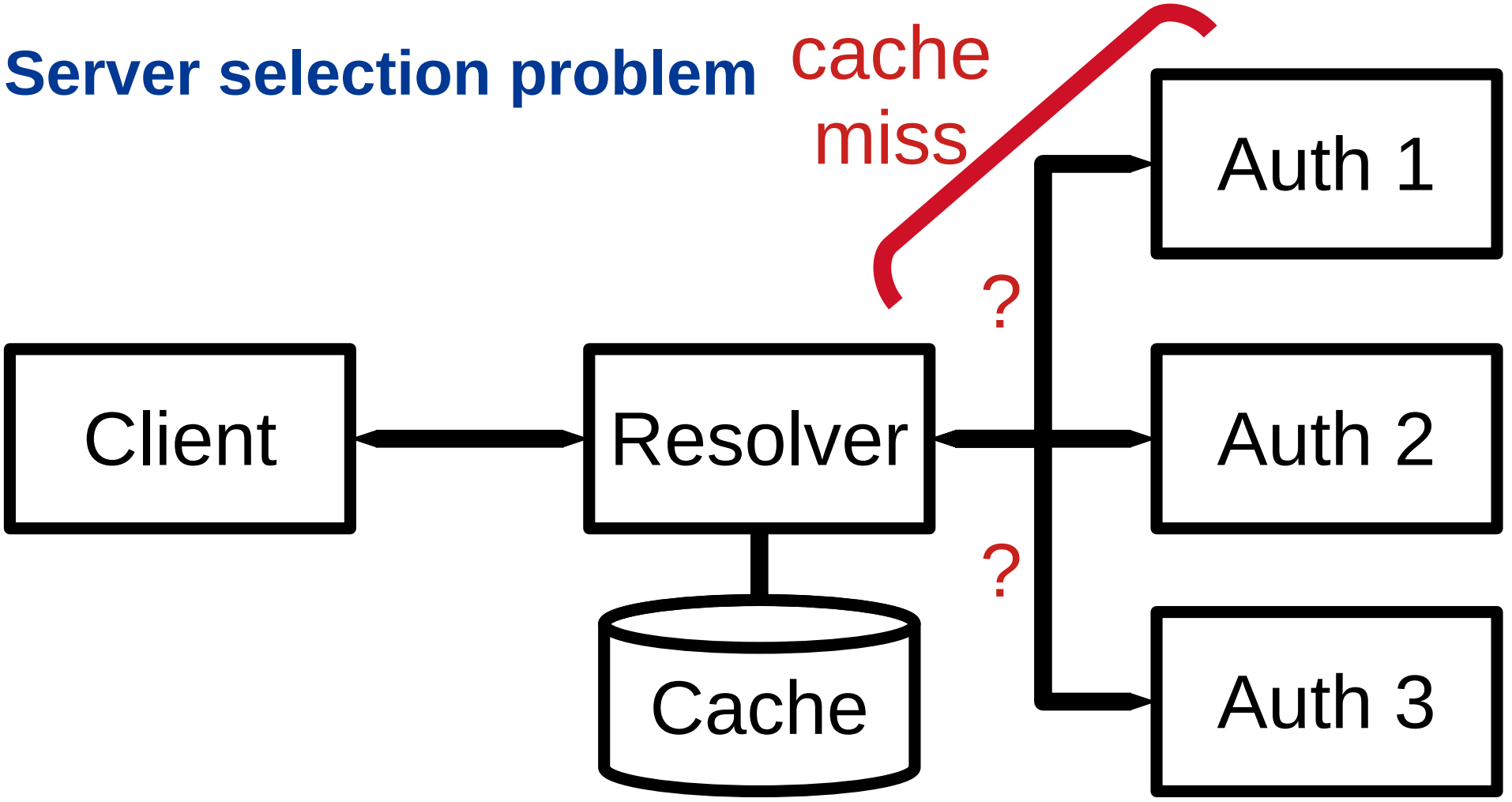
DNS Maze:
Testing server selection

Petr Špaček • petr.spacek@nic.cz • (speaking)
Štěpán Balážik • stepan.balazik@nic.cz • (working)

CZ DOMAIN REGISTRY

Server selection problem



Server selection problem

- “Which server should I ask next?”
  - Resolve missing glue? Throttle? Backoff? Serve-stale?
- Optimizing for
  - # of packets
  - Non-predictability
  - RTT for clients
- Messy code
  - Hard to understand (sometimes), arbitrary (?) choice of constants

Server selection testing

- Two sub-problems
- Phase 1 – reproducible test environment
- Phase 2 – test scenarios & data analysis

Part 1

Reproducible test environment

DNS Maze: new tool!

- Open-source and **Work-in-Progress**
  - https://gitlab.labs.nic.cz/knot/maze/
- Mock (smallish) Internet!
  - Network environment (IP addrs, delay, loss)
  - Authoritative servers
  - Resolver
  - Client
  - Time – if necessary for DNSSEC tests

DNS Maze: design principles

- Minimalistic
- Fast
  - 30 servers in 5 seconds on a laptop
- Flexible
  - Simulating broken behavior
- Mix-and-match existing SW
  - Systemd unit generator
  - Flexible config templating

DNS Maze: scenarios

auth

```
name: "e-root-server"
service: "knot"
config: "knotd.conf"
address: ["1.0.0.50"]
netem: "delay 50ms"
zonefile: "../root.zone"
domain: "."
```

resolver

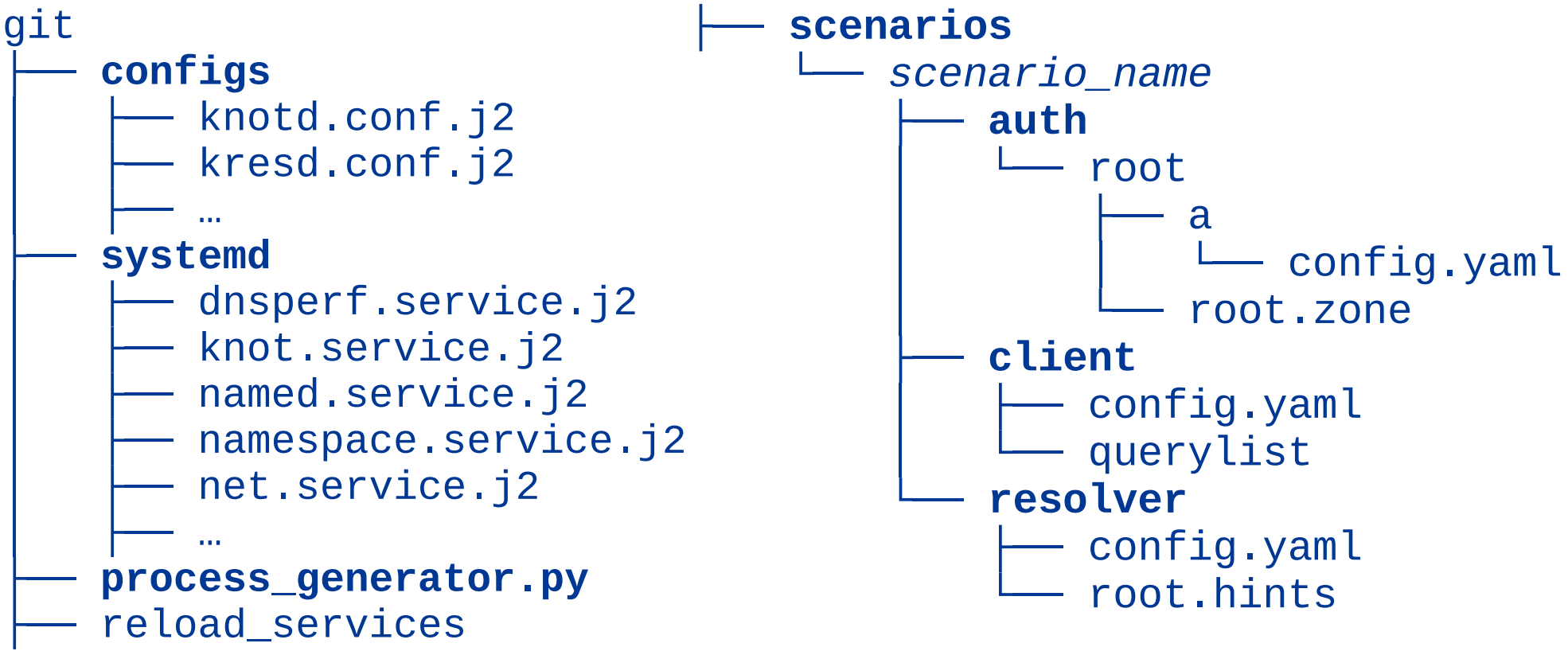
```
name: "resolver"
service: "named"
address: ["1.0.0.1"]
config: "named.conf"
roothints: "root.hints"
```

client

```
name: "dnsperf"
service: "dnsperf"
address: ["2.0.0.1"]
dnsperf: -d querylist
```

- Config file templates
- Systemd services
- Traffic => PCAP

DNS Maze: Maze of config files :-)



IP network simulation

- Linux namespaces – Systemd orchestration
  - `[Service]`
    `PrivateNetwork=yes`
    `ExecStartPre=/sbin/ip link add dummy0 type dummy`
  - `[Unit]`
    `JoinsNamespaceOf=MAZE_NAMESPACE_SERVICE.service`
- NetEm – Network Emulator
  - `$ tc qdisc ... dev dummy0 add netem OPTIONS`
  - Gory implementation details

Auth server simulation

- Knot DNS authoritative server
  - Config templates
  - User-provided zones
- Special "test server"
  - (qname, qtype) => an answer packet
- Latency/jitter/loss => NetEm rule
  - `delay 68ms 4ms loss random 1.48%`

DNS client simulation

- dnsperf
  - Config template
  - User-provided query list
    - 0.aaa A
    - 1.aaa A
    - 2.aaa A
    - 3.aaa A
    - …

Resolver = software-under-test

- Pick your victim :-)
  - Config template
  - libfaketime for DNSSEC tests
- Tested with
  - BIND, Knot Resolver, PowerDNS Recursor,
  - Teachable DNS Resolver ***, Unbound

Running through a maze

- # generate systemd unit files

  ```
  $ ./process_generator.py /tmp/scenario-bp-simple
  ```

  - => output /tmp/bp-simple-xe7mwcxf

- # start test (inside systemd machinery)

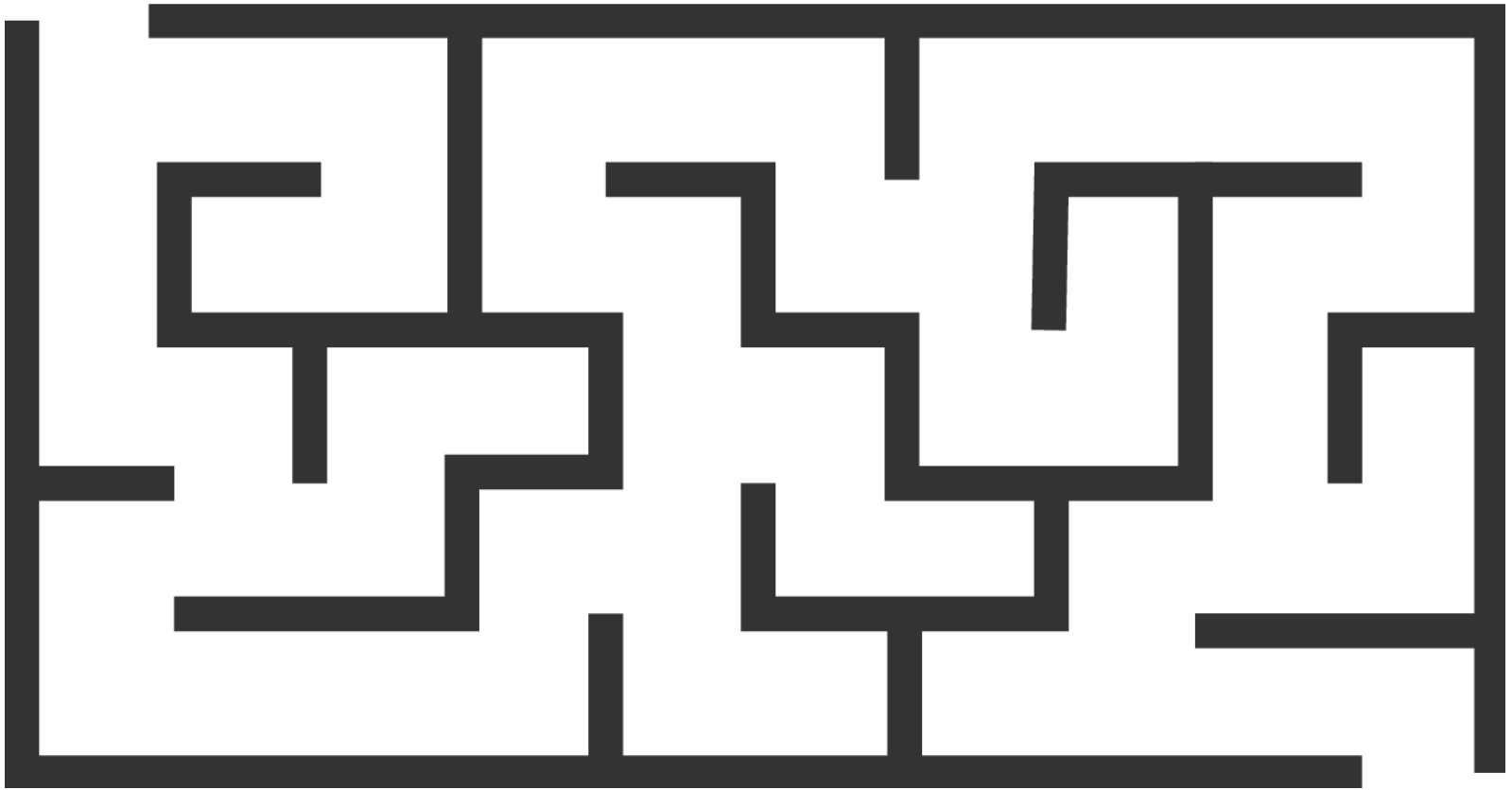
  ```
  $ sudo ./reload_services /tmp/tmpname bp-dnsperf
  ```

  - wait for /tmp/dump.pcap

Part 2

Test scenarios & data analysis

Server selection: theory



Server selection: reality

Source: https://commons.wikimedia.org/wiki/File:Labyrinth.svg CC BY SA 2.5, modified

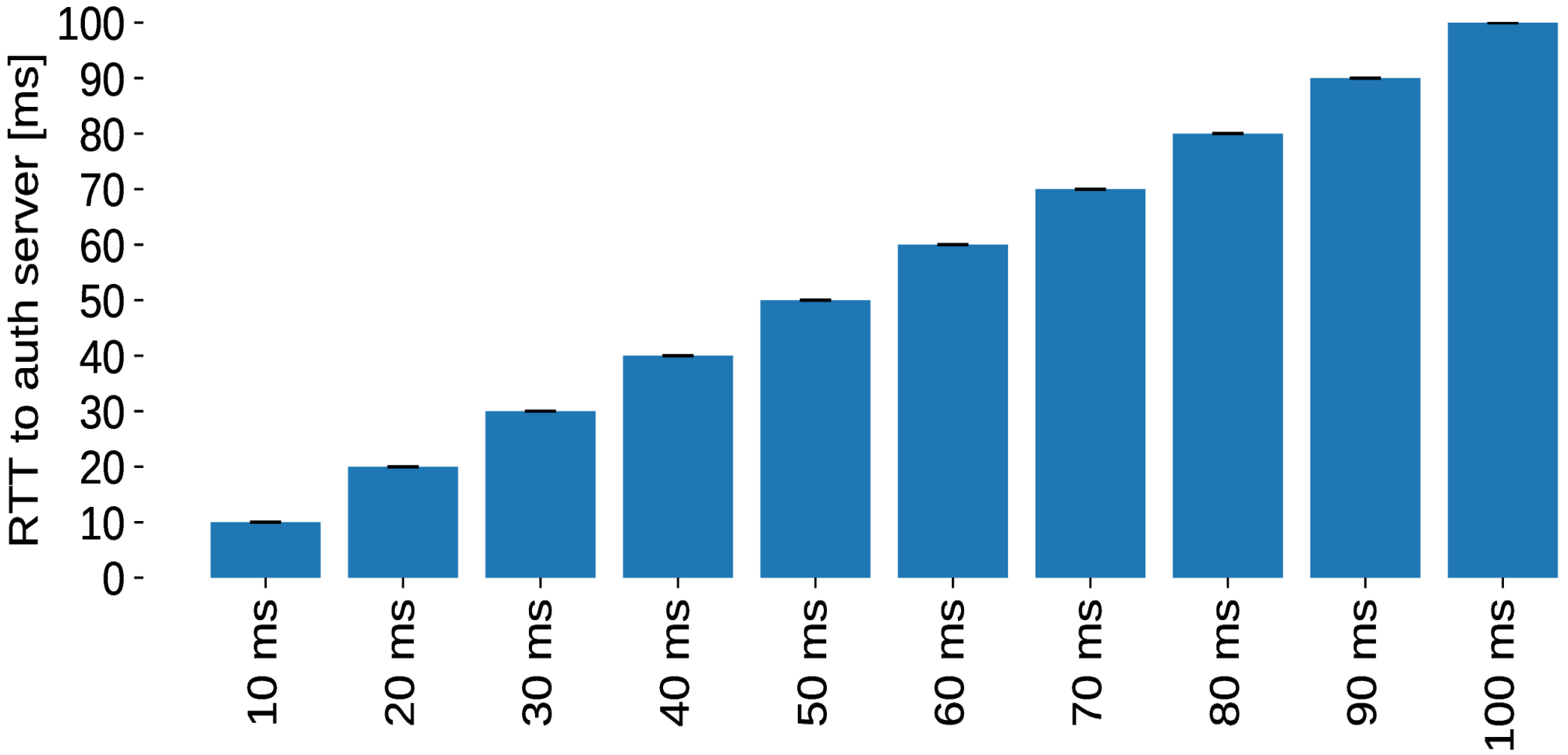
Test scenarios

- Laboratory
  - Hand-crafted
  - Simulating various errors
- Real-life
  - Generated from RIPE Atlas DNS measurements
  - *"Resolver running in networks around the world"*

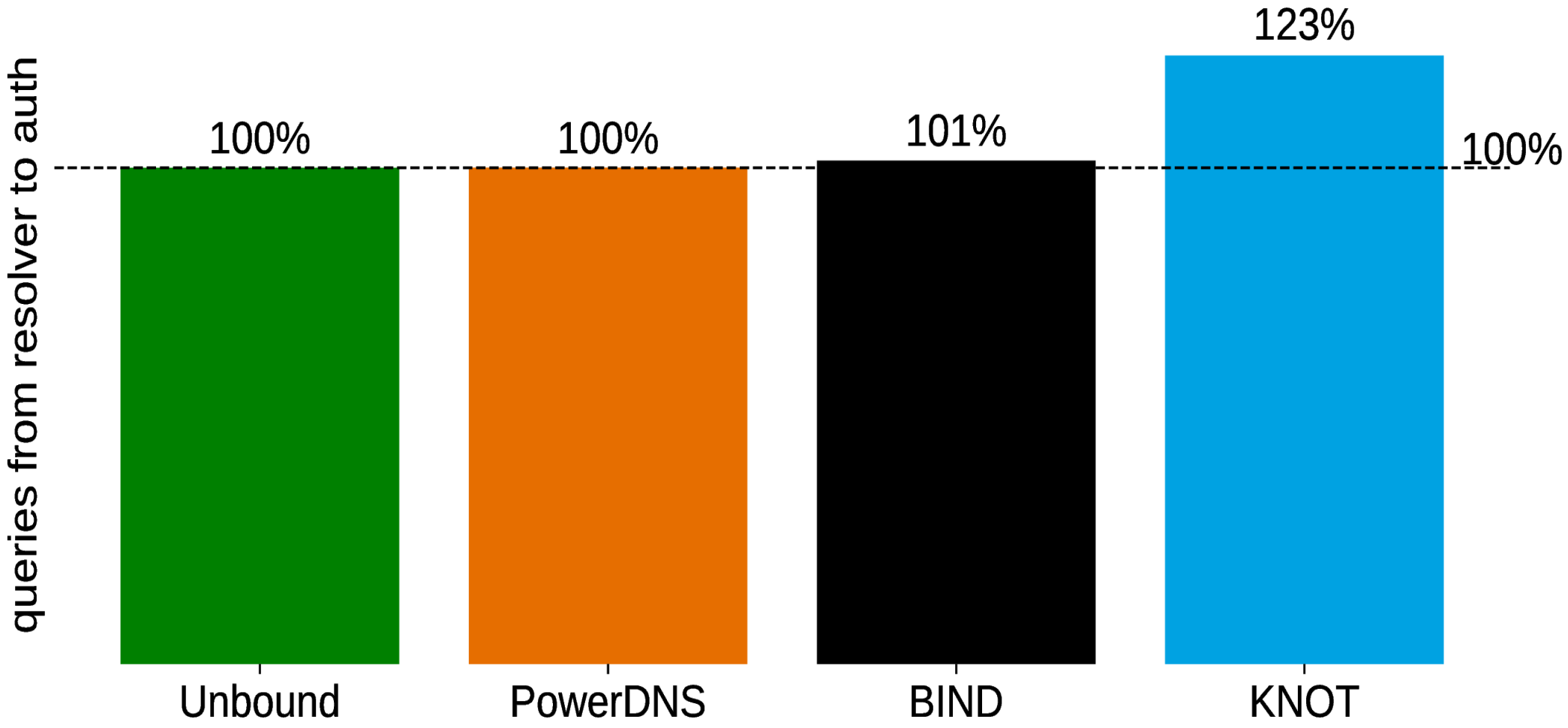
Part 2a

Laboratory

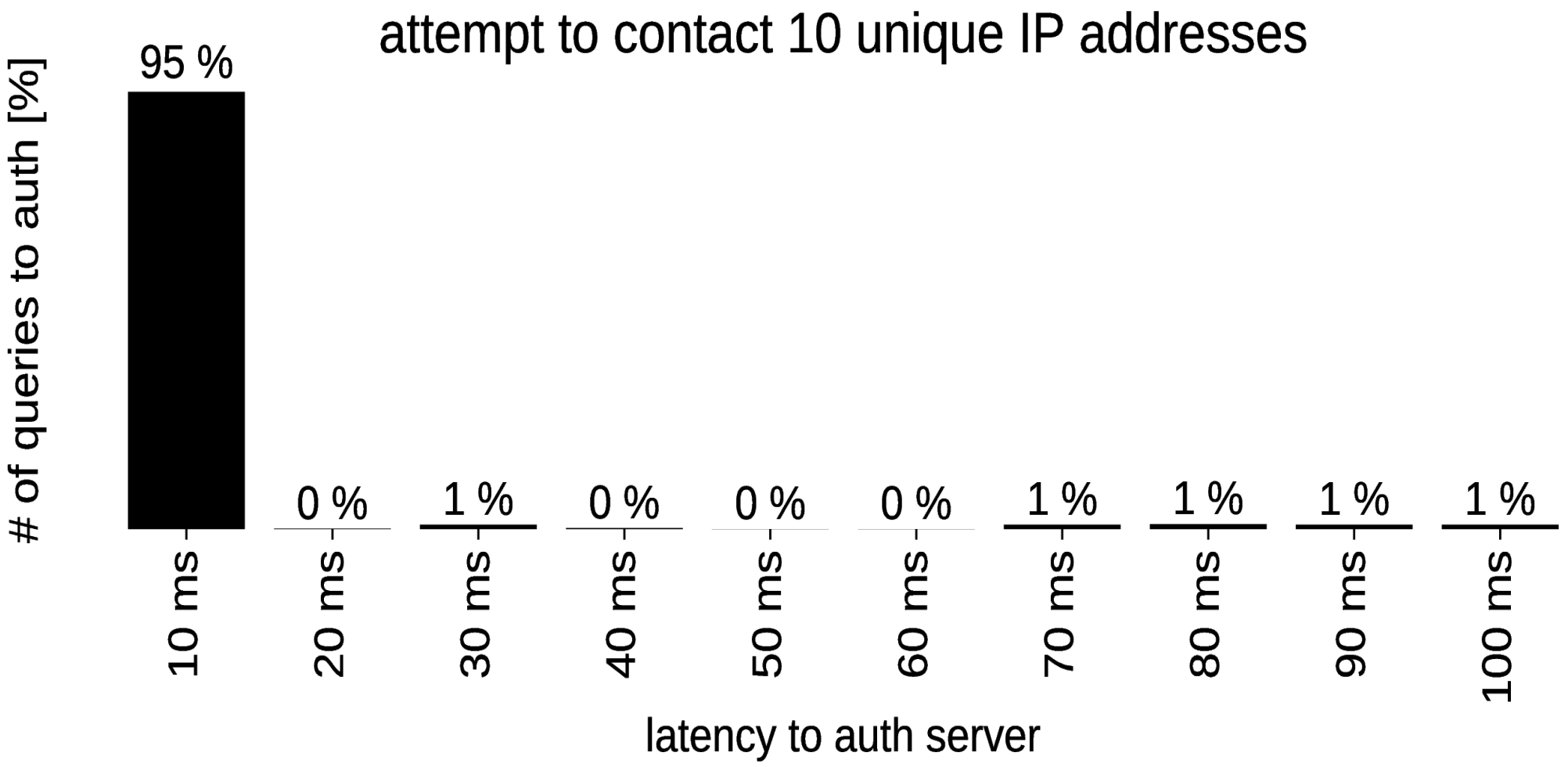
10 x 10 ms steps: auth servers



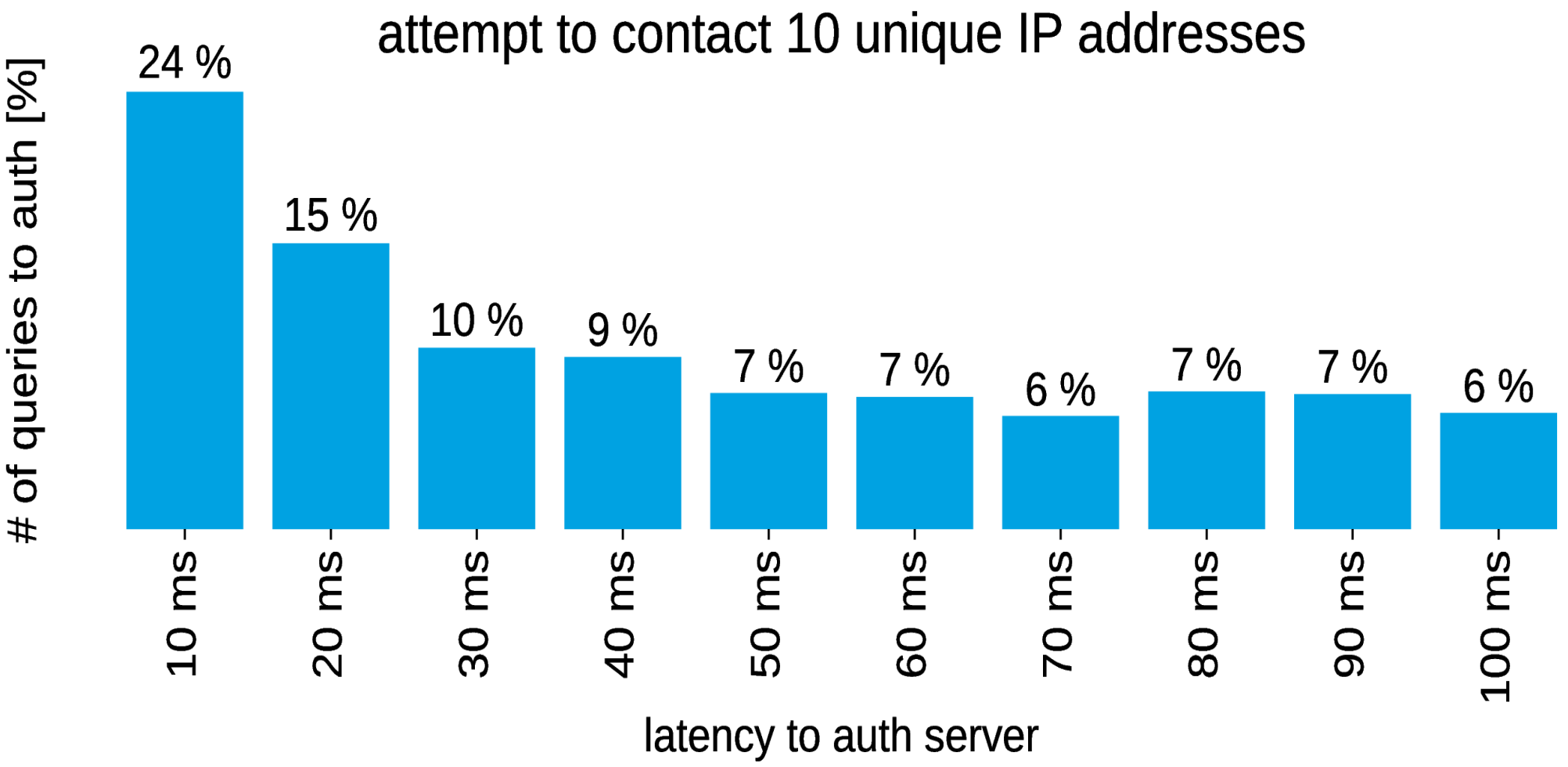
10 x 10 ms steps: # of queries to auth



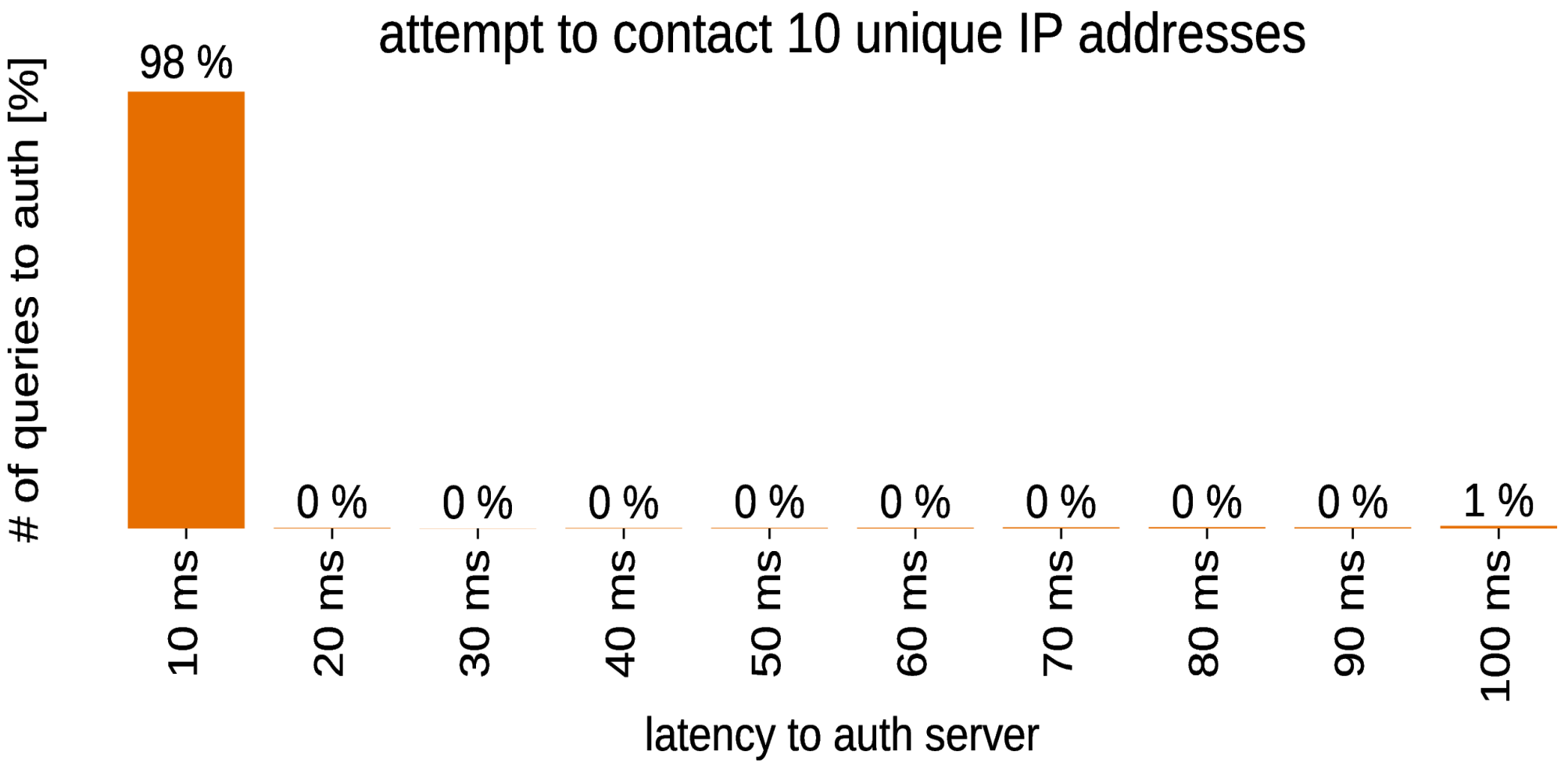
10 x 10 ms steps: BIND packets to auth



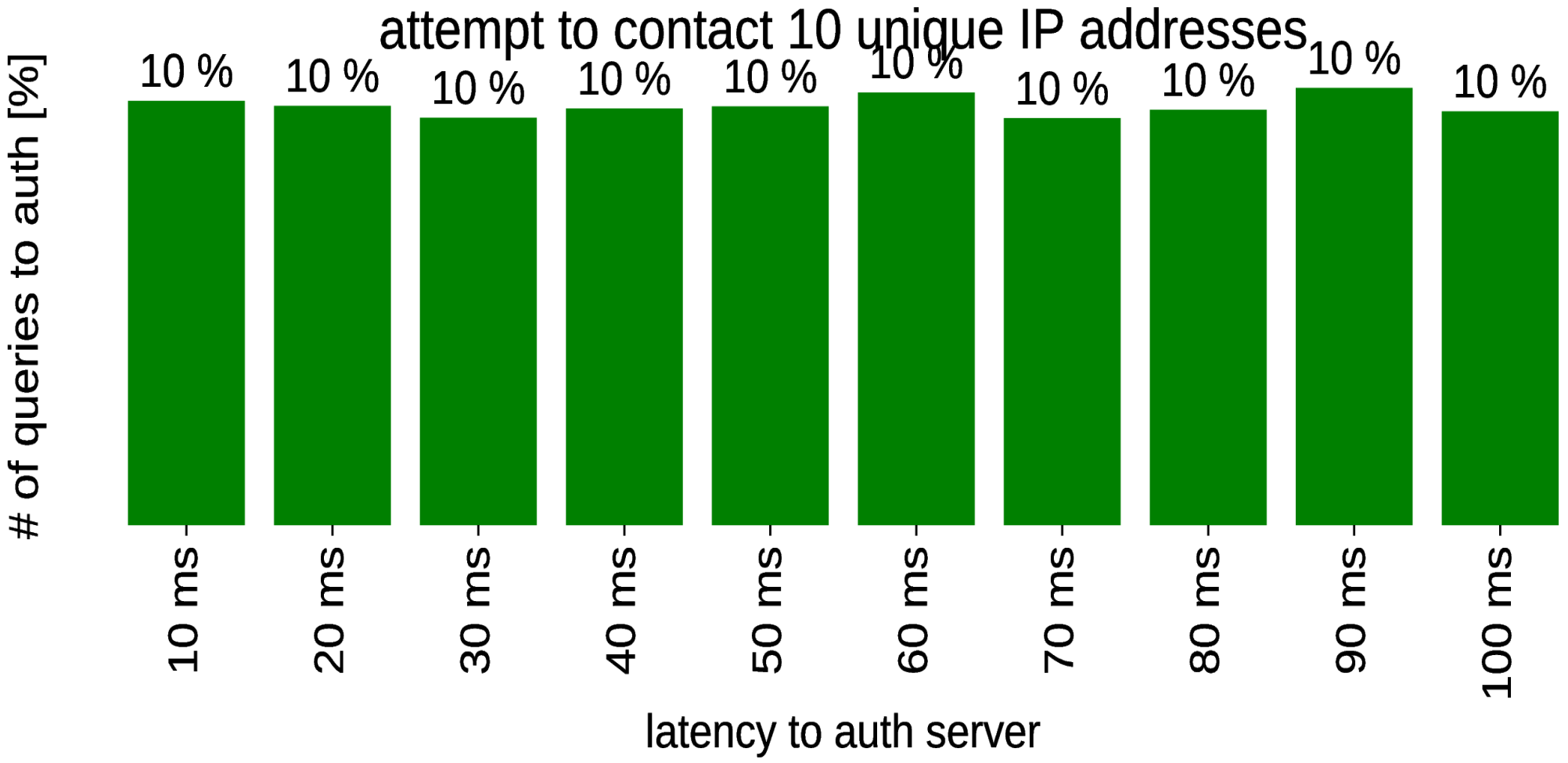
10 x 10 ms steps: KNOT packets to auth



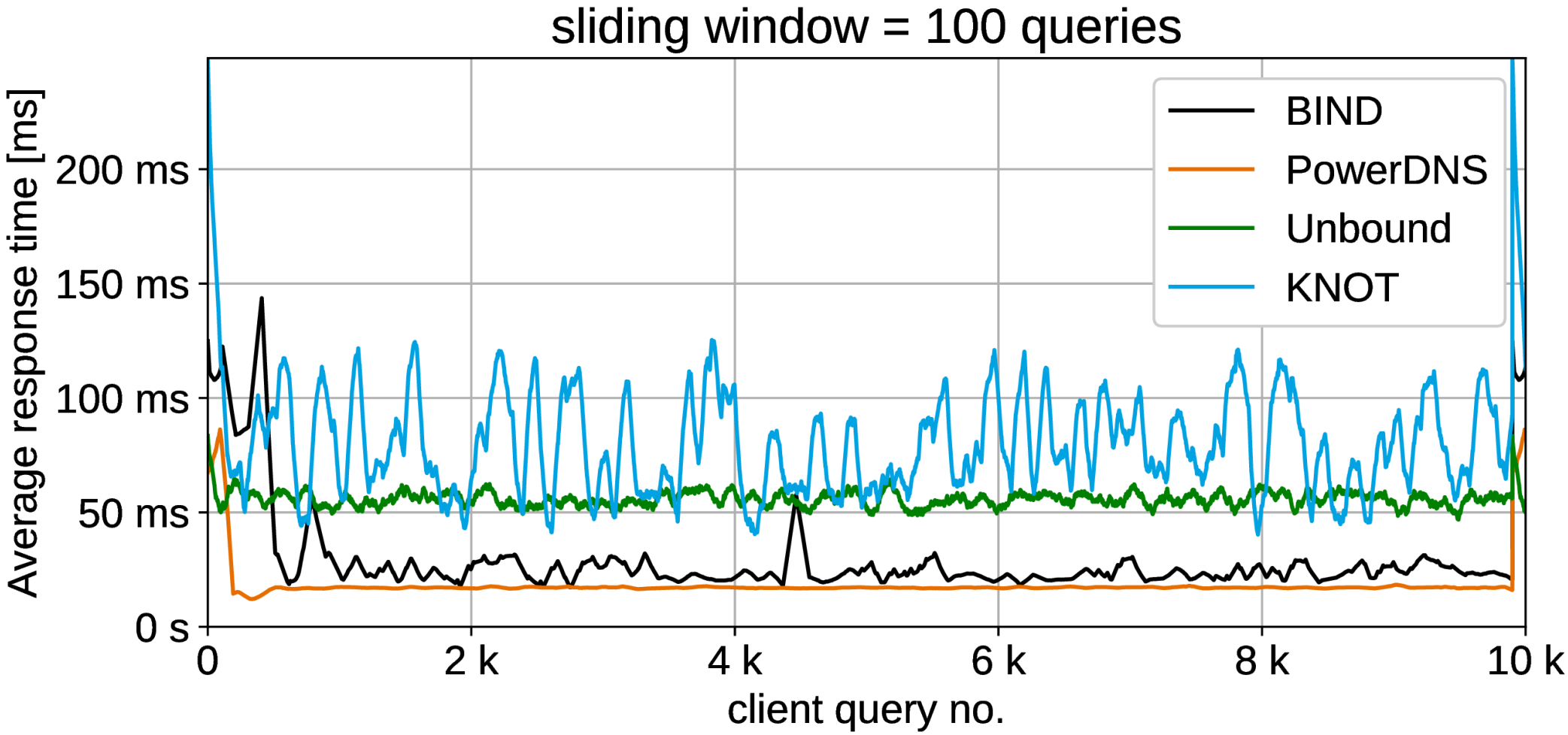
10 x 10 ms steps: PowerDNS packets to auth



