IPv6
Faster, Further, FreeBSD

BSDCan 2012
Bjoern A. Zeeb (bz@zabbadoz.com)

TOC

- Multi-FIB IPv6 changes
- Performance improvements
- Security
- Compliance
- No-inet ("IPv6-only")
- Other “requests”
- World IPv6 Launch Day?

Multi-FIB IPv6

- IPv4 support was committed in 2008.
- Feature Parity!
- Sponsored by Cisco Systems, Inc.

Multi-FIB IPv6

- Up to 16 distinct "routing tables" in kernel (FIB).
- setfib(2) and setfib(8) to set it per process.
- ifconfig, ipfw, and pf (prepared) to set per packet.
- setsockopt(..SO_SETFIB..) to set per socket.

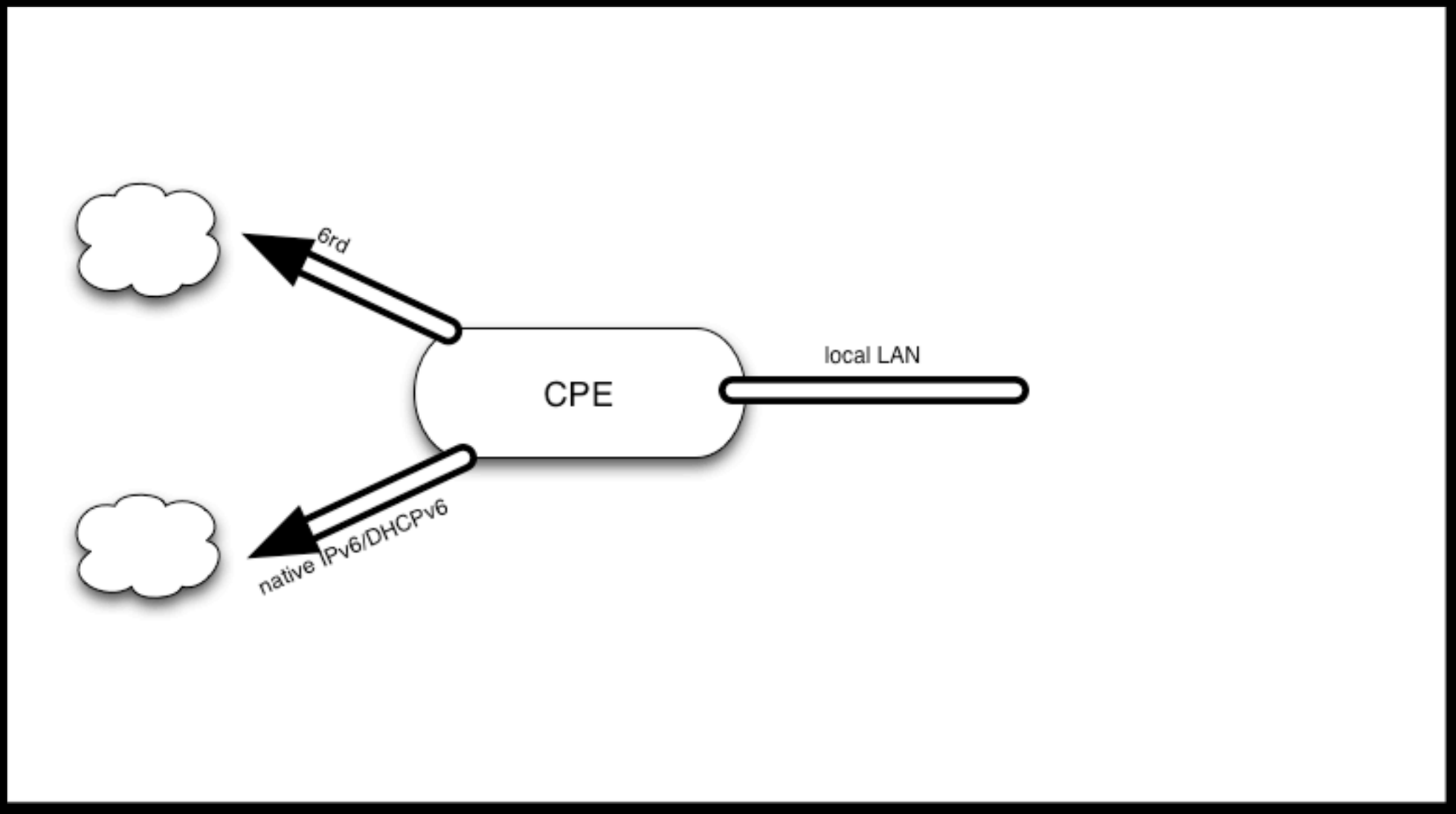
Multi-FIB IPv6

- netstat extended
- pfctl adjustments
- netcat had support
- regression tests

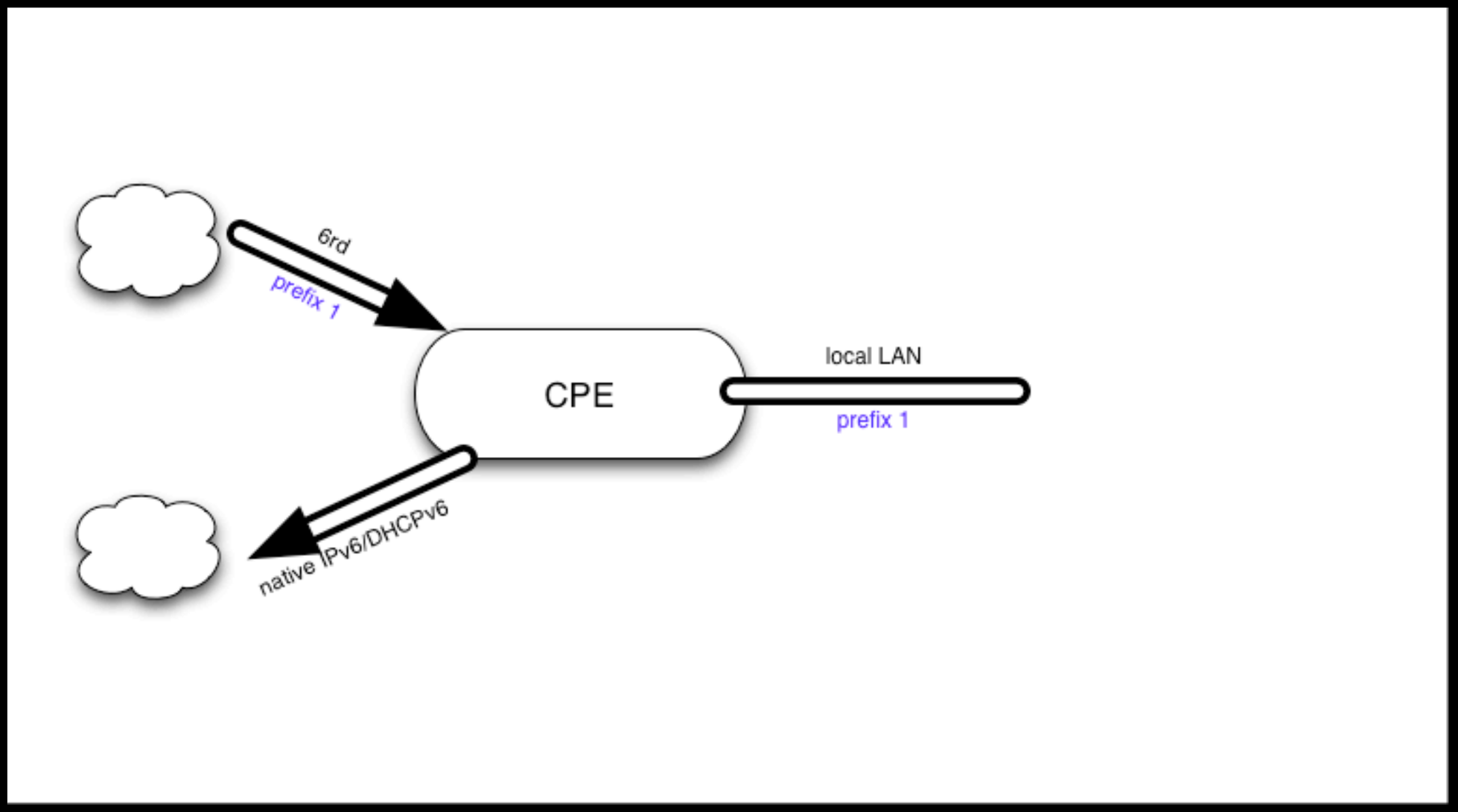
Multi-FIB IPv6

- Use with whatever you did in IPv4.
- Use with “IP-jails”.
- Useful in multi-exit gateway/prefix setups.
- Possibly useful for IPv6 CPEs (draft-townsley-troan-ipv6-ce-transitioning).

