

Open Source Router @ 10g

Talk at KTH/Kista 2007-04-29

Robert Olsson/Uppsala University
Olof Hagsand/KTH

Open Source Router @ 10g

Motivation

Breakthrough

Multi core CPU

Buses

NIC with multiple queue's

Operating systems
prime time for

Commercial interest
www.vyatta.com

Open Source Router @ 10g

Team

Bengt Görden/KTH

Olof Hagsand/KTH

Robert Olsson/Uppsala University

Challenges

Packet budget, bandwidth to reach 10g

PCIe (PCI-Express)

Open Source Router @ 10g

Hardware selection (More later)

Several vendors have boards

We tested some boards.

Good connections to Intel who
testing

Neterion (s2io) has boards

SUN's has new interesting boards.

Open Source Router @ 10g

Software selection

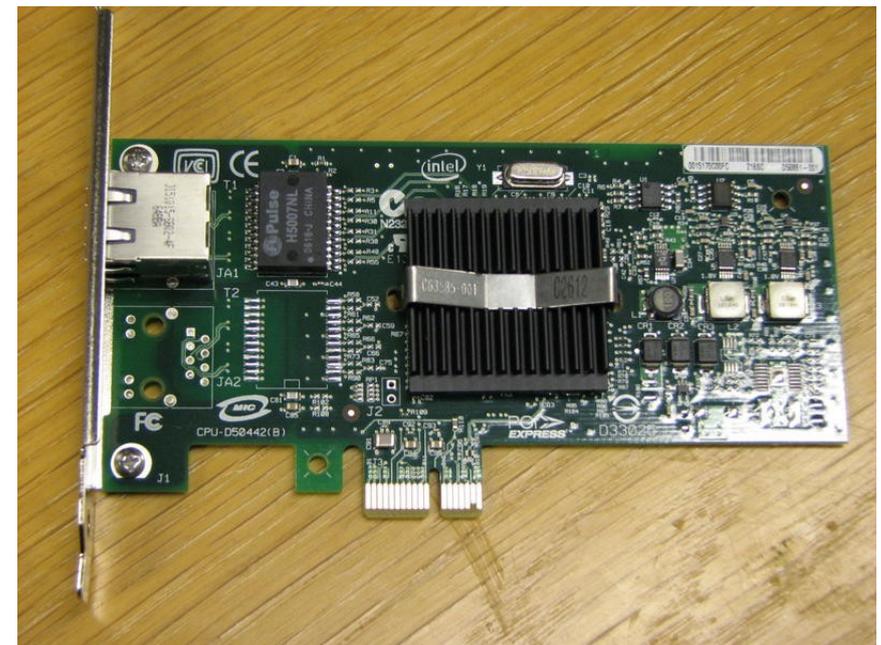
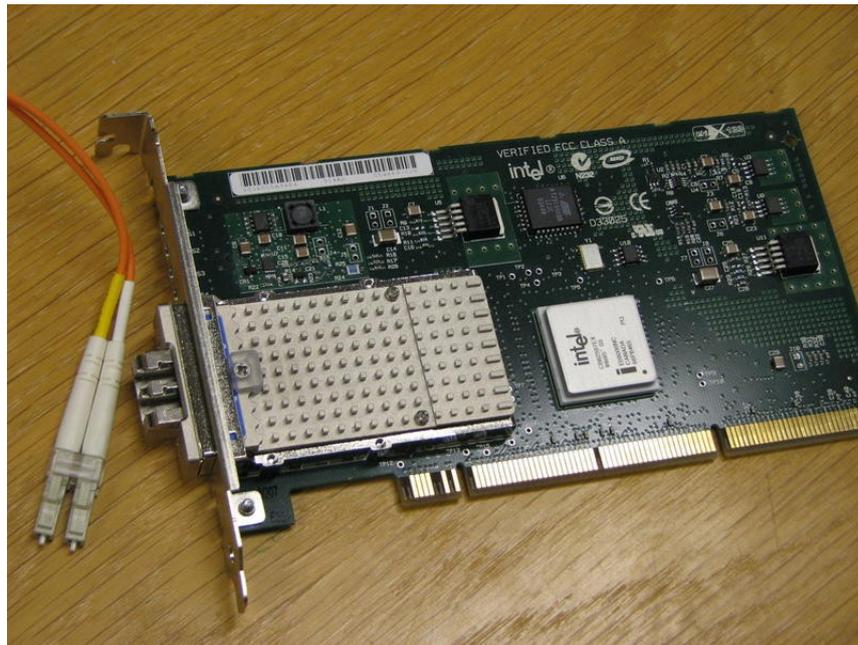
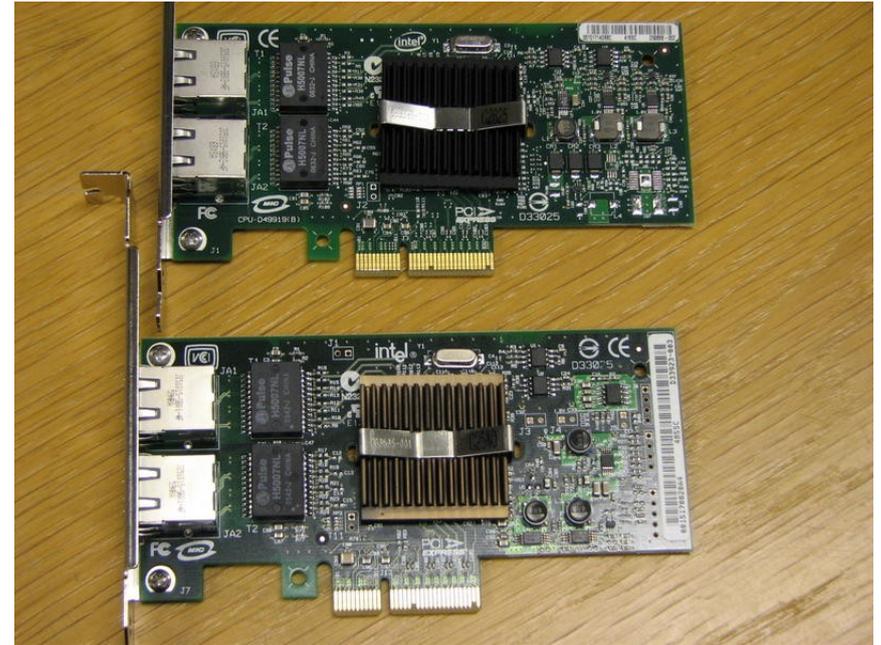
Linux/bifrost

tuning for ip-forwarding
device drivers are crucial

Open Source Router @ 10g

Lab work. Setup etc.