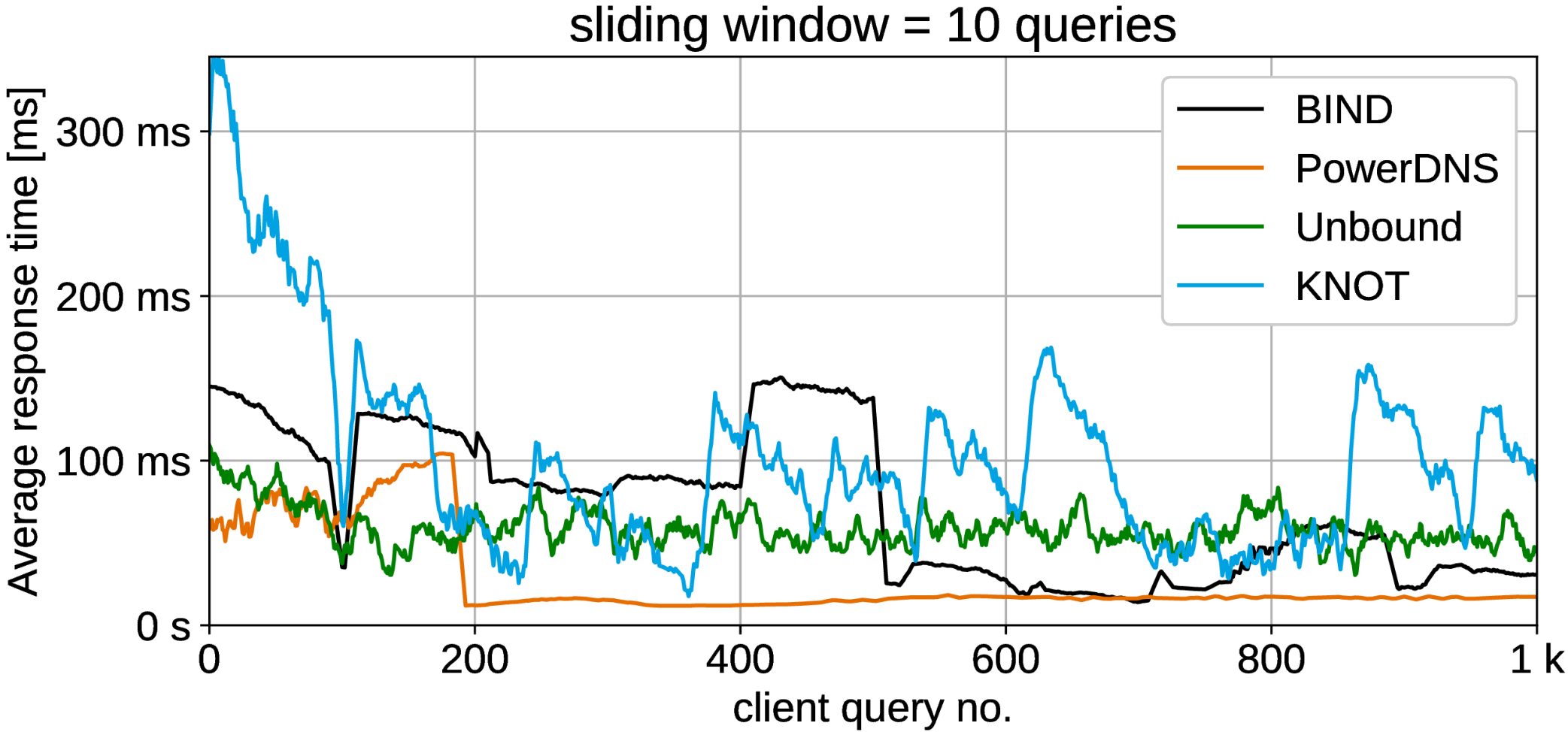
10 x 10 ms steps: Unbound packets to auth



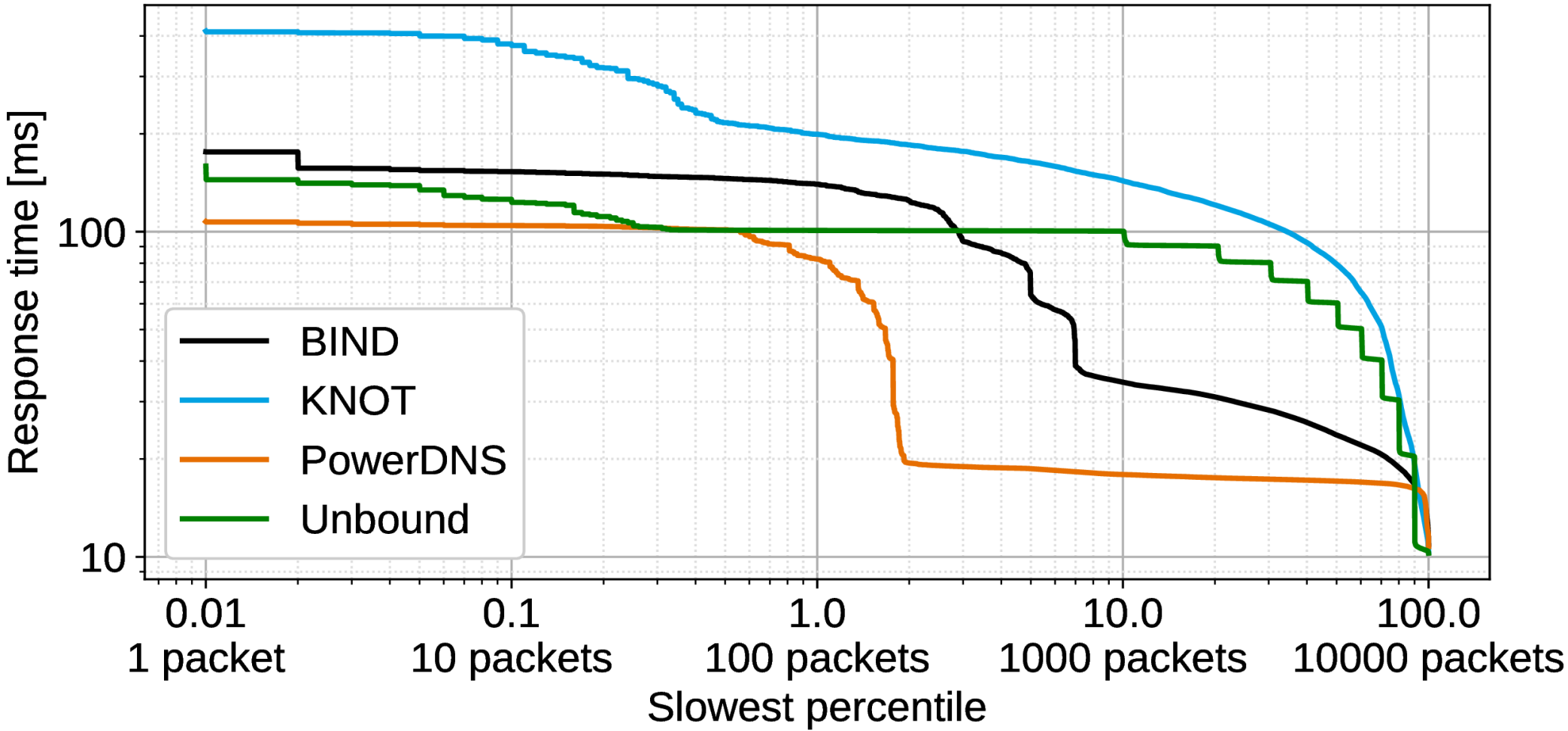
10 x 10 ms steps: latency over time



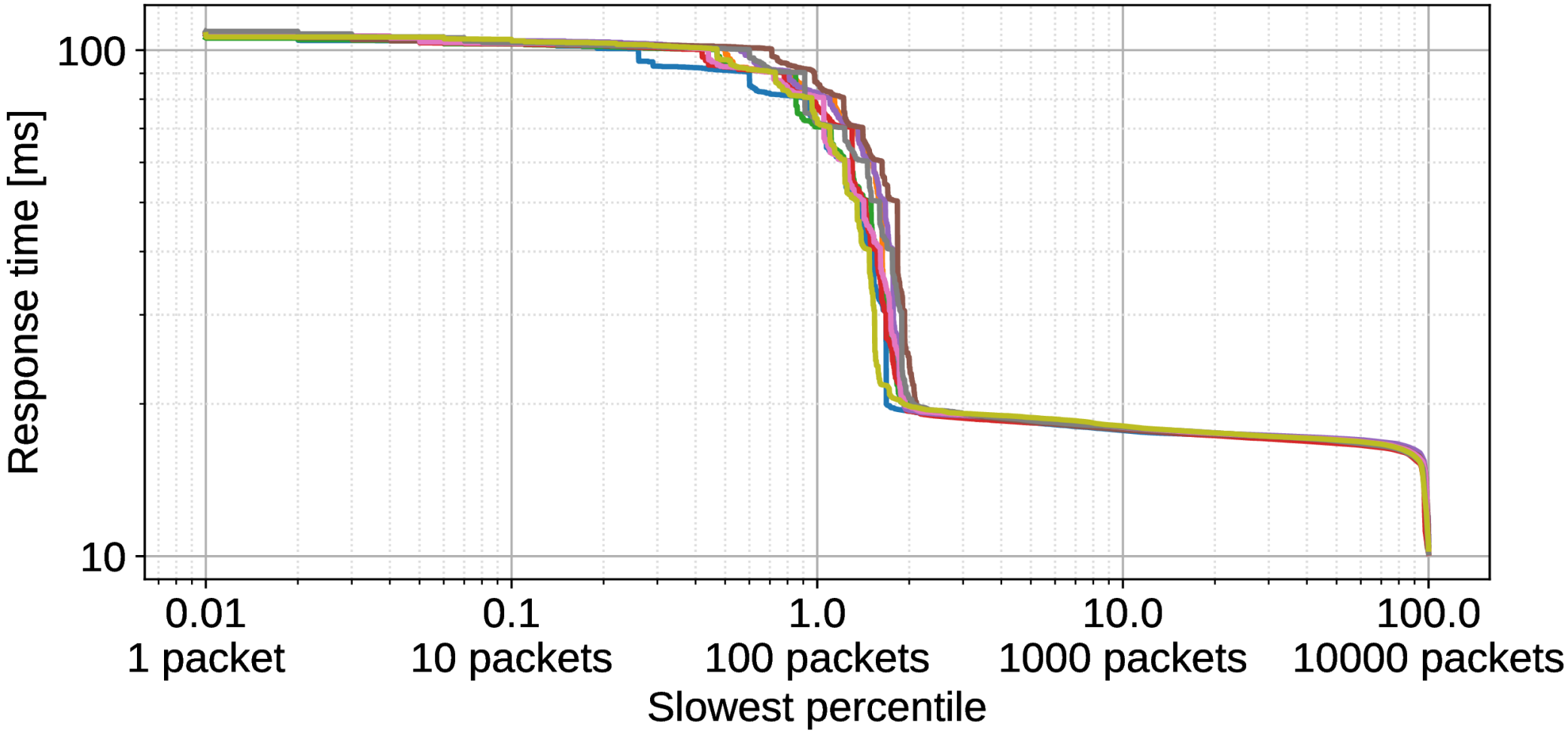
10 x 10 ms steps: latency over time (zoom)