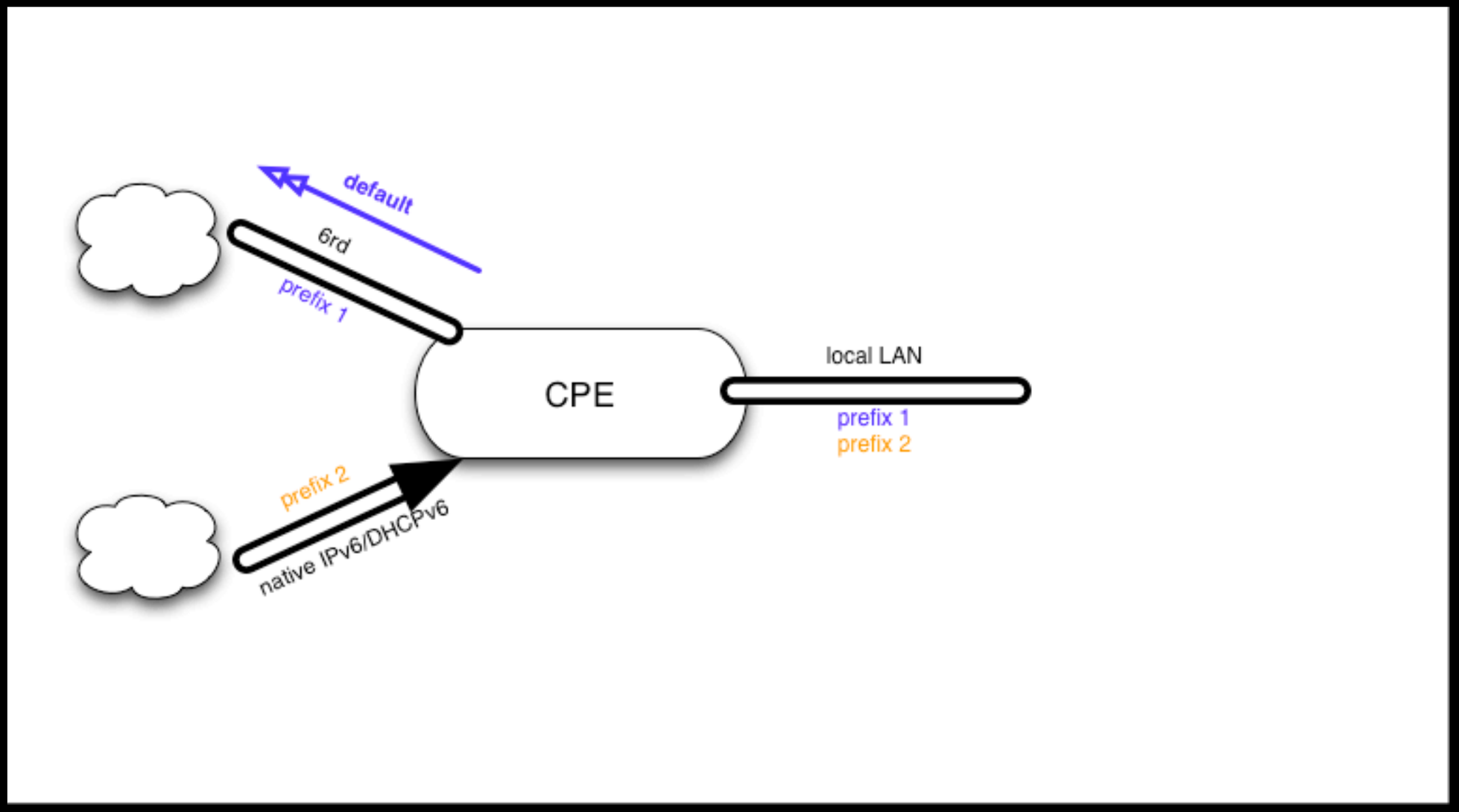
Multi-FIB IPv6, CPE



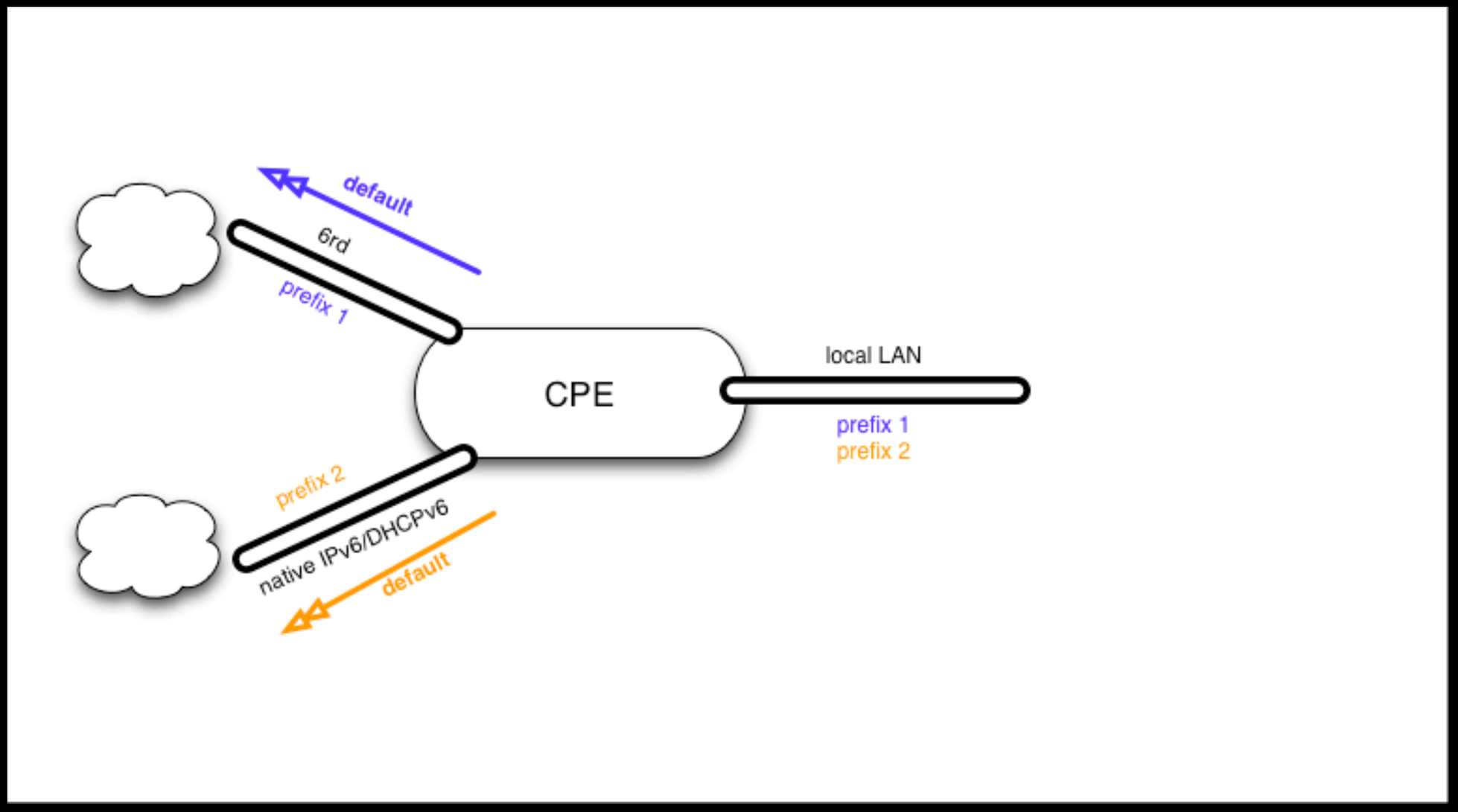
Multi-FIB IPv6, CPE



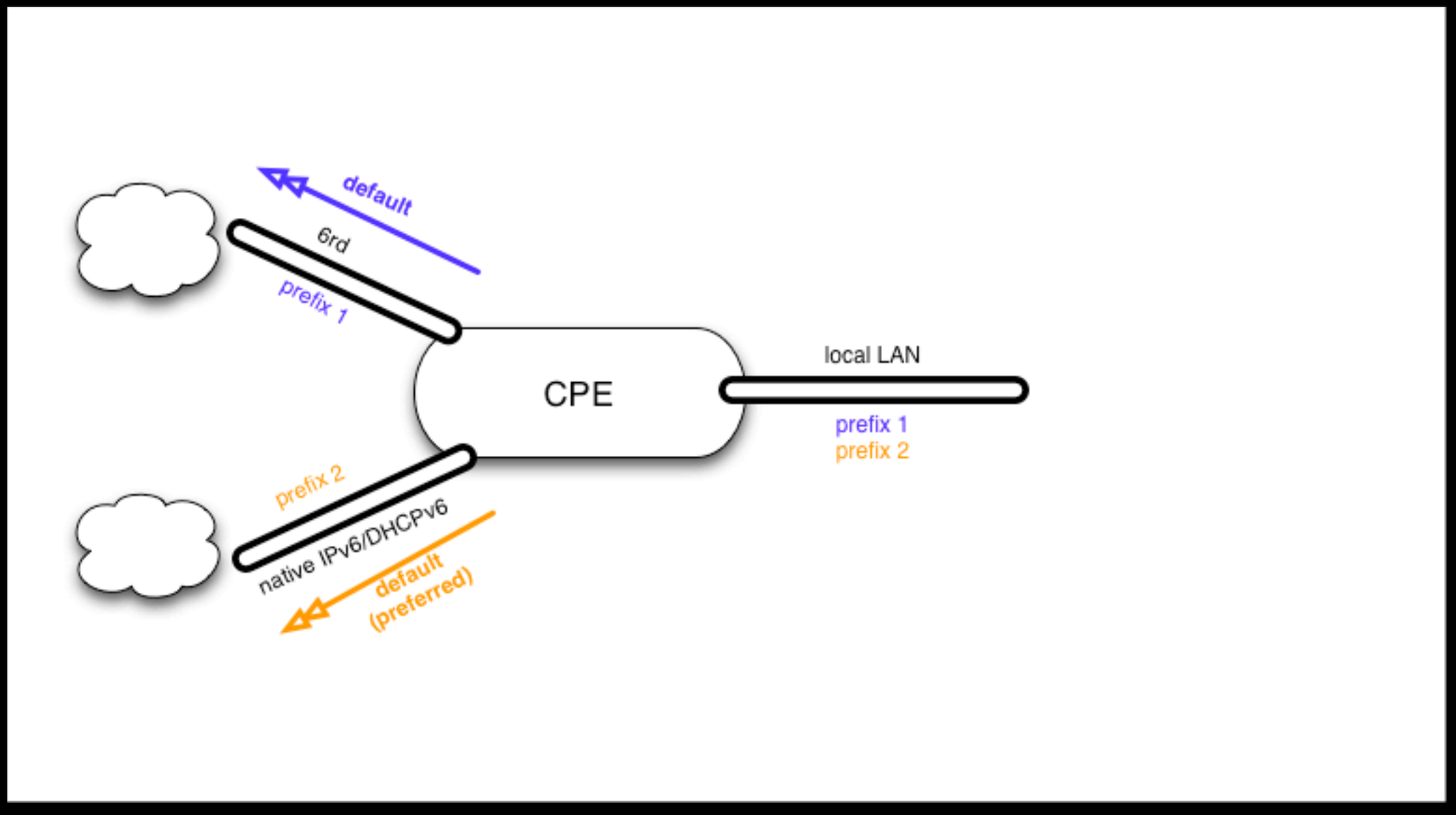
Multi-FIB IPv6, CPE



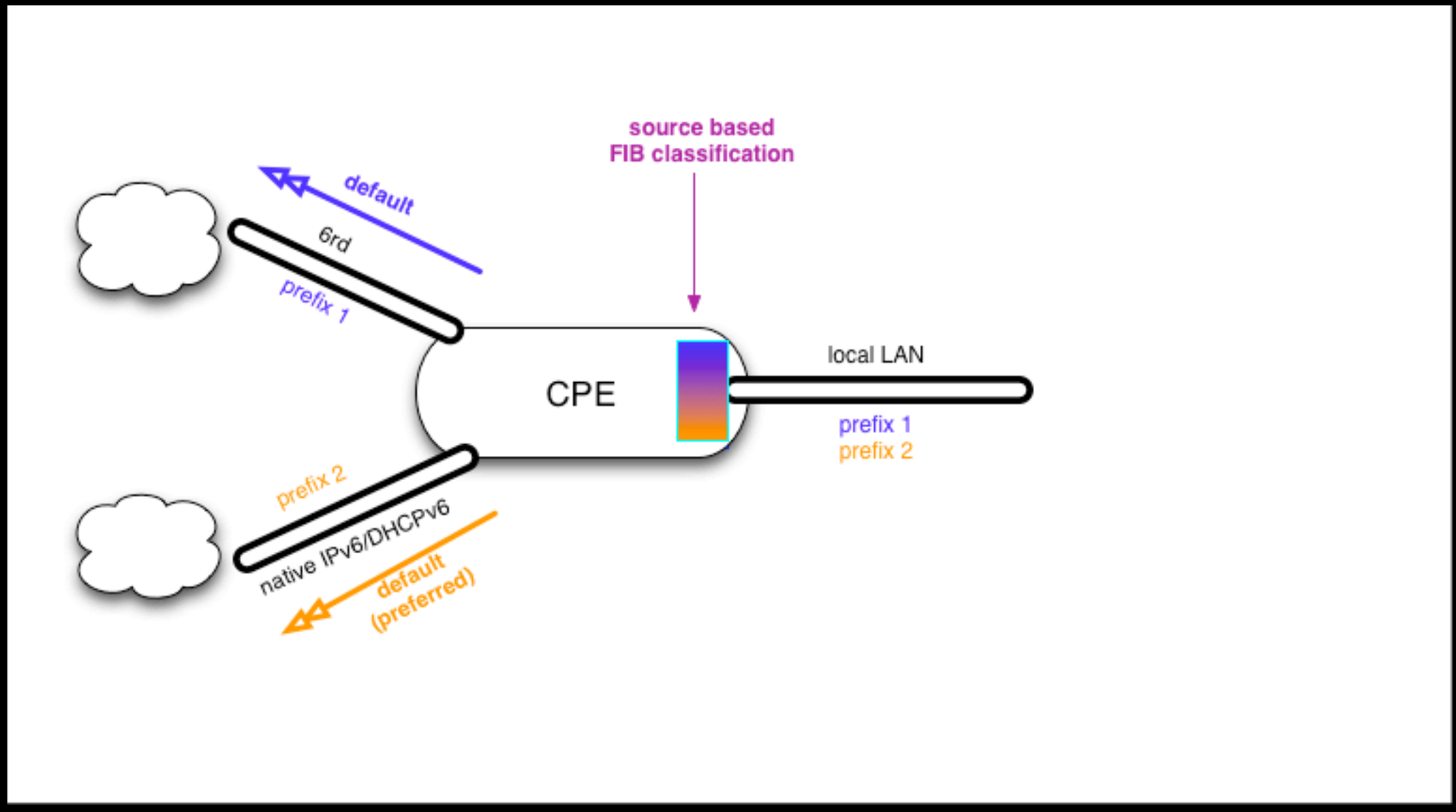
Multi-FIB IPv6, CPE



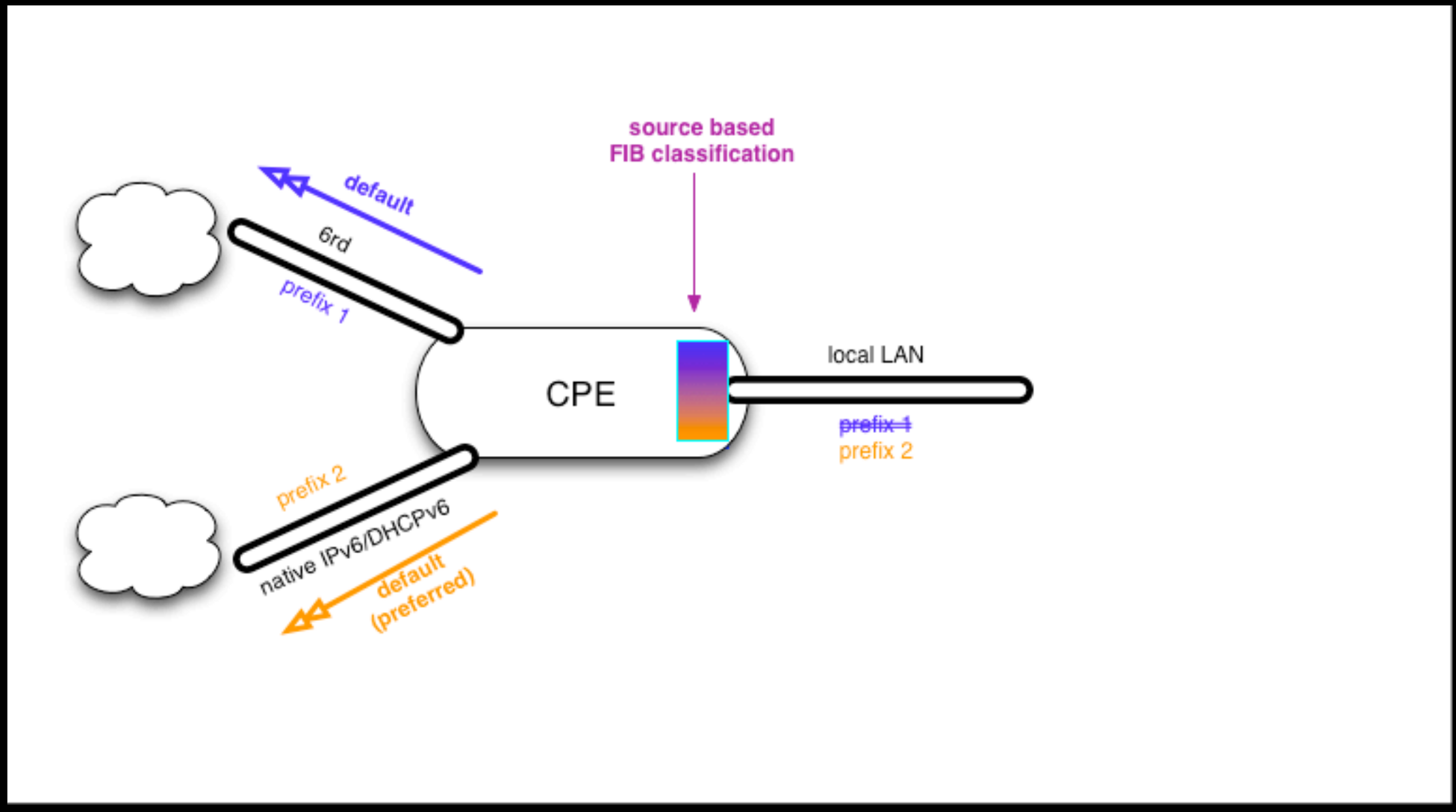
Multi-FIB IPv6, CPE



Multi-FIB IPv6, CPE



Multi-FIB IPv6, CPE



Multi-FIB

- Not special to IPv6 but in HEAD, stable/[98] multi-FIBs are now available in GENERIC with just the loader tunable.

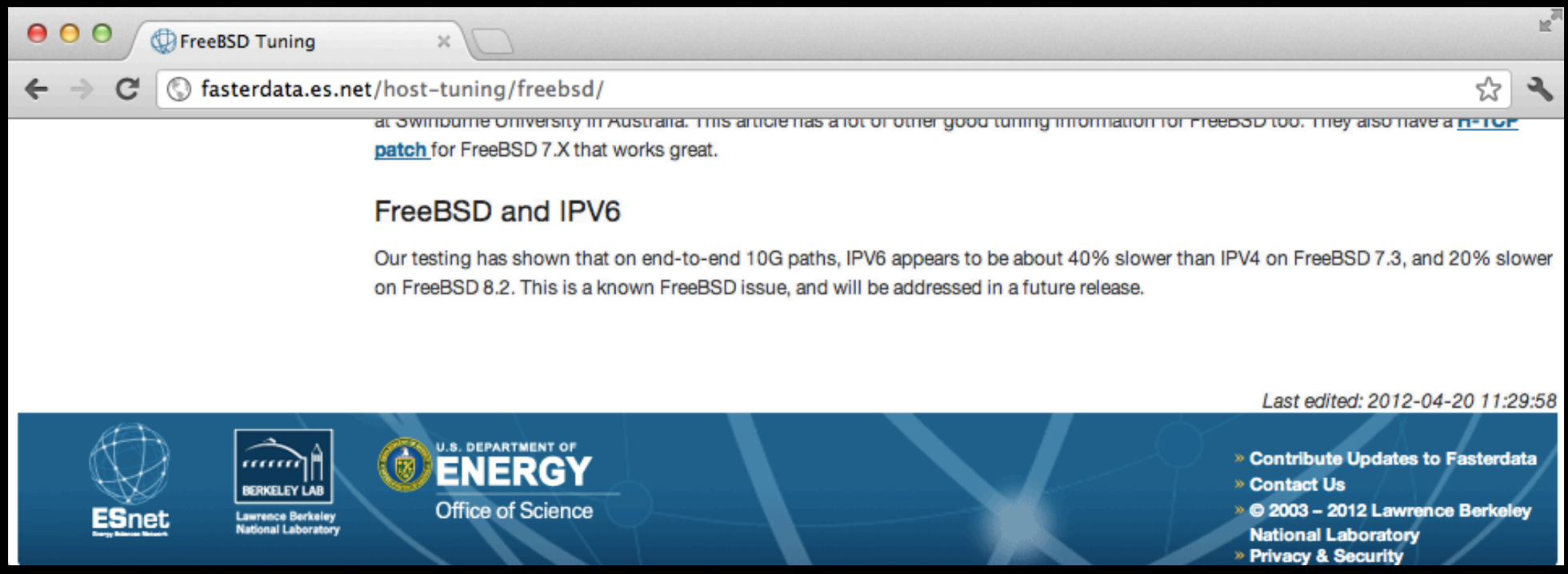
IPv6 Performance

- Sponsored by the FreeBSD Foundation and iXsystems, Inc.

IPv6 Perf-what?

