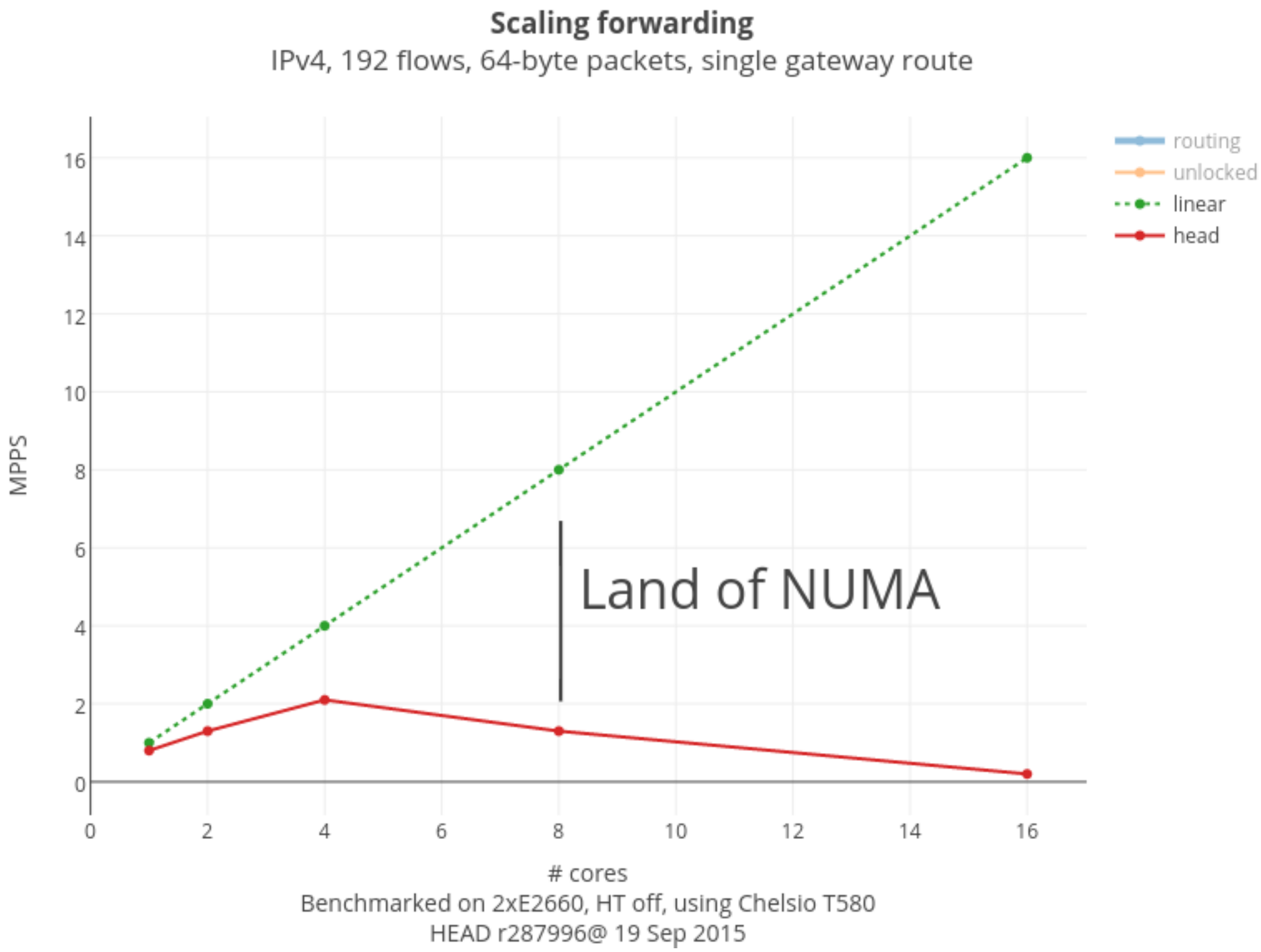
Scaling routing stack

Agenda

- Routing does not scale well
- projects/routing branch description
- Next steps

Head performance



Current locking

- per-rte mutex
- per-radix rwlock
- per-lle rwlock
- per-interface rwlock
- ifa refcounting

New locking

- ~~per-rte mutex~~
- per-radix ~~rwlock~~ rmlock
- ~~per-llc rwlock~~
- ~~per-interface rwlock~~ per-lltable rmlock
- ~~ifa refcounting~~

Implementation: routing

- Opaque routing (fib[46]_lookup_prepend, ..)
- Export nexthop info, hide rte
- API for different consumers (fwd, rpf, control)