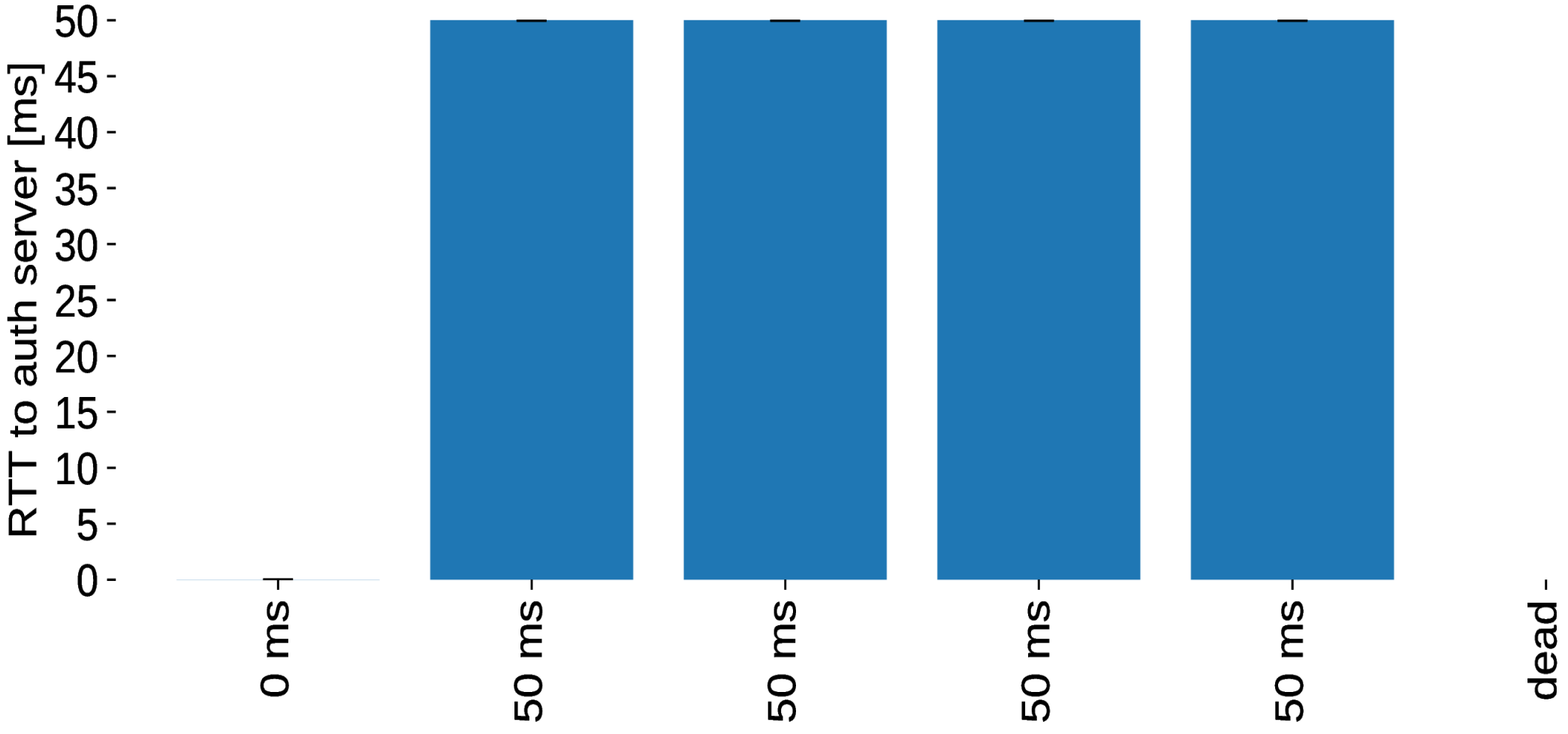
10 x 10 ms steps: latency histogram



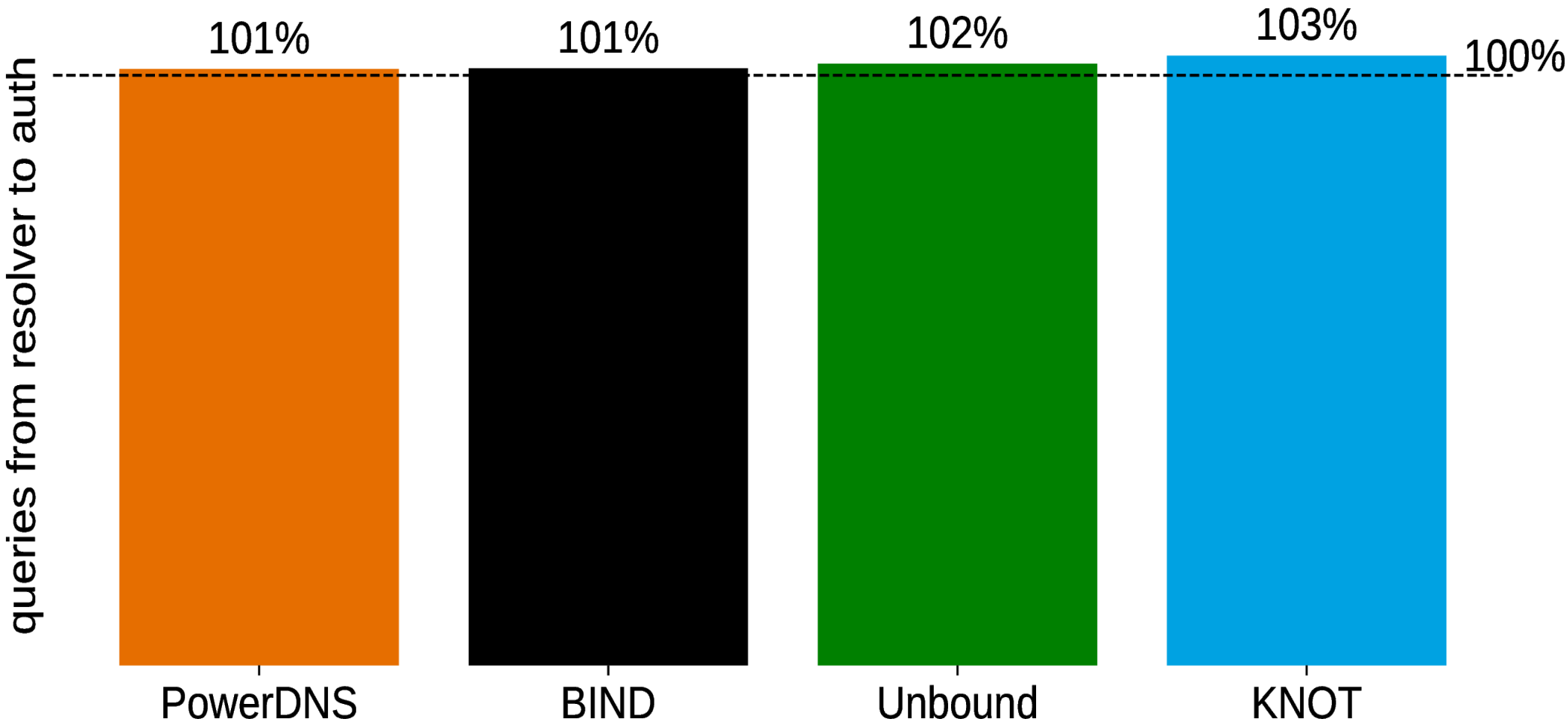
Test stability? 10 rounds of PowerDNS



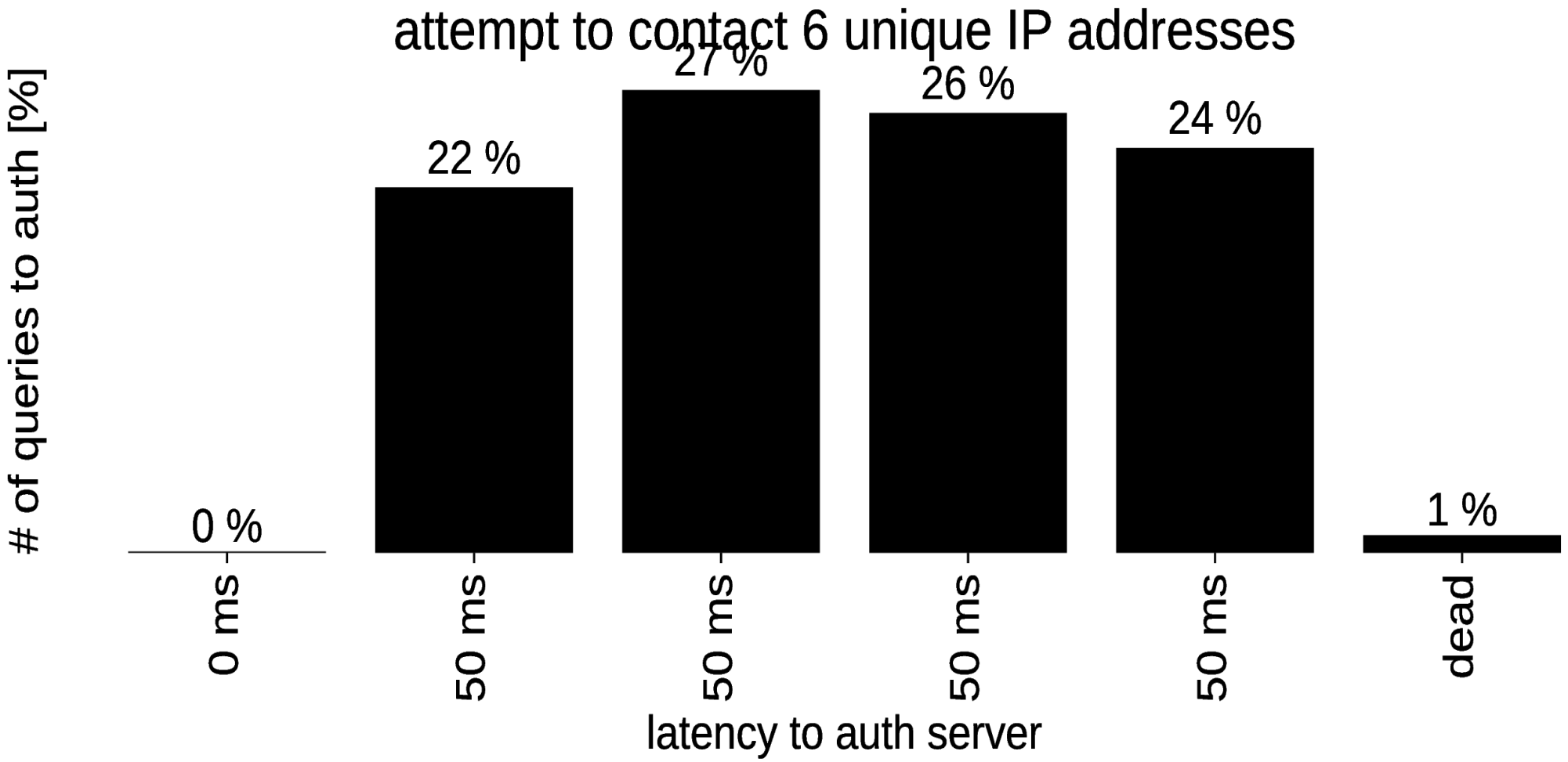
1 NS timeouts: auth servers



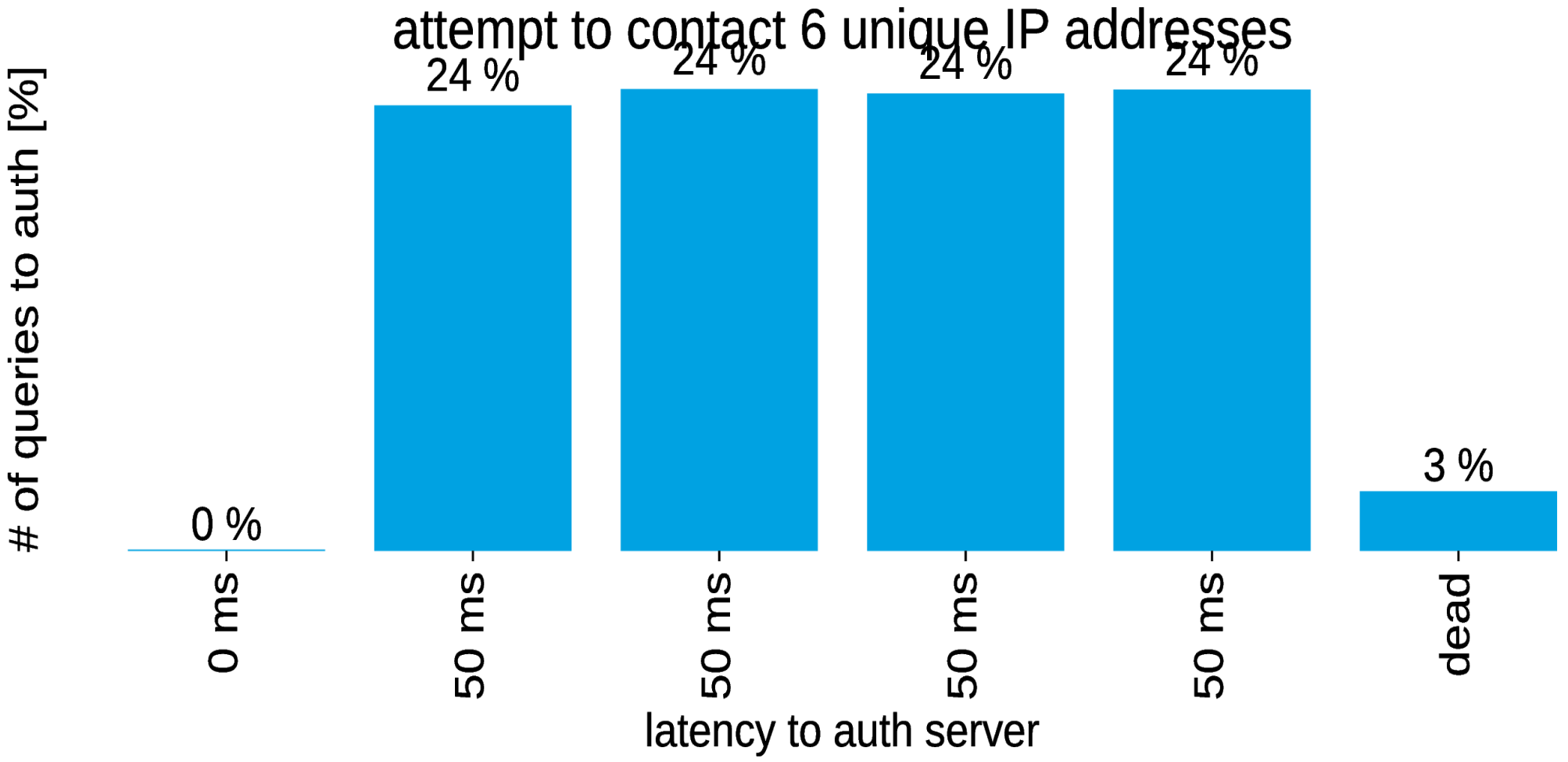
1 NS timeouts: # of queries to auth



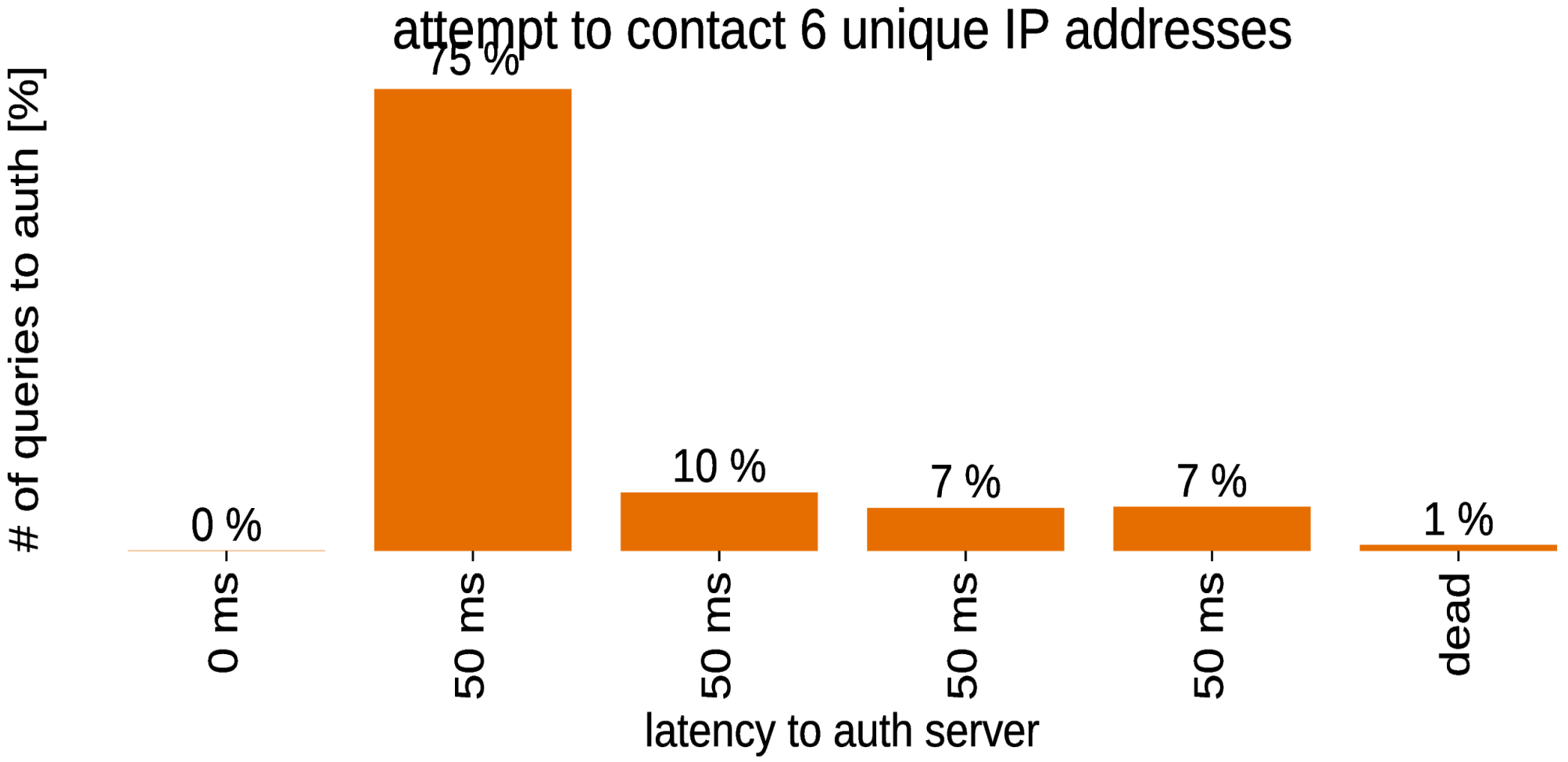
1 NS timeouts: BIND packets to auth



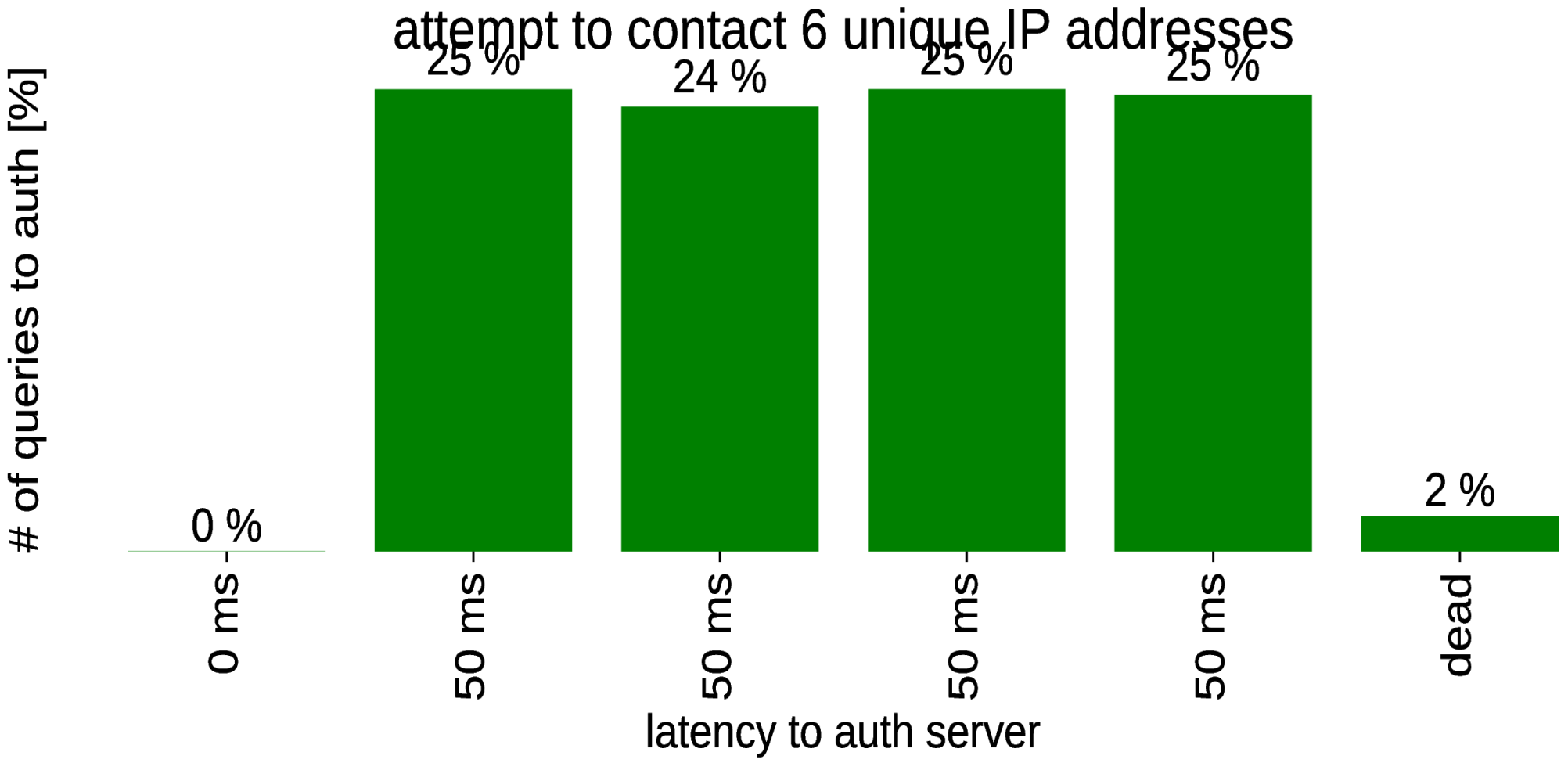
1 NS timeouts: KNOT packets to auth



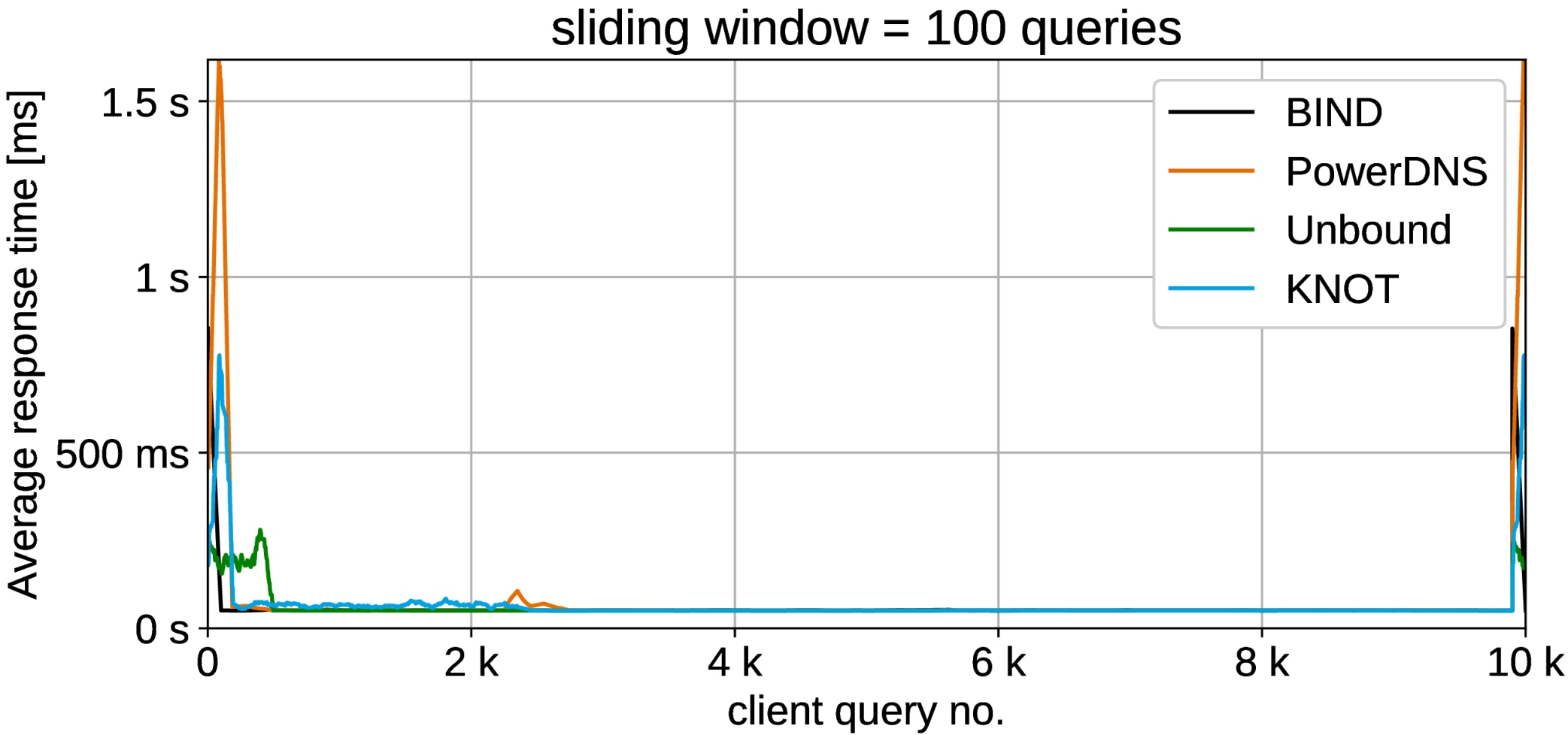
1 NS timeouts: PowerDNS packets to auth



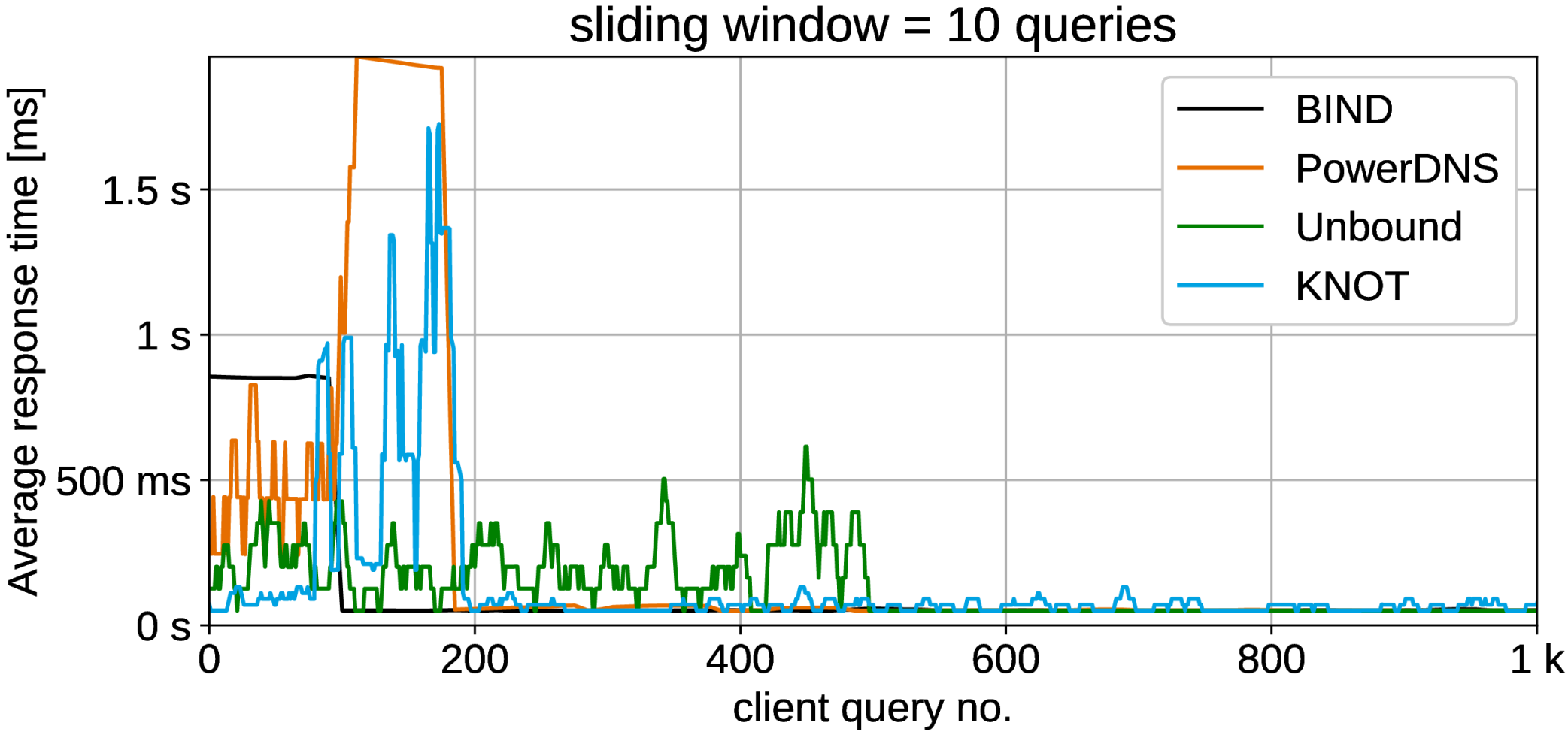
1 NS timeouts: Unbound packets to auth



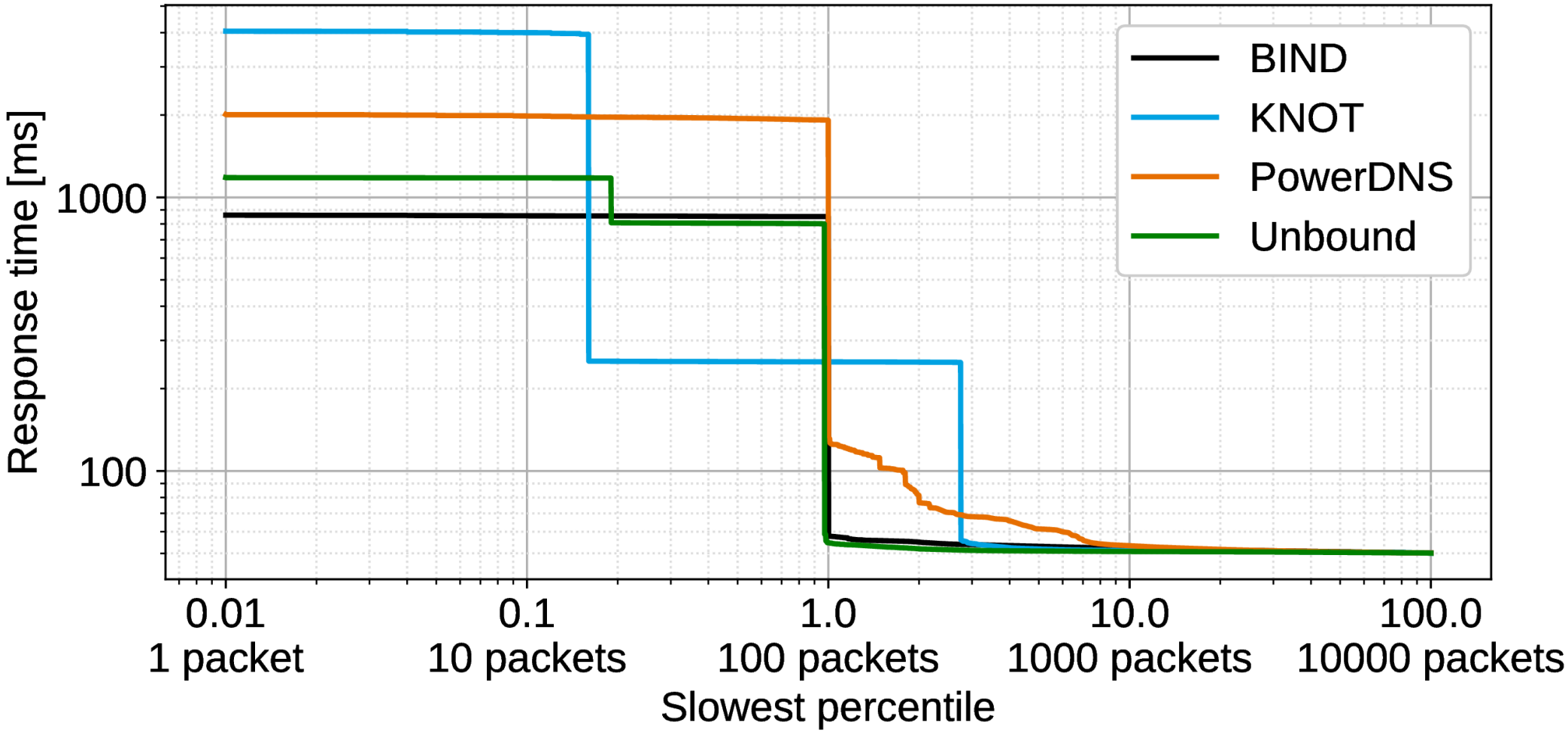
1 NS timeouts: latency over time



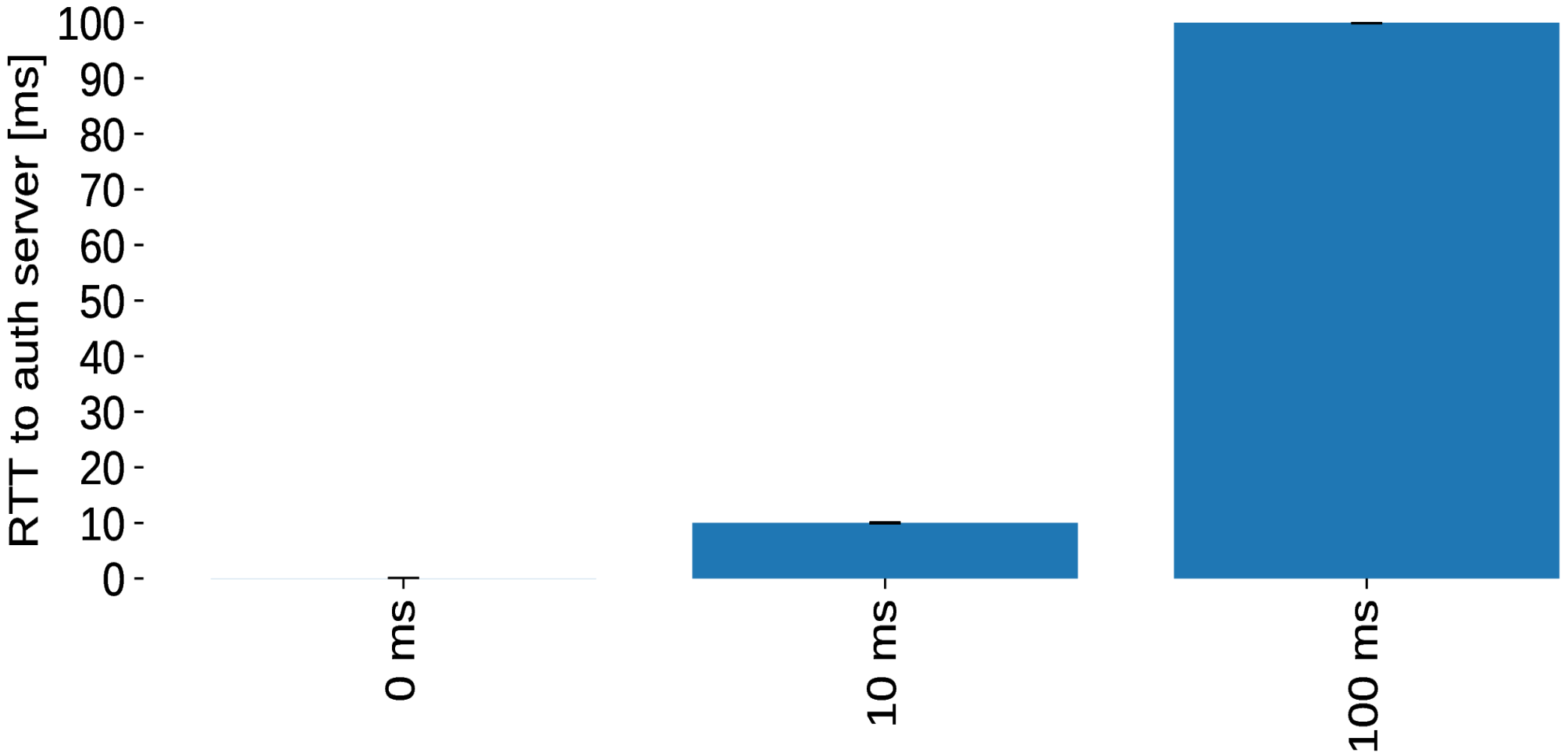
1 NS timeouts: latency over time (zoom)



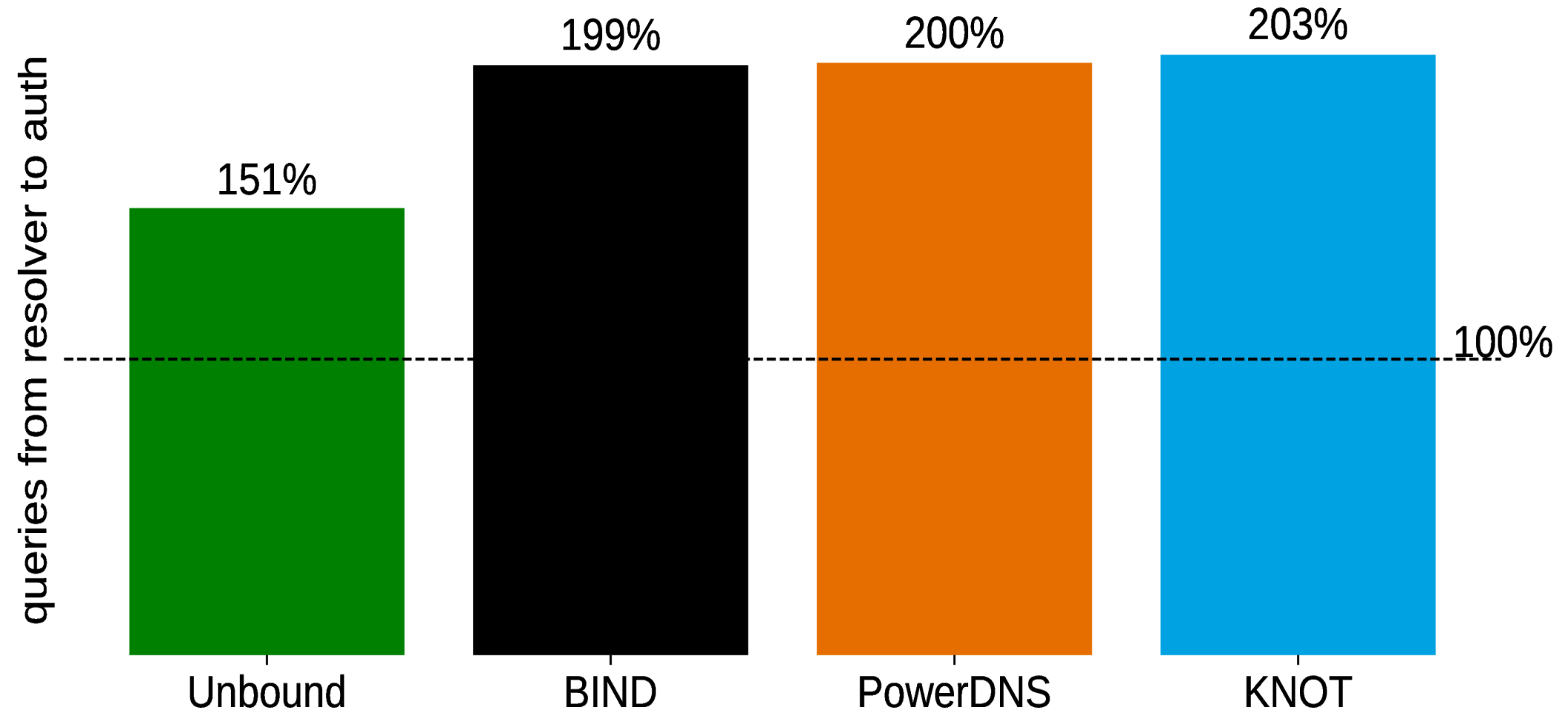
1 NS timeouts: latency histogram



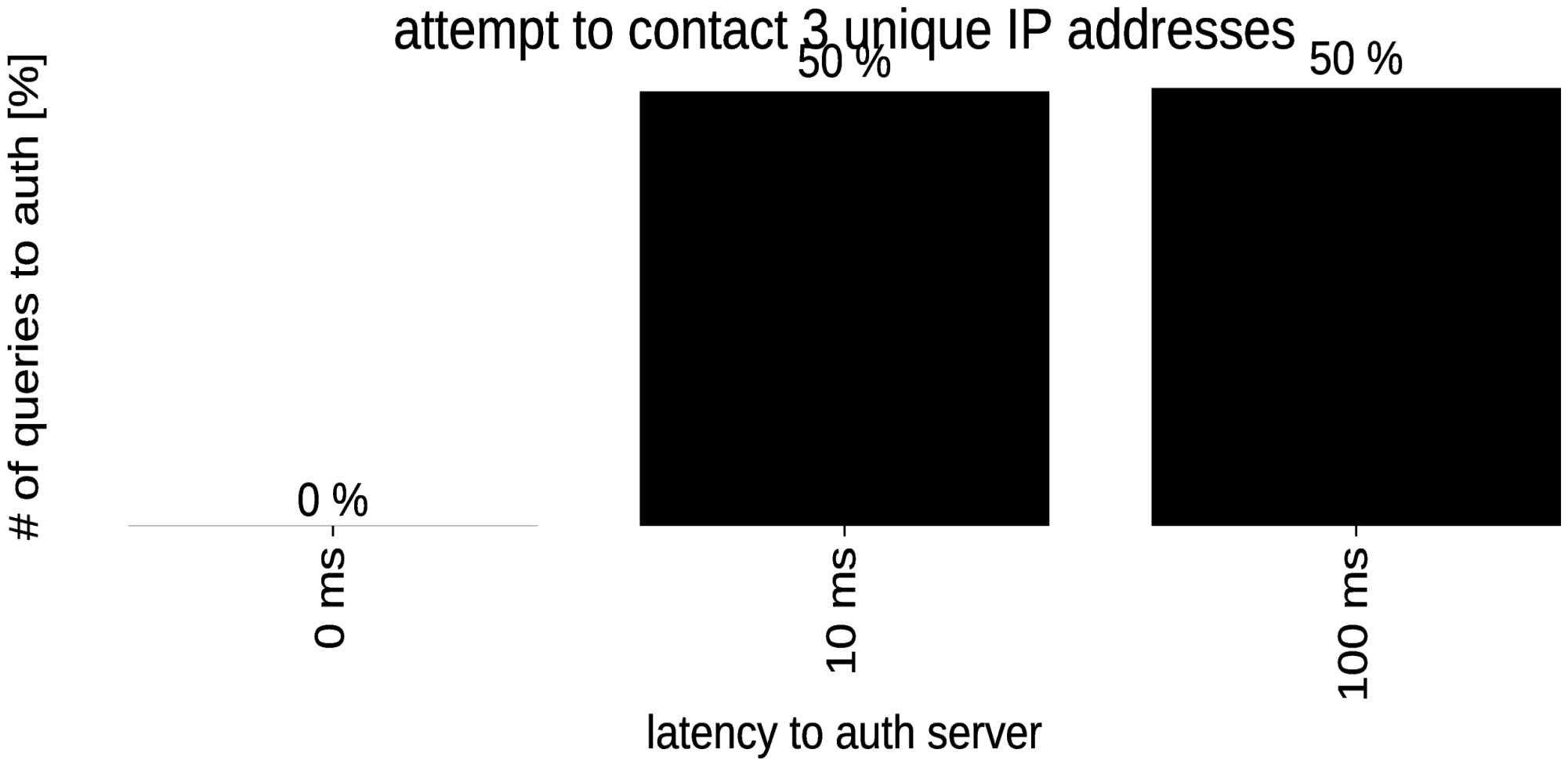
fast DNSSEC expired, slow OK: auth servers



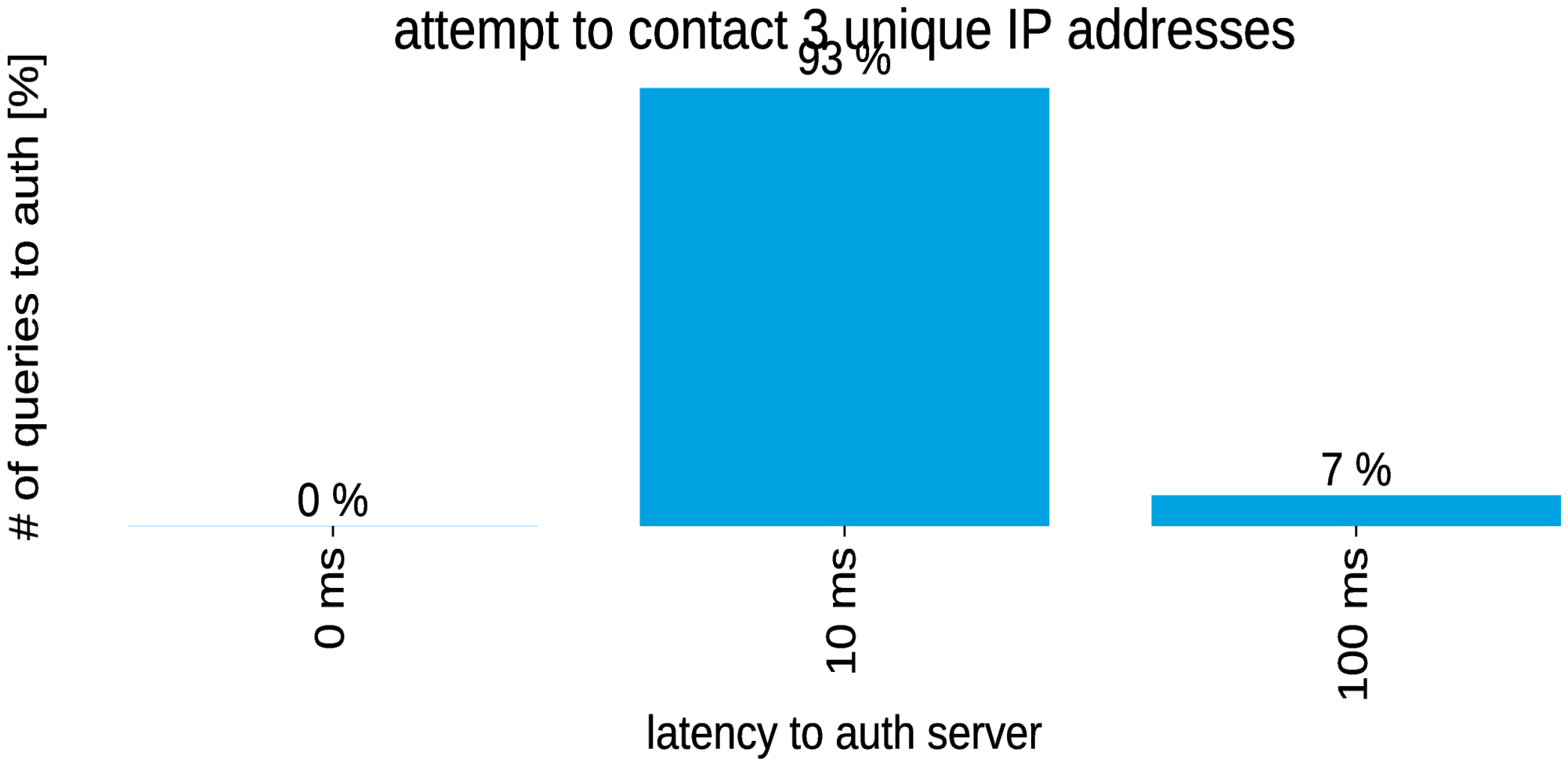
fast DNSSEC expired, slow OK: # of queries to auth



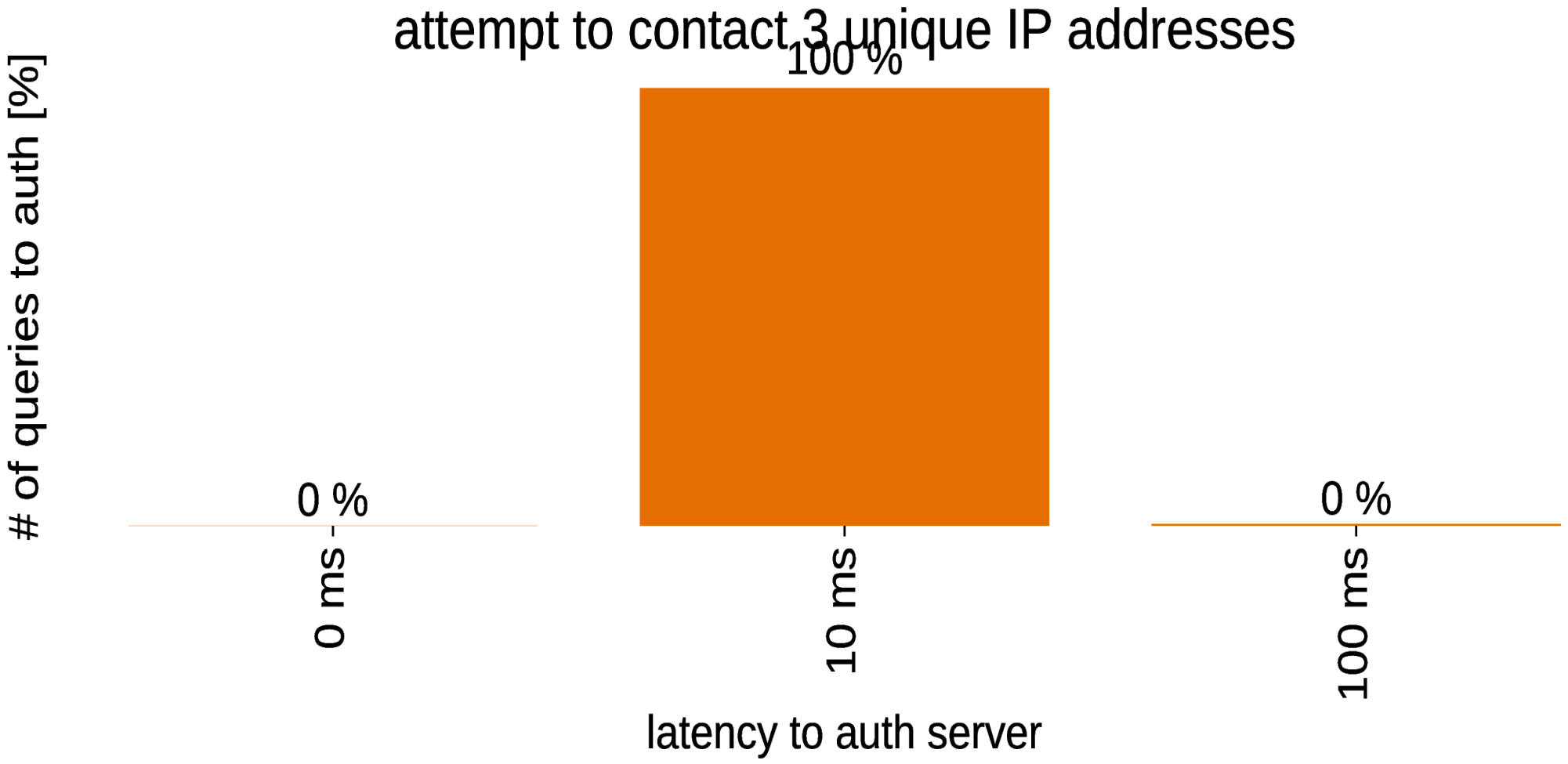
fast DNSSEC expired, slow OK: BIND packets to auth



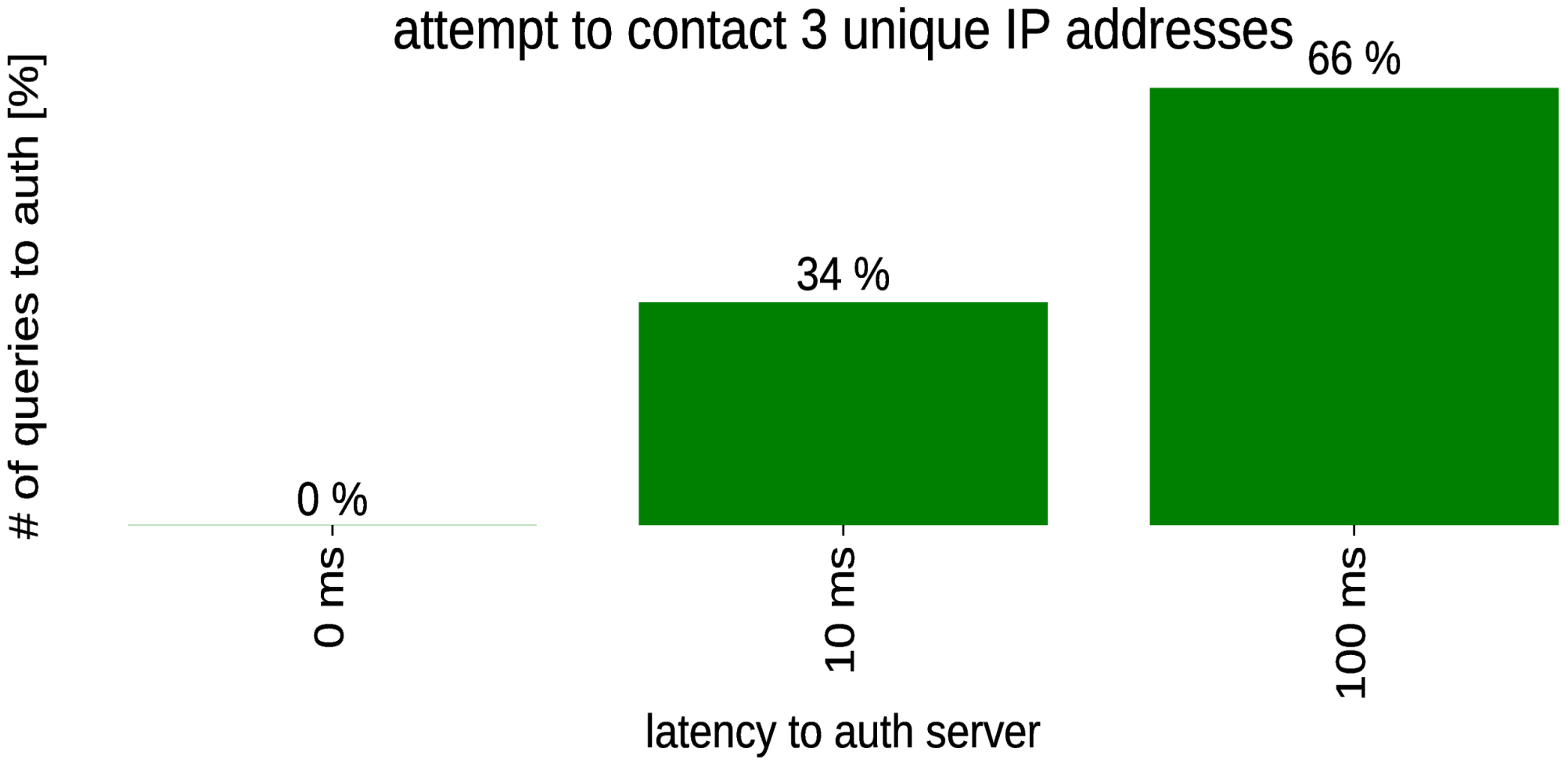
fast DNSSEC expired, slow OK: KNOT packets to auth