Cannot prefer IPv6 due to bad loopback performance.

IPv6 Perf-what?



Used to say:

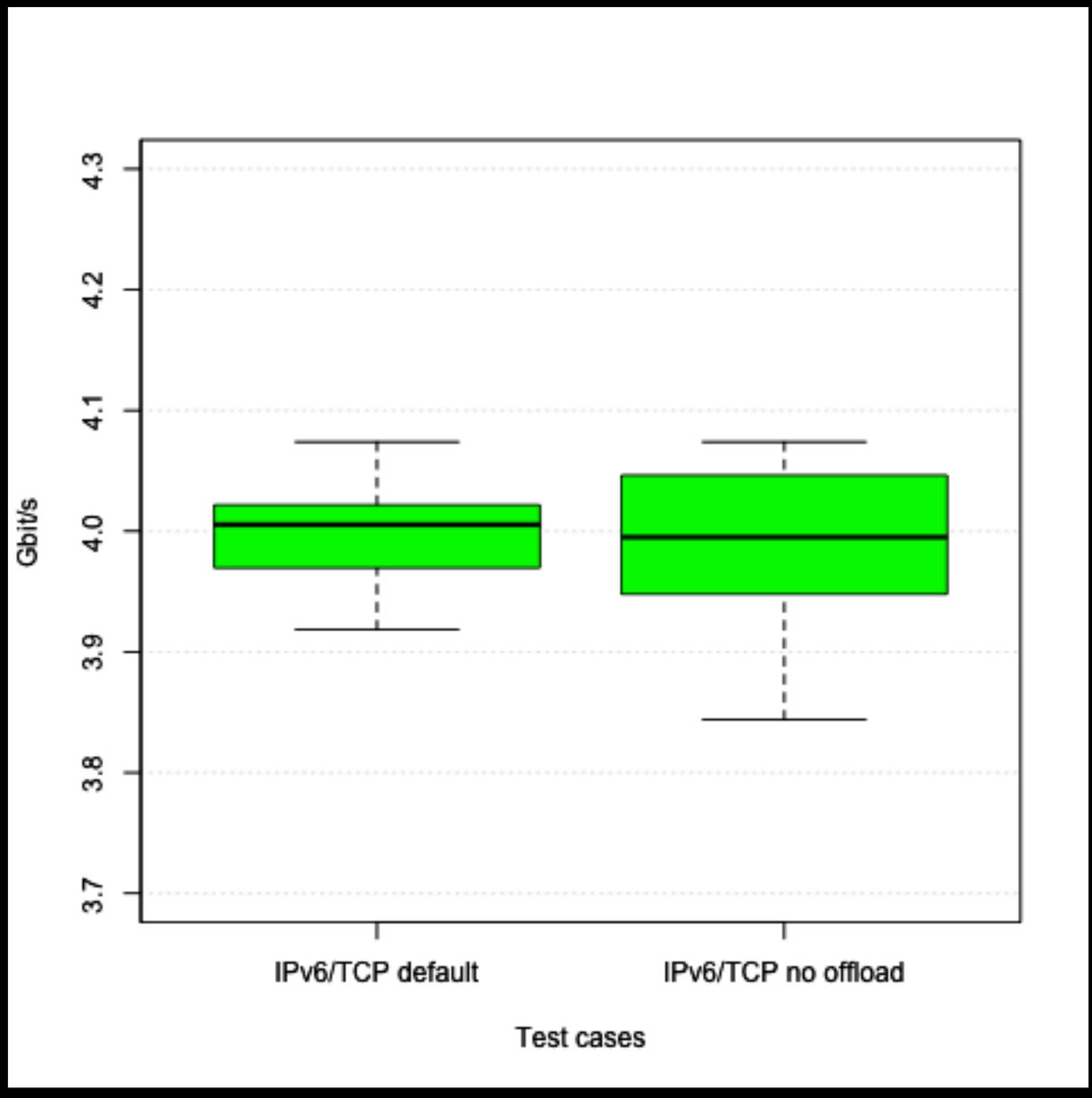
``Our testing has shown that on end-to-end 10G paths, IPV6 appears to be about 40% slower than IPV4 on FreeBSD 7.3, and 20% slower on FreeBSD 8.2. We are not yet sure why this is, as other OSes do not exhibit this behavior."

So where are we?

- Who did benchmarks on IPv6 the last years to give us numbers?
- NOTE WELL: all of the following numbers are to compare IPv4 to IPv6 not to get optimal performance with minimal effort!
- Tests done in the netperf cluster at Sentex. See http://people.freebsd.org/~bz/bench/ .