Currently Intel NIC's

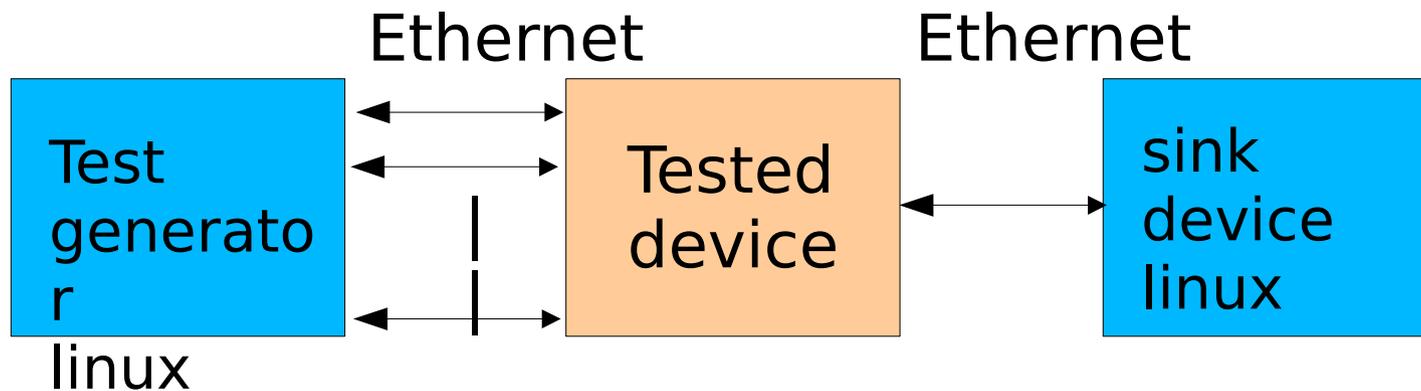


Not or were blessed...



Flexible netlab at Uppsala University

El cheapo-- High customizable -- We write code :-)



- * Raw packet performance
- * TCP
- * Timing
- * Variants

Lab



High-end Hardware

Bifrost components 2008-01-03

=====

This represents equipment which have "passed" our tests at some point of time.

NOTE. This implies no guarantee whatsoever.

Hi-End Opteron system:

TYAN Thunder n5550W (S2915-E)	1 *
256 MB Reg ECC PC3200 (400 MHz)	4 *
Opteron dual core 2222 processor	2 *
4U Chassi	1 *
USB memory stick	1 *

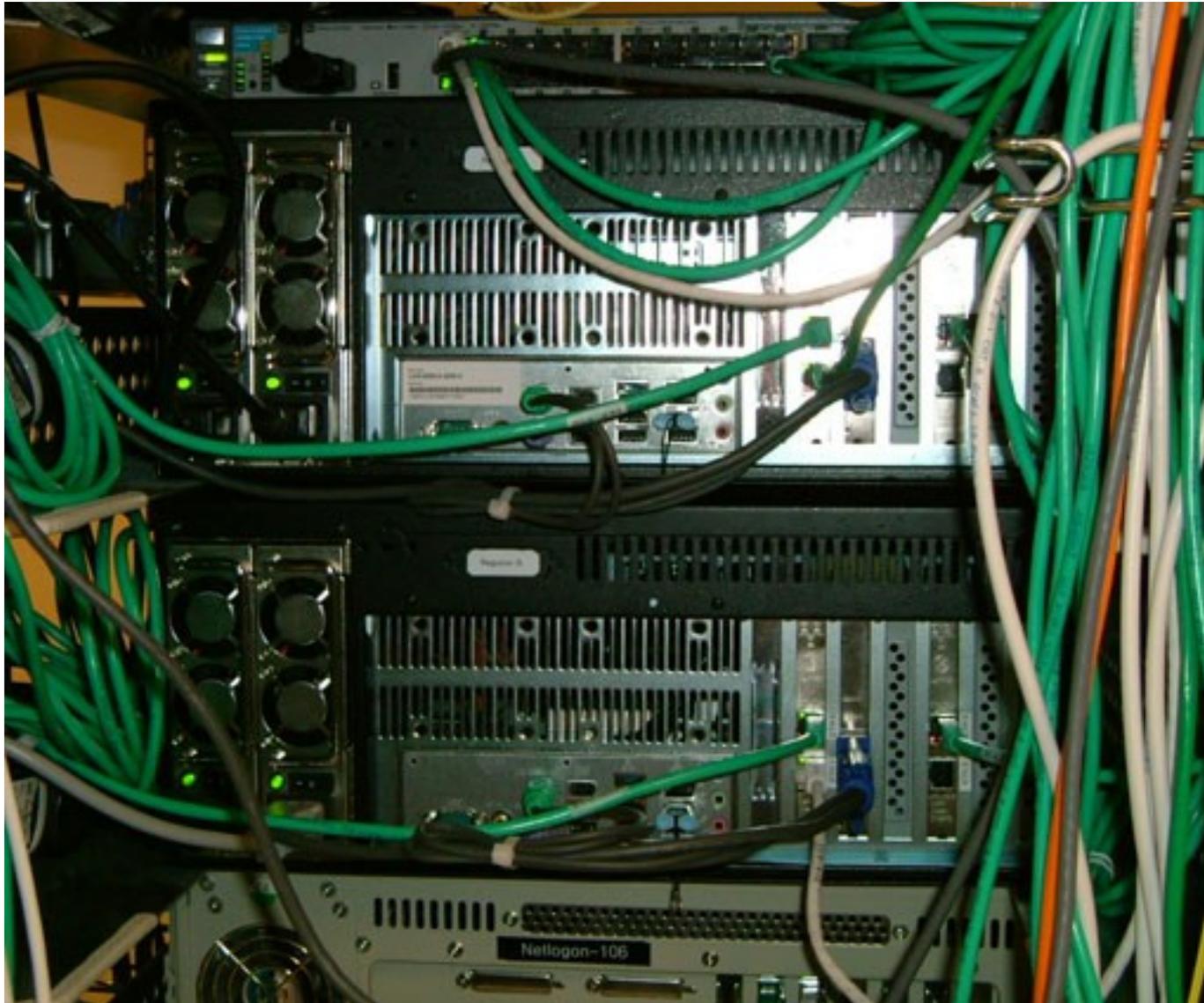
Option:

Redundant power supply

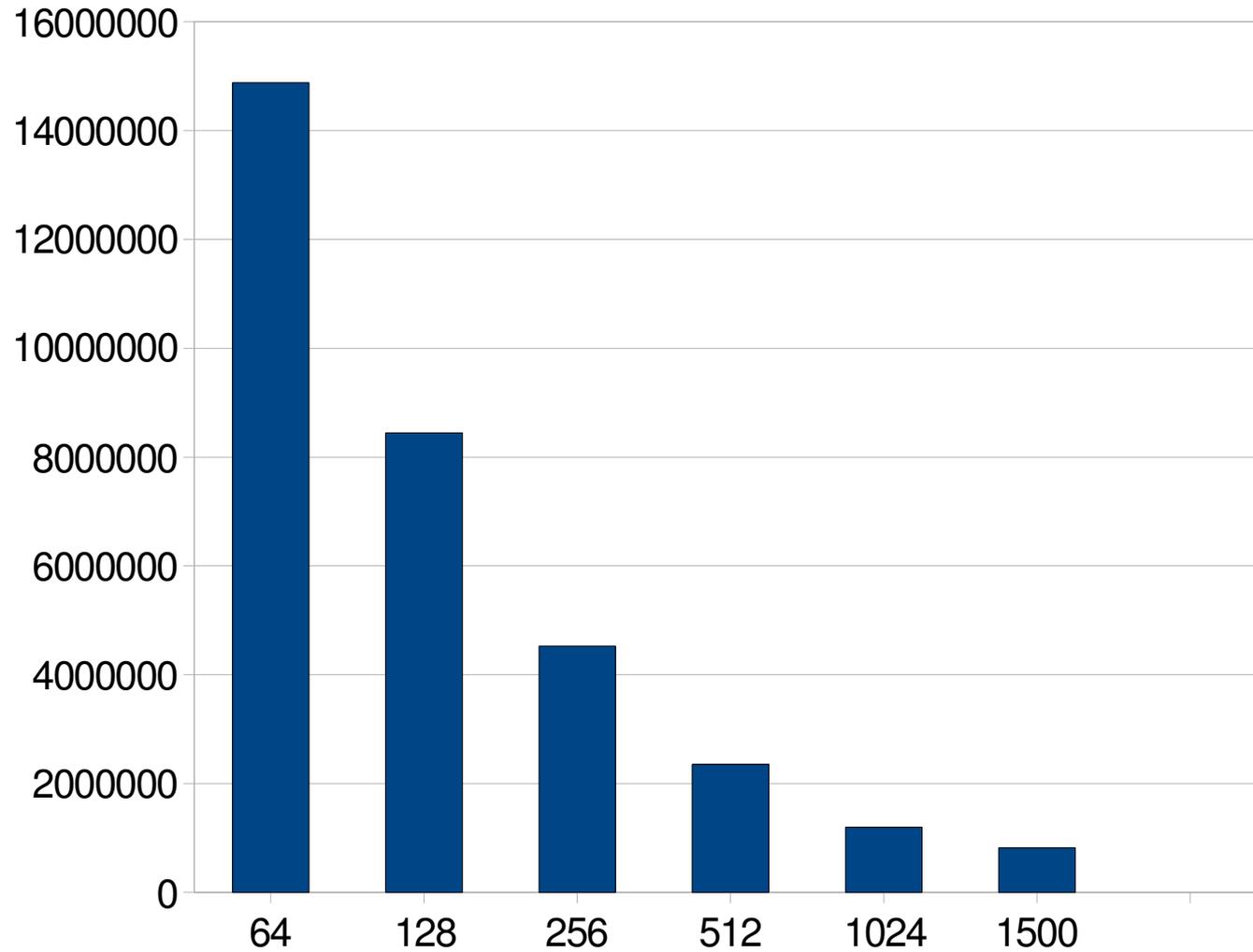
EMACS MRG-6500P Redundant PSU 2x500W

Intel Corporation 82571EB GIGE based Network cards

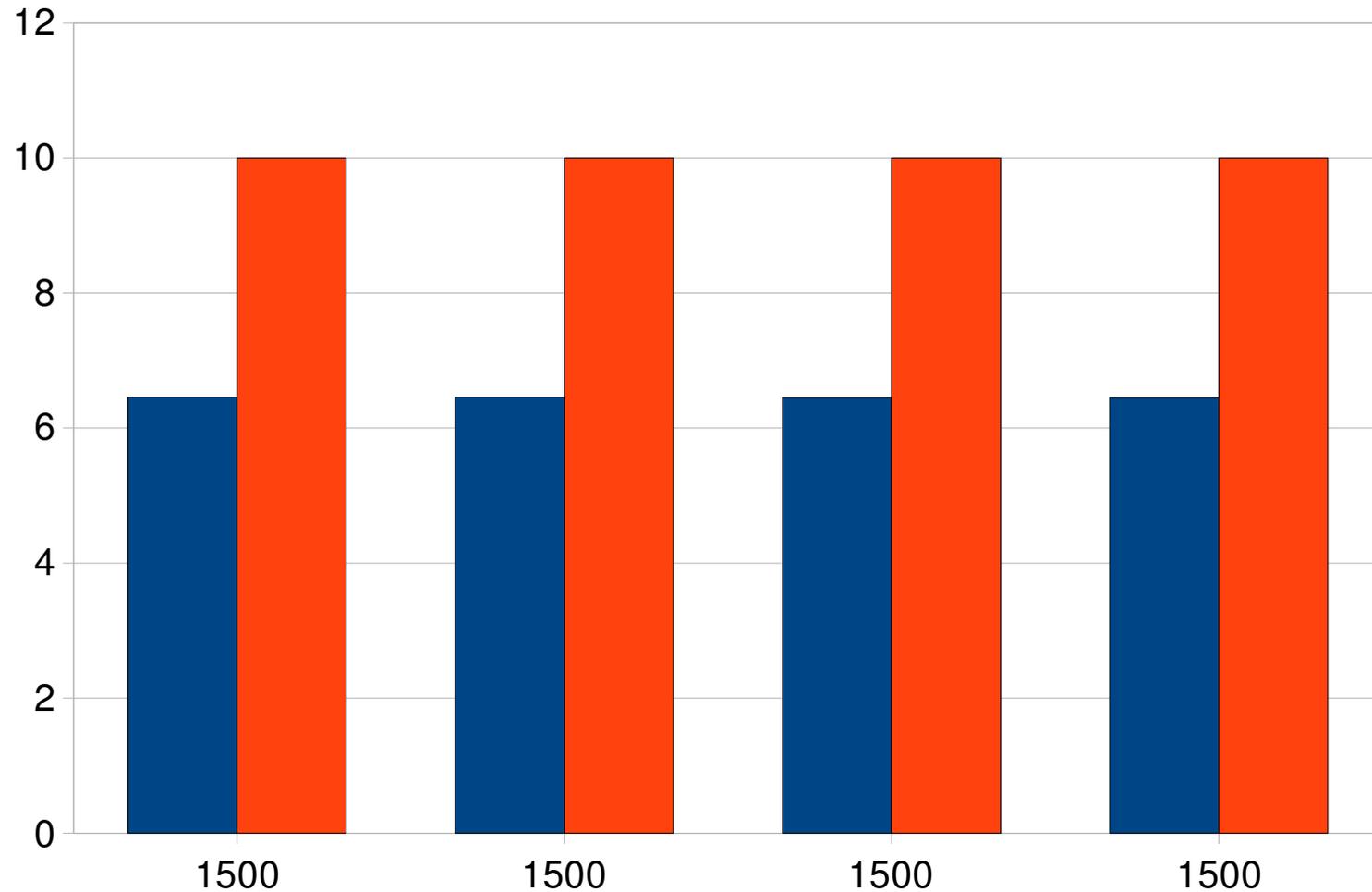
HW in use @ UpUnet-S



Theoretical pps (packet per second) numbers to reach wire speed at 10 Gbit/s