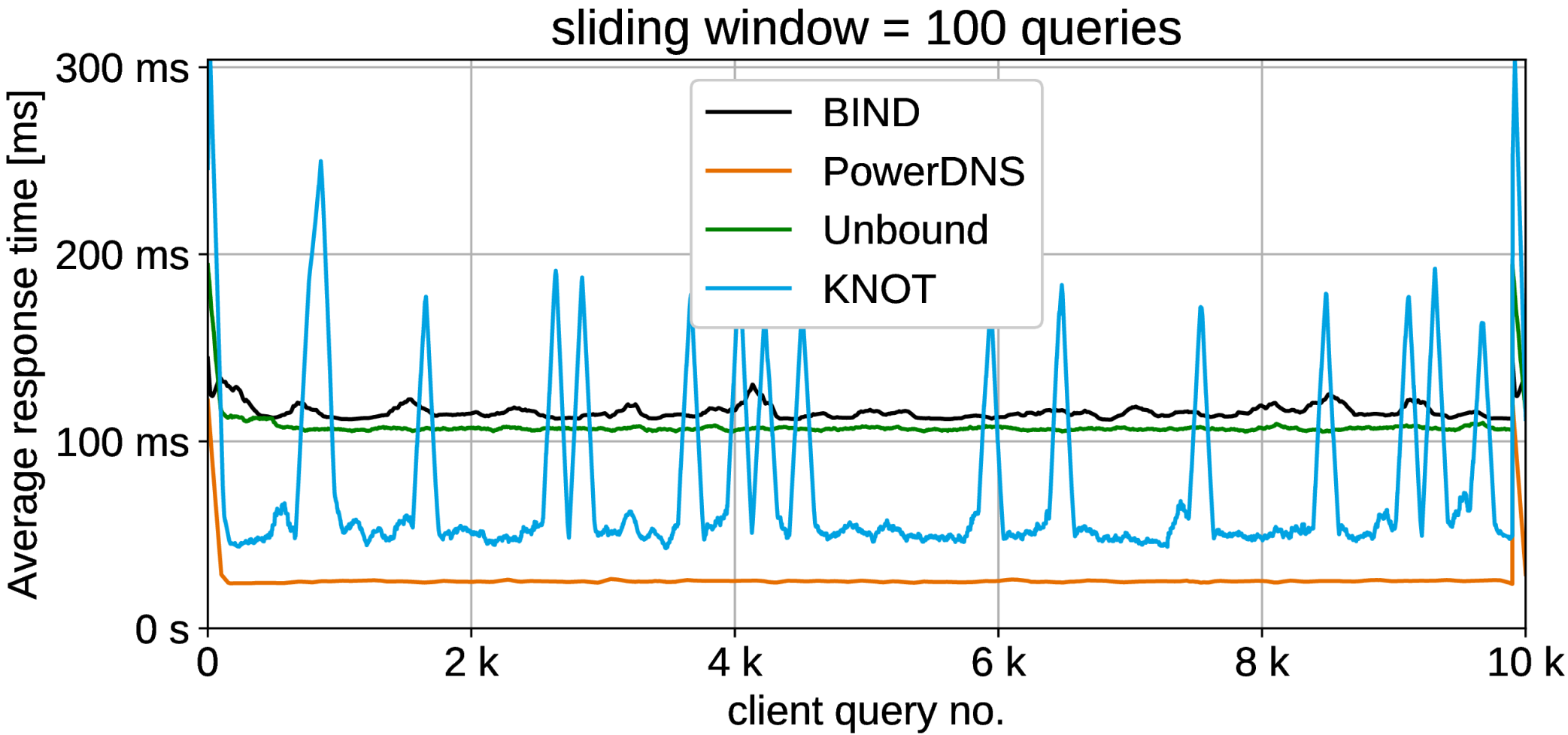
fast DNSSEC expired, slow OK: PowerDNS packets to auth



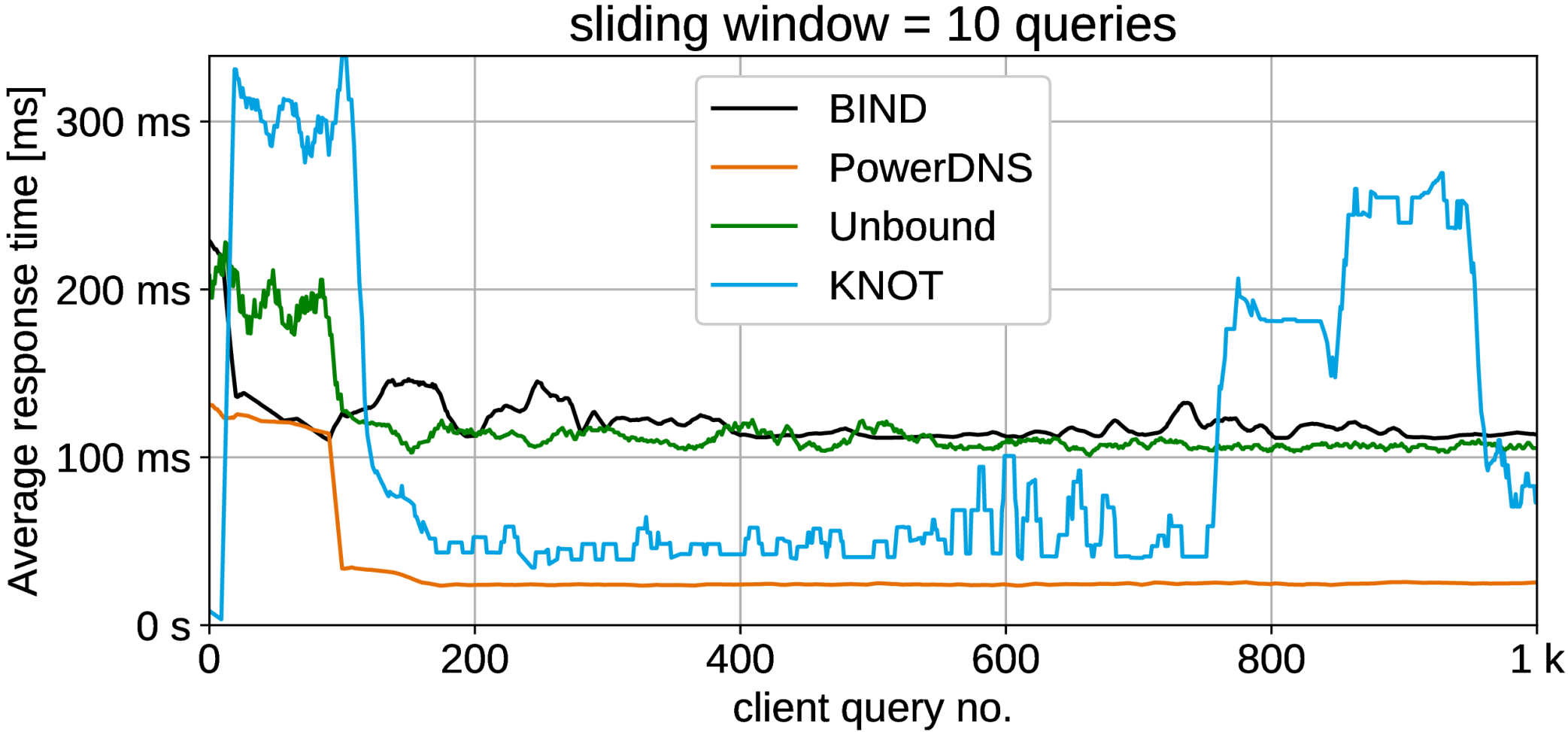
fast DNSSEC expired, slow OK: Unbound packets to auth



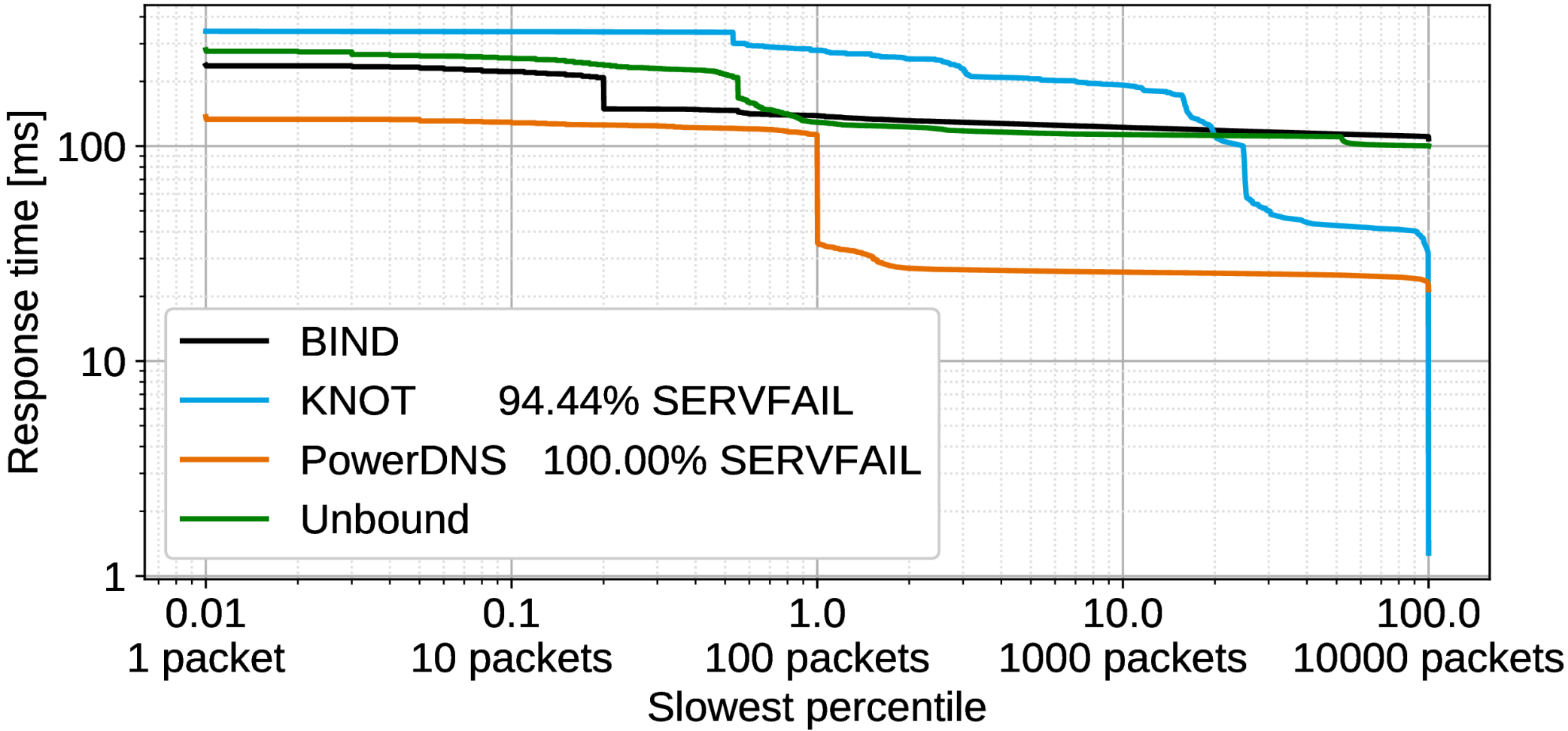
fast DNSSEC expired, slow OK: latency over time



fast DNSSEC expired, slow OK: latency over time (zoom)



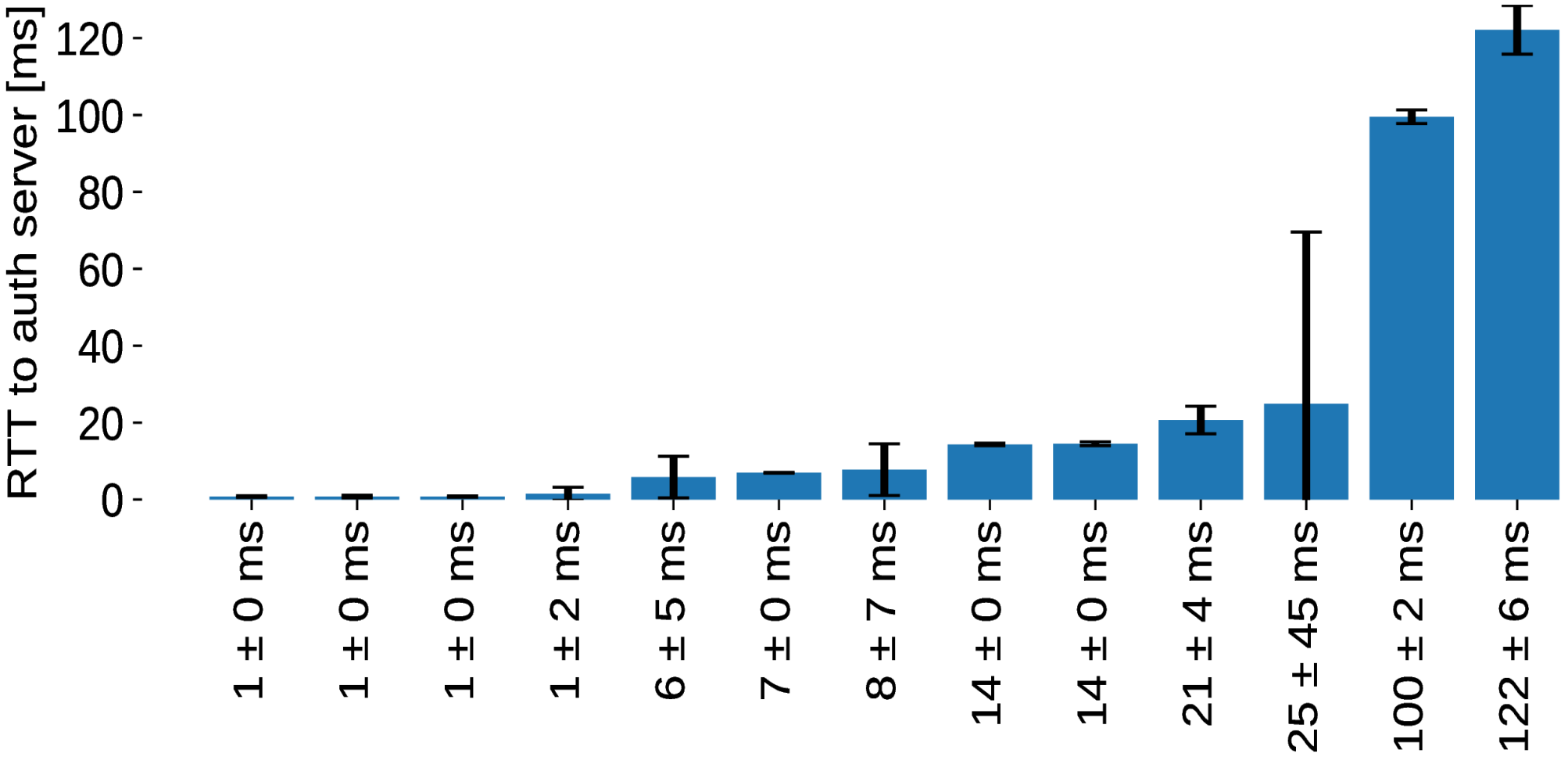
fast DNSSEC expired, slow OK: latency histogram