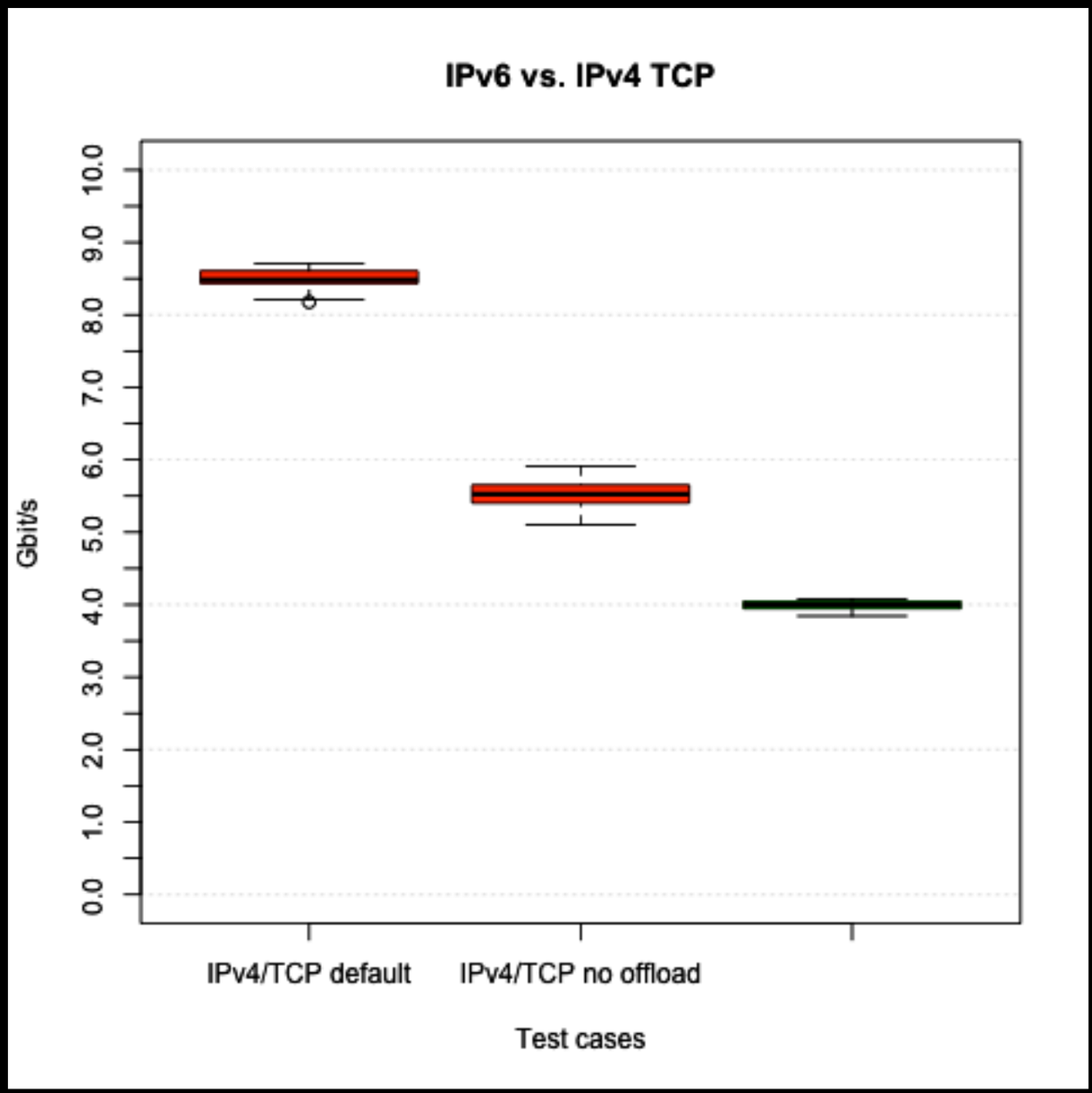
Initial numbers

TCP/IPv6



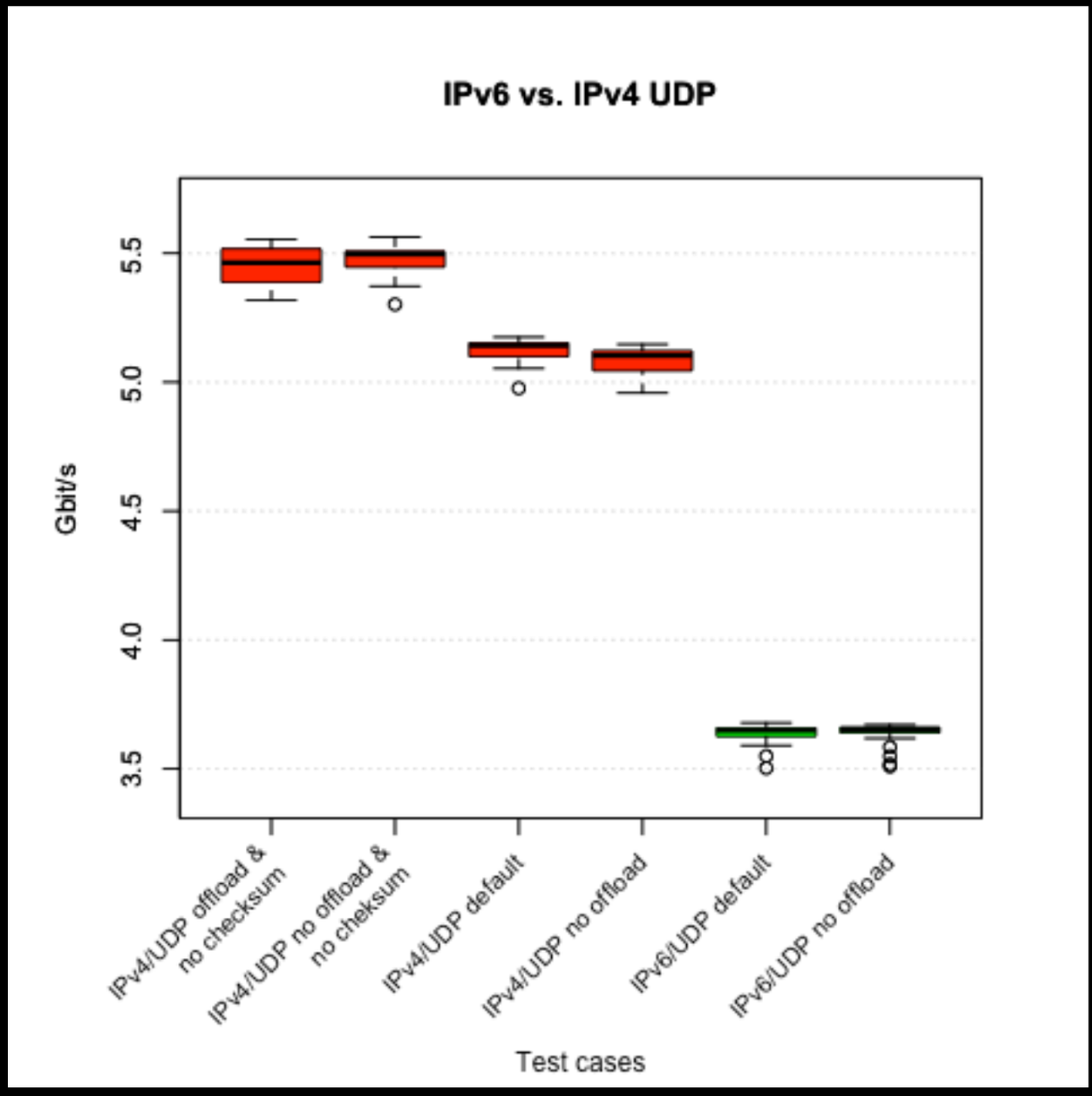
Initial numbers

TCP on IPv4(red) and IPv6 (green)



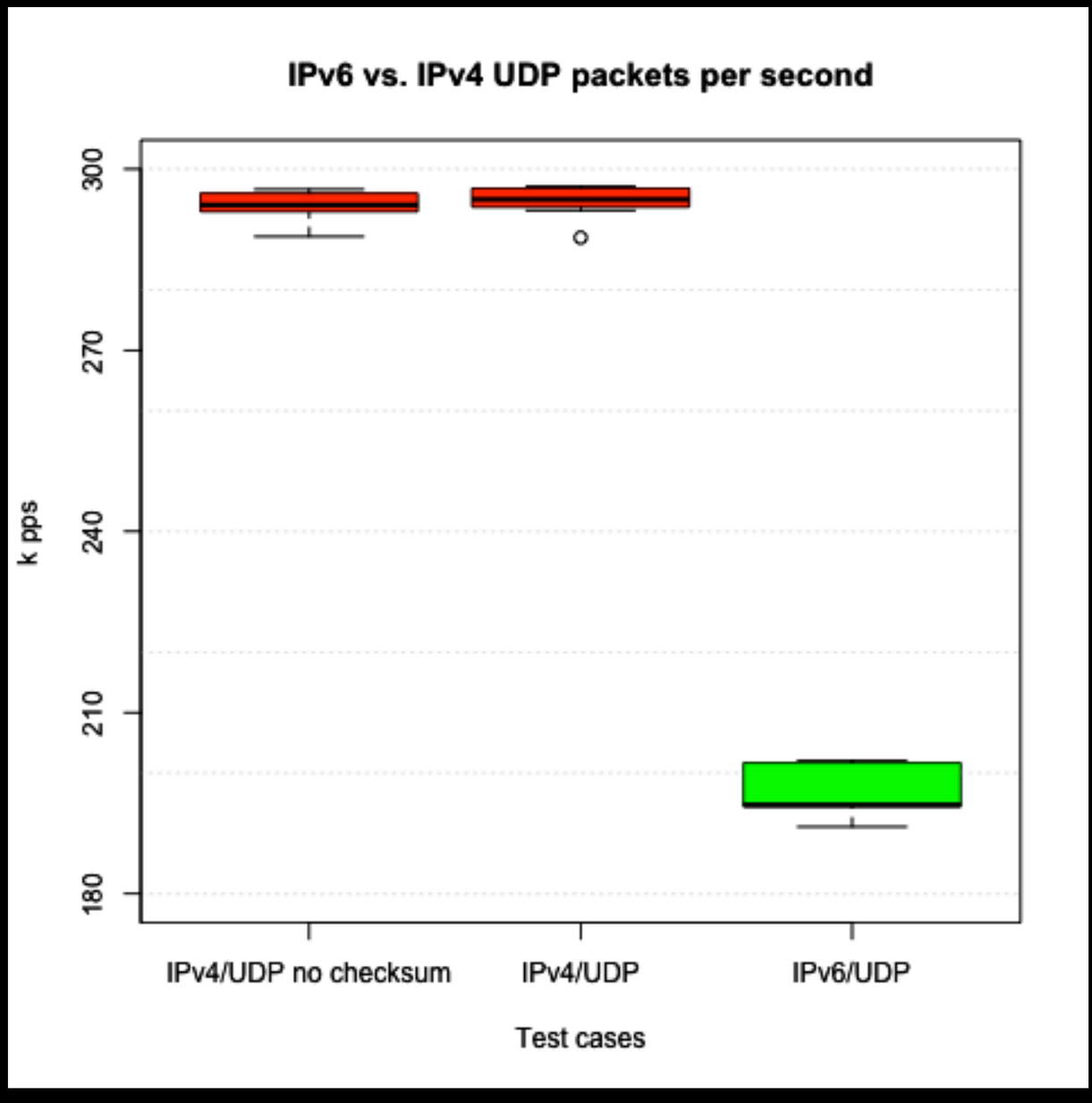
Initial numbers

UDP on IPv4(red) and IPv6 (green)



