MP-safe Networking in NetBSD

Masanobu SAITOH (msaitoh@netbsd.org)
Ryota Ozaki (ozaki-r@{netbsd.org,iij.ad.jp})
Kengo Nakahara (knakahara@netbsd.org,k-nakahara@iij.ad.jp)

BSDTW 2017
2017/11/12

Current Status of the Project

- Many components of Layer 3 and below are MP-safe and scalable
  - src/doc/TODO.smpnet list what are already MP-safe and what's not.
- Stable enough for daily use as a router.
  - Kernels with "options NET_MPSAFE"

Why NET_MPSAFE is disabled by default?

- Because some non-MP-safe components are enabled in conf/GENERIC file and they may cause panic.
  - See src/doc/TODO.smpnet for the detail

MP-safe Network Components (1/3)

- Layer 2
  - Ethernet
  - bridge(4)
  - Fast forward
- Layer 3
  - Routing table, IP addresses, ARP/ND, etc.
  - Except for MPLS and some options such as MROUTING

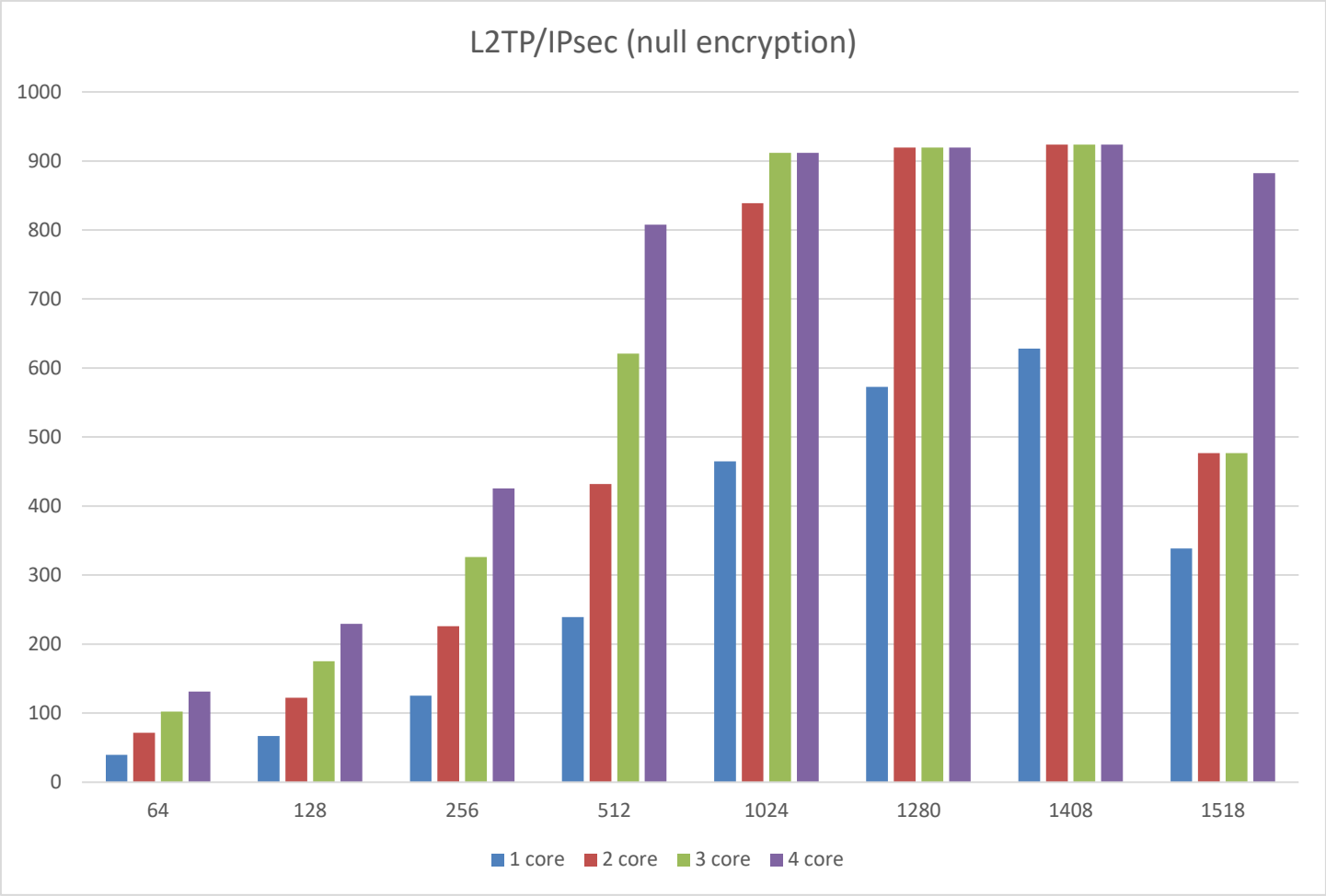
MP-safe Network Components (2/3)

- Network device drivers
  - wm(4), vioif(4), vmx(4), ixg(4), ixv(4)
  - Hardware multi-queue support
    - Except for vioif(4)
- Pesuedo interfaces
  - gif(4), l2tp(4), pppoe(4), tun(4) and vlan(4)

MP-safe Network Components (3/3)

- Others
  - pfill(9), npf(7) and bpf(4)
  - opencrypto(9) and ipsec(9) New!
    - Need more work for device driver under opencrypto(9)
      - e.g. qat(4) and hifn(4)
  - Added 10G support into ipgen
    - An Interactive Packet Generator using with netmap
    - Support RFC 2544 test
    - https://github.com/iij/ipgen

Performance at glance



- 2 DUTs
  - Atom C2558
    - 4 core
    - 2.4 GHz
  - 8GB memory
  - Intel GbE
    - I354
- four L2TP/IPsec connections
  - null encryption

Ongoing Works

- ipsec(4)
  - Add routing based IPsec interface
- qat(4)
  - Intel QuickAssist driver.
- agr(4)
- Adding ATF tests (using with rump kernel)
- dogfooding

Remaining works (1/2)

- Improve single thread performance
  - We have worked for scalability so far
- Common functions
  - RSS hash stuff
  - Jumbo buffer allocation
- ipsec(9)
  - Scalability in terms of the number of SA (>1000)
- In-kernel AES-NI
- ppp(4), pipex(4), vxlan(4)
- Rework for mii(4)

Remaining Works (2/2)

- Layer 4
- Layer 2 other than Ethernet
- Many pseudo interfaces such as gre(4)
- Packet filters: ipf and pf

Any question?