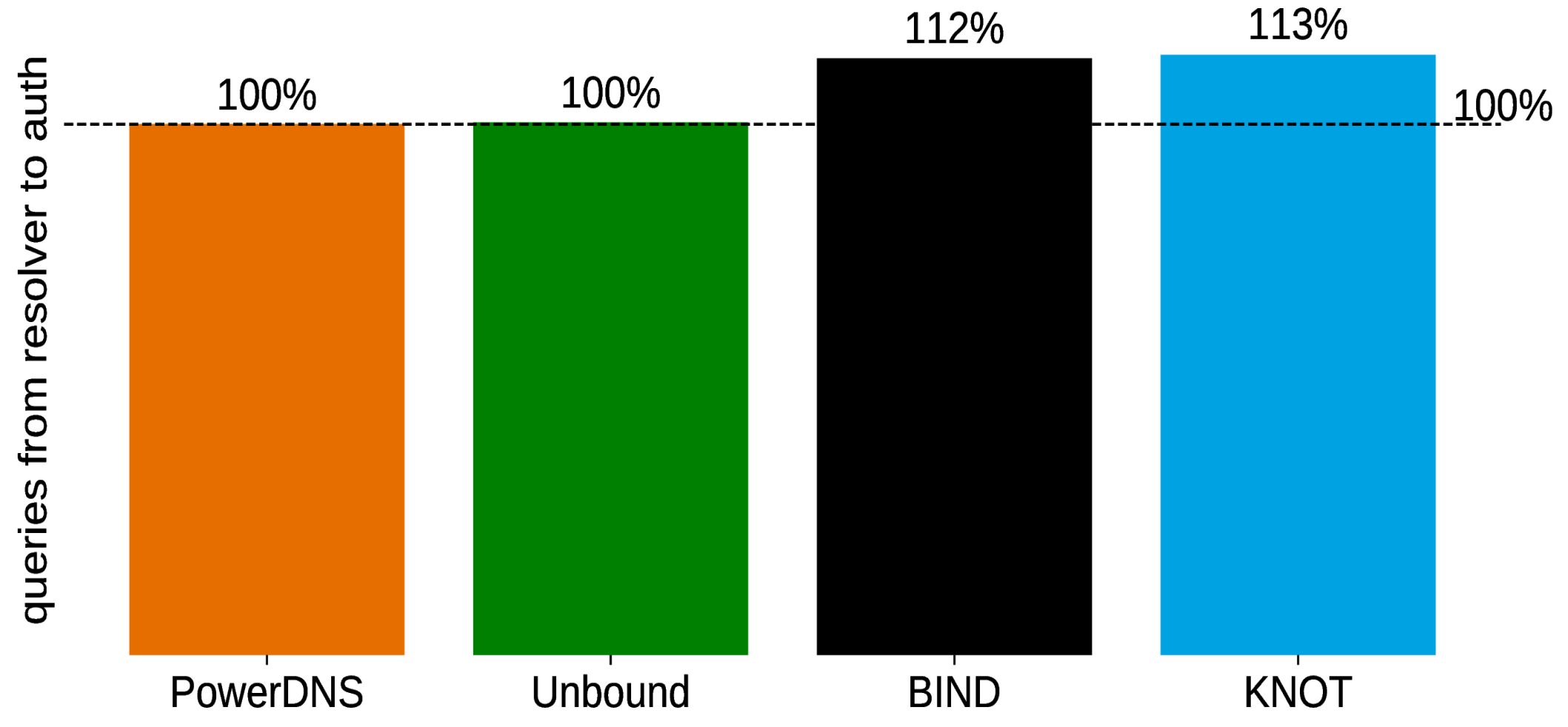
Part 2b

Real world

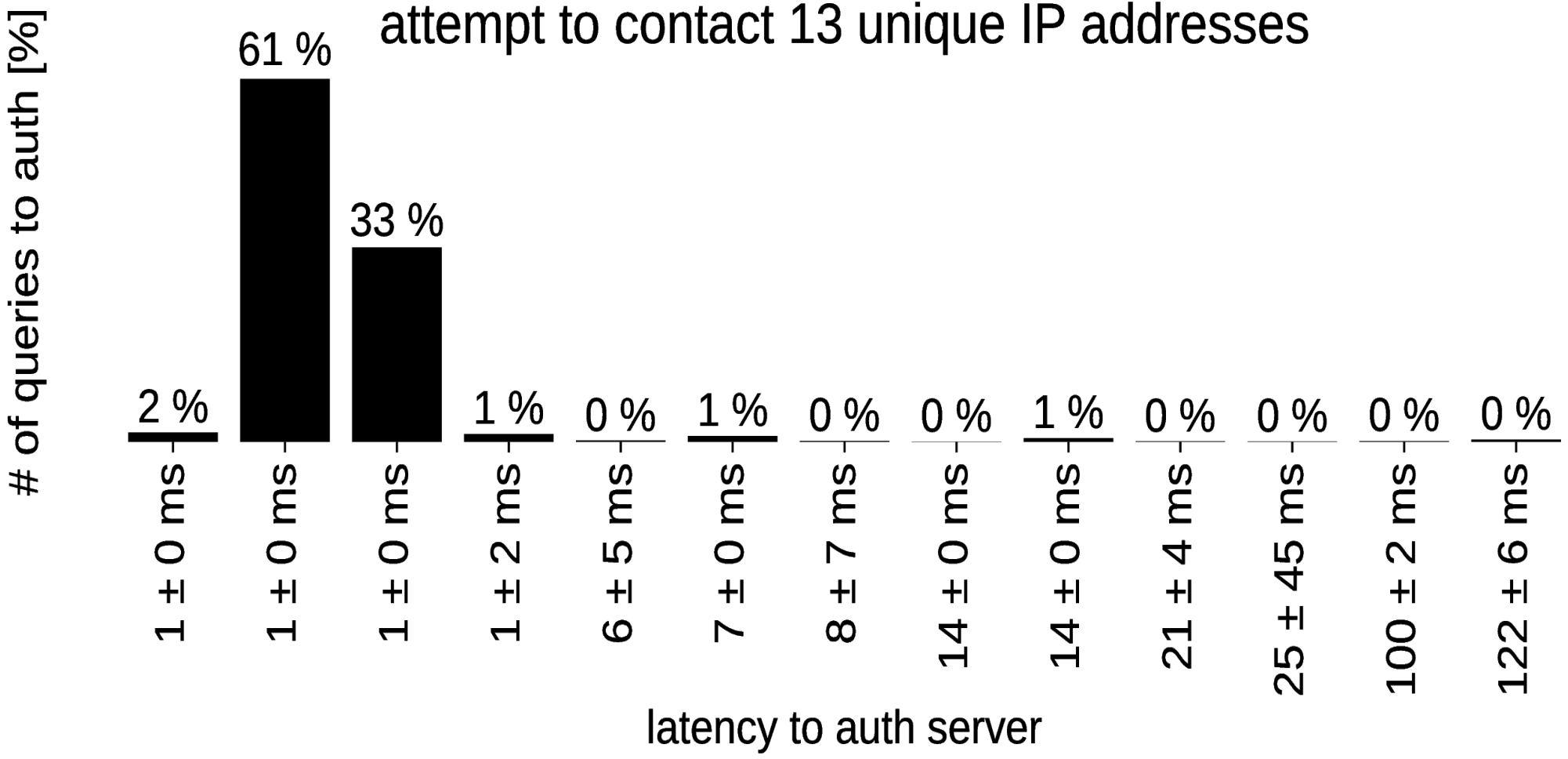
Prague: auth servers



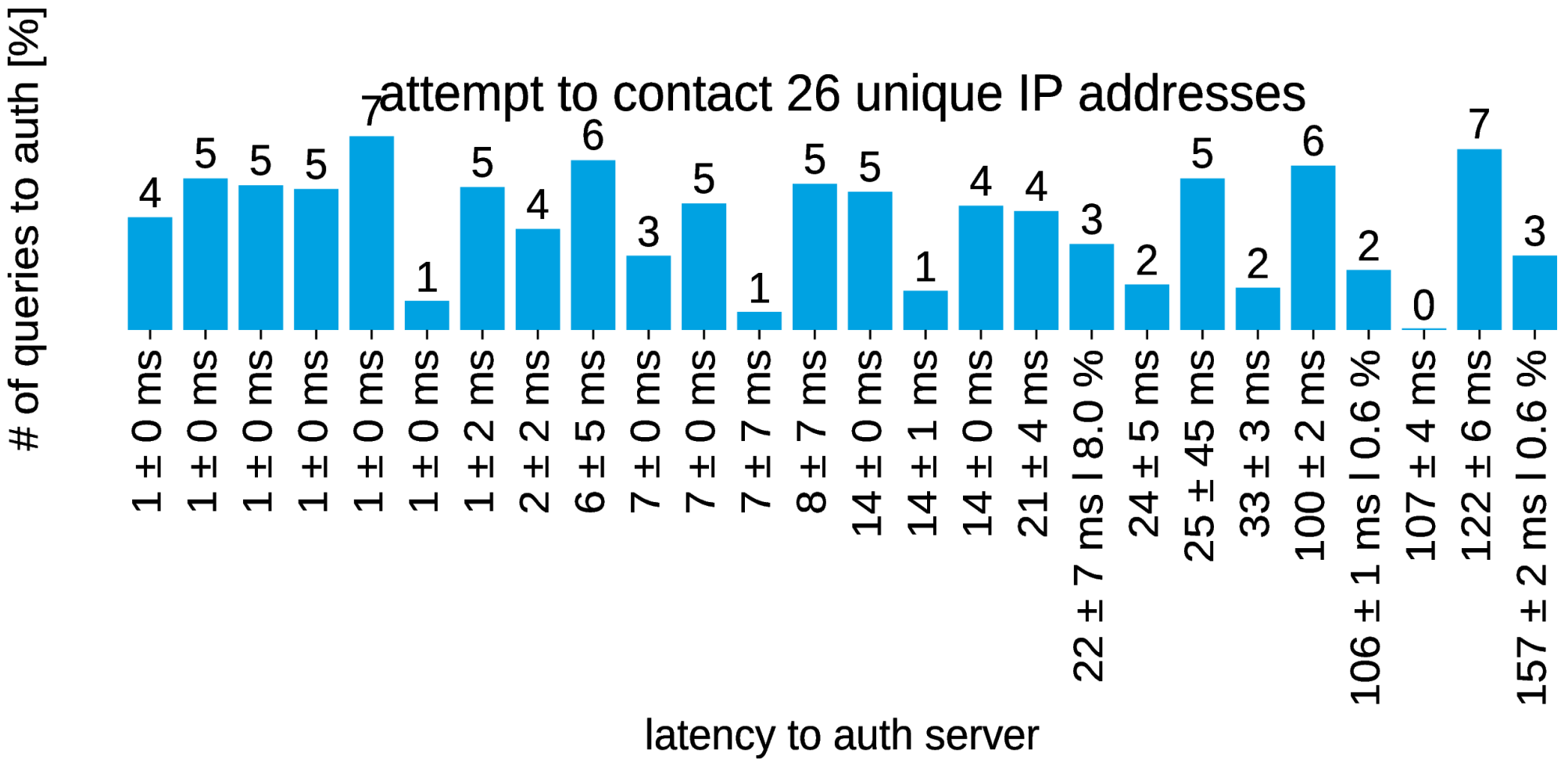
Prague: # of queries to auth



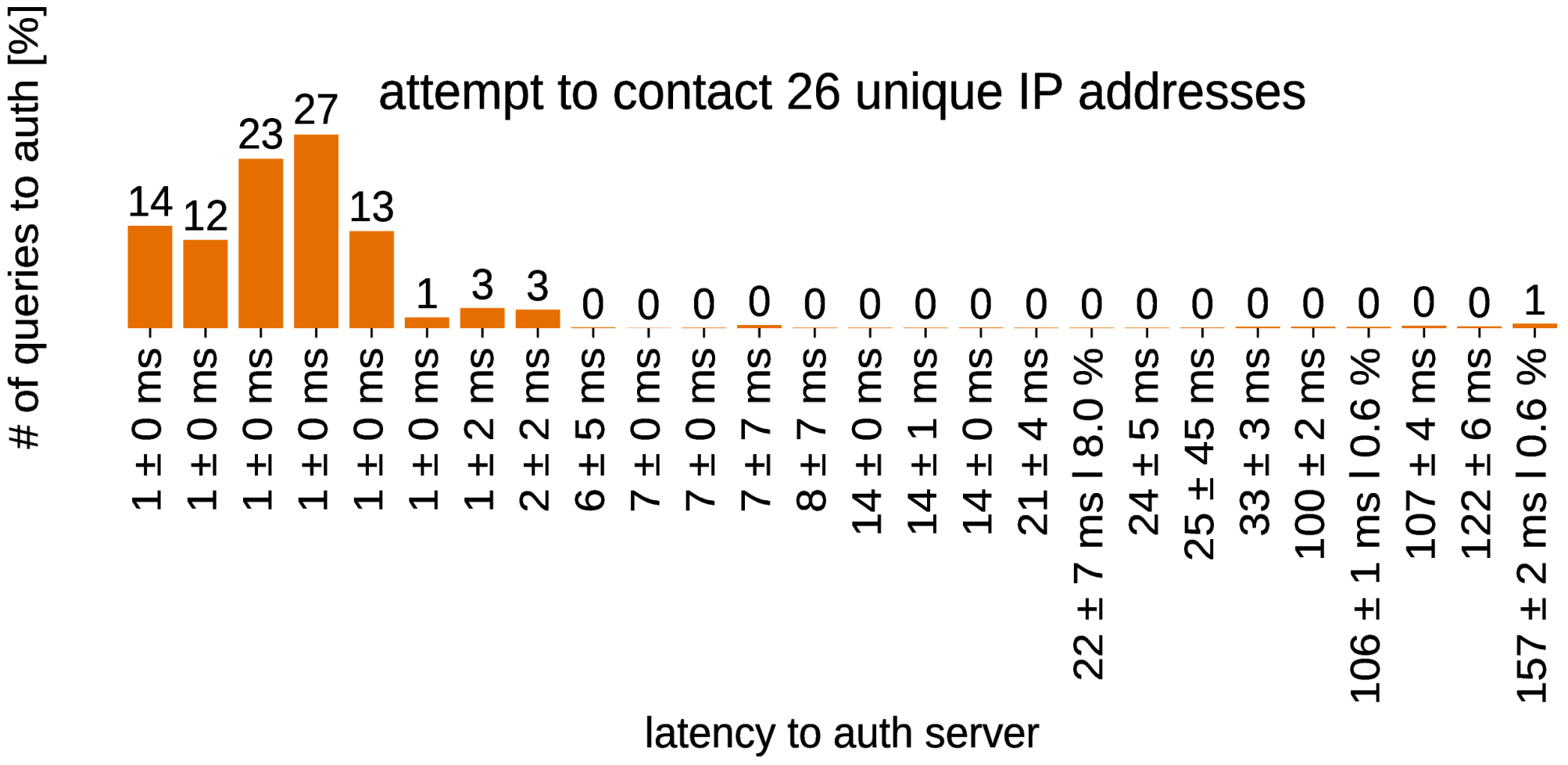
Prague: BIND packets to auth



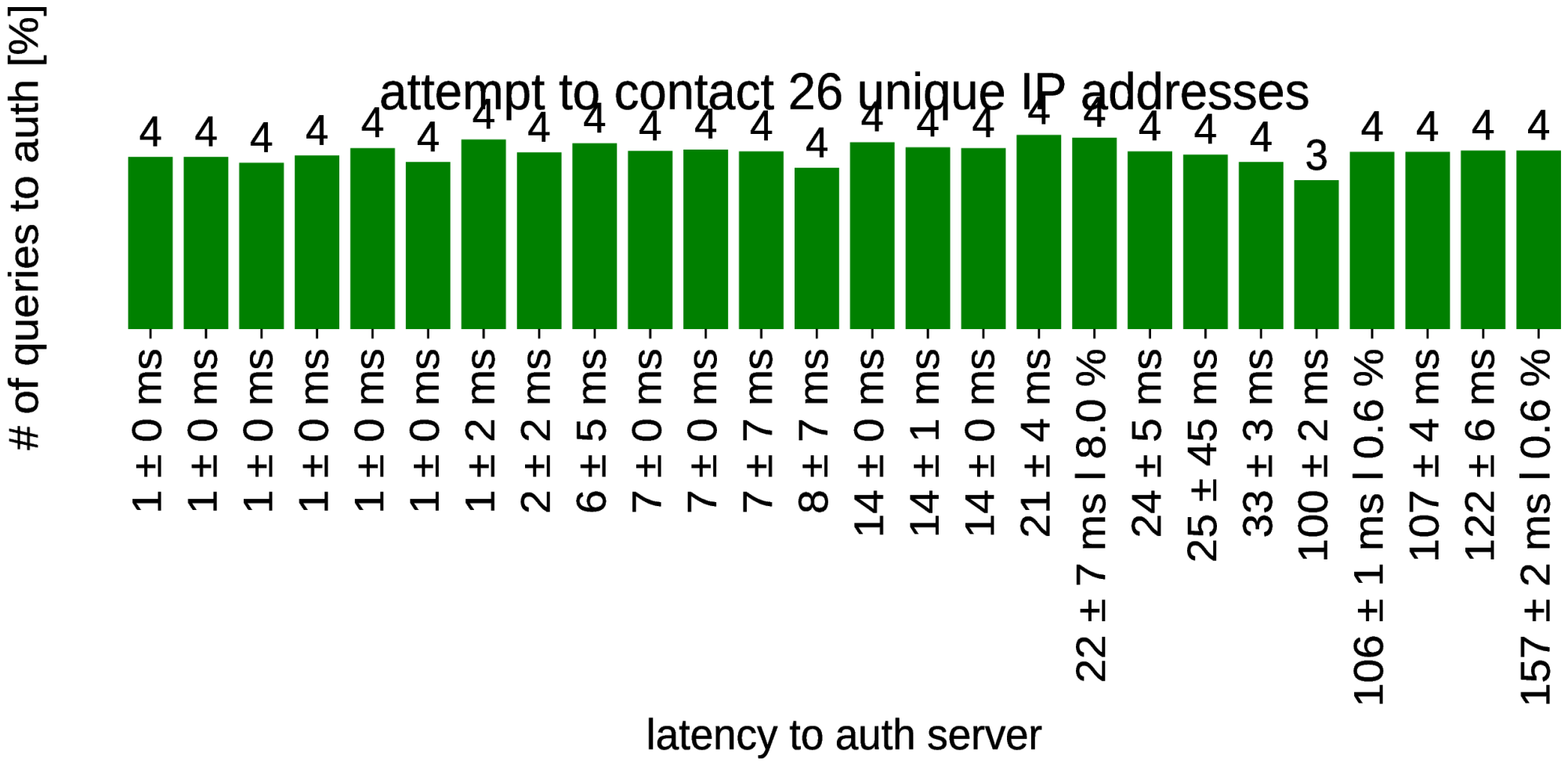
Prague: KNOT packets to auth



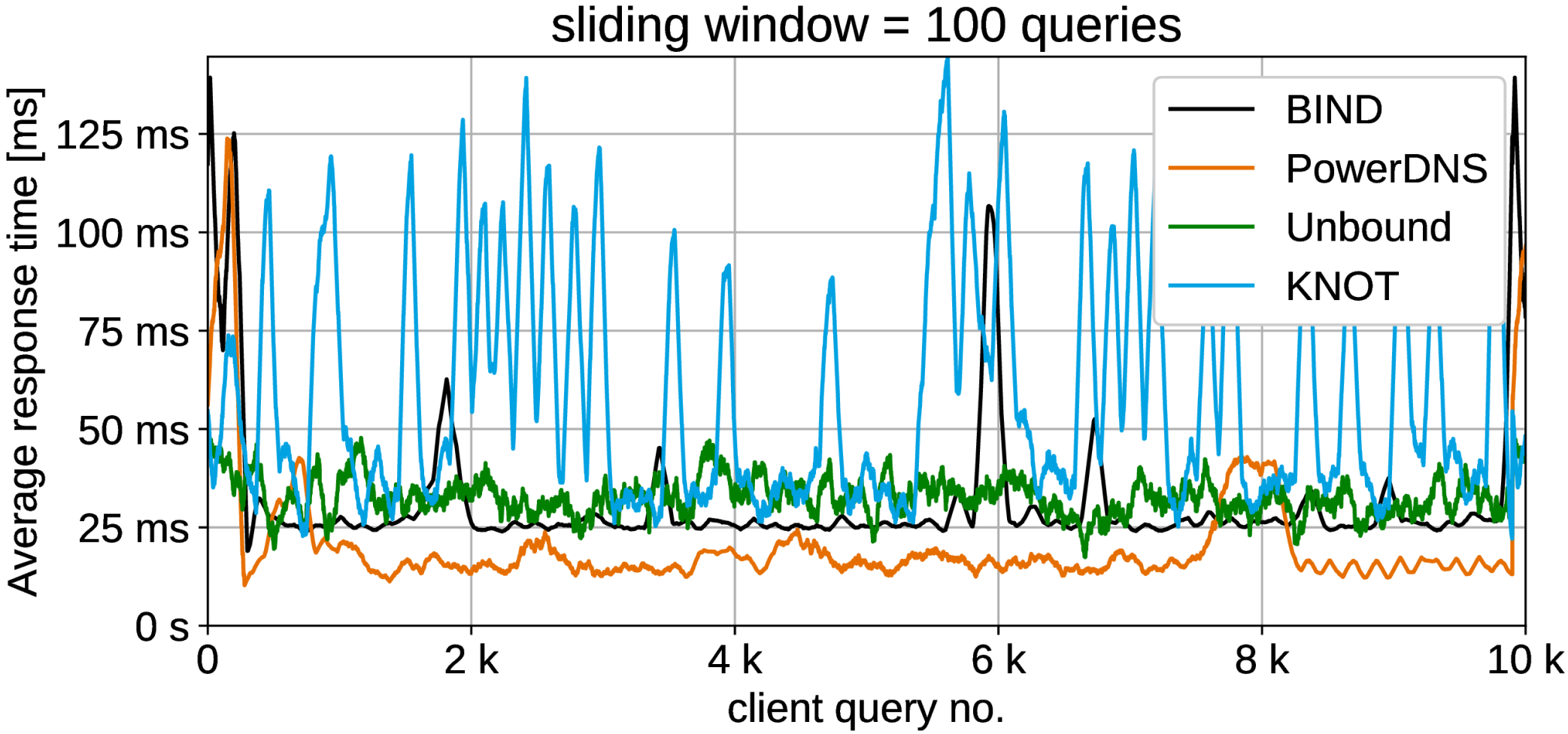
Prague: PowerDNS packets to auth



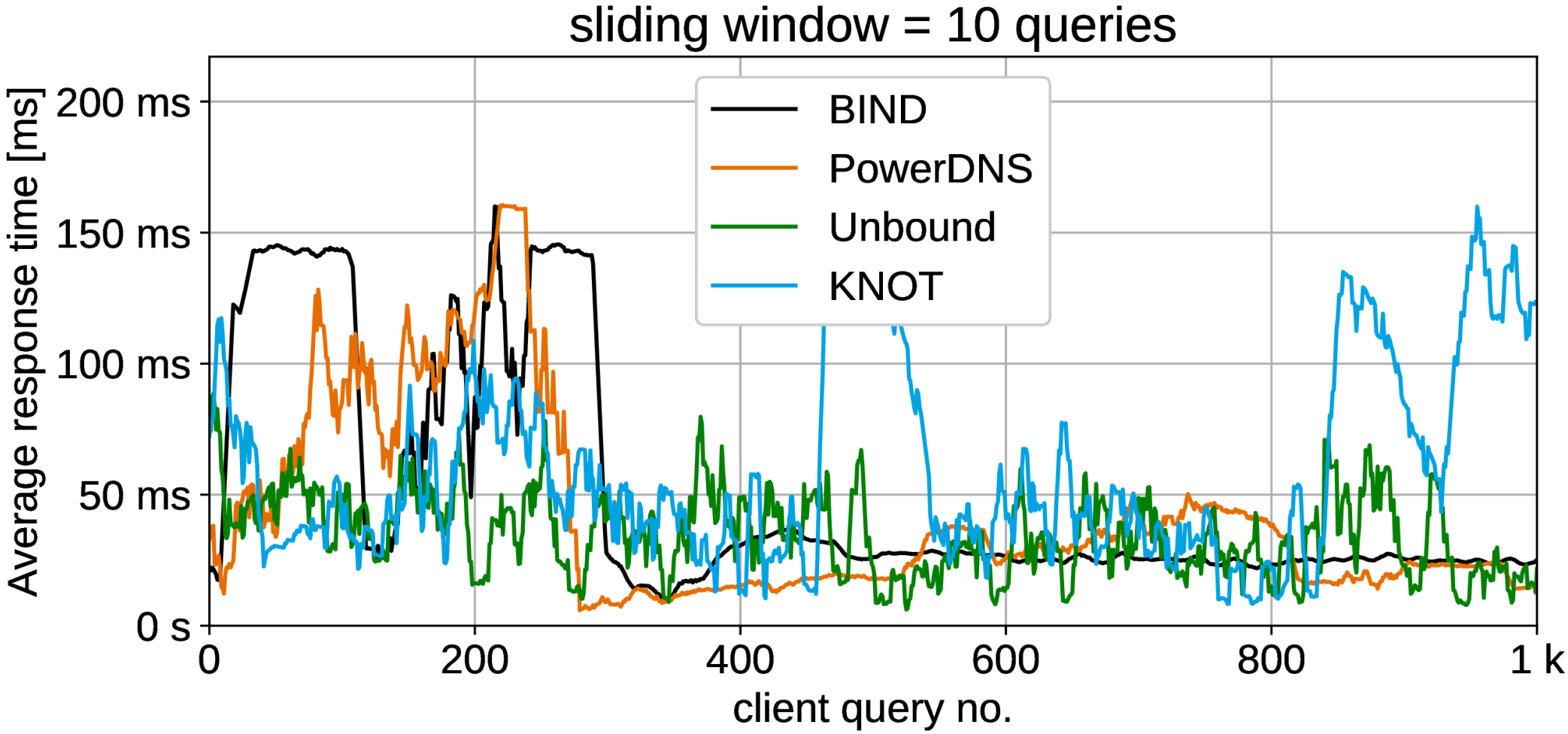
Prague: Unbound packets to auth



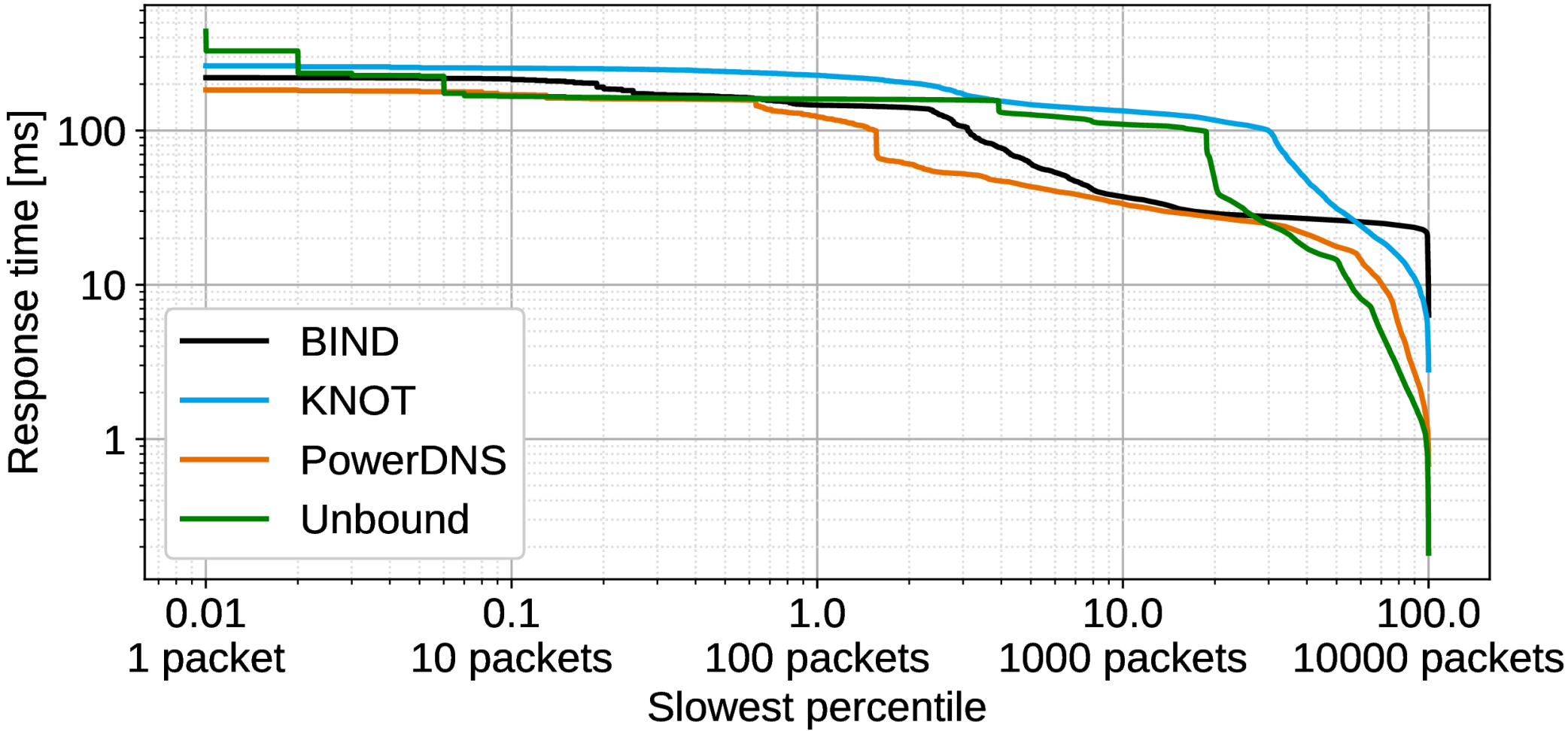
Prague: latency over time



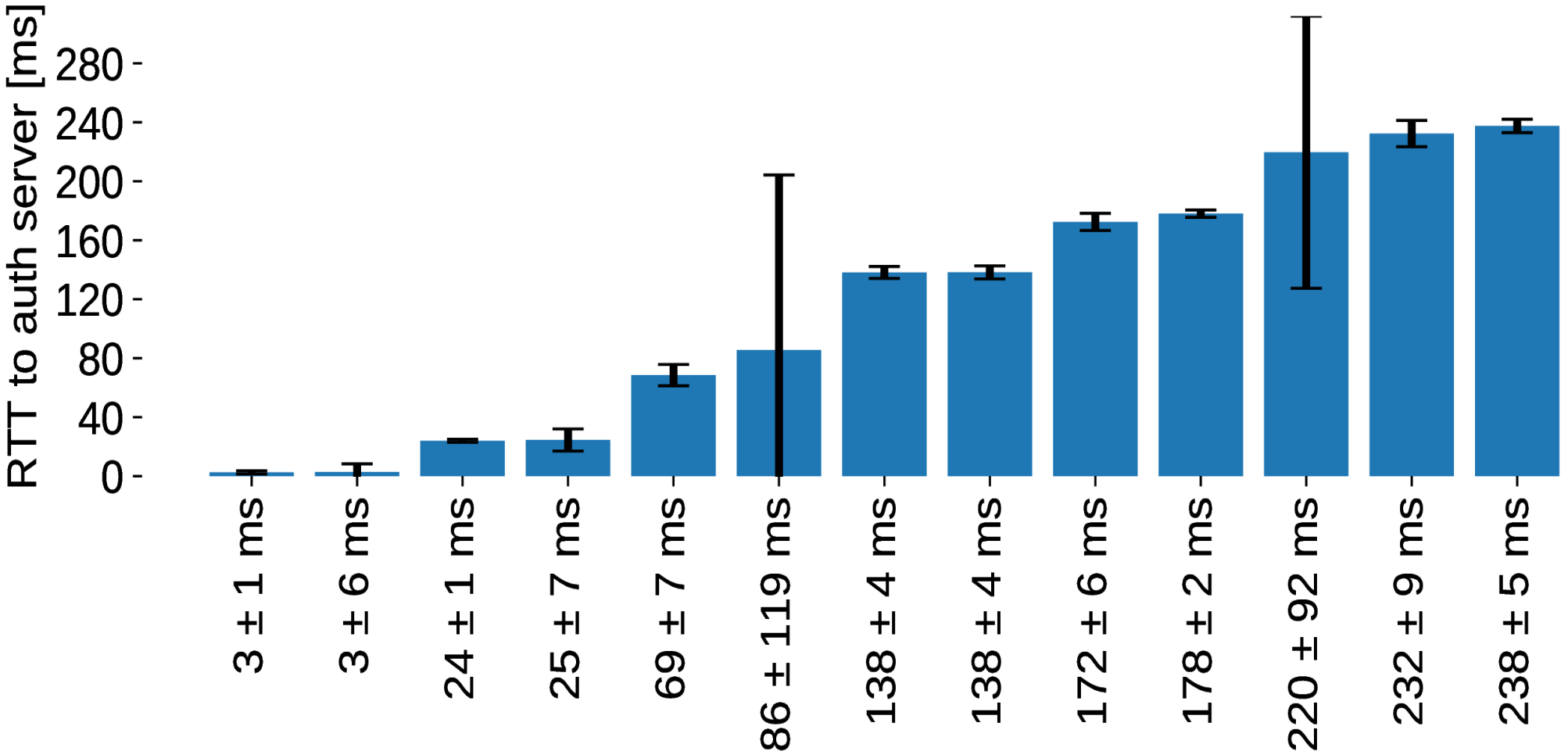
Prague: latency over time (zoom)