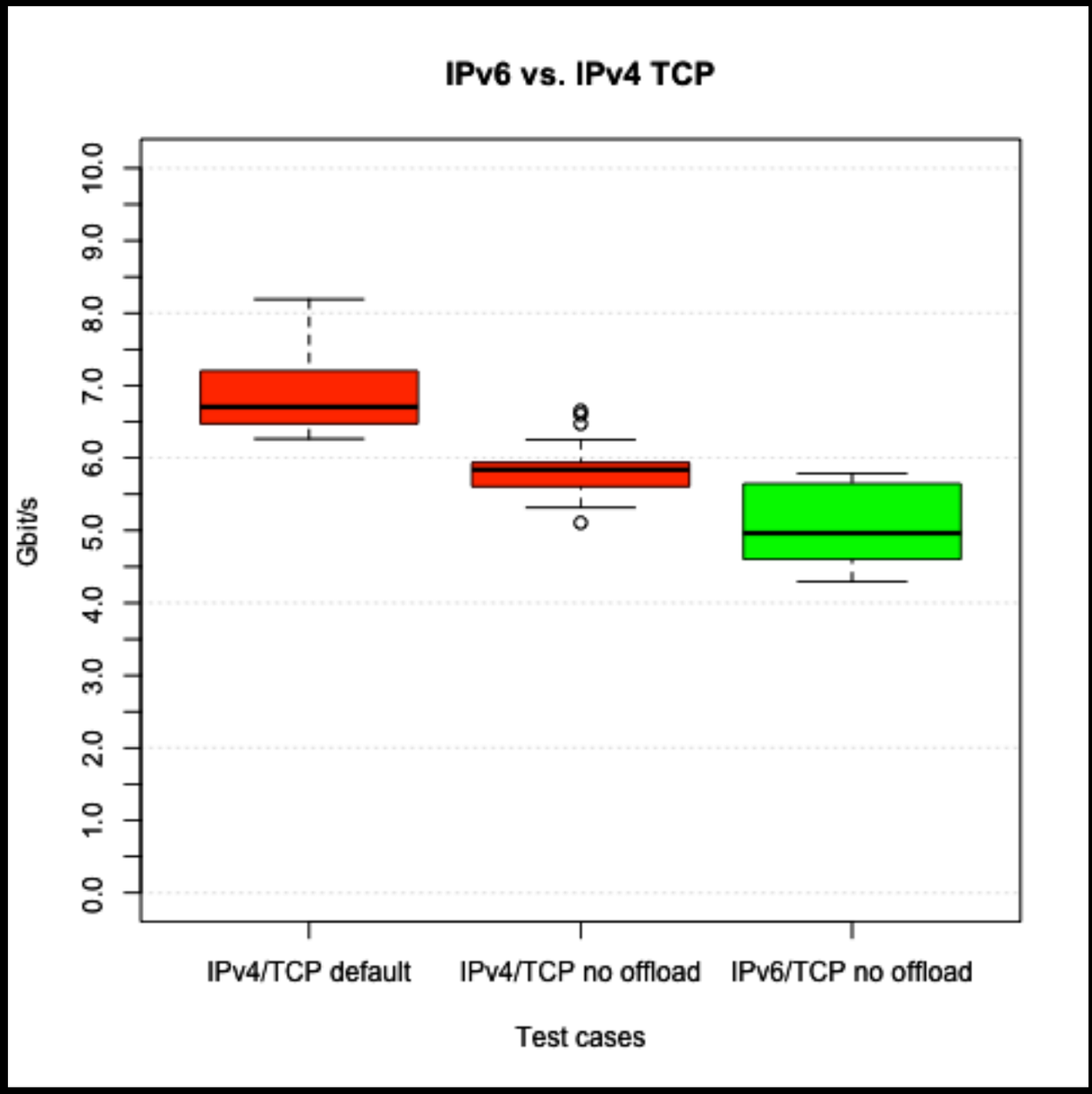
Initial numbers

UDP pps on IPv4(red) and IPv6 (green)



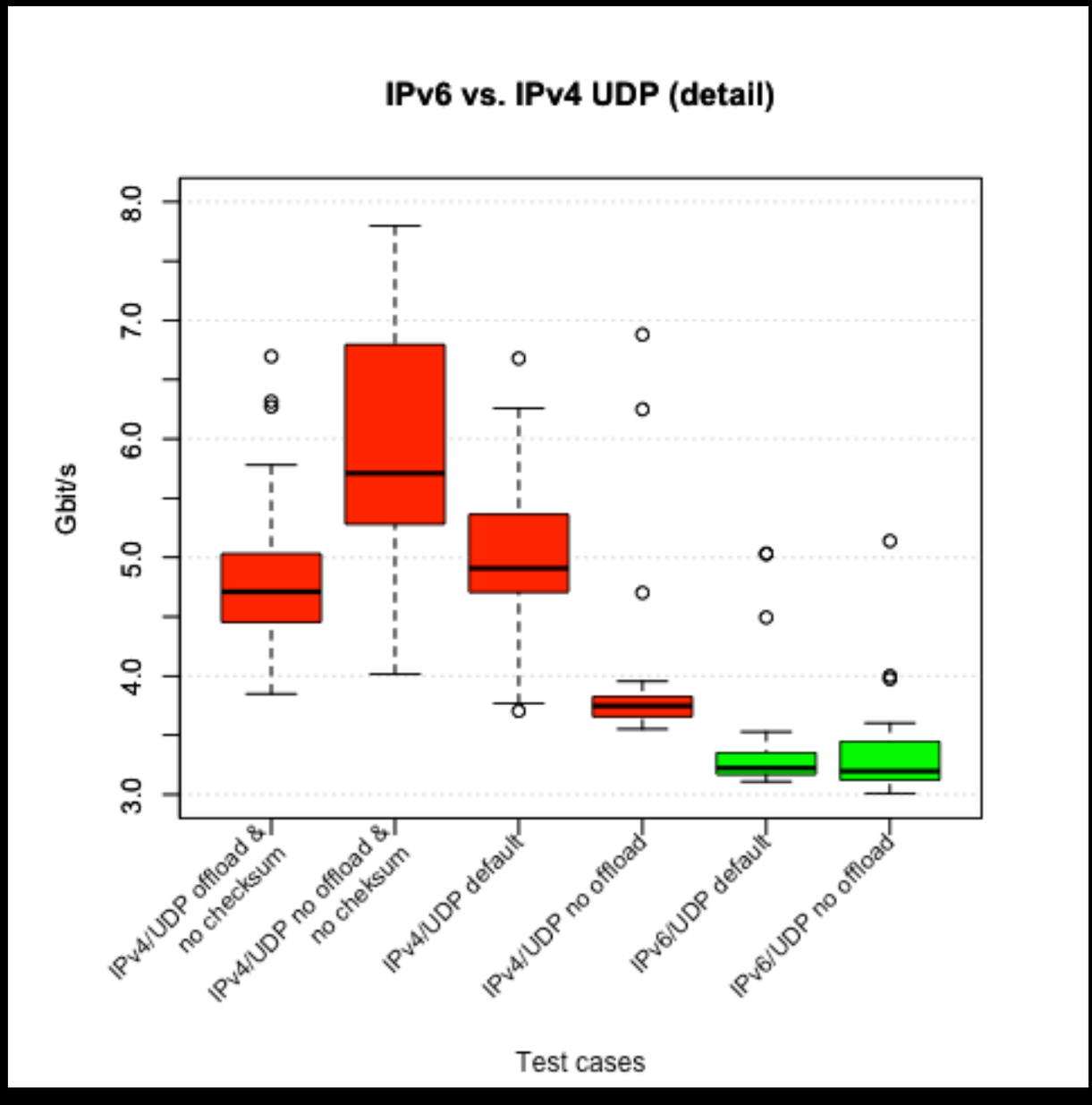
Initial lo0 numbers

TCP on IPv4(red) and IPv6 (green)



Initial lo0 numbers

UDP on IPv4(red) and IPv6 (green)



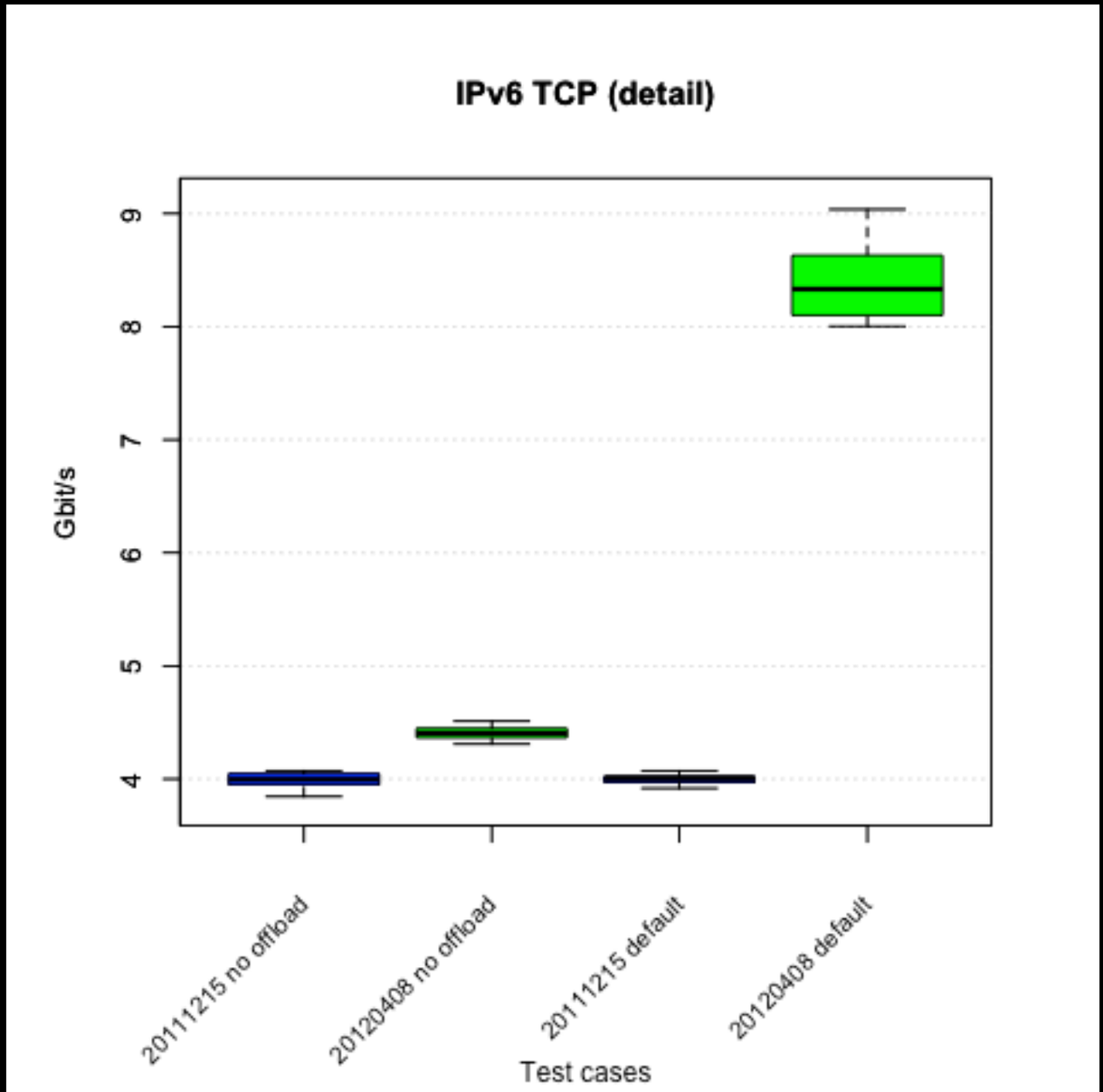
Results

- Where are we now?

- blue: initial numbers (left of current)
  red: current IPv4
  green: current IPv6