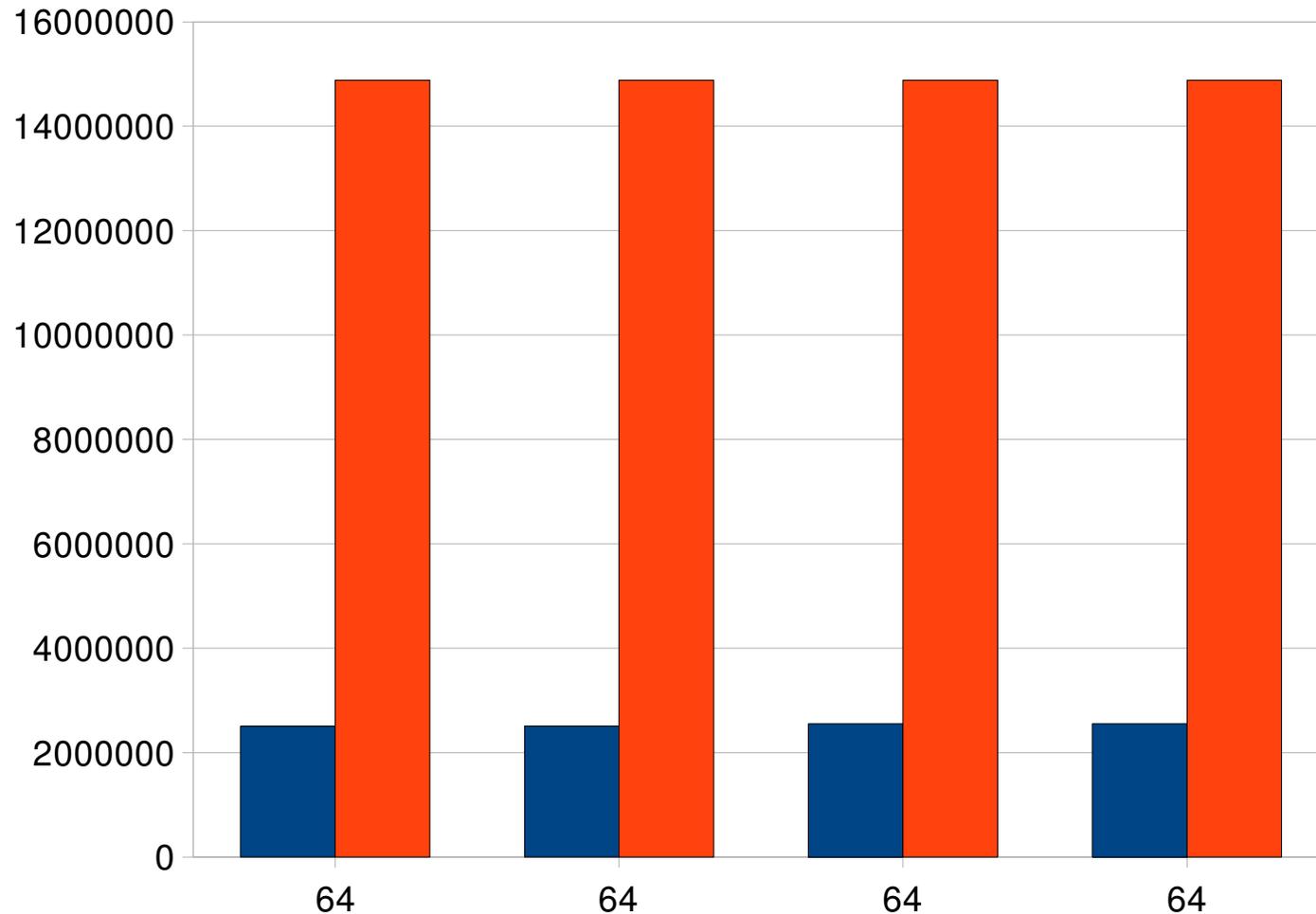
TX performance with 1500 byte pkts Sending on 4 interfaces simultaneously. We reach 25.8 Gbit/s. Theoretical 4*10 Gbit/s



Distribution is fair and we reach 25.8 Gbit/s...

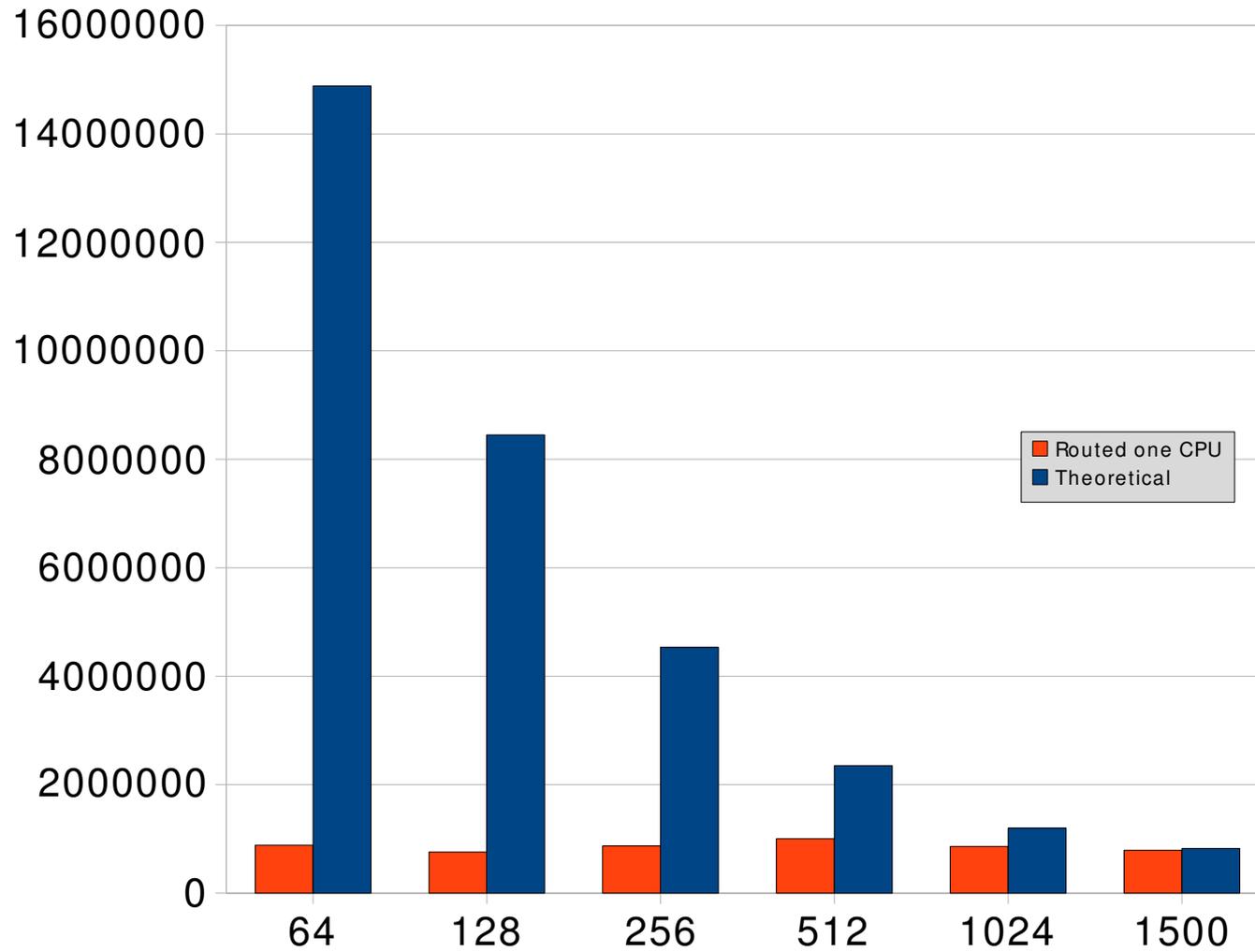
Exercise bus, bandwidth etc.