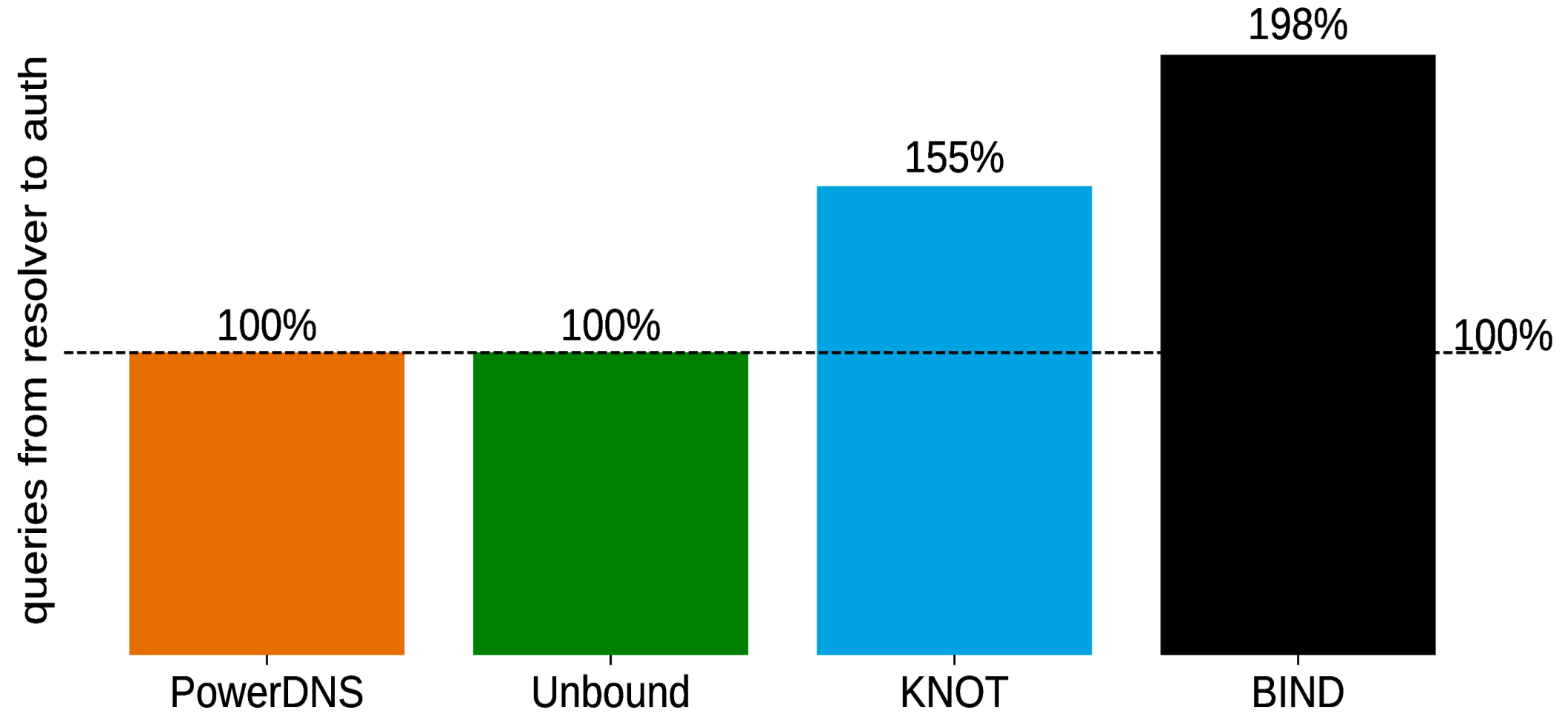
Prague: latency histogram



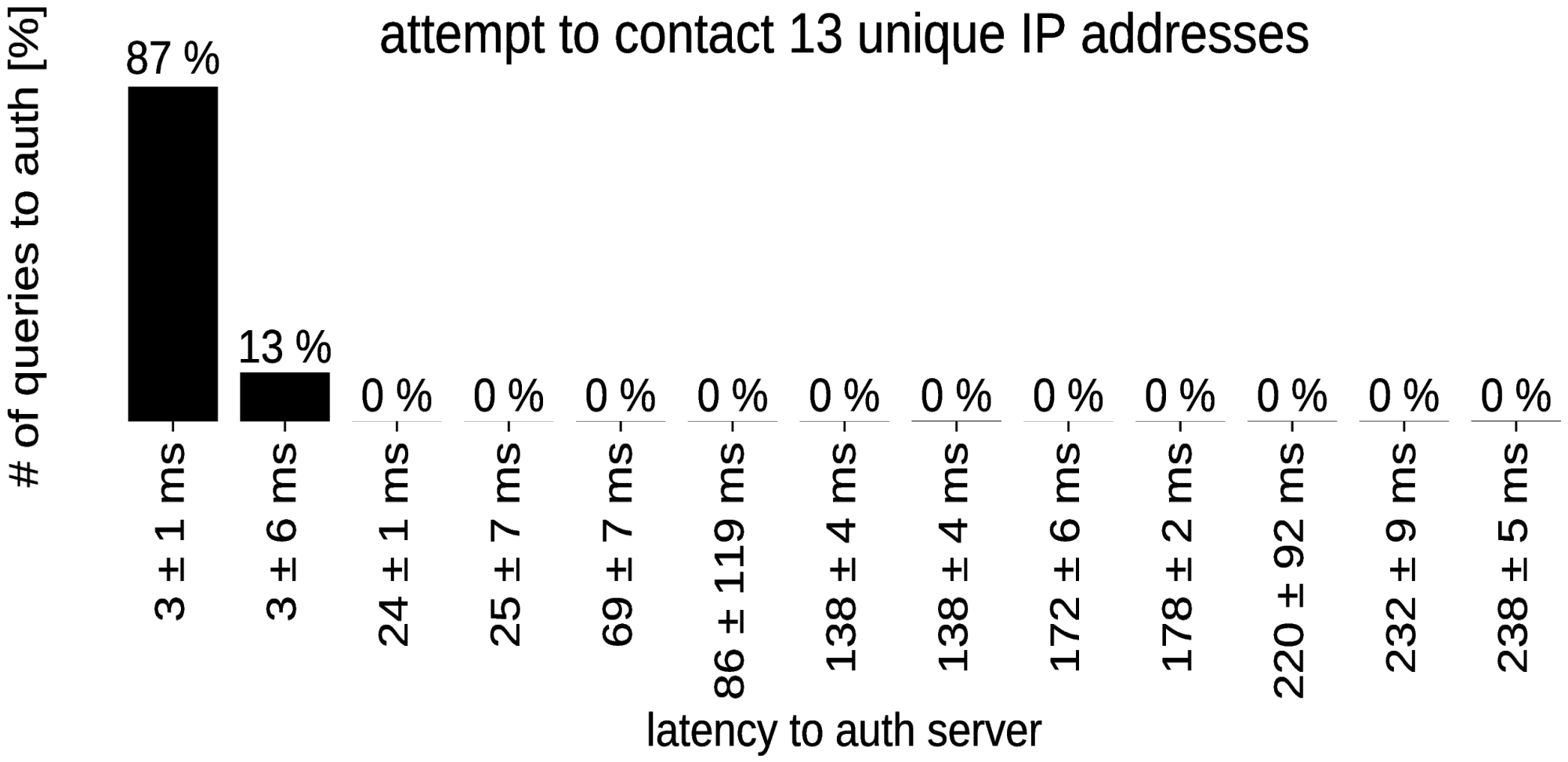
New Caledonia: auth servers



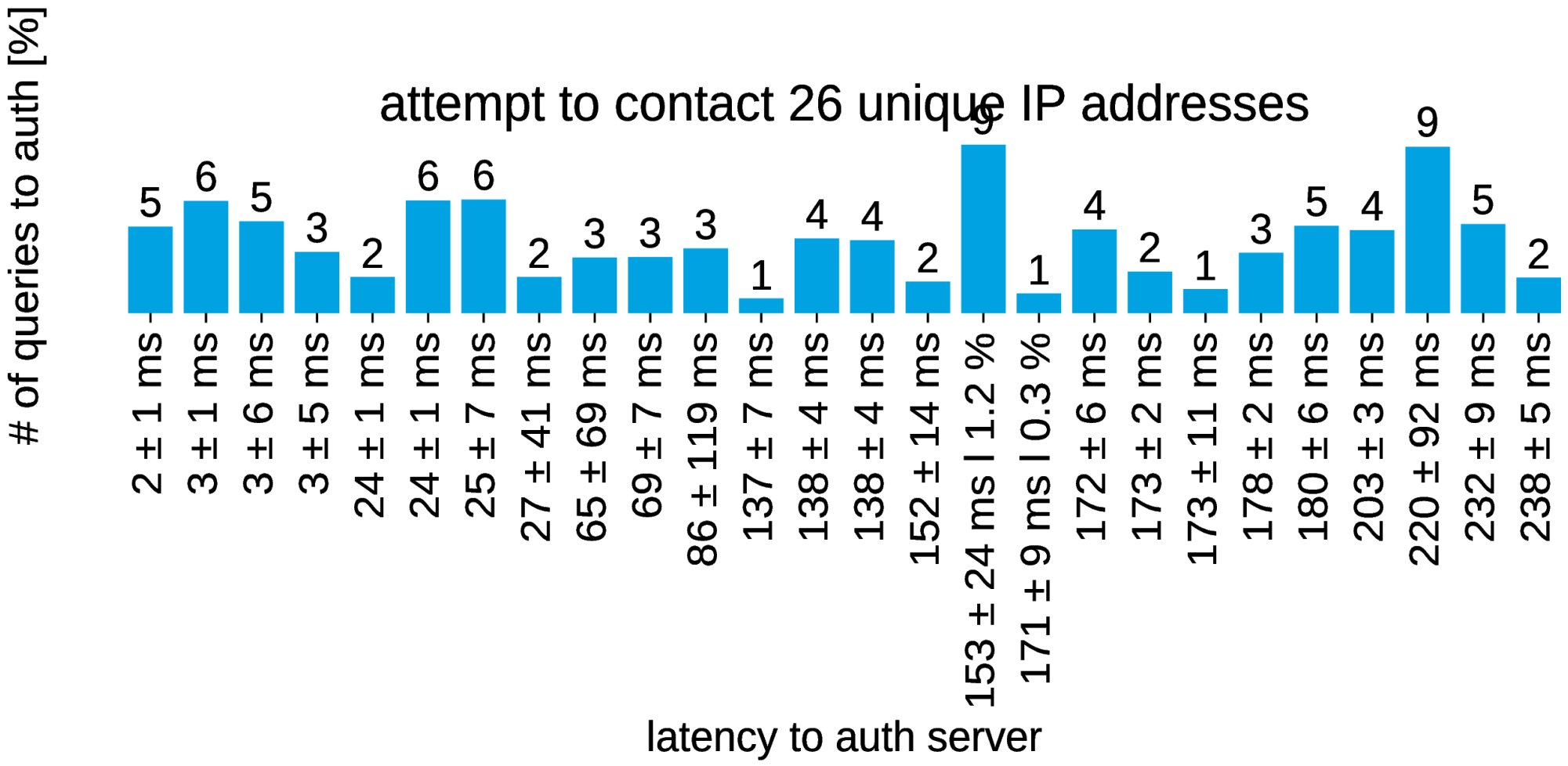
New Caledonia: # of queries to auth



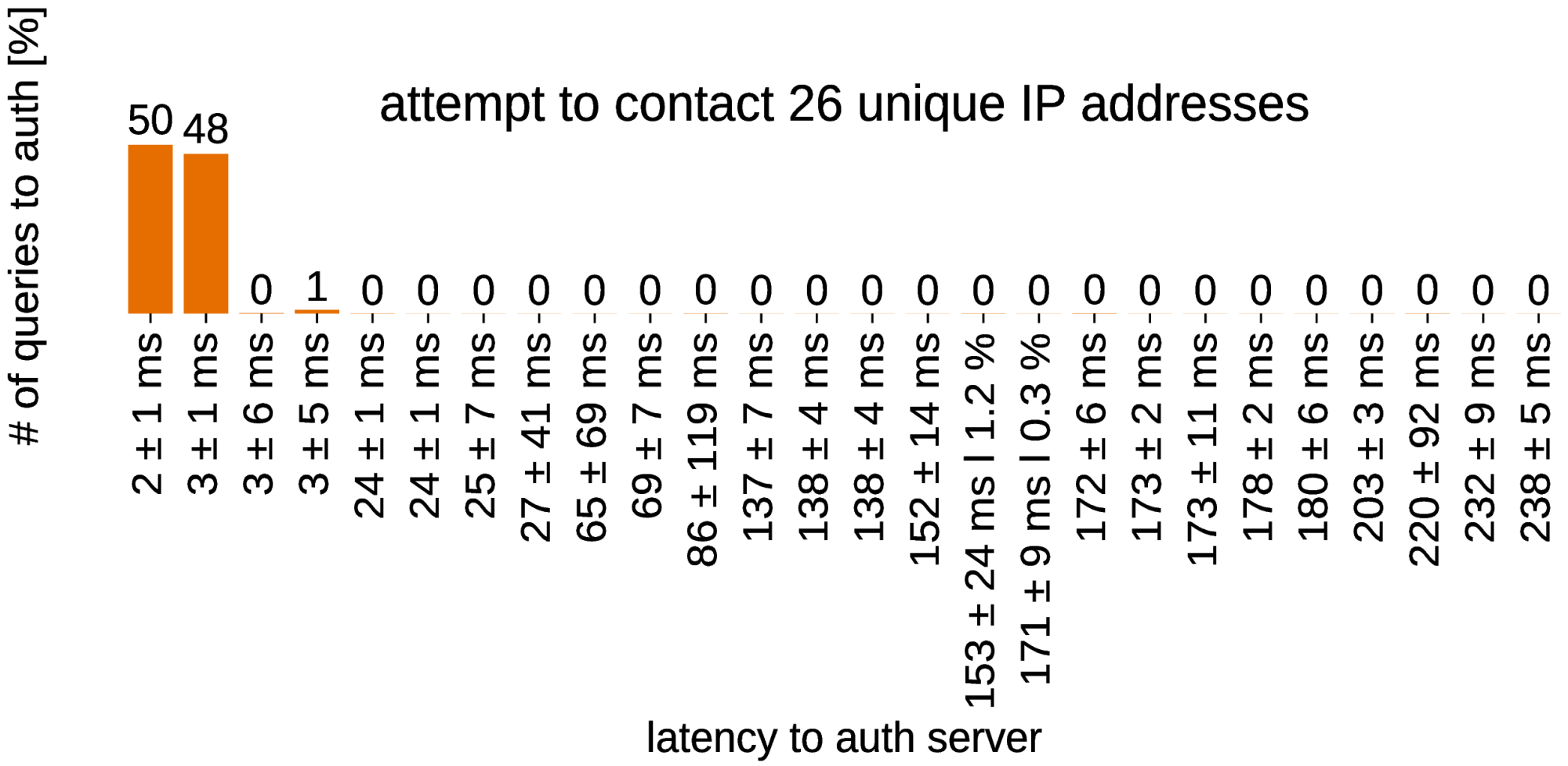
New Caledonia: BIND packets to auth



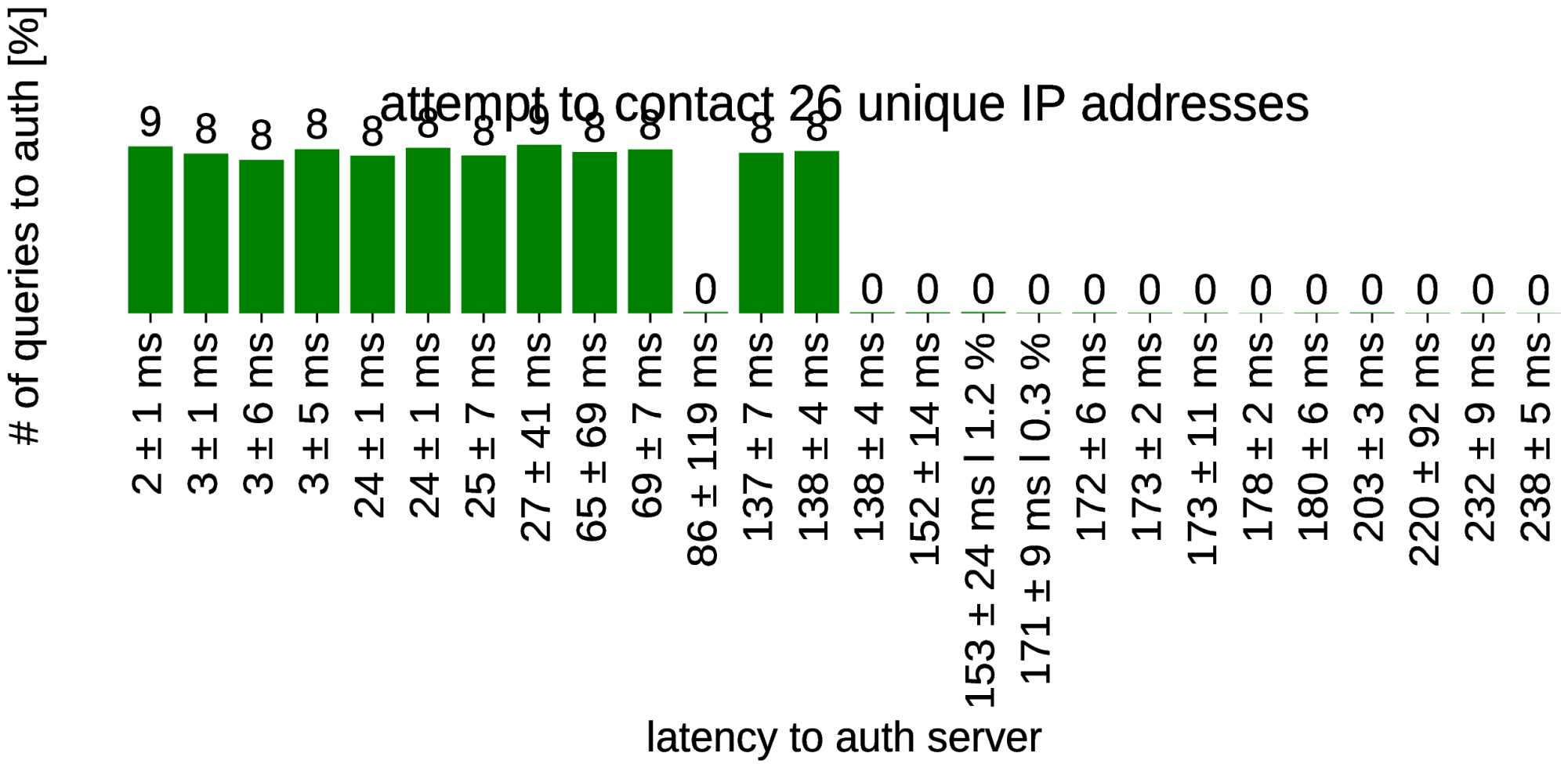
New Caledonia: KNOT packets to auth



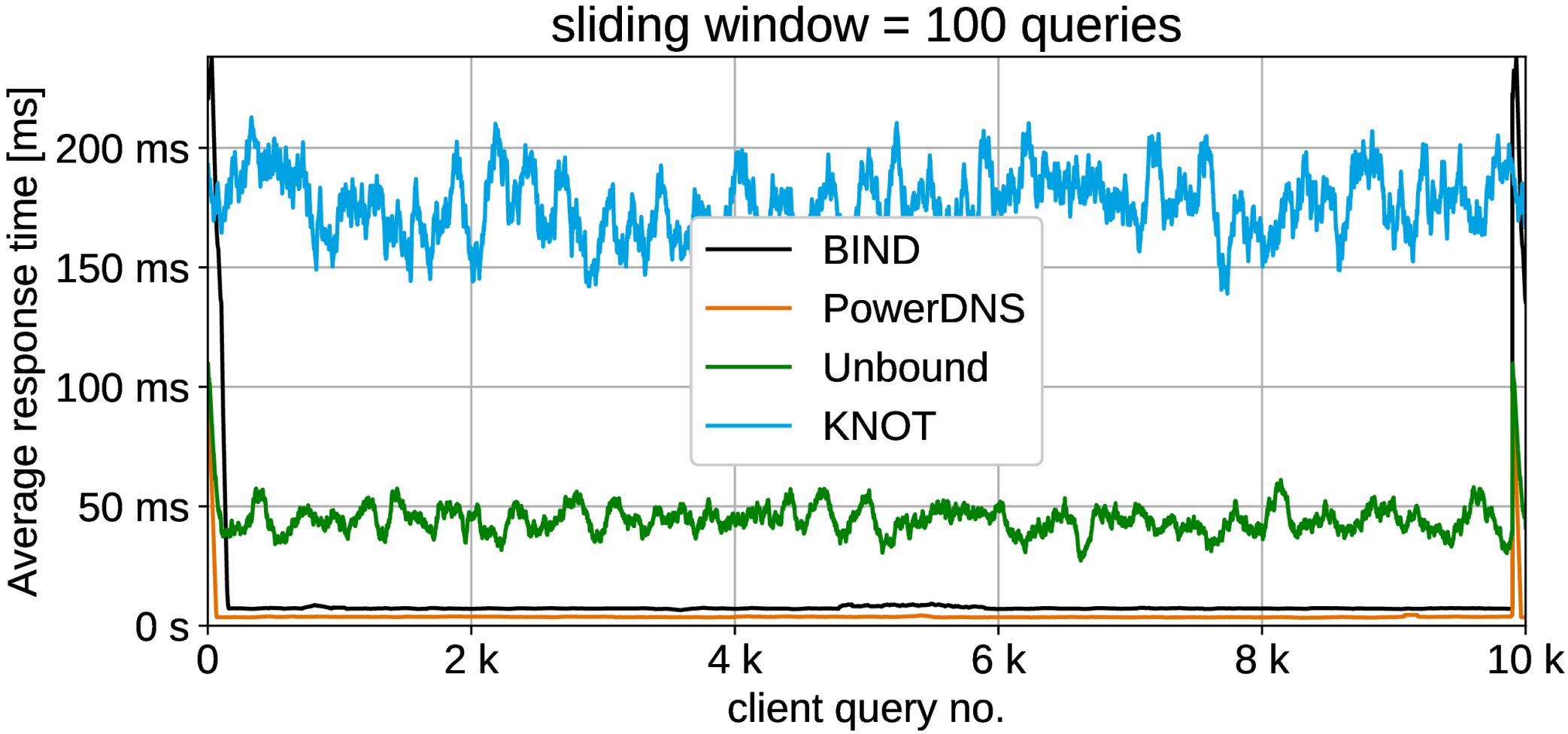
New Caledonia: PowerDNS packets to auth



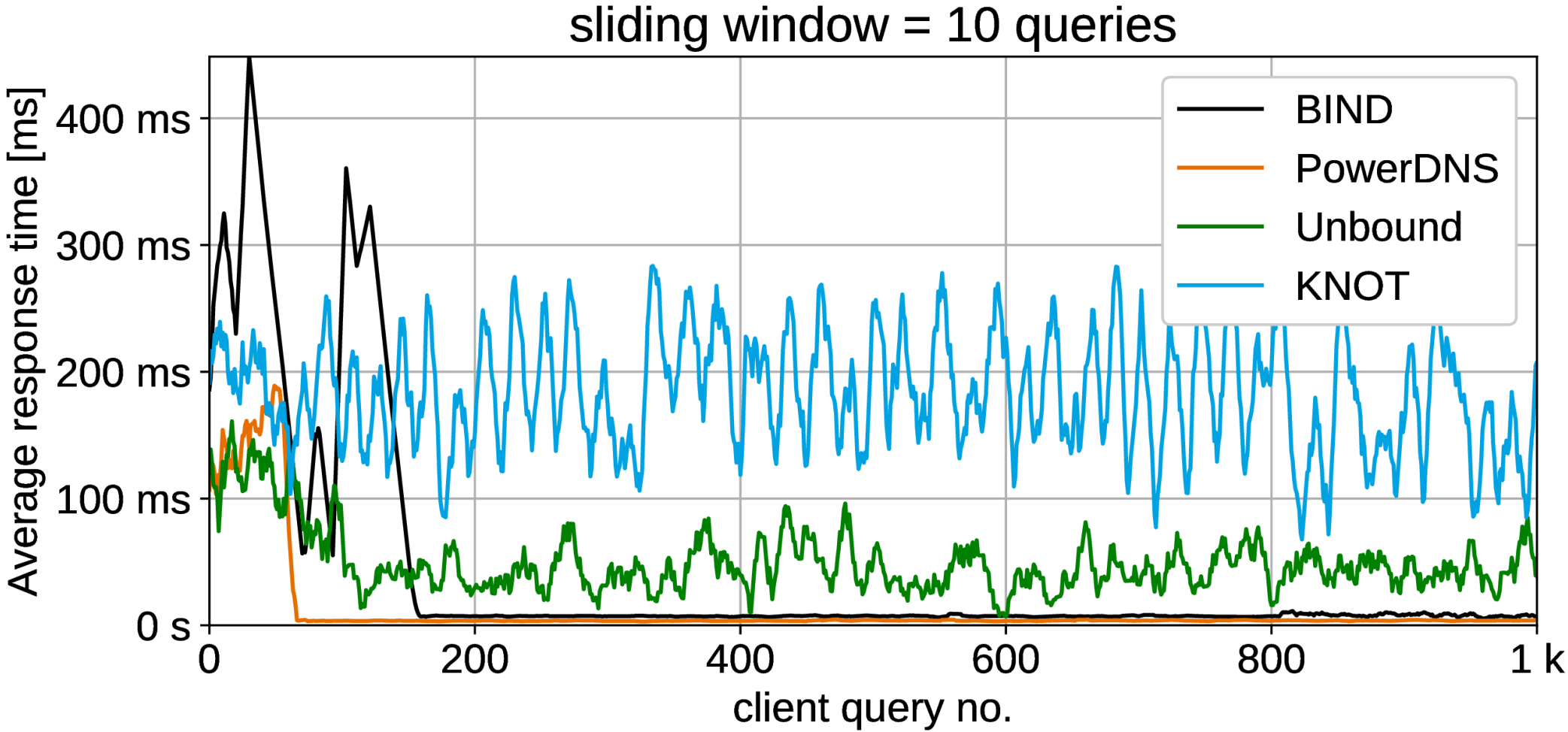
New Caledonia: Unbound packets to auth



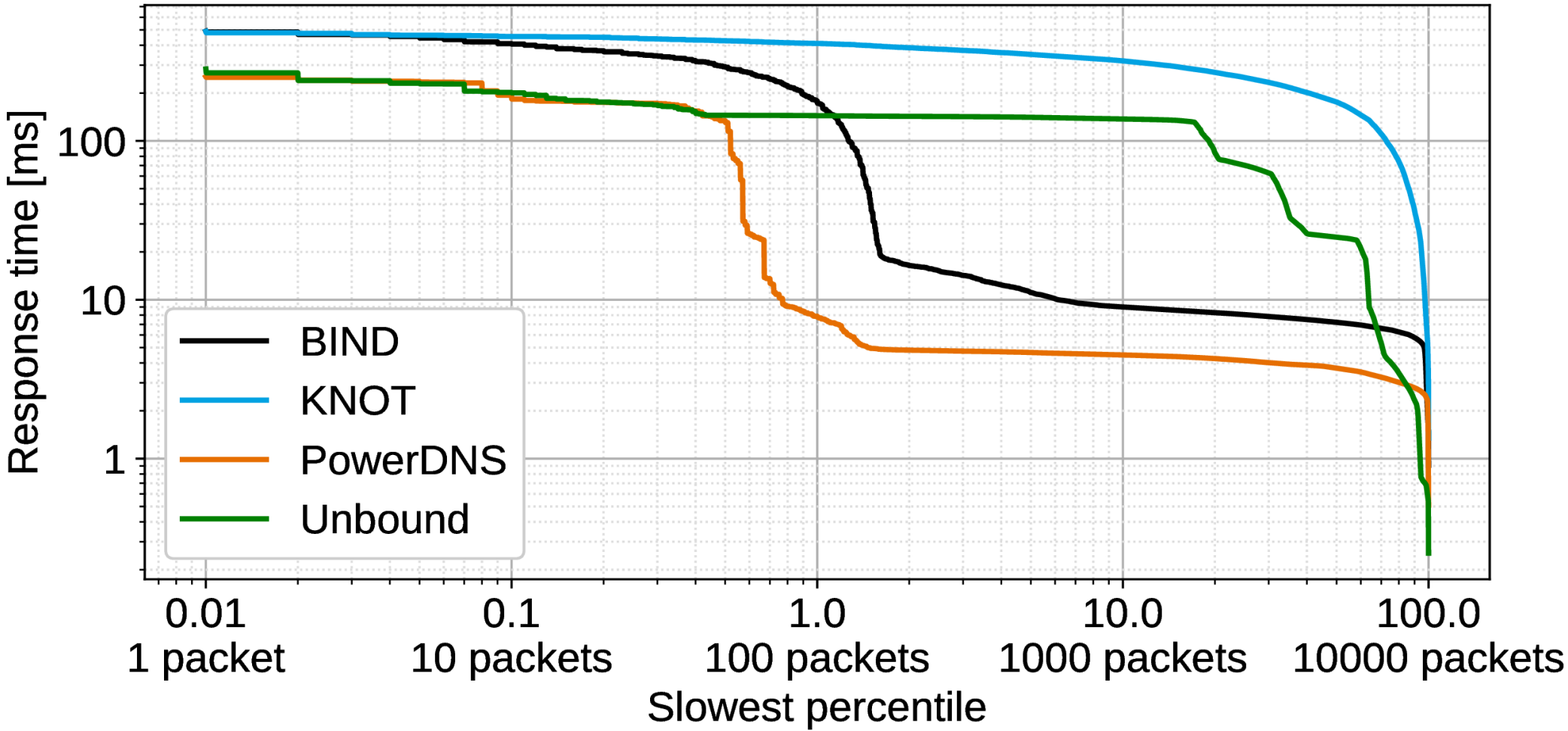
New Caledonia: latency over time



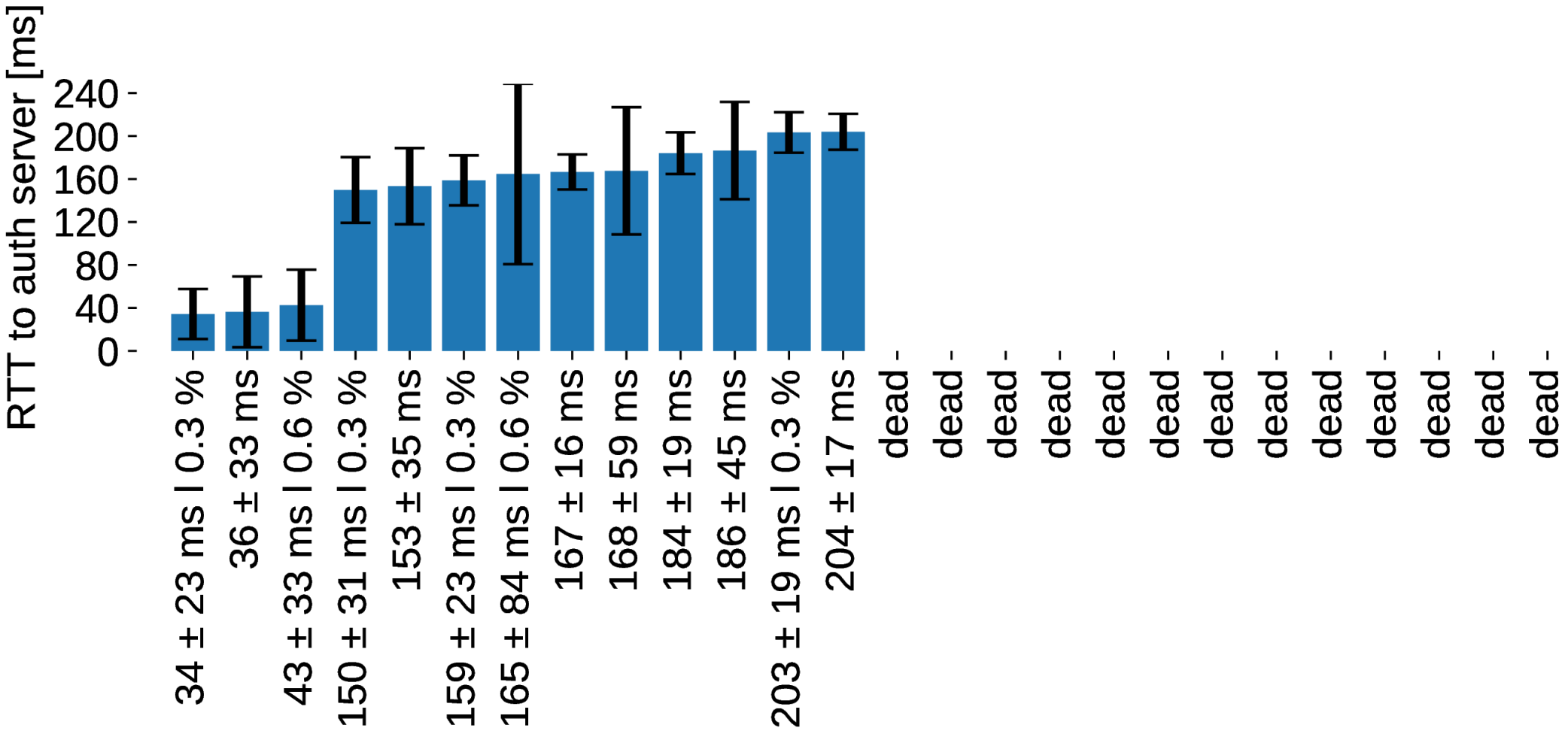
New Caledonia: latency over time (zoom)