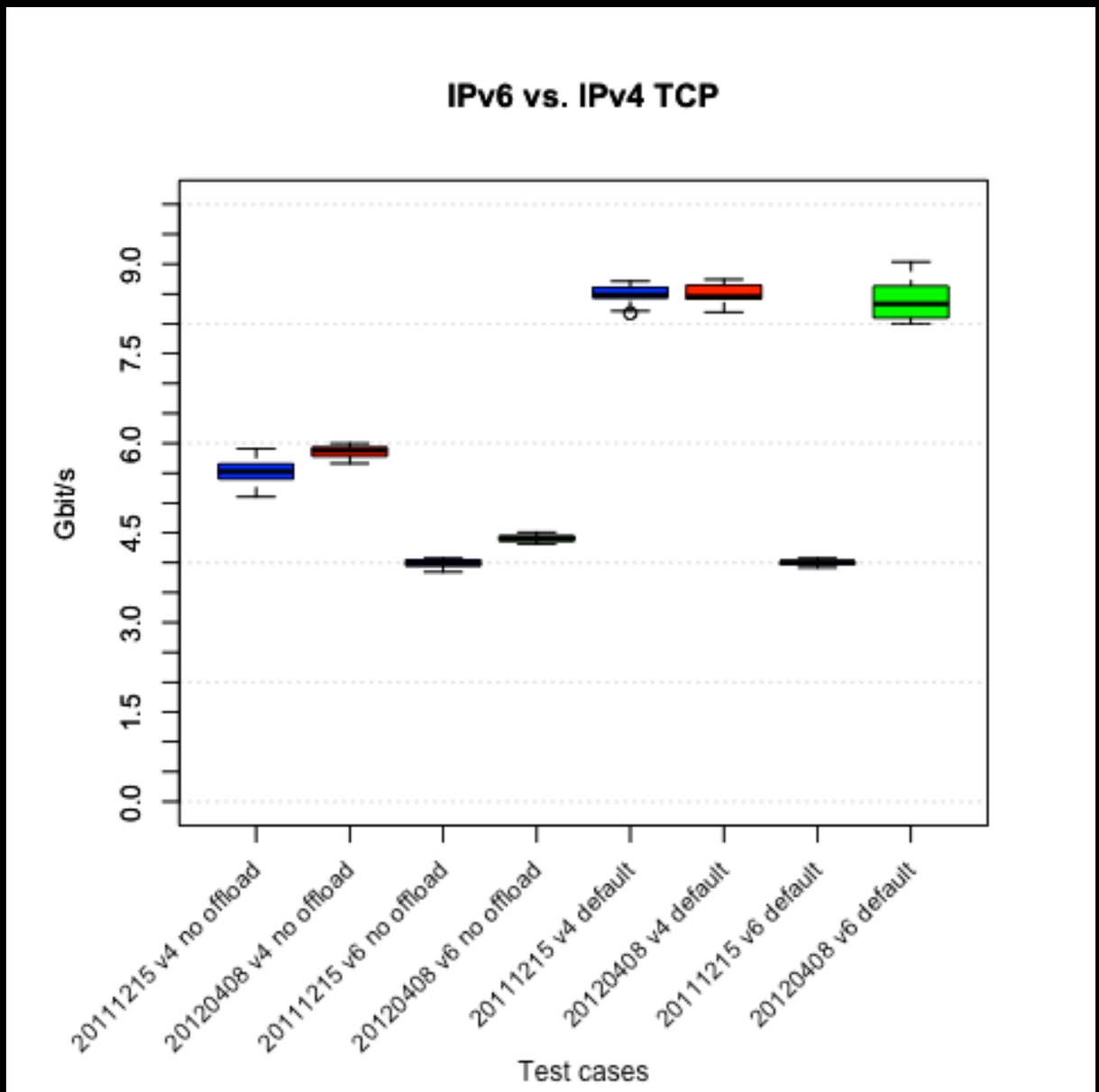
Current numbers

TCP/IPv6



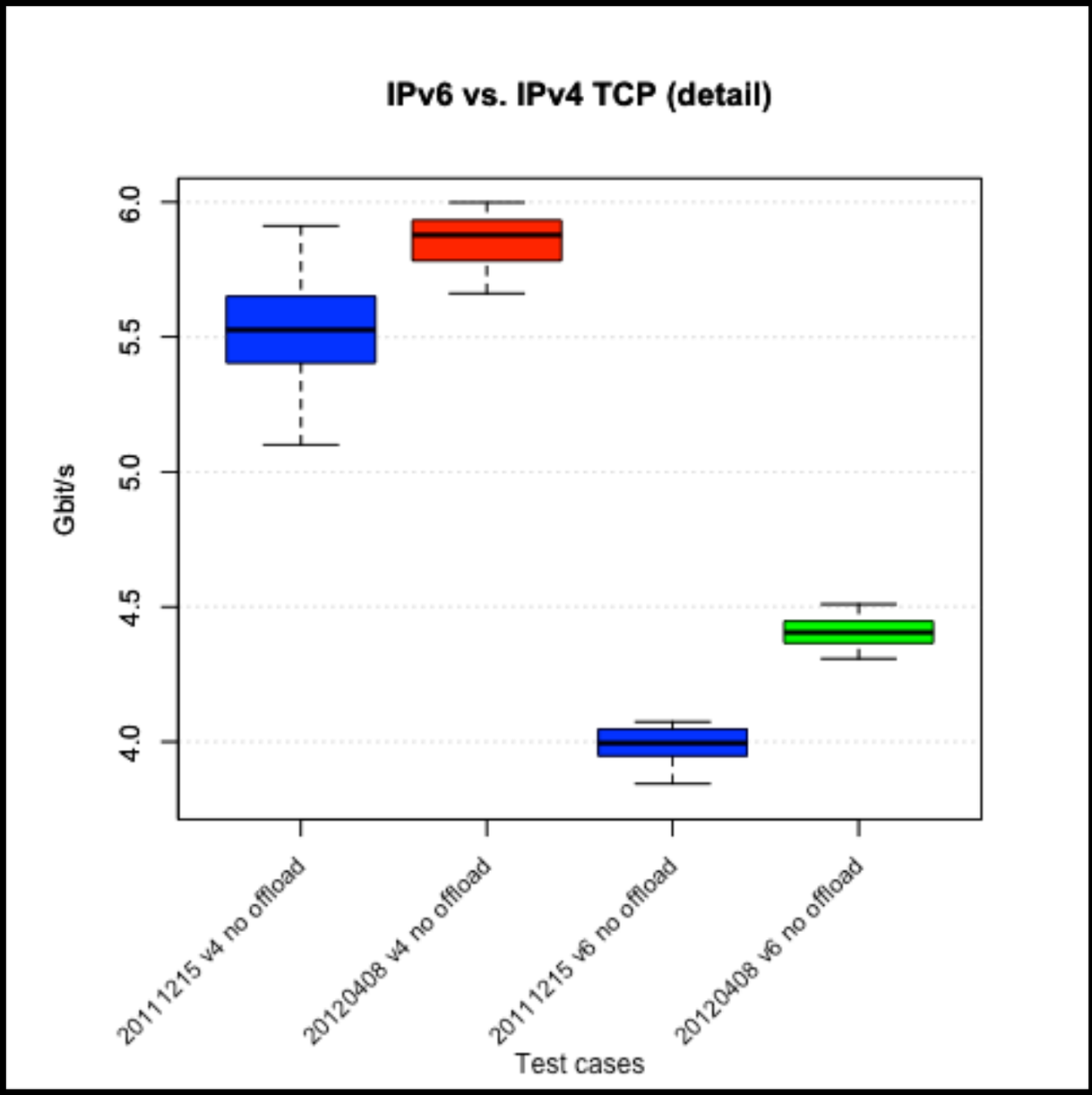
Current numbers

TCP on IPv4 and IPv6



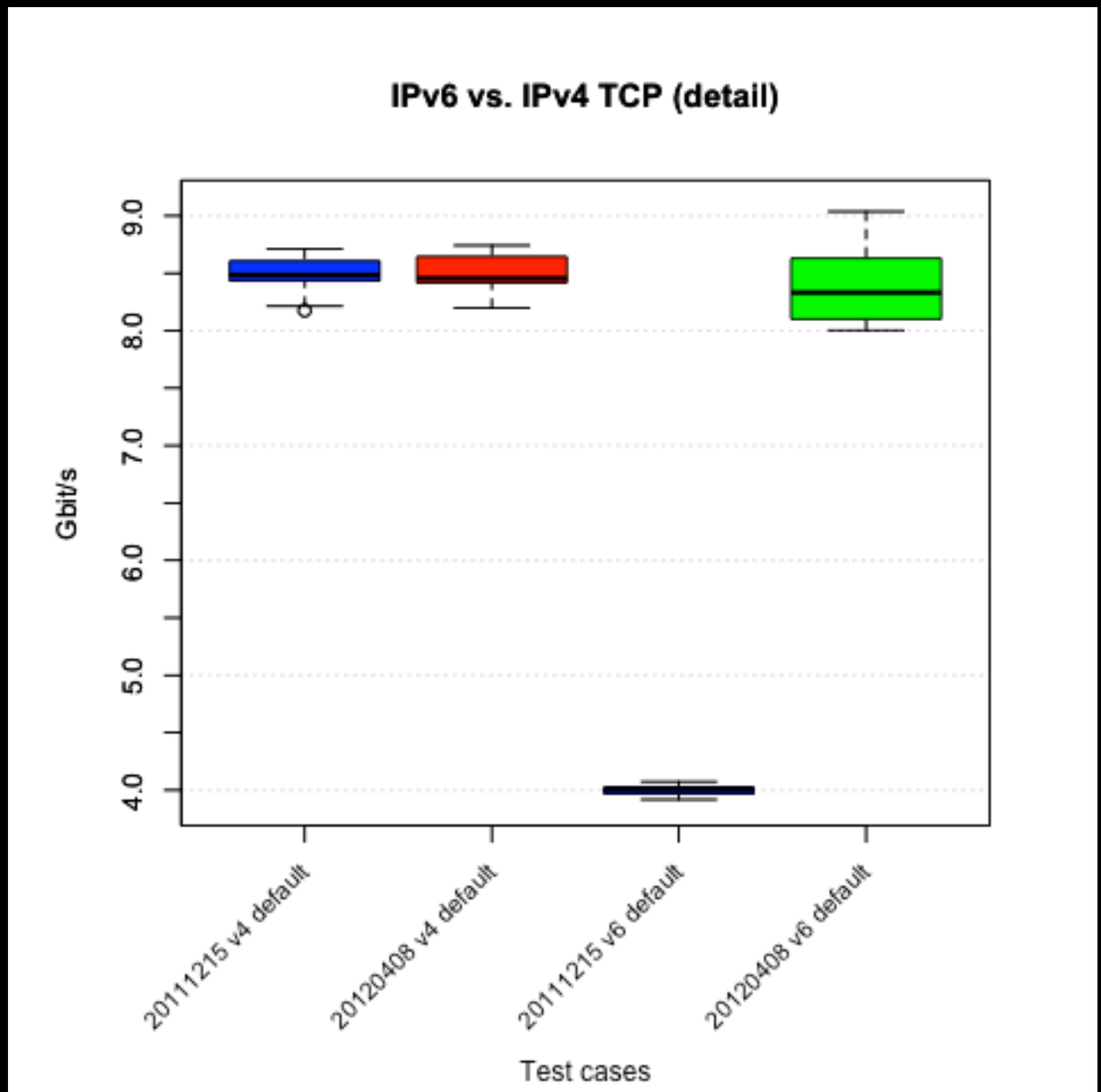
Current numbers

TCP on IPv4 and IPv6 (no offload)



Current numbers

TCP on IPv4 and IPv6 (default)

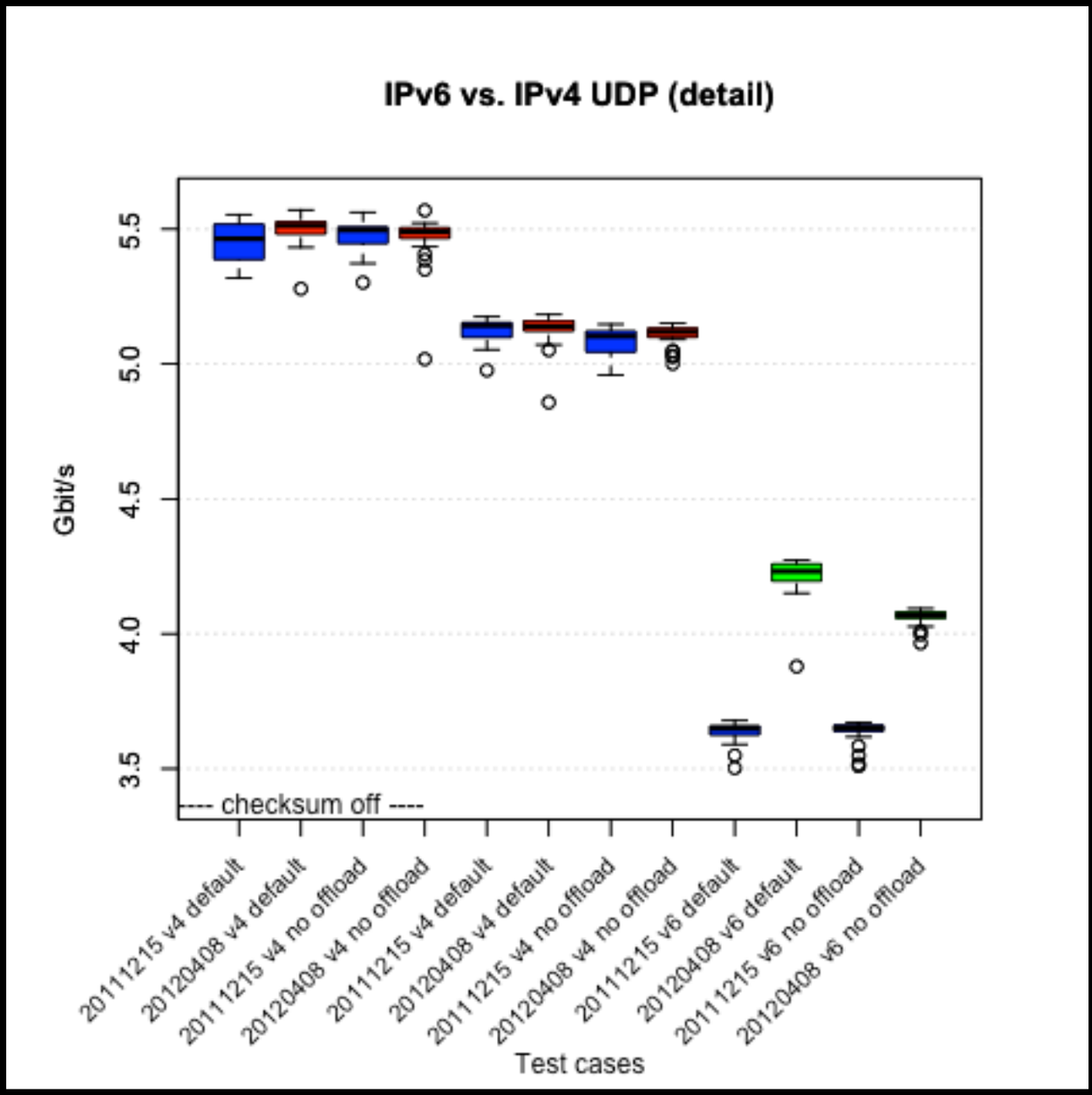


Conclusion TCP

- TCP with a NIC doing offloading is close to no difference anymore on IPv4 vs. IPv6. :)
- TCP in no-offloading case:
  - can be improved on both IPv4 and IPv6
  - IPv6 case wants to be improved

