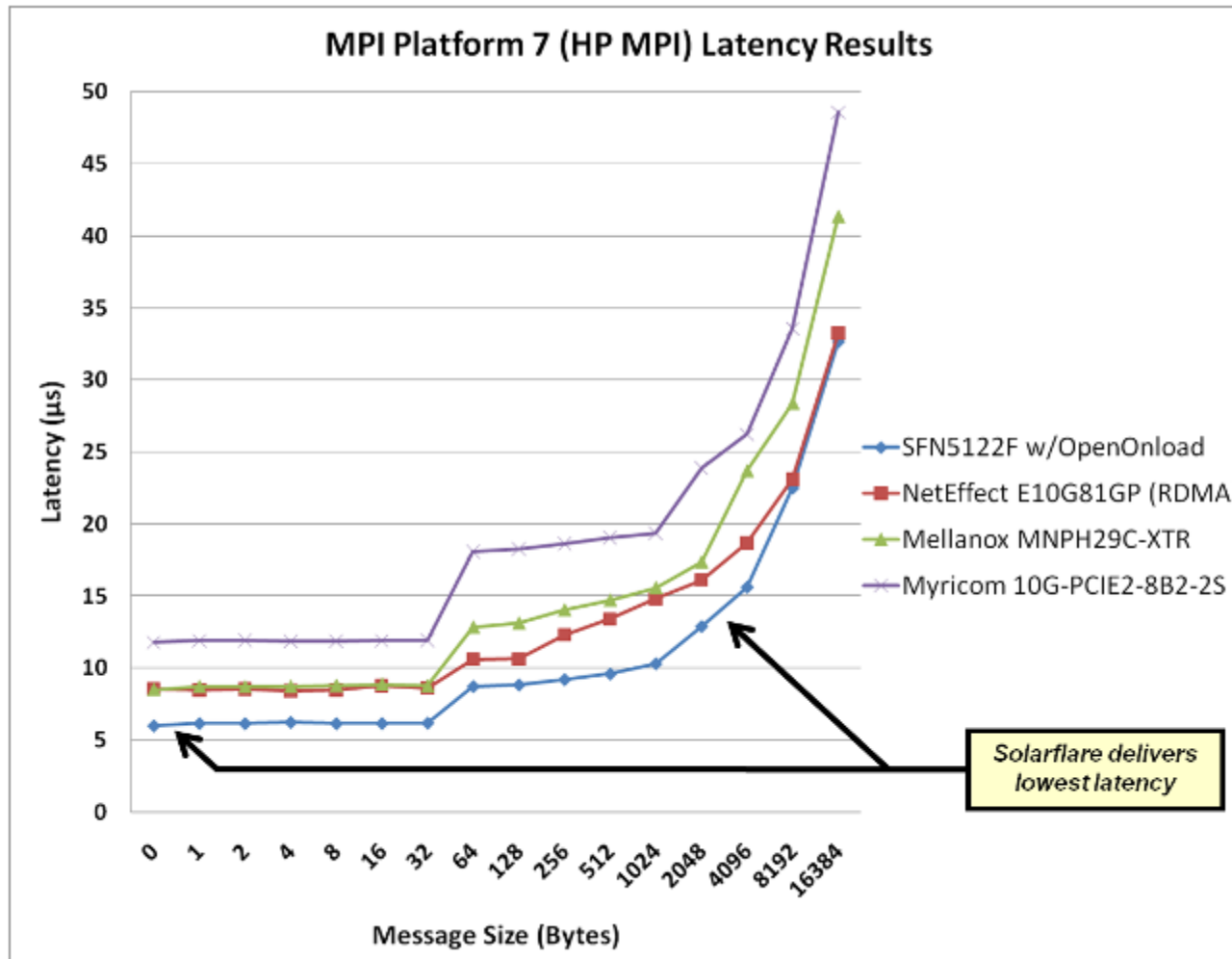


Solarflare for Ultra Low Latency





Low Latency HPC



2x CPU E5620 @ 2.4GHz 8 Cores, 8, 6GB RAM, Intel 5500 chipset
SFP+, Back-to-back, 1500MTU - RHEL 5.5 x86_64 HP MPI MPI + IMB / Pallas

- Solarflare: the gold-standard for high performance Ethernet
- Excellent support for: Linux, Windows, Solaris, VMware, BSD, ...
 - OpenOnload for even higher performance on Linux
- No “B.S.”
 - Customers achieve the numbers we show in presentations like this!
 - See the “Low Latency Quick Start Guide” at <http://download.solarflare.com/>