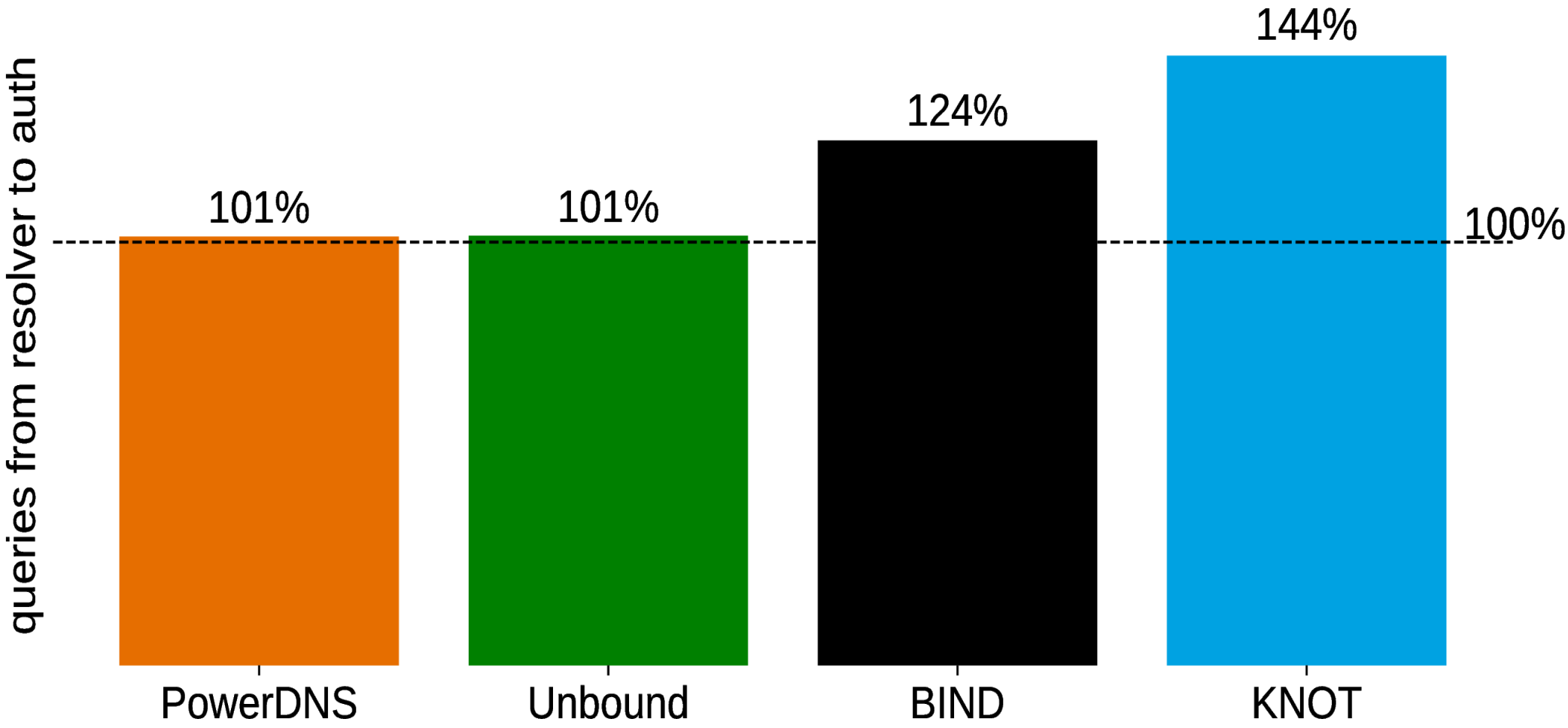
New Caledonia: latency histogram



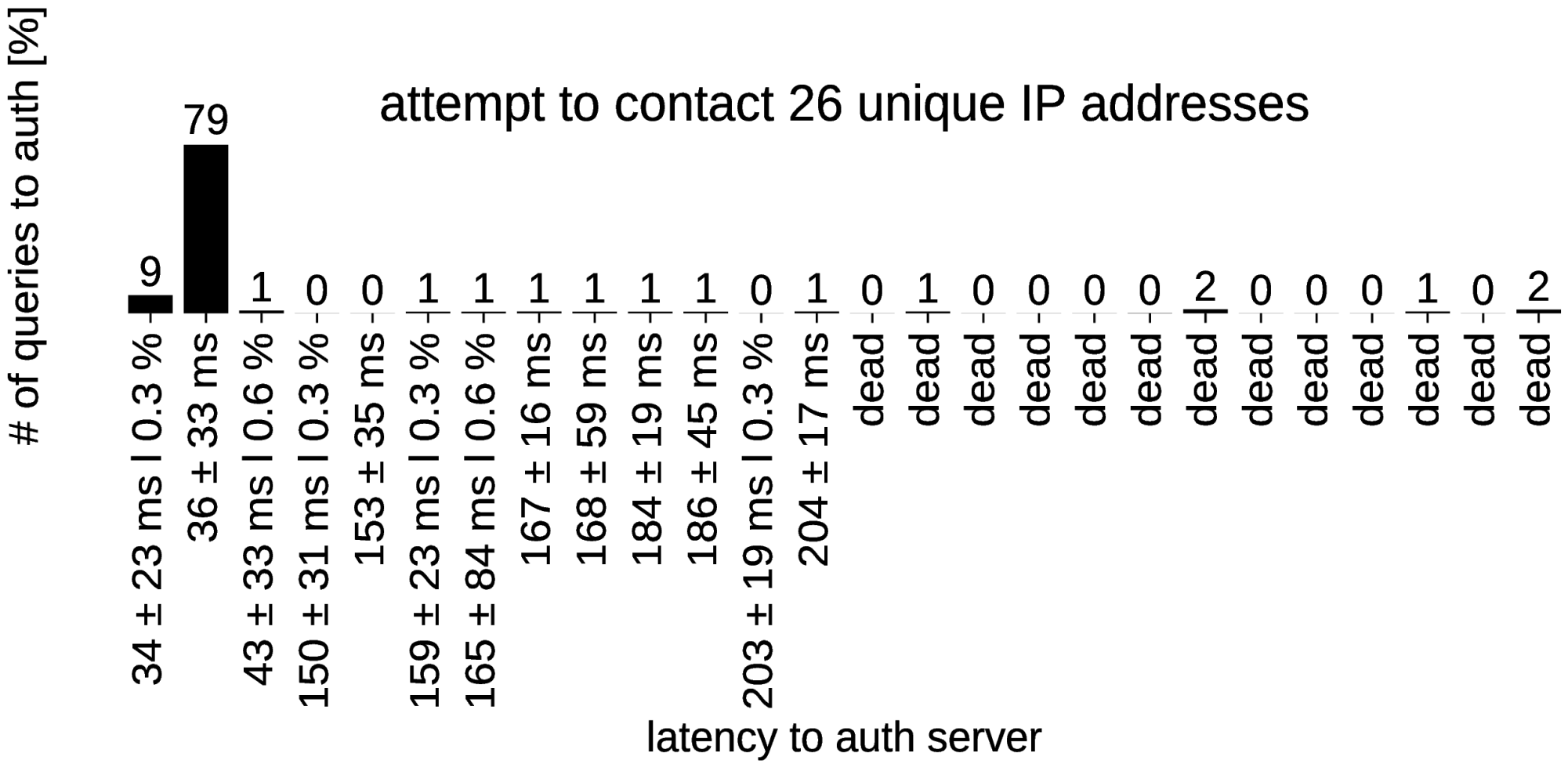
Brazil: auth servers



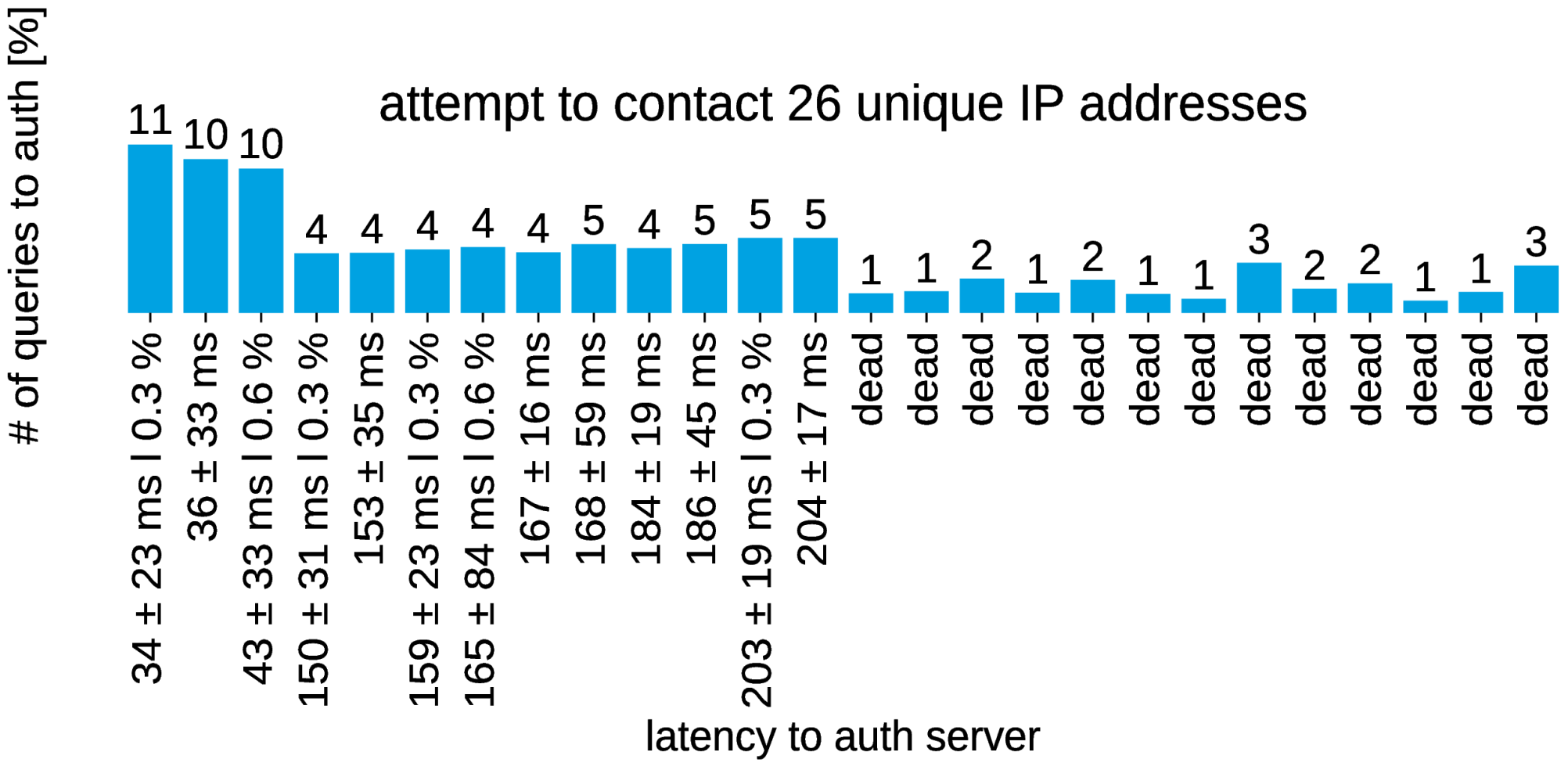
Brazil: # of queries to auth



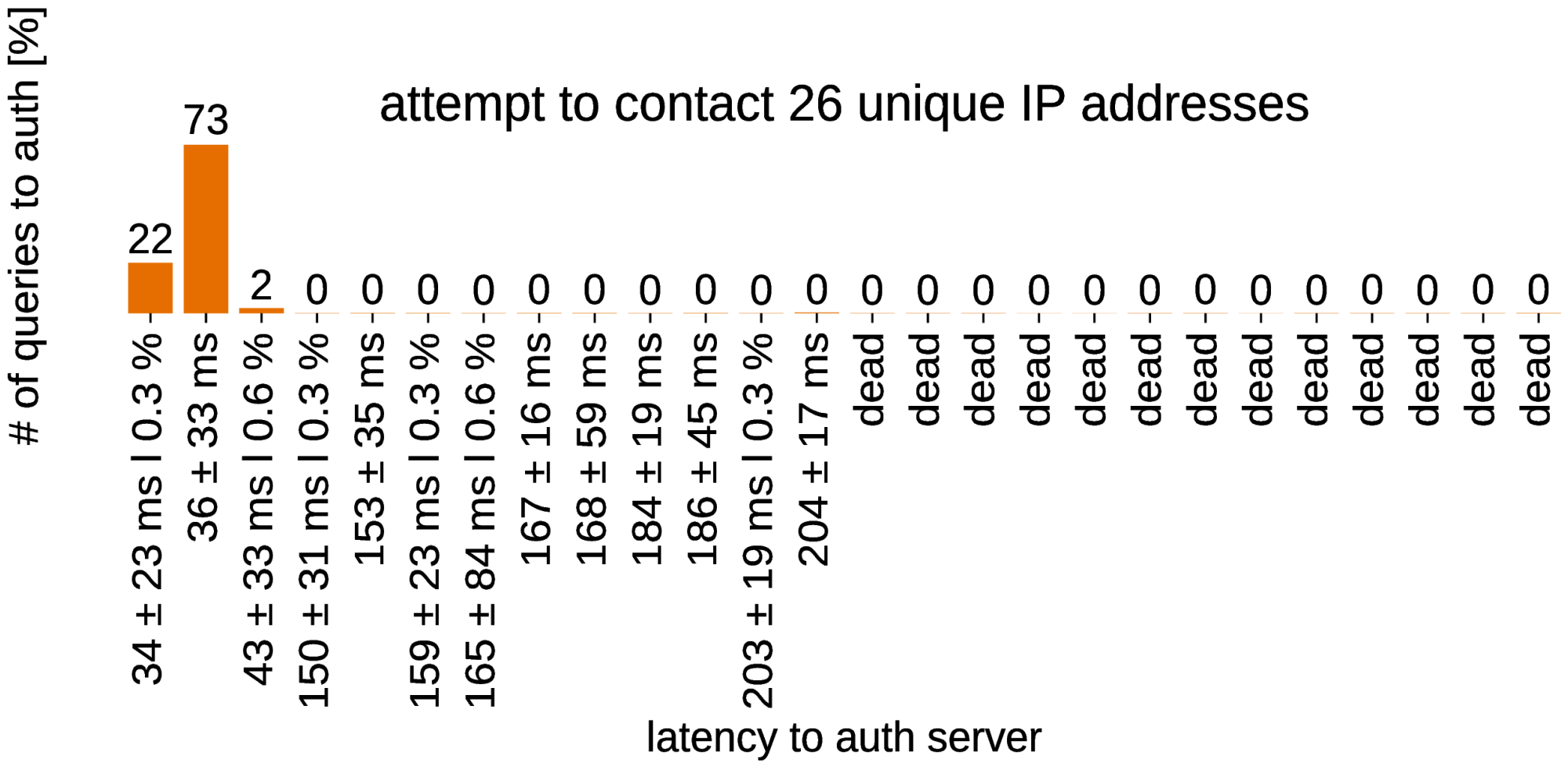
Brazil: BIND packets to auth



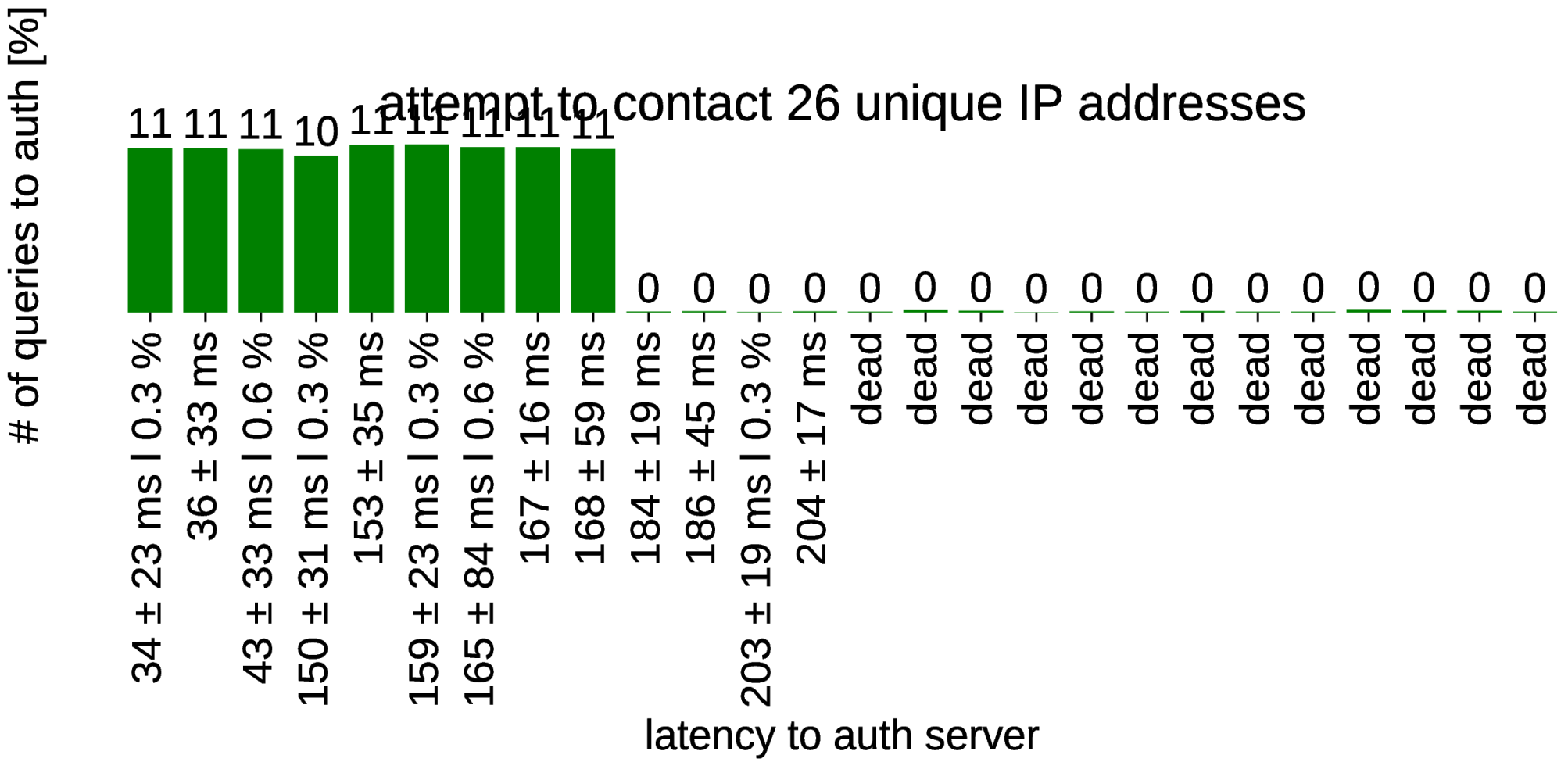
Brazil: KNOT packets to auth



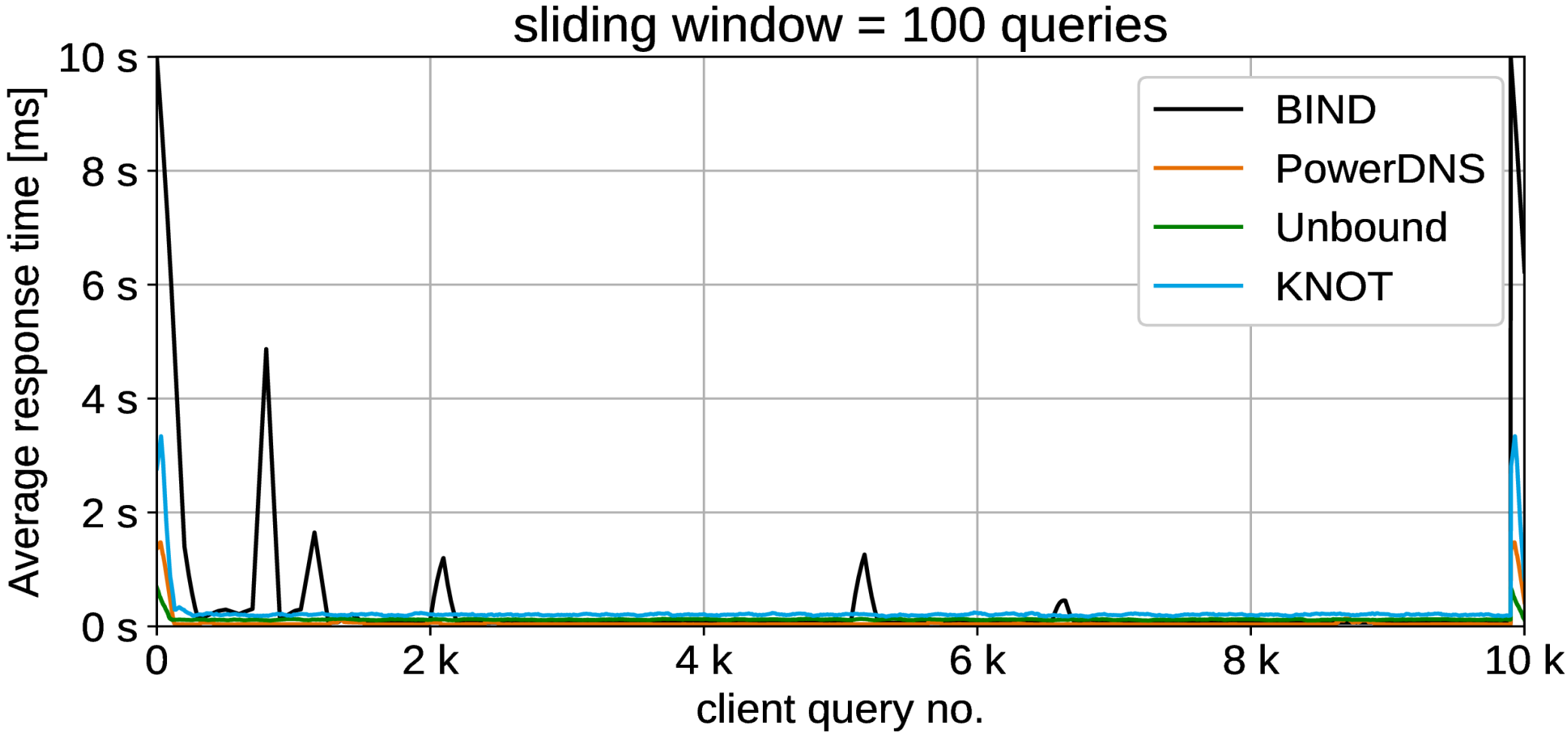
Brazil: PowerDNS packets to auth



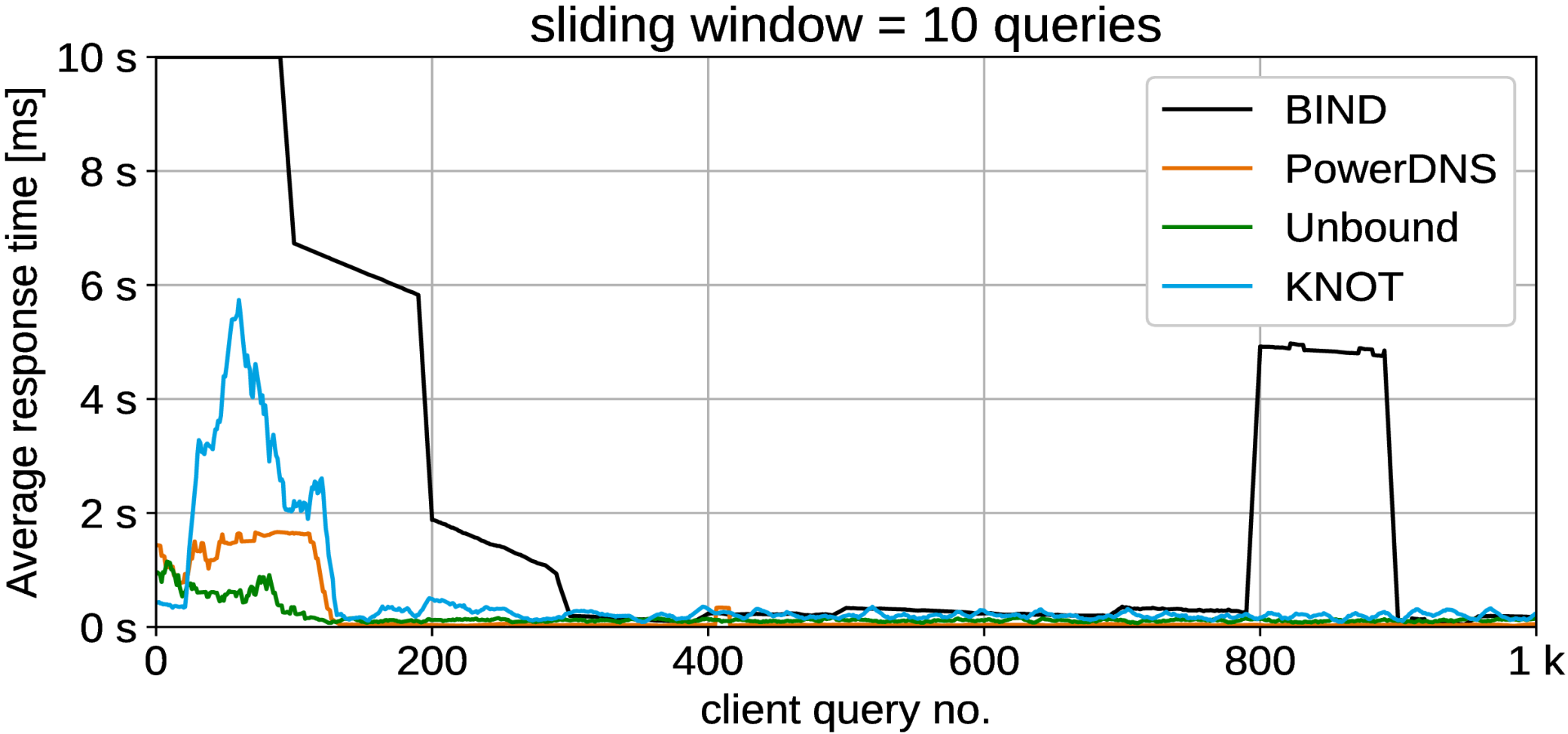
Brazil: Unbound packets to auth



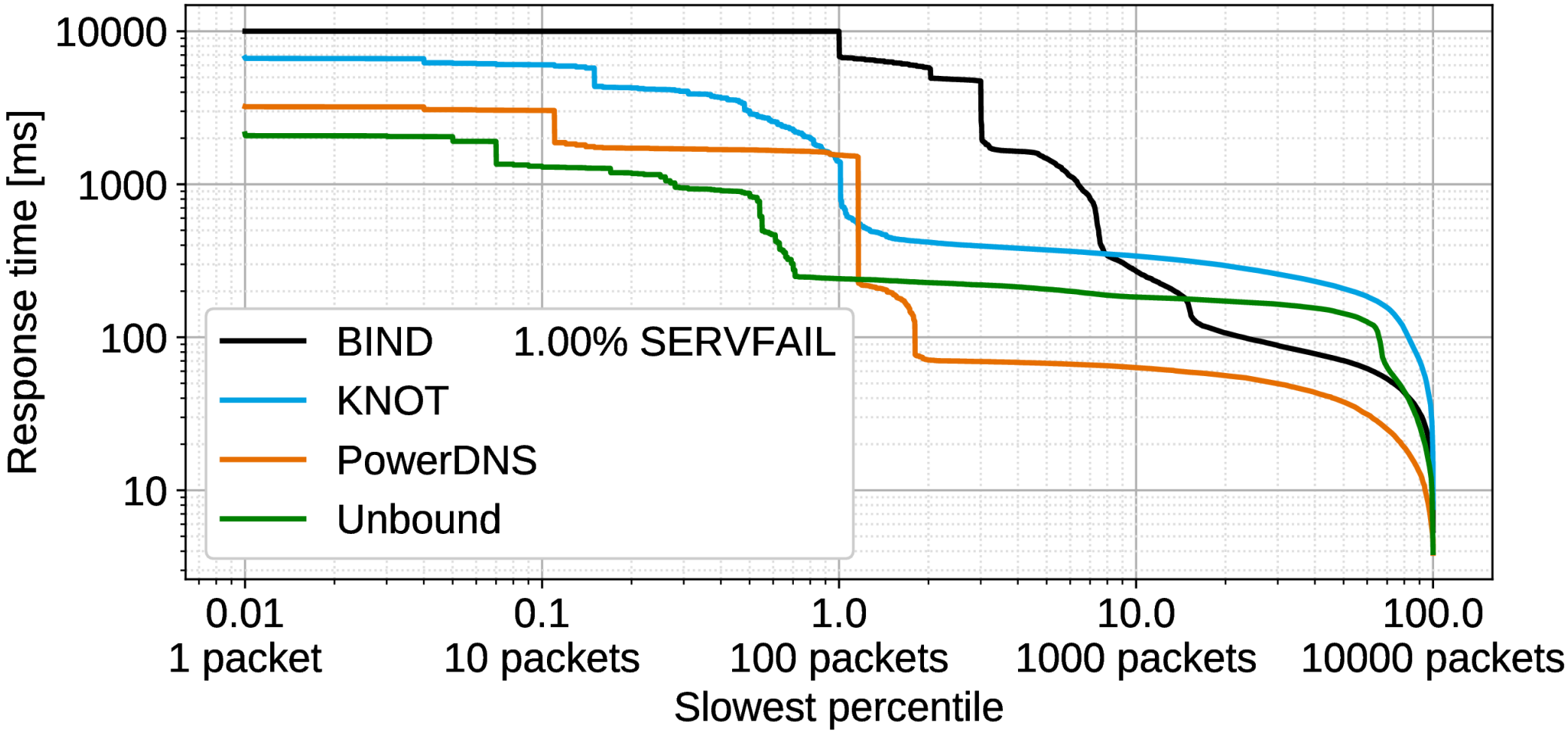
Brazil: latency over time



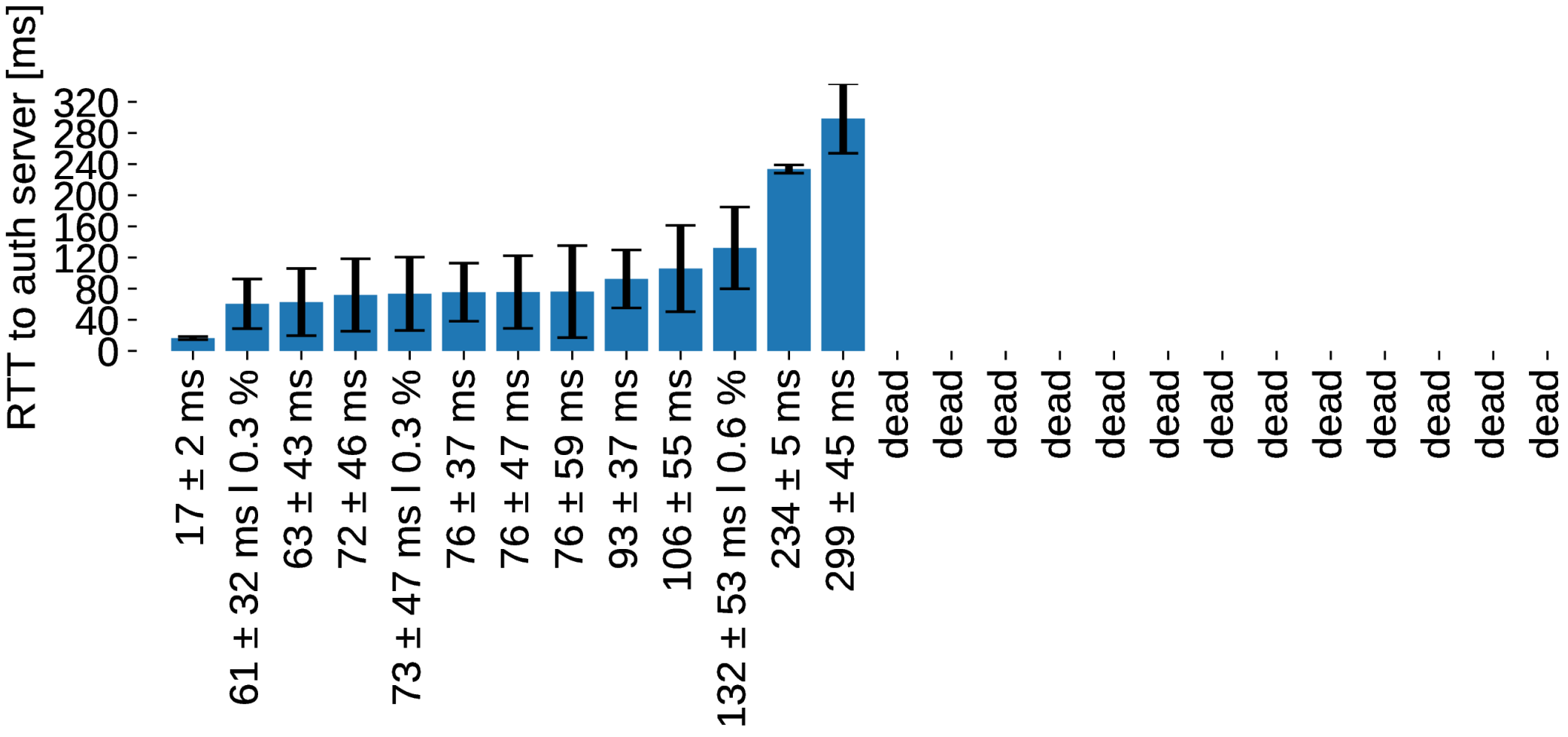
Brazil: latency over time (zoom)



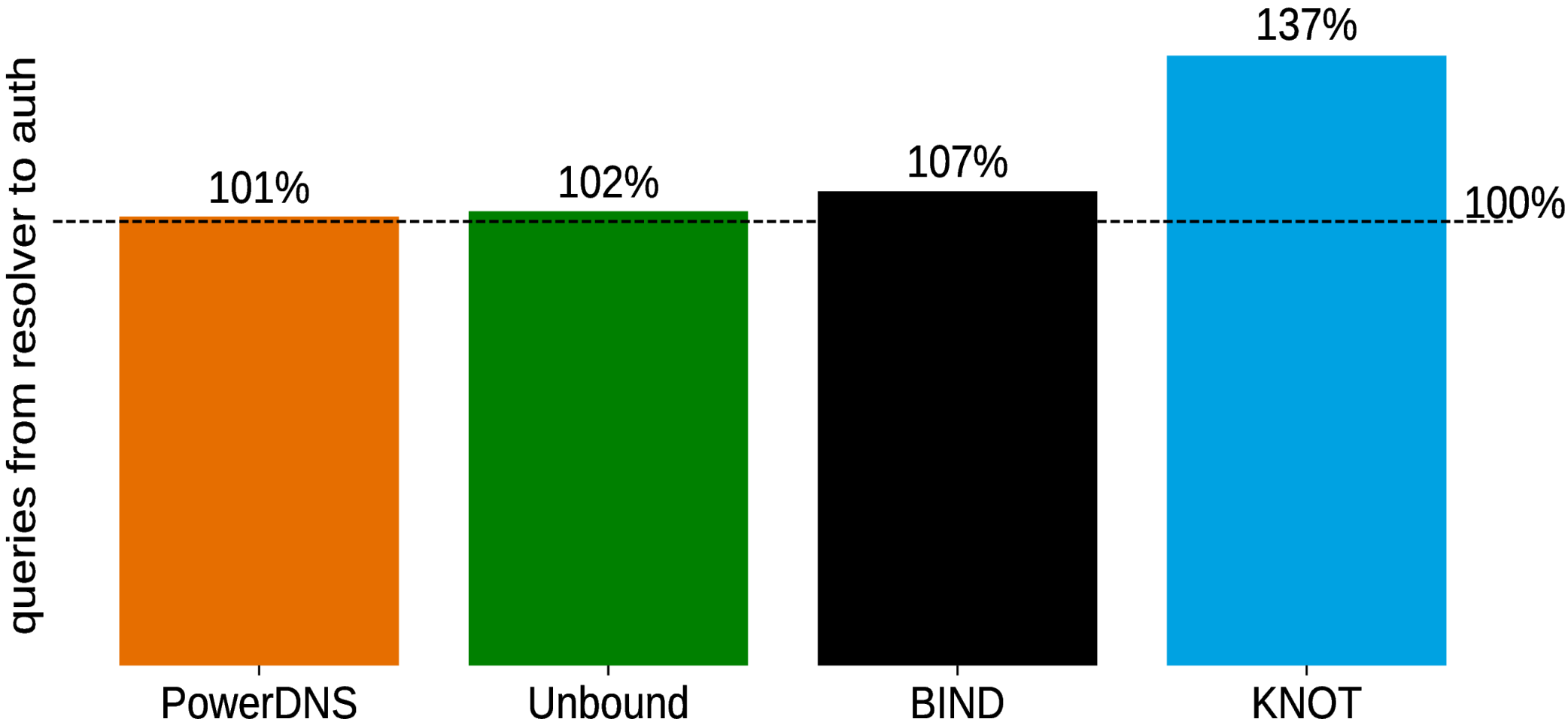
Brazil: latency histogram



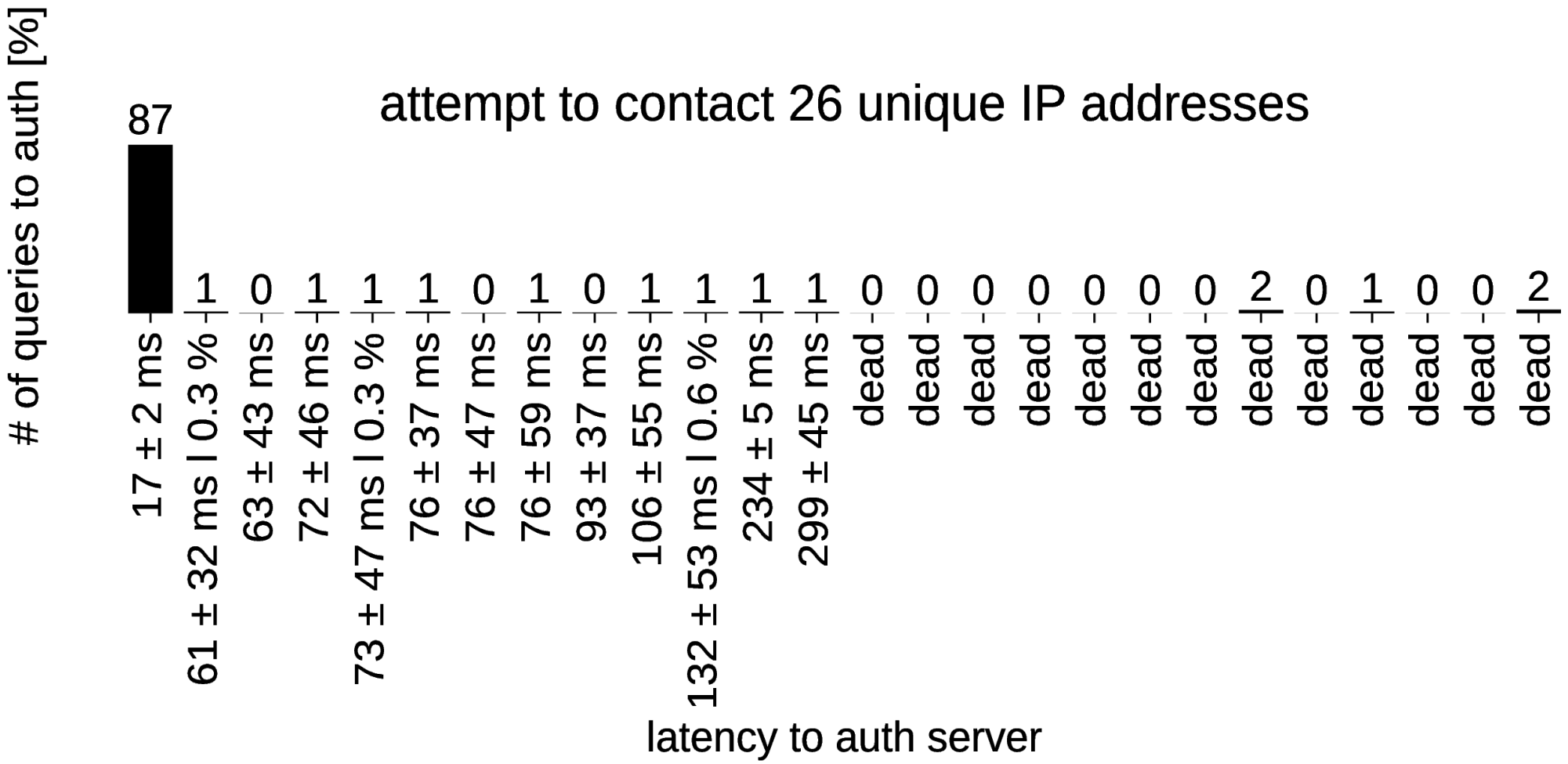
Hawai: auth servers



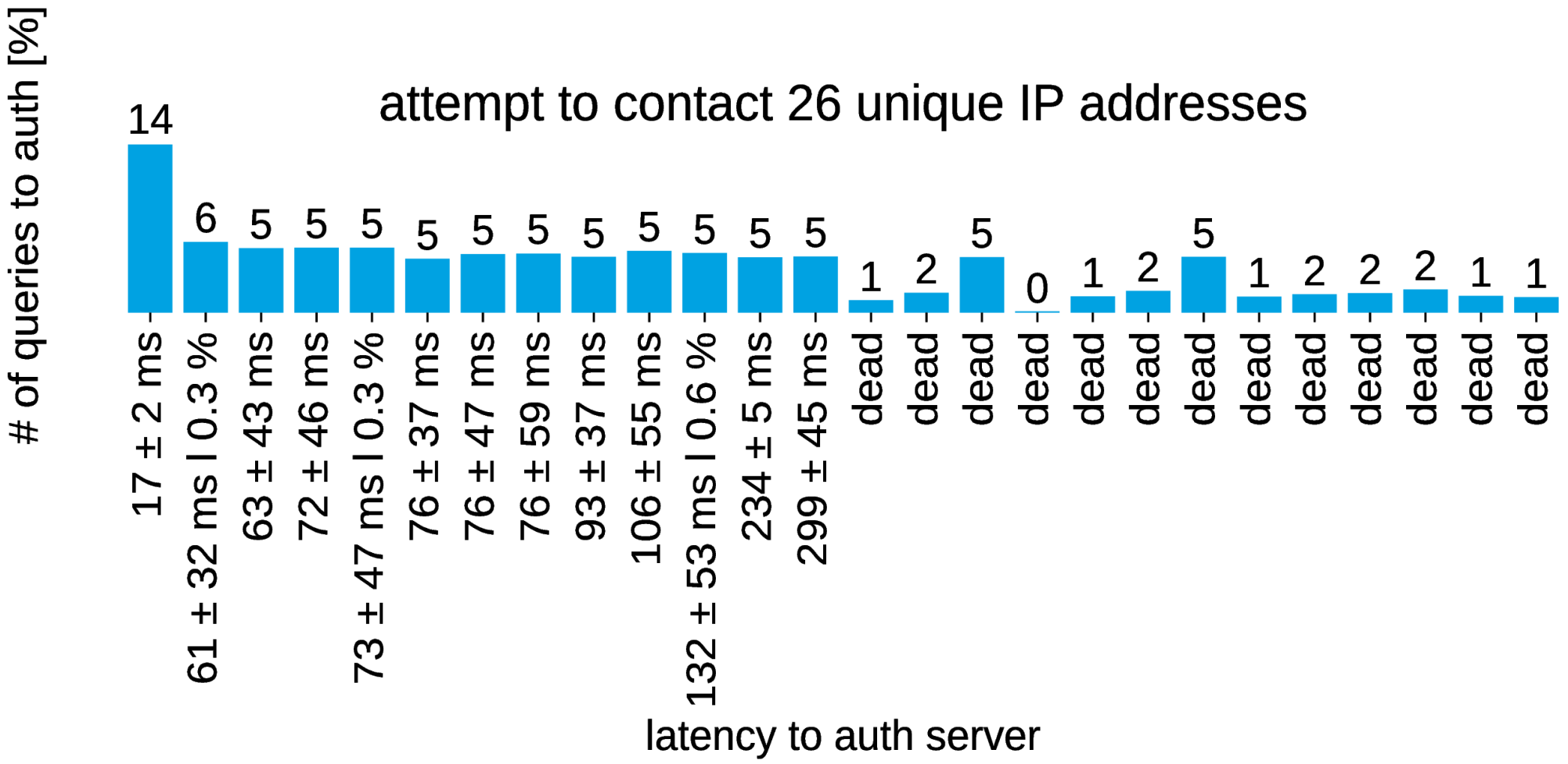
Hawai: # of queries to auth



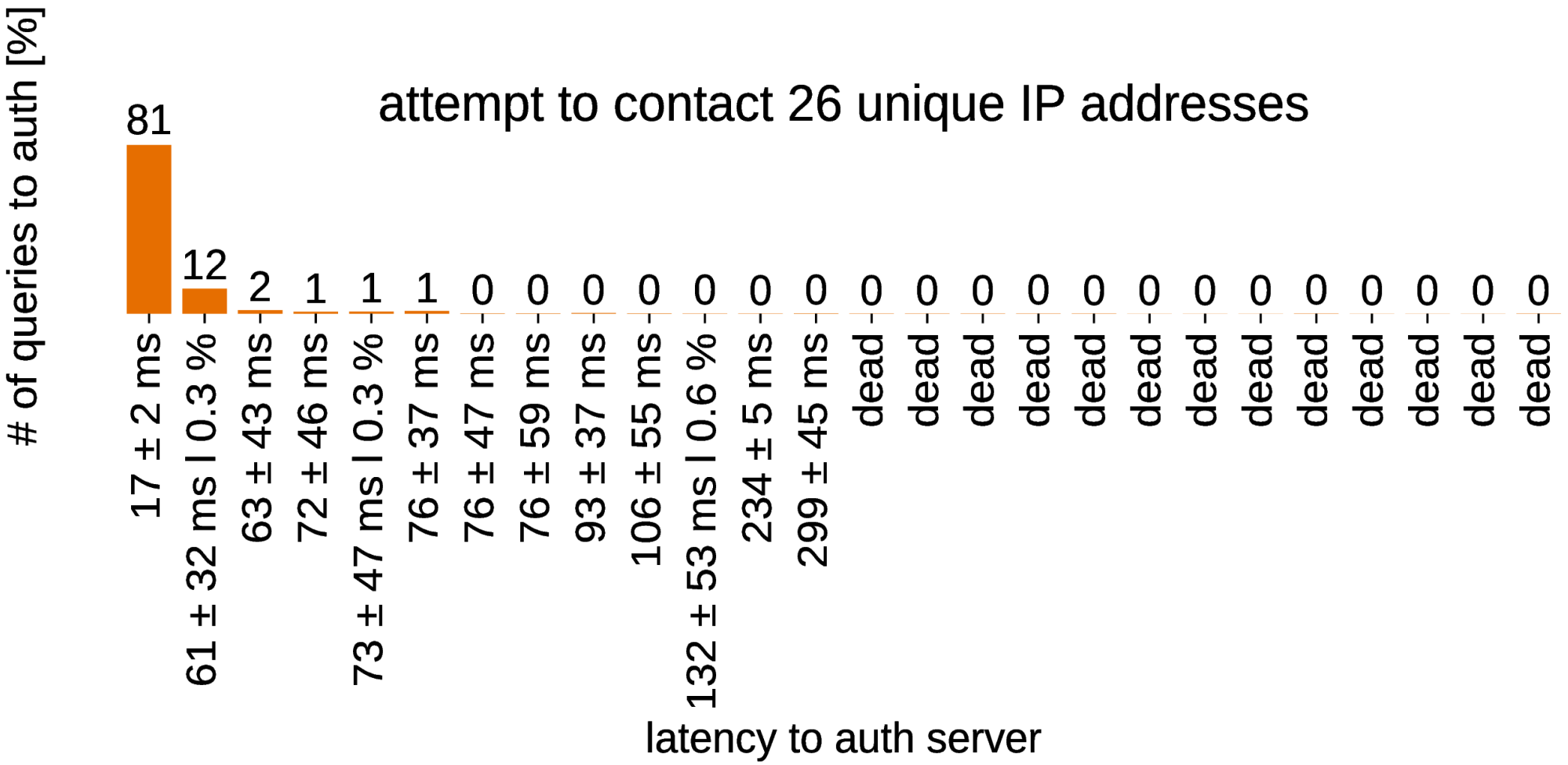
Hawai: BIND packets to auth



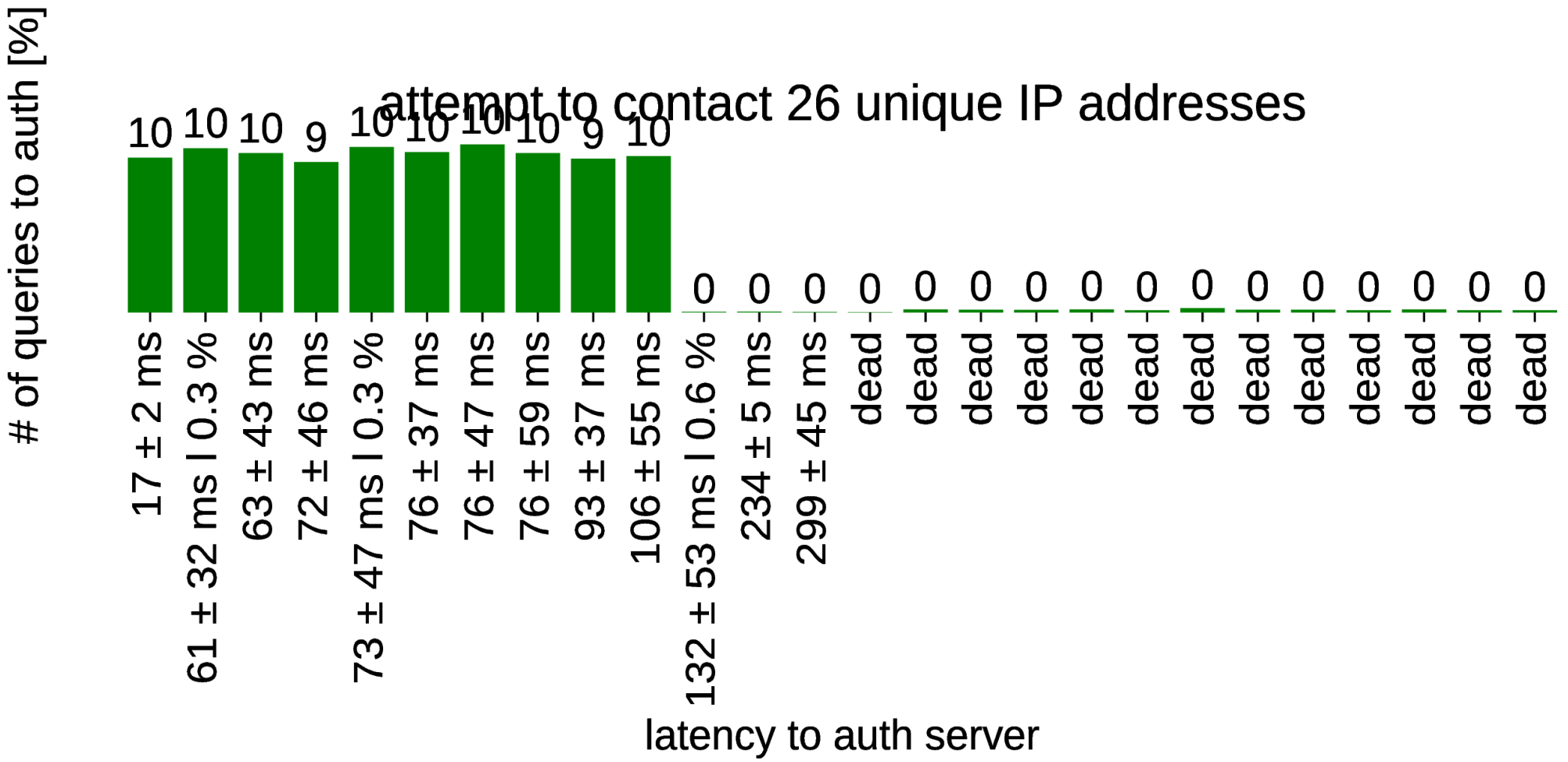
Hawai: KNOT packets to auth



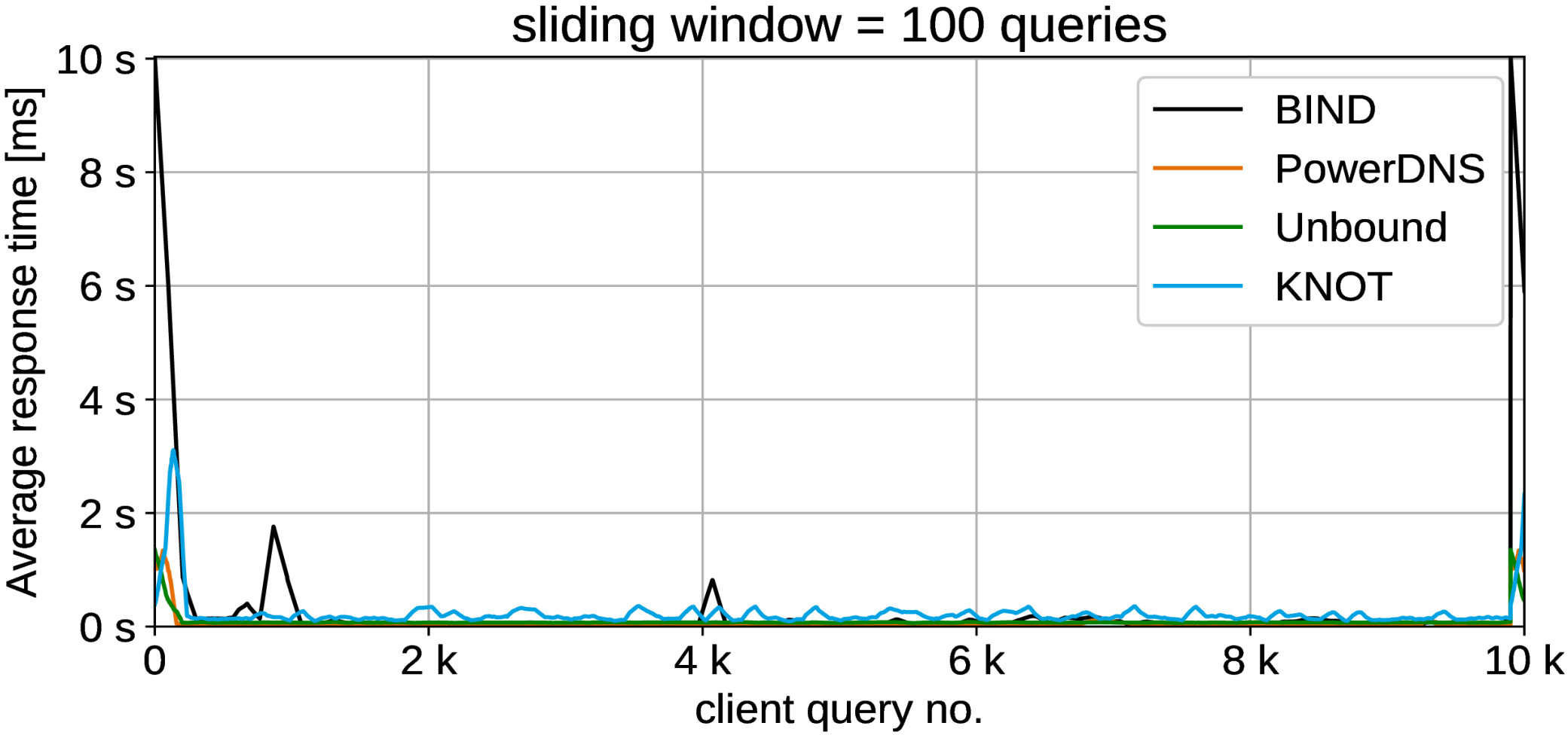
Hawai: PowerDNS packets to auth



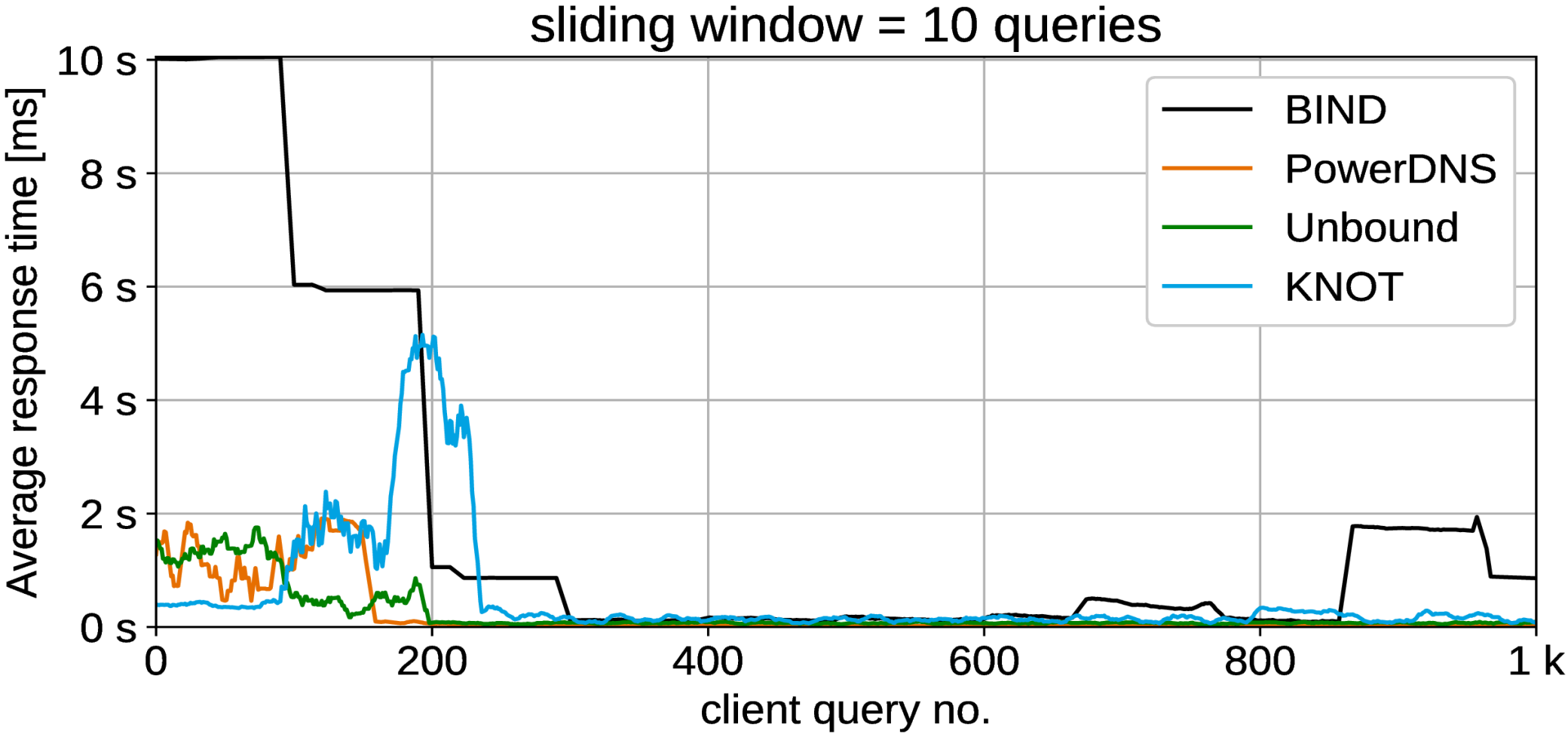
Hawai: Unbound packets to auth



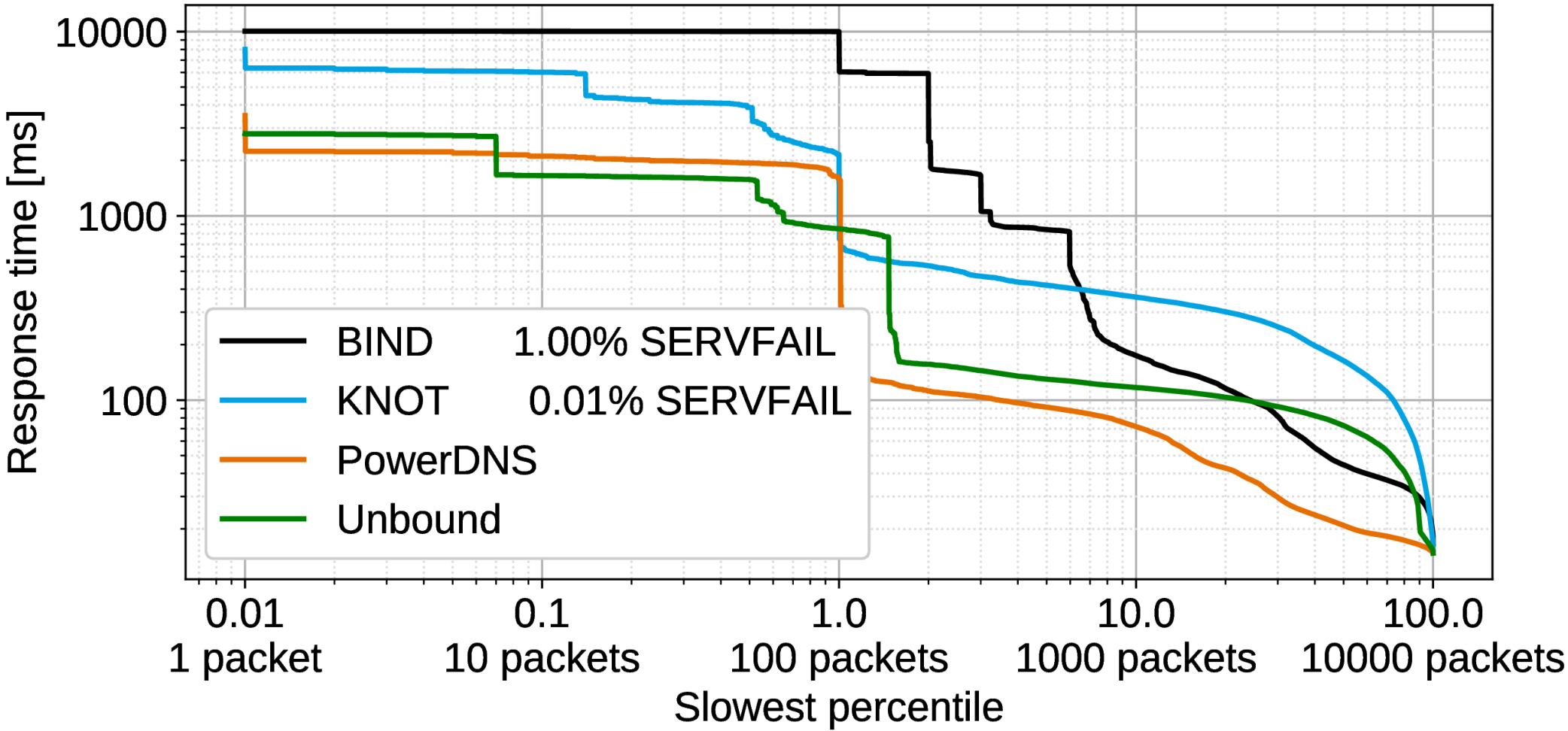
Hawai: latency over time



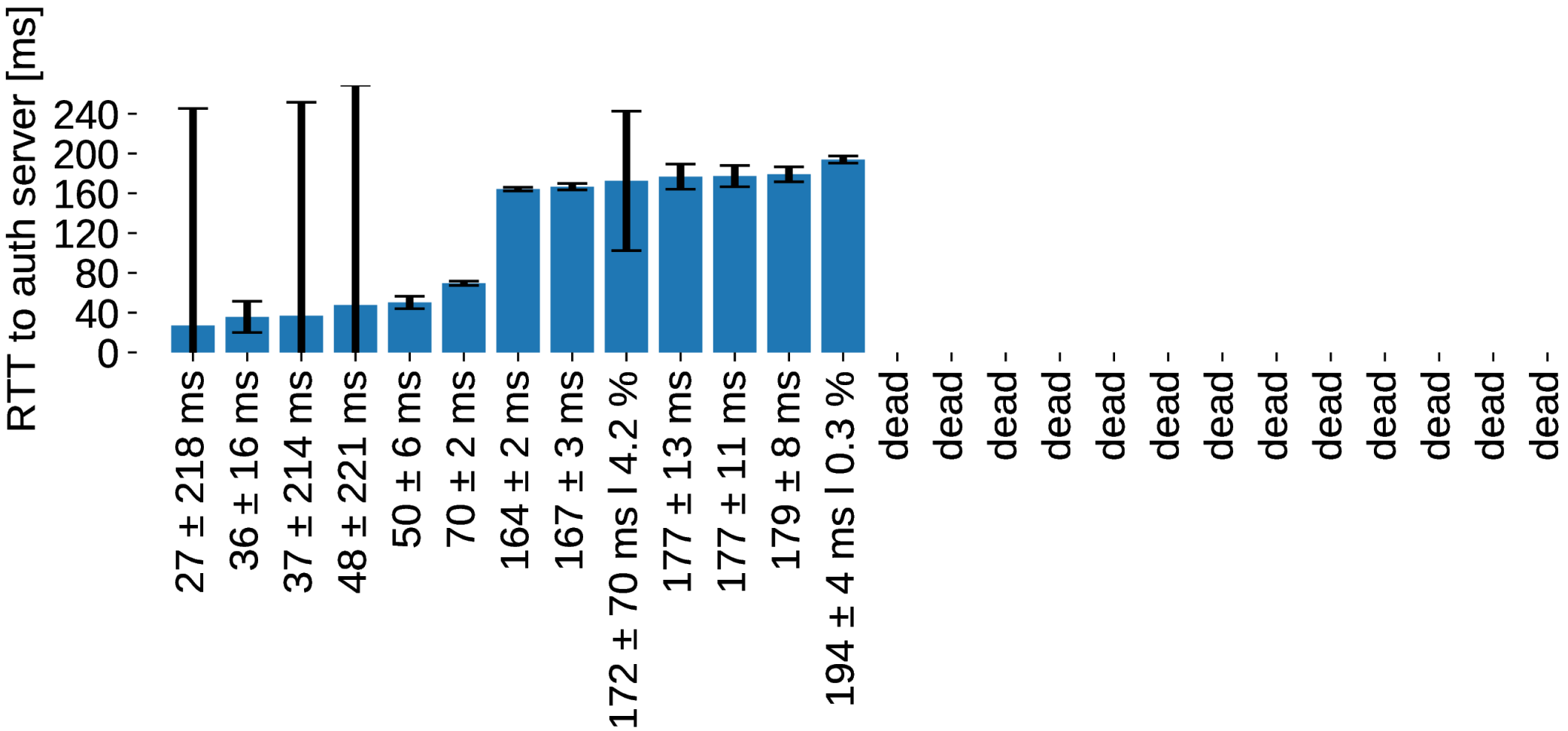
Hawai: latency over time (zoom)