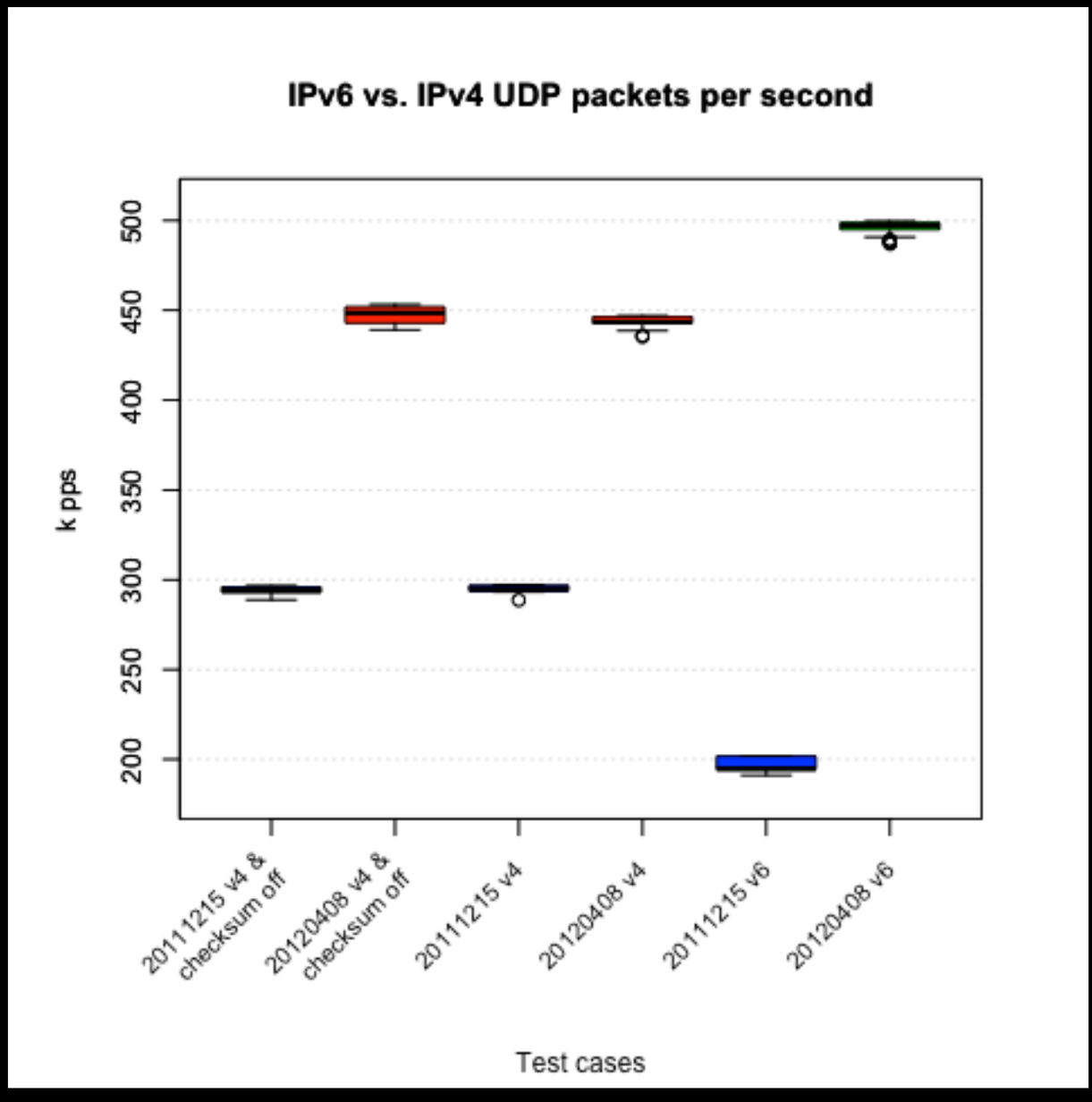
Current numbers

UDP on IPv4 and IPv6



Current numbers

UDP pps on IPv4 and IPv6

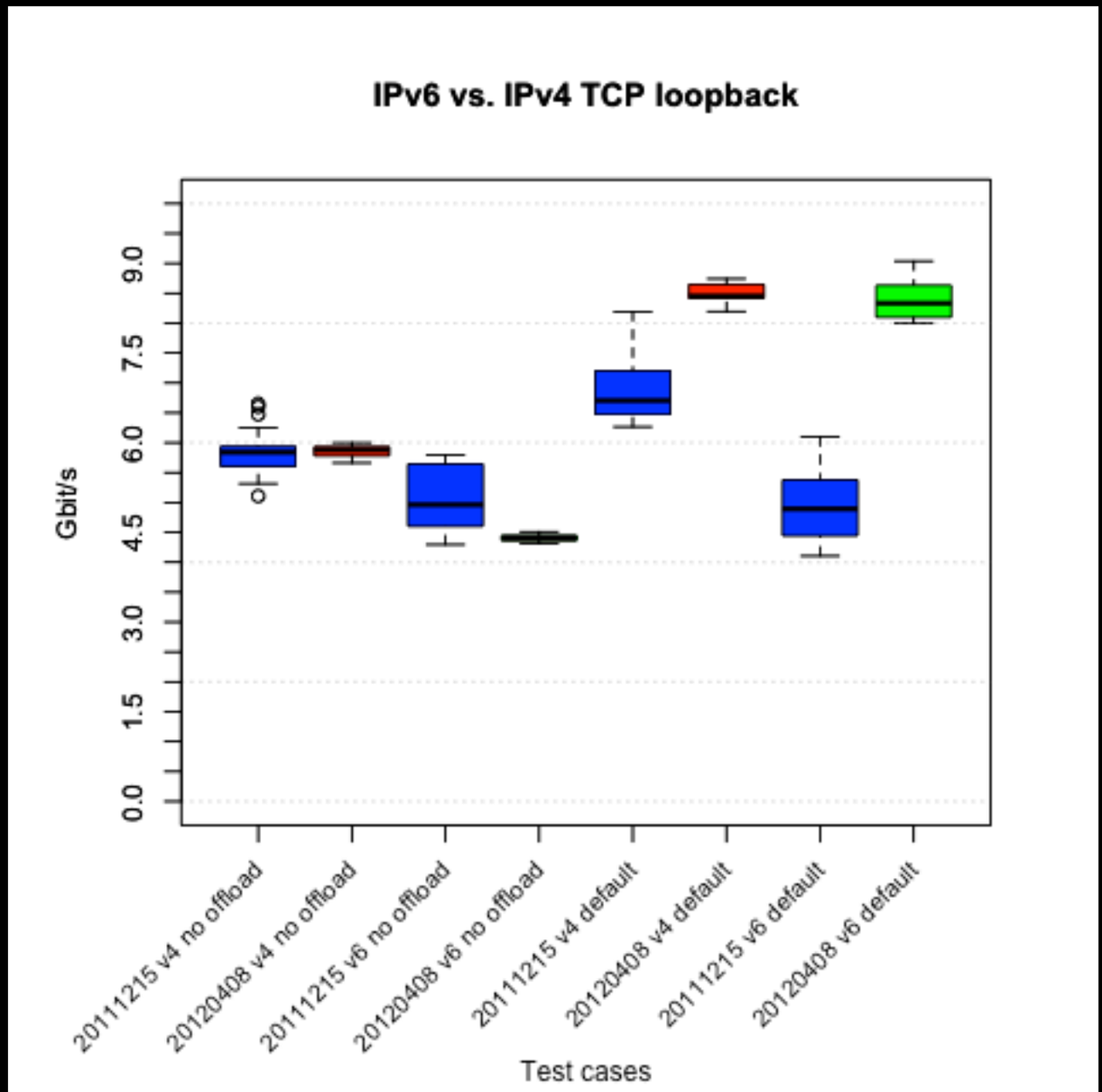


Conclusion UDP

- We can do more pps in parallel on IPv6. :)
- Pushing data, we do noticeably better on IPv6 now :)
- However IPv6 still needs to be investigated more.