TX performance with 64 byte pkts Sending on 4 interfaces simultaneously. We reach 10 Mpps . (Packets Per Second)



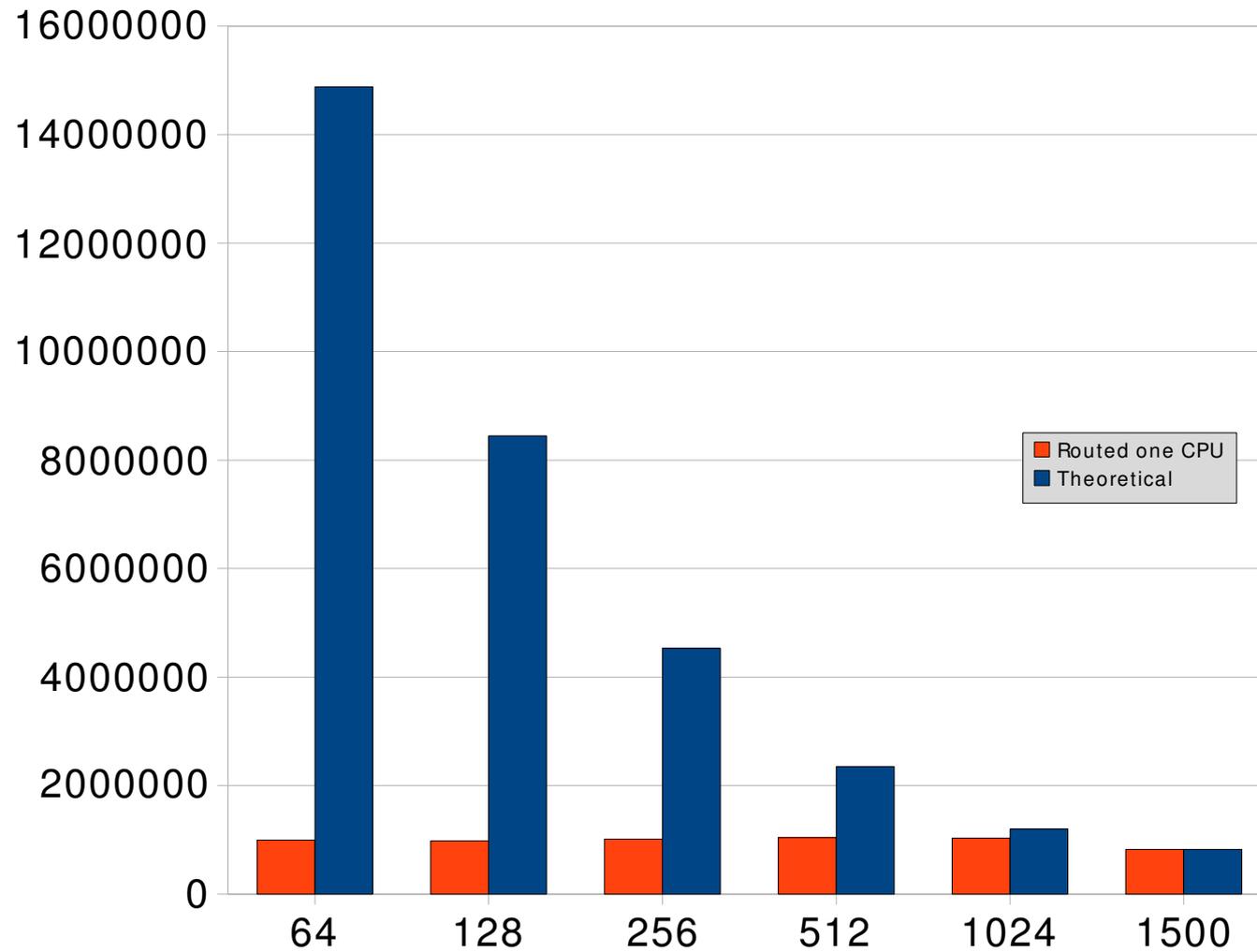
Distribution is fair and we reach about 10 Mpps in total. Exercise for latency etc.

Routing performance using one CPU vs Theoretical numbers to reach wire speed at 10 Gbit/s