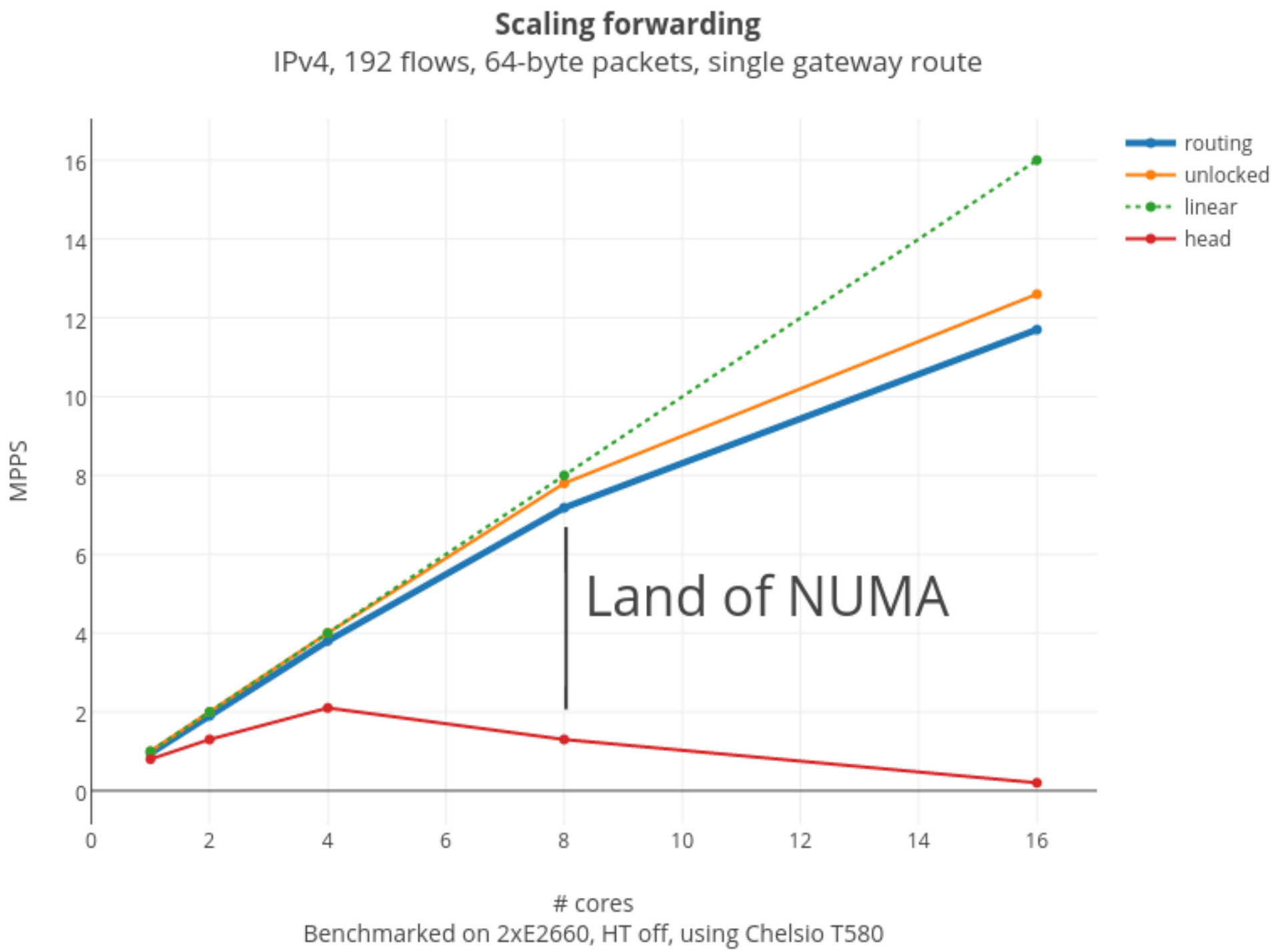
L3 nexthops

```
struct nhop_prepend {
        uint16_t        nh_flags;       /* NH flags */
        uint8_t         nh_count;       /* Number of nexthops or data length */
        uint8_t         spare0;
        uint16_t        nh_mtu;         /* given nhop MTU */
        uint16_t        lifp_idx;       /* Logical interface index */
        union {
                uint16_t        ifp_idx;        /* Transmit interface index */
                uint16_t        nhop_idx;       /* L2 multipath nhop index */
        } i;
        uint16_t        aifp_idx;       /* Interface address index */
        uint16_t        spare1[2];
        union {
                char    data[MAX_PREPEND_LEN];  /* data to prepend */
#ifdef INET
                struct in_addr  gw4;            /* IPv4 gw address */
#endif
#ifdef INET6
                struct in6_addr gw6;            /* IPv4 gw address */
#endif
        } d;
};
```

Implementation: lltable

- Opaque lltable (arpresolve, nd6_resolve())
- Do not modify lle in fast path
- Provide “feedback” that entry is really used

Projects/routing performance



Next steps

- Real nexthops with tracking
- Full prepend in single lookup (GW routes)
- Proper multi-path (special “recursive” nexthops)
- Per-AF per-fib lookup algos (DIR24-8, DXR, hash,..)

Next steps

- Modular routing lookup engines
- lookup(IP) -> nhop_id
- Modular lookup algos (DIR24-8, DXR, hash,..)
- Per-AF per-table algo

More info

- https://wiki.freebsd.org/ProjectsRoutingProposal
- projects/routing branch
- Phabric: D3688, D3780 (more will follow)