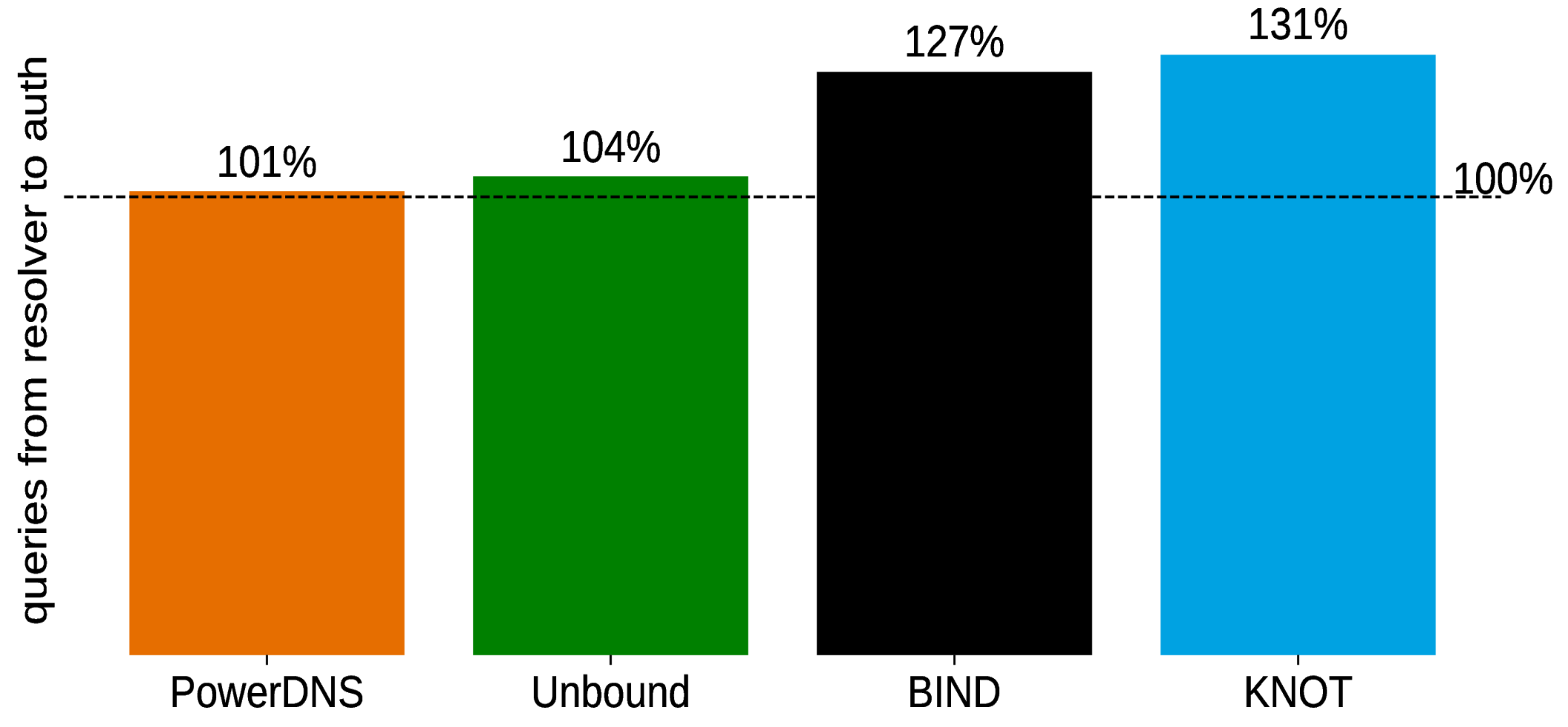
Hawai: latency histogram



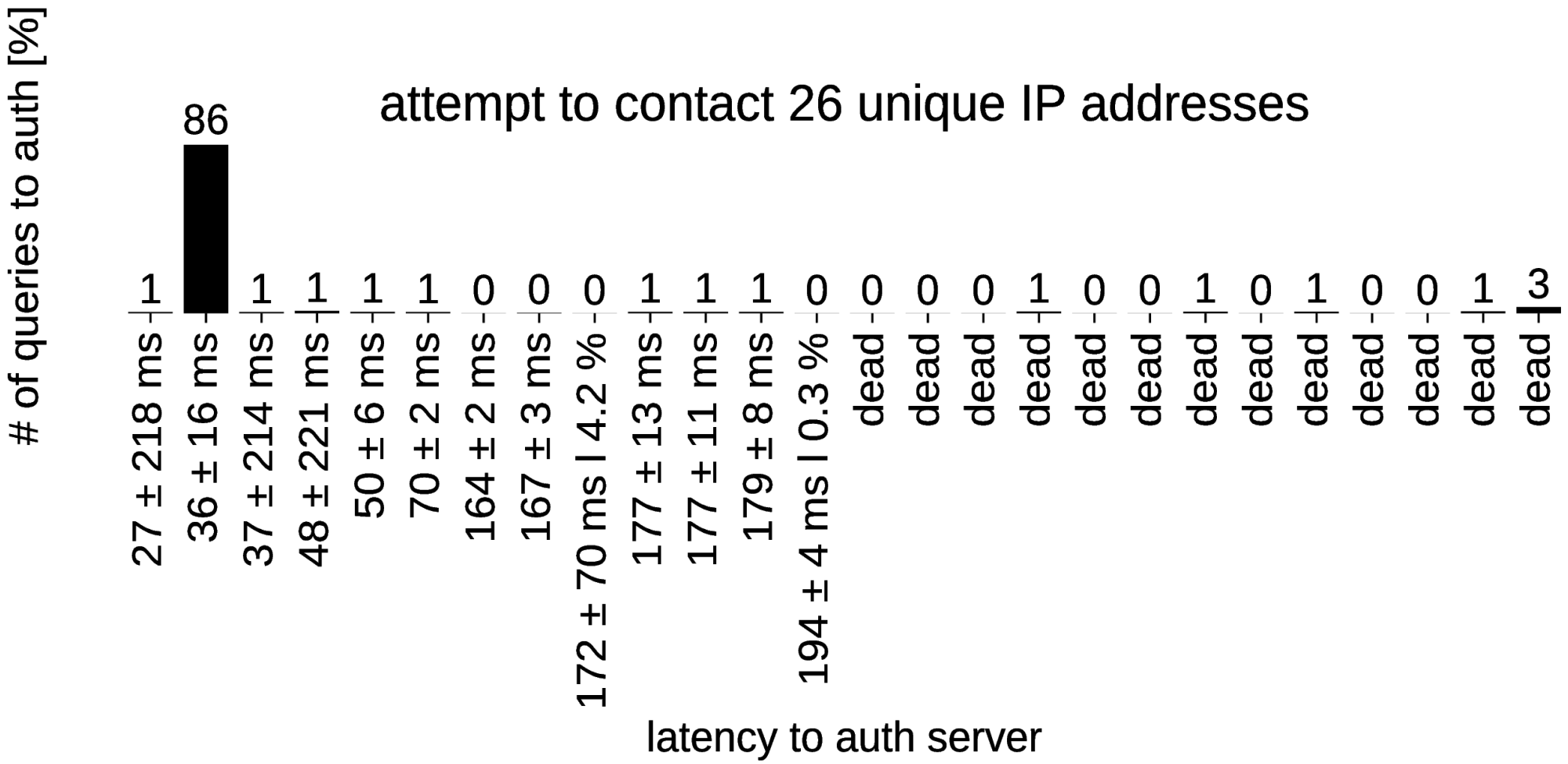
Tasmania: auth servers



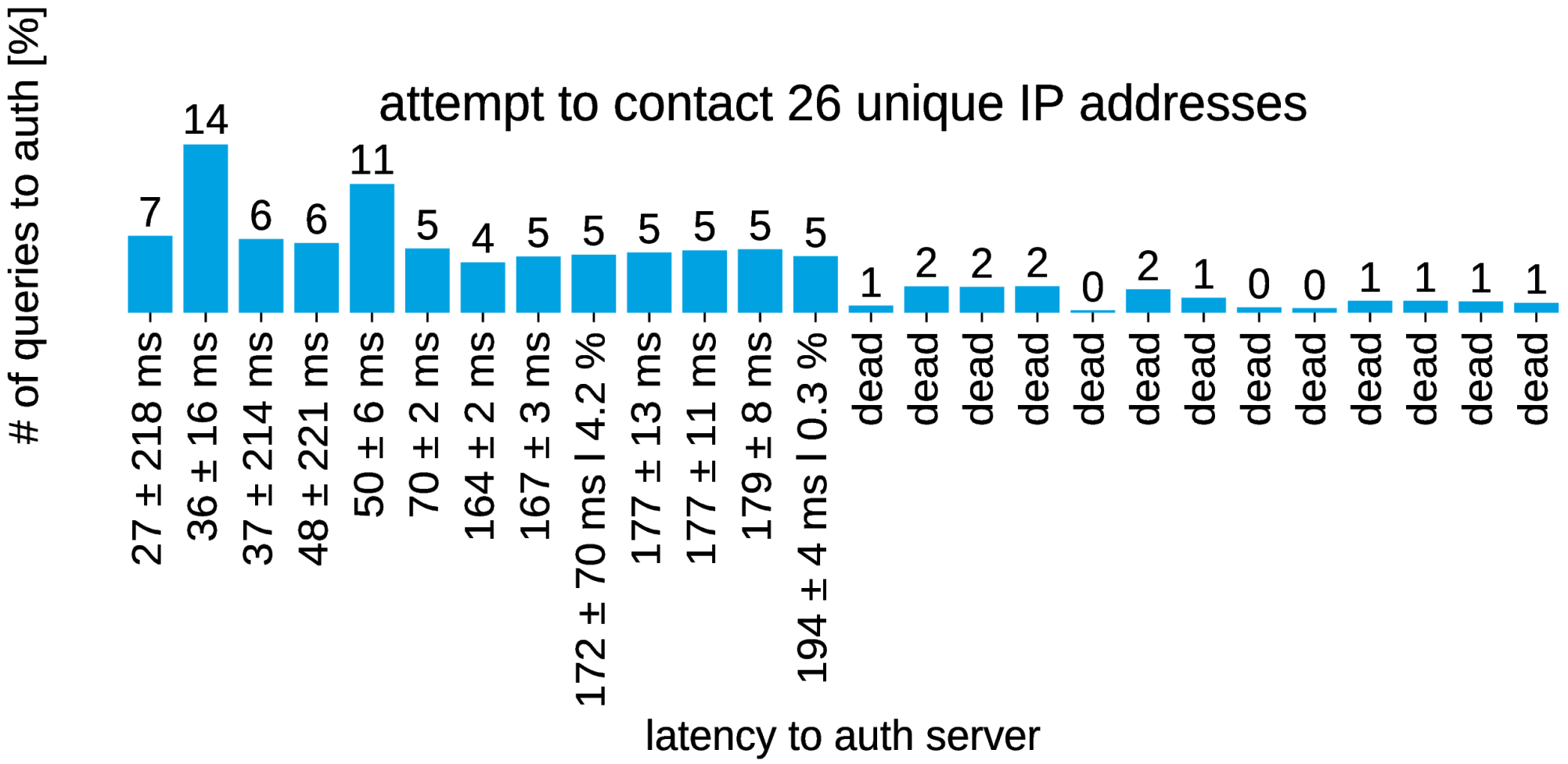
Tasmania: # of queries to auth



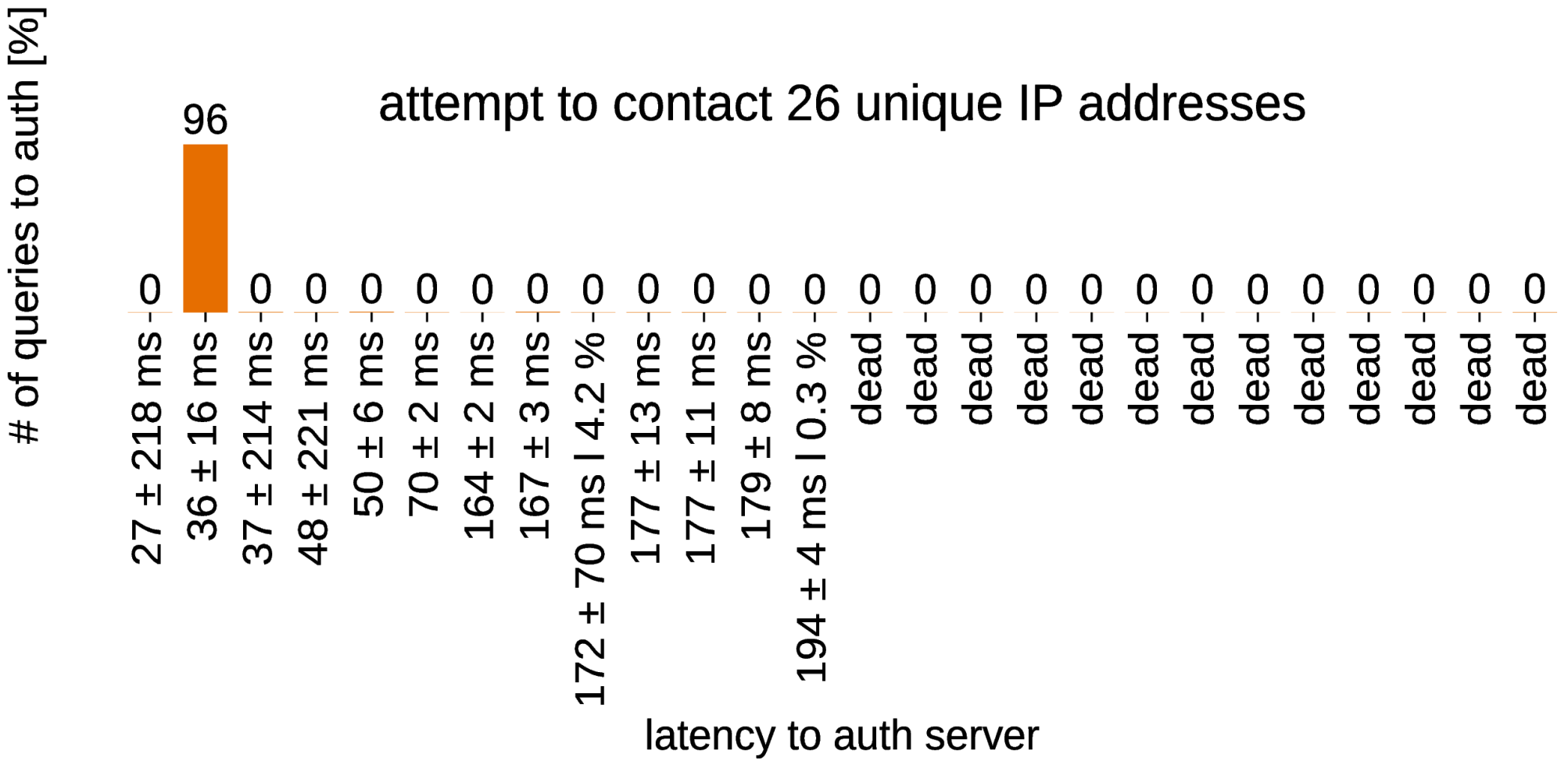
Tasmania: BIND packets to auth



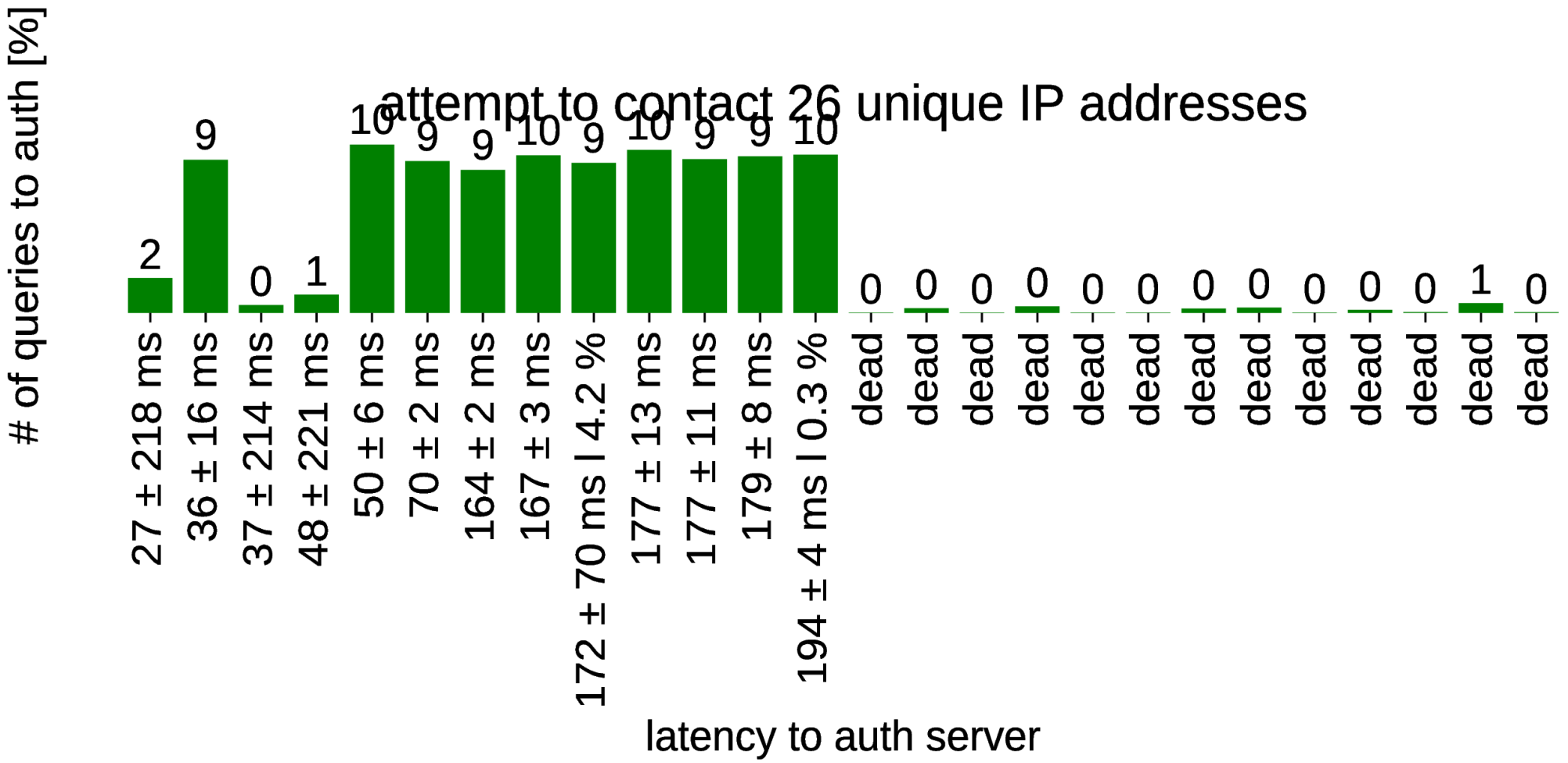
Tasmania: KNOT packets to auth



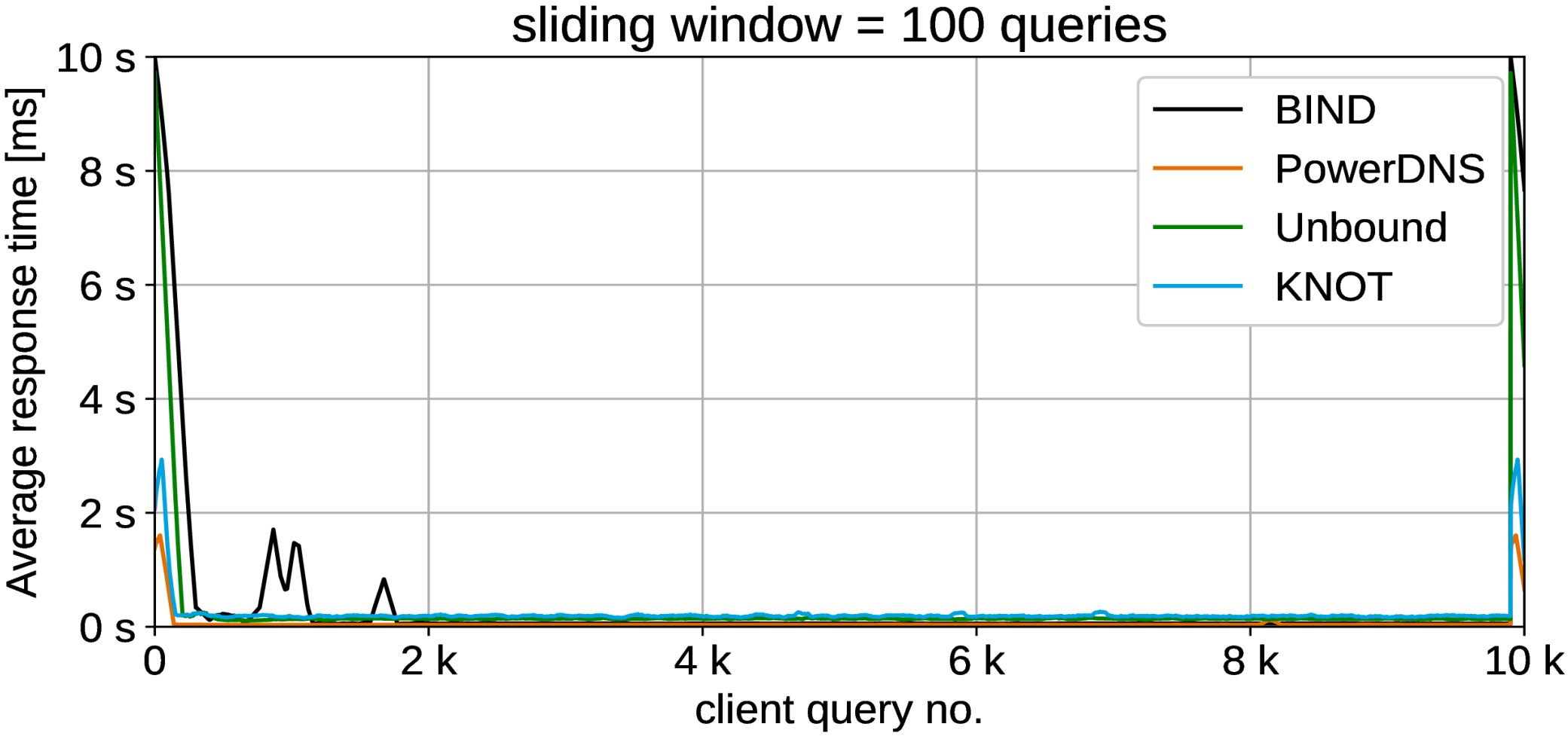
Tasmania: PowerDNS packets to auth



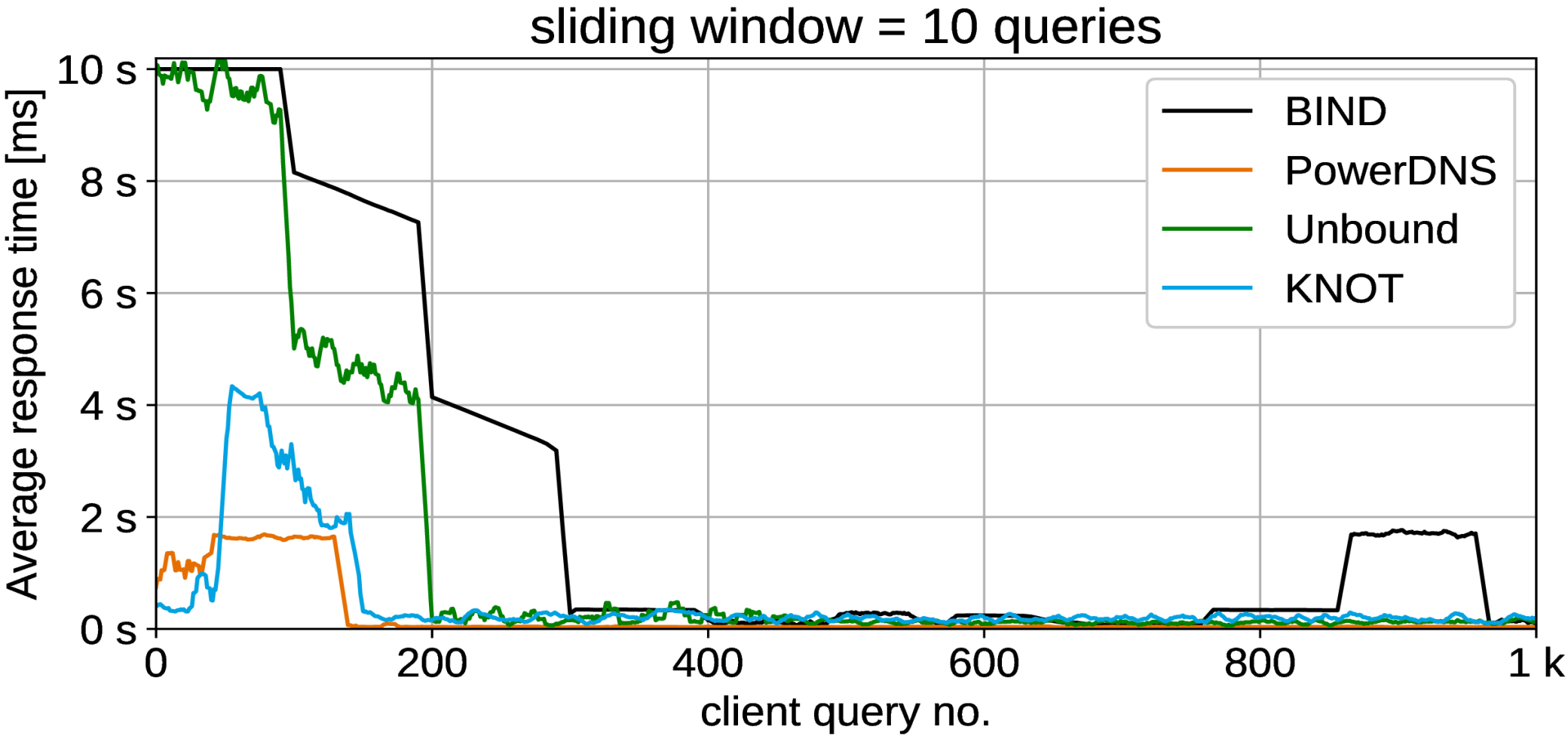
Tasmania: Unbound packets to auth



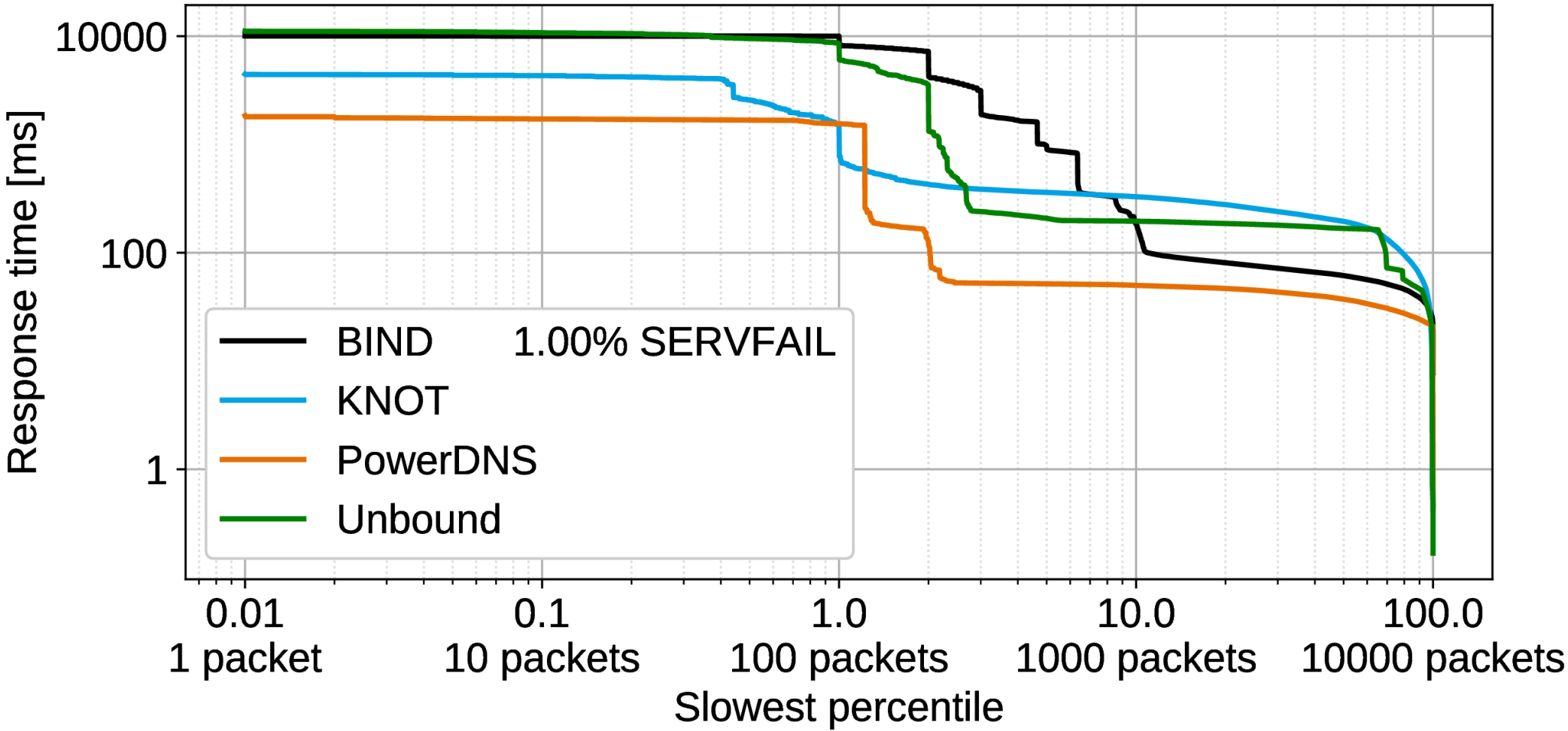
Tasmania: latency over time



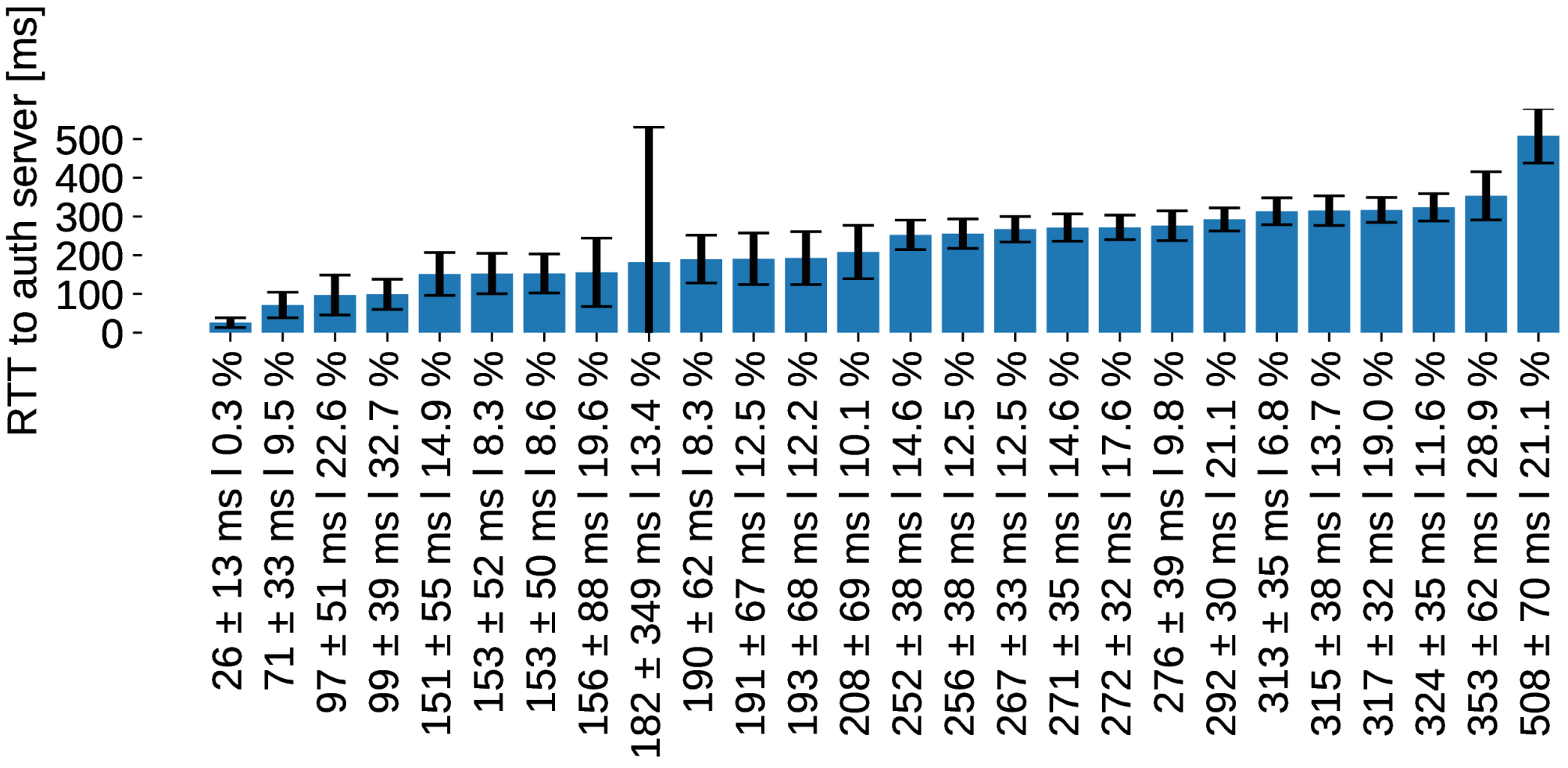
Tasmania: latency over time (zoom)