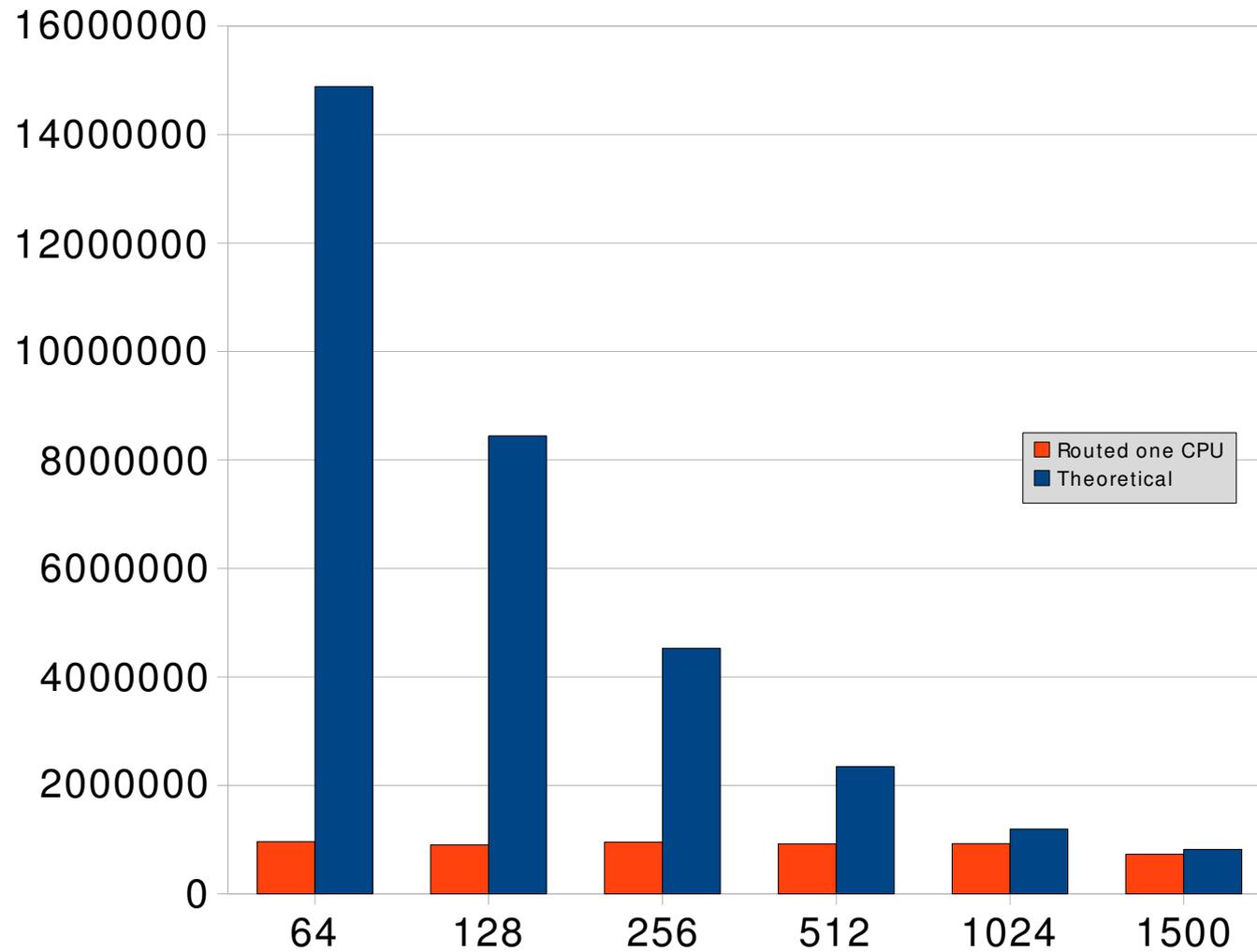
Single flow. No modules. IN 1-PORT -> OUT 2-Port

Routing performance using one CPU vs Theoretical numbers to reach wire speed at 10 Gbit/s



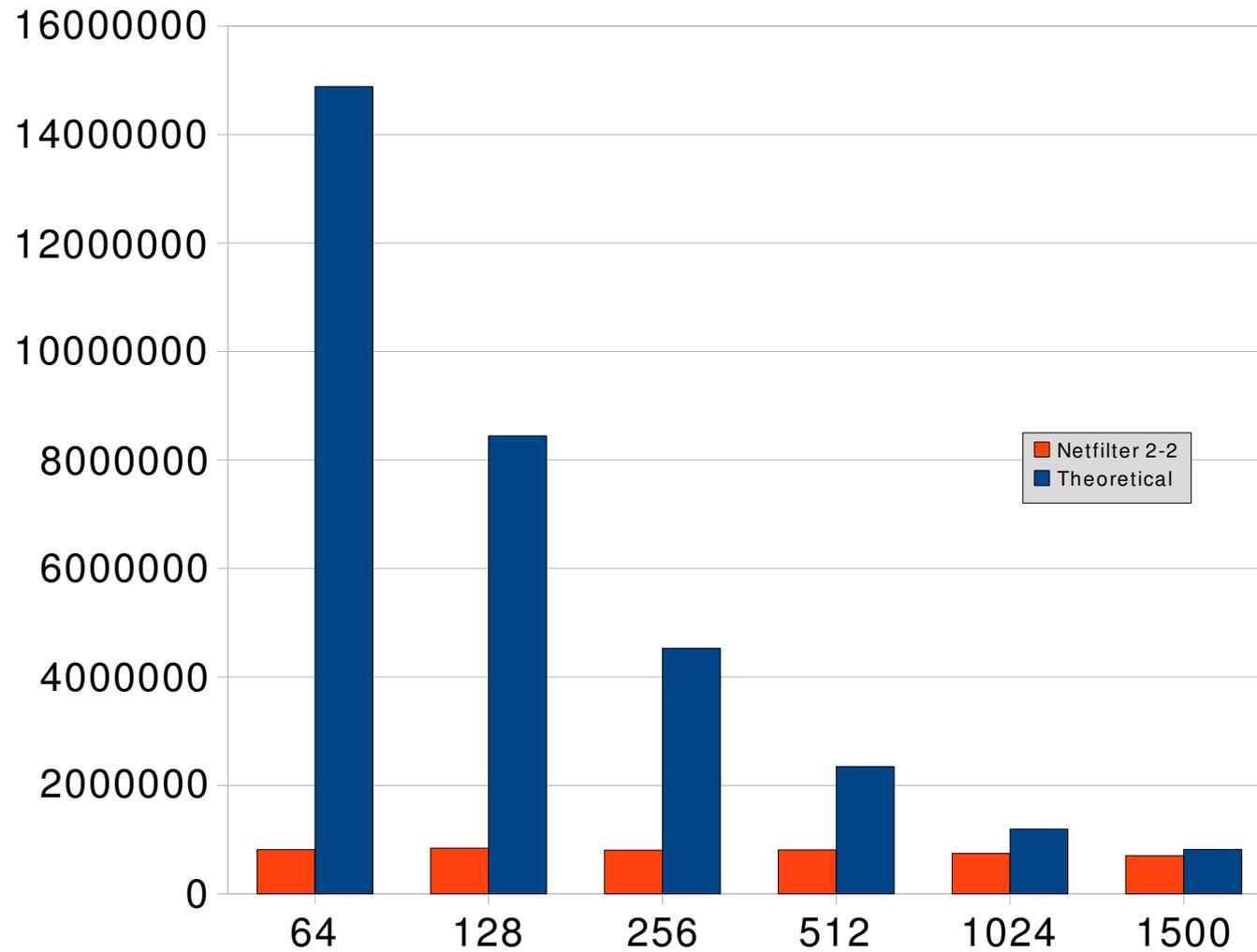
Single flow. No modules. IN 2-PORT -> OUT dummy0

Routing performance using one CPU vs Theoretical numbers to reach wire speed at 10 Gbit/s