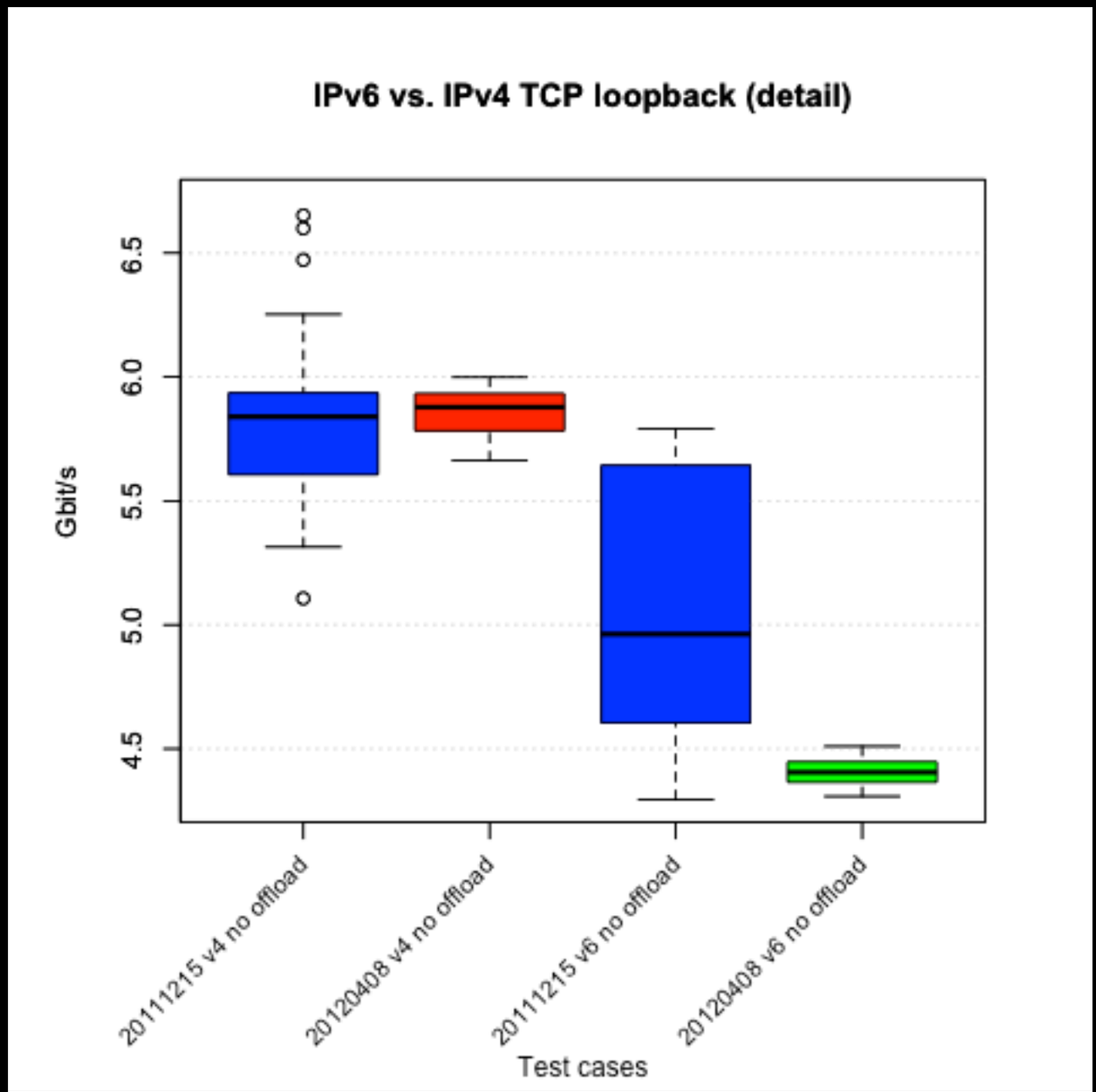
Current lo0 numbers

TCP on IPv4 and IPv6



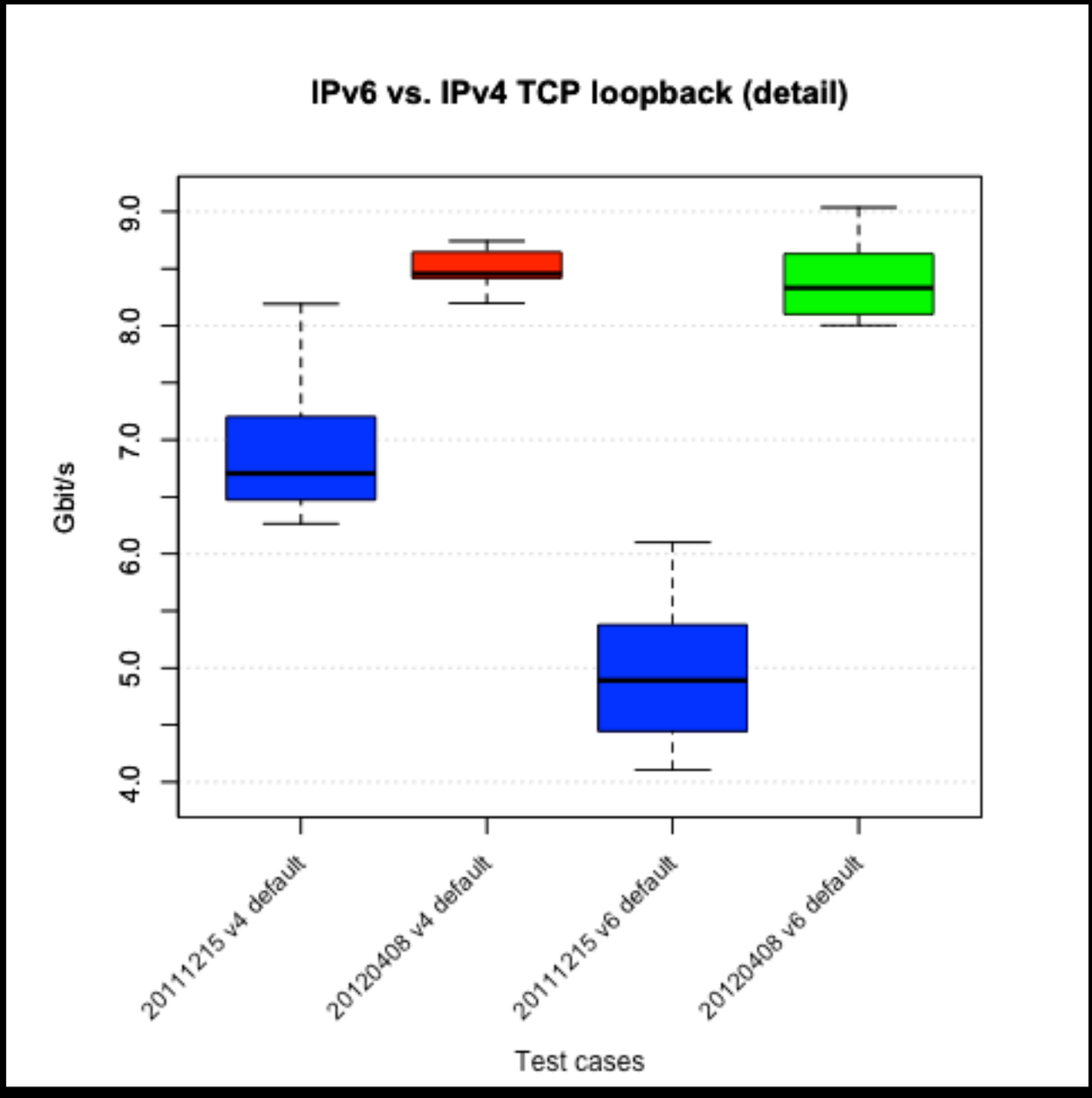
Current lo0 numbers

TCP on IPv4 and IPv6 (no offload)



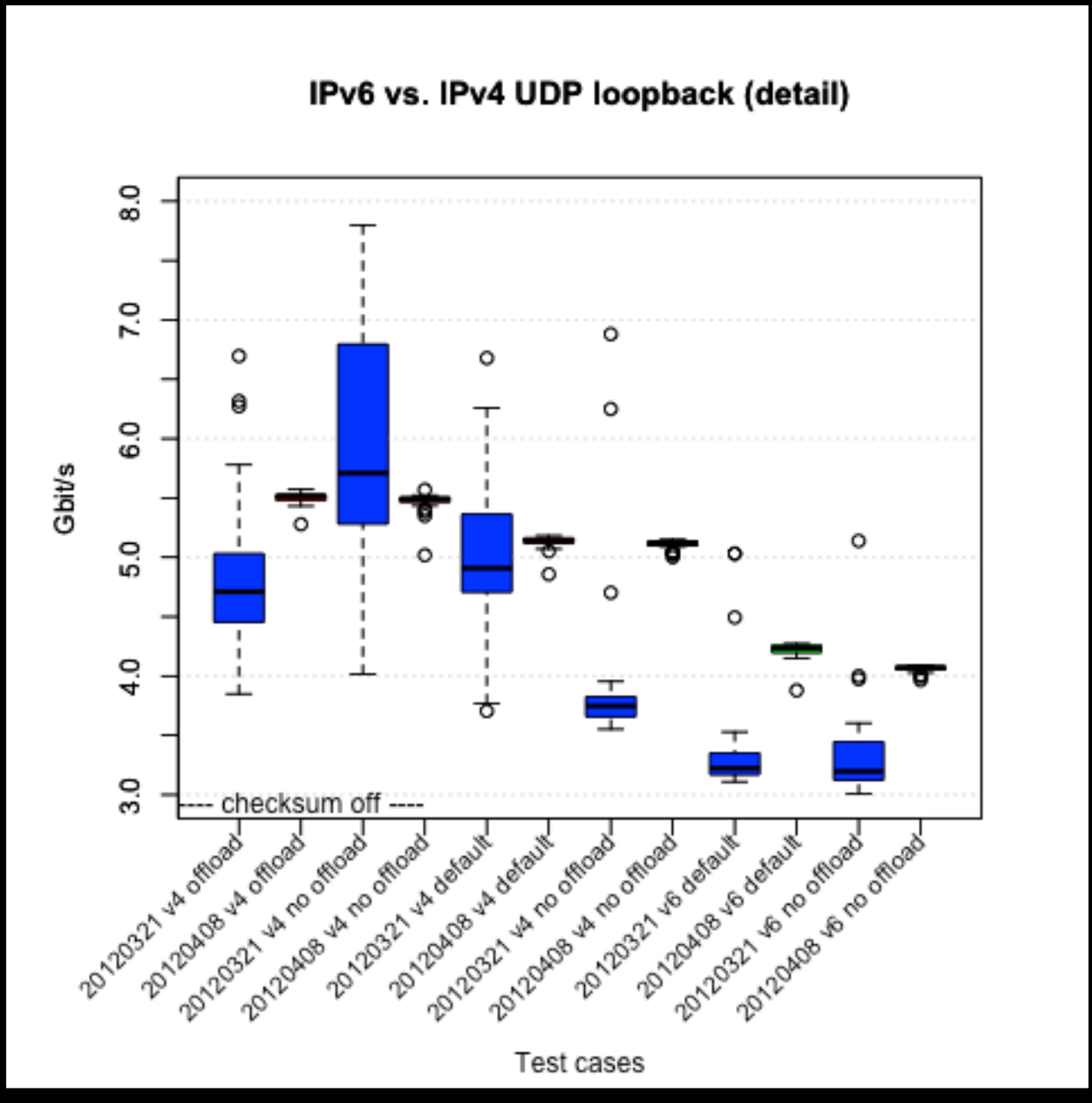
Current lo0 numbers

TCP on IPv4 and IPv6 (default)



Current lo0 numbers

UDP on IPv4 and IPv6



Conclusion loopback

- Relevant TCP case as good on IPv6 as on IPv4 now. :)
- UDP can be improved in general. IPv6 still behind IPv4 but doing better as well.
- SCTP improvement to come as well.
- Generally need to “fix” loopback.

Changes

- TSO6, LRO for IPv6. Delayed checksums.
- NIC driver adjustments.
- UDP/IPv6 locking. Scope locking.
- Early route lookup.
- Cache footprint, cache misses, bzero, compile out, initialize when needed,

Performance

- Where are we now?
- Generally better now than before with IPv6. :)
- Plan to commit/put out patches the week after BSDCan.
- More to come.

Security

- Fernando will tell you more about it.
- Work in progress for the easy items incl. regression tests.
- Some really nice junior kernel hacker tasks. If you are interested or have a student to work on Internet drafts/RFCs and IPv6, ping6 me.

Compliance

- TAHI - yes we ... for 4-5 releases not fixing.
- Various parties involved:
  - Vendors, hello.
  - Contact with two test centers.
- As we proceed hands would be helpful.

no-inet ("IPv6-only")

- Release builds (9.0) and snapshot (HEAD) builds.
- Not officially supported. No security updates currently.
- Plan to do a private freebsd-update.
- Would be interested in collaborated testing.

no-inet ("IPv6-only")

- GSoC student, Jonathan Calmels, to work on “IPv6 userland improvements”
- If you want to fix gre(4) or other things, would be helpful, again ping6 me.
- The project was mentioned on /. as part of a wider announcement:

  http://tech.slashdot.org/story/12/01/13/2348206/ipv6-only-is-becoming-viable

Other “requests”

- 6rd
- pf NAT64 (currently not planning others)
- pf frag6
- IVI (stateless) Translation
- DHCPv6
  (once done, also a compliance thing)
- dummynet and IPv6
- http://wiki.freebsd.org/IPv6TODO

World IPv6 Launch Day

- www.freebsd.org is signed up.
- Not sure beyond that yet.
- Certainly hang out, making it a “answer support questions day”.

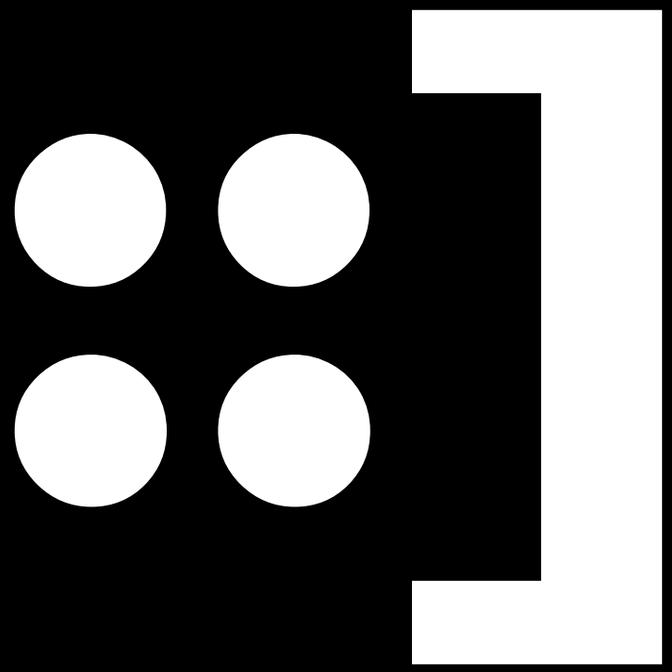
Call for hands

- We need more help!

Questions?

Feedback:
bz@FreeBSD.org

Thanks and happy IPv6ing!



(IPv6 smiley)