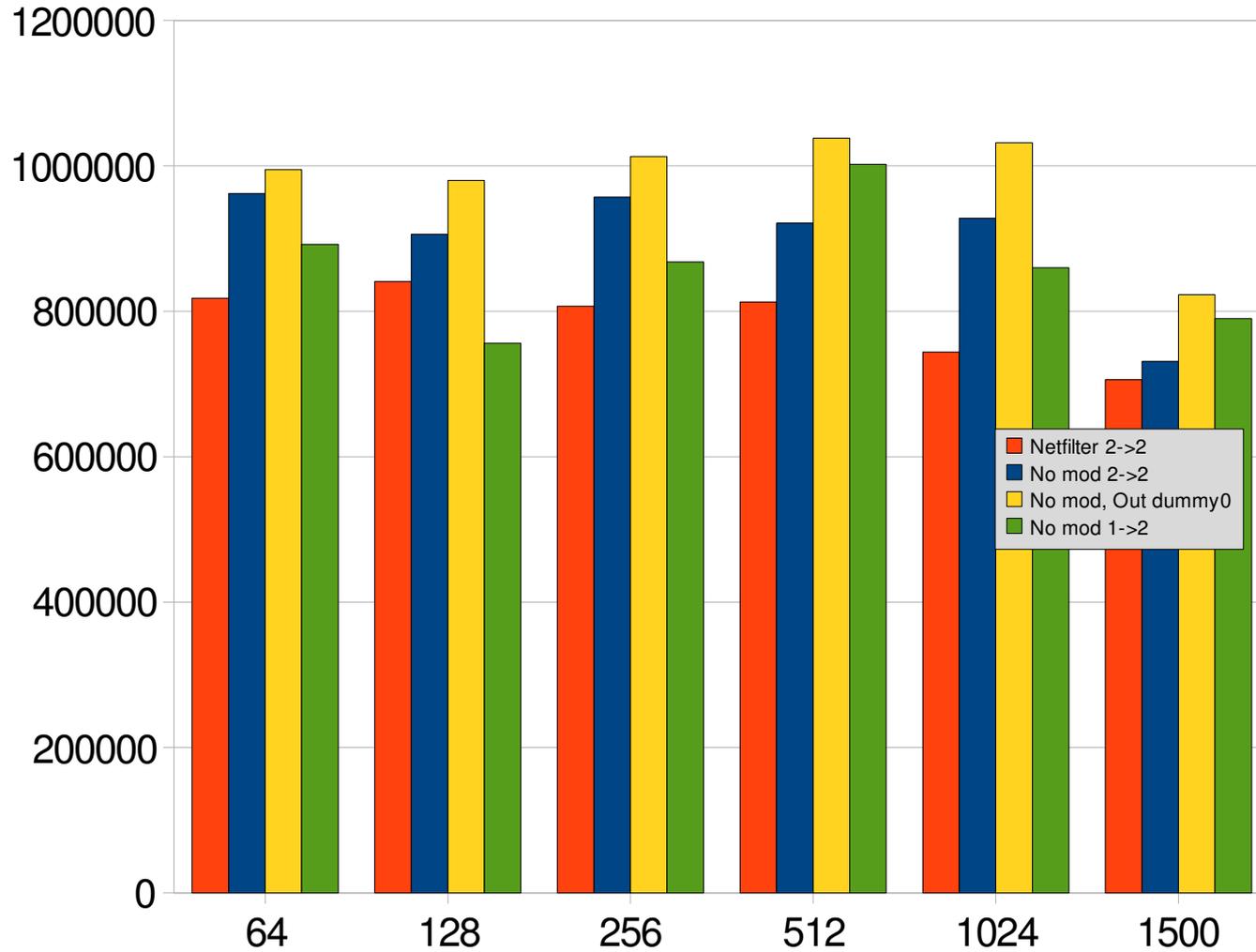
Single flow. No modules. IN 2-PORT -> OUT 2-PORT (Same NIC)

Routing performance using one CPU vs Theoretical numbers to reach wire speed at 10 Gbit/s



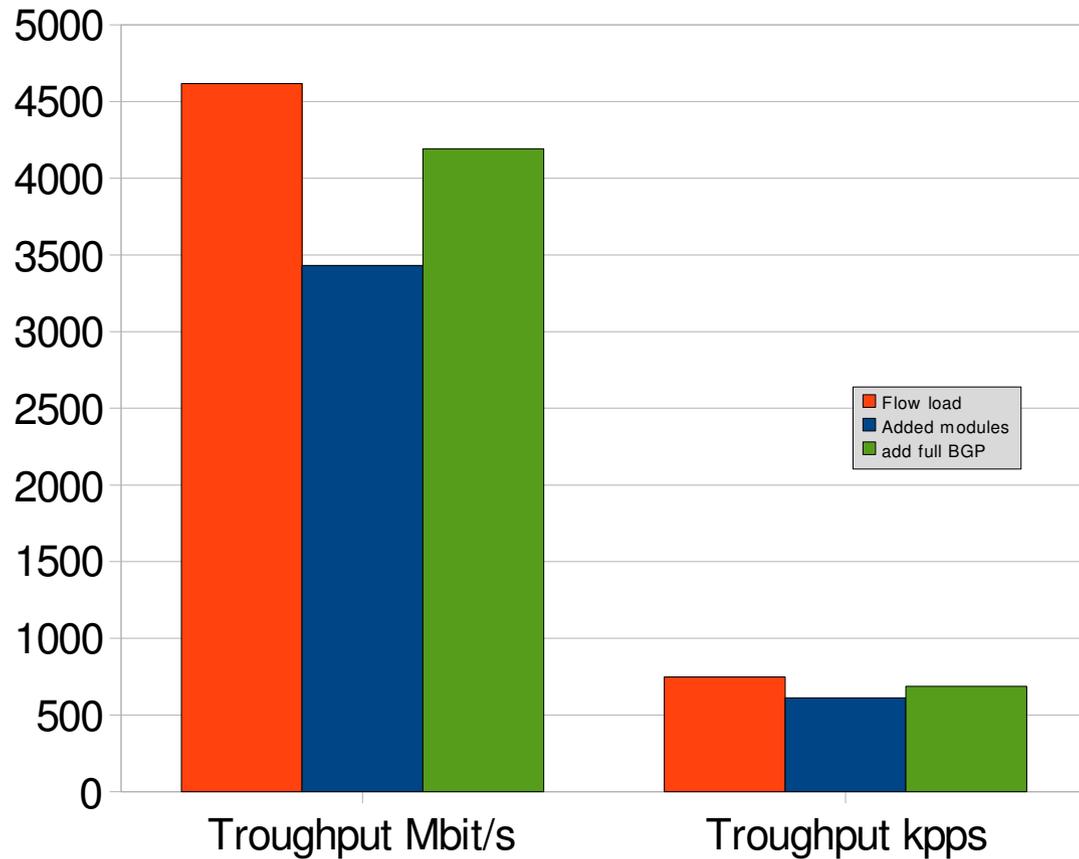
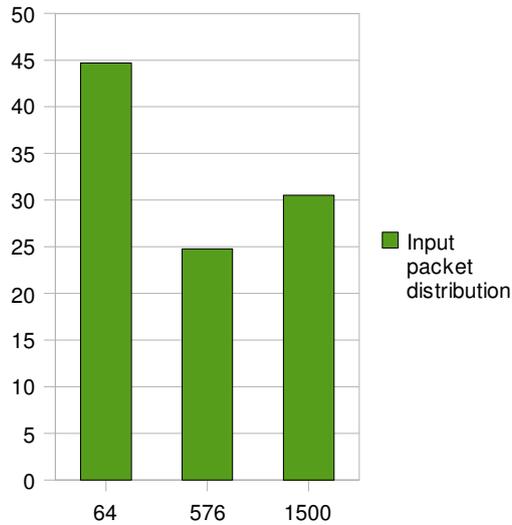
Single flow. Netfilter modules. IN 2-PORT -> OUT 2-PORT (Same NIC)

Routing performance comparison using one CPU different bus setups and modules



Non-multiqueue driver. Packet budget approx 900 kpps using one CPU
So we route close to 10g wire speed with large packet sizes.

Routing performance comparison using one CPU
flow load (forcing hash and fib lookups) and
input 64, 512, 1500 bytes according to graph



Non-multiqueue driver.

What is the Multi-queue stuff?

Part of virtualization

Ability to share load among different CPU's

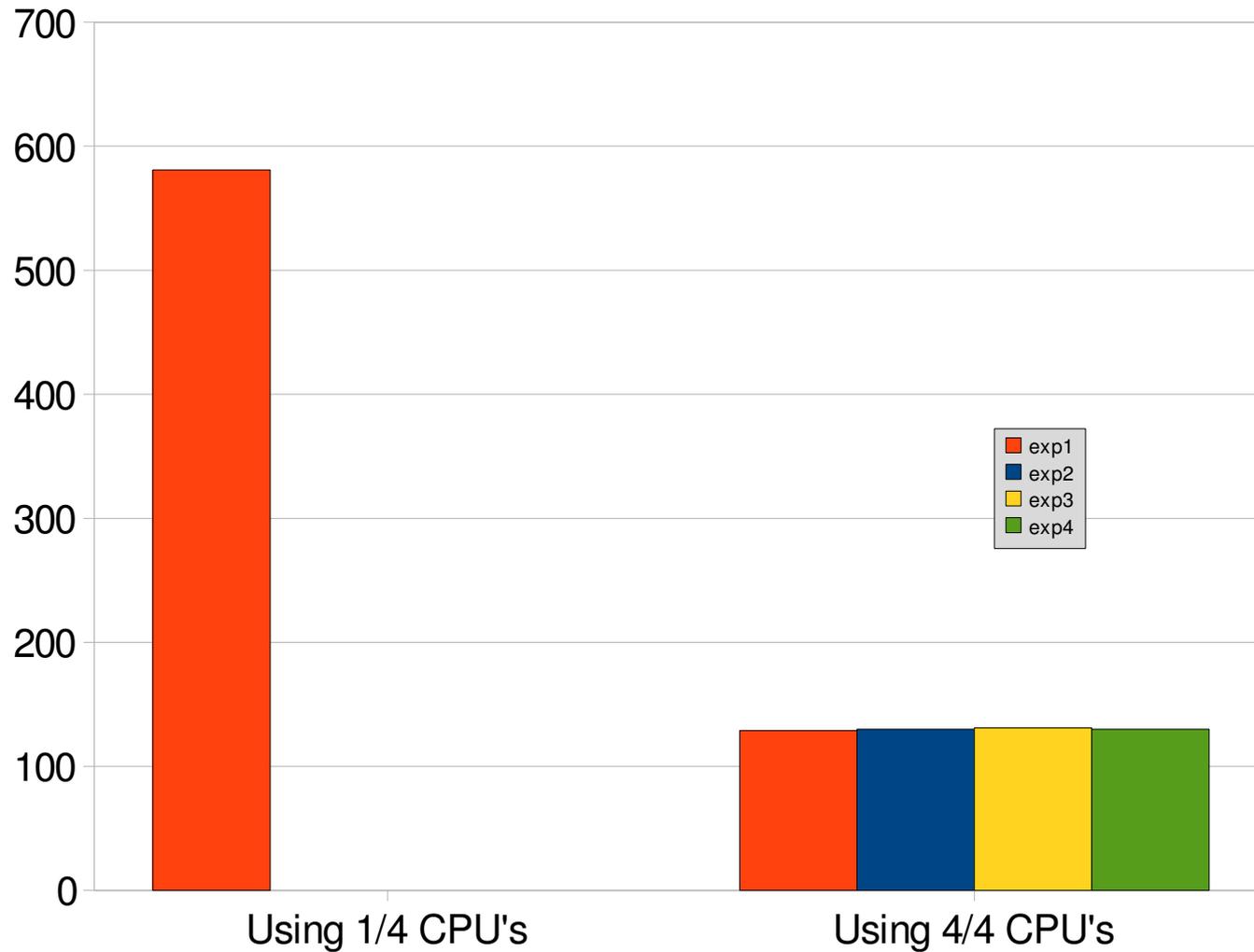
Ability classify and control network load at very high speed

Opens many new possibilities

Needs hardware support by Interface boards/chips

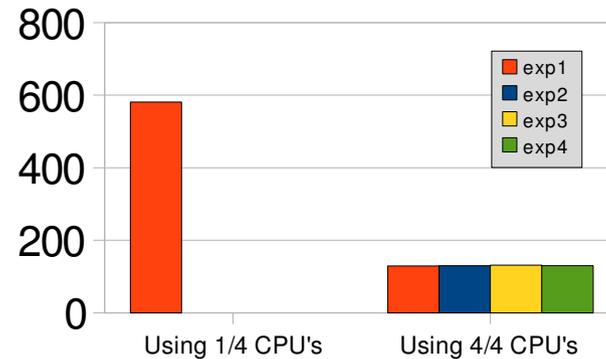
Needs software support by Operating System

Routing performance comparison using one CPU
flow load (forcing hash and fib lookups) and
pktsize 64, 512, 1500 bytes



MultiQ vs no MultiQ [2048@30](#) with 64 byte. Pkts
Less performance with 4 CPU's!!

Routing performance comparison using one CPU flow load (forcing hash and fib lookups) and pktsize 64, 512, 1500 bytes



All very exciting we run per-CPU parallel until we hit
TX. `dev_queue_xmit`, `__qdisc_run`

Virtual TX is needed

More and exciting work needed.... ;)

Open Source Router @ 10g

Questions Please....

A new network symbol has been seen...

The Penguin Has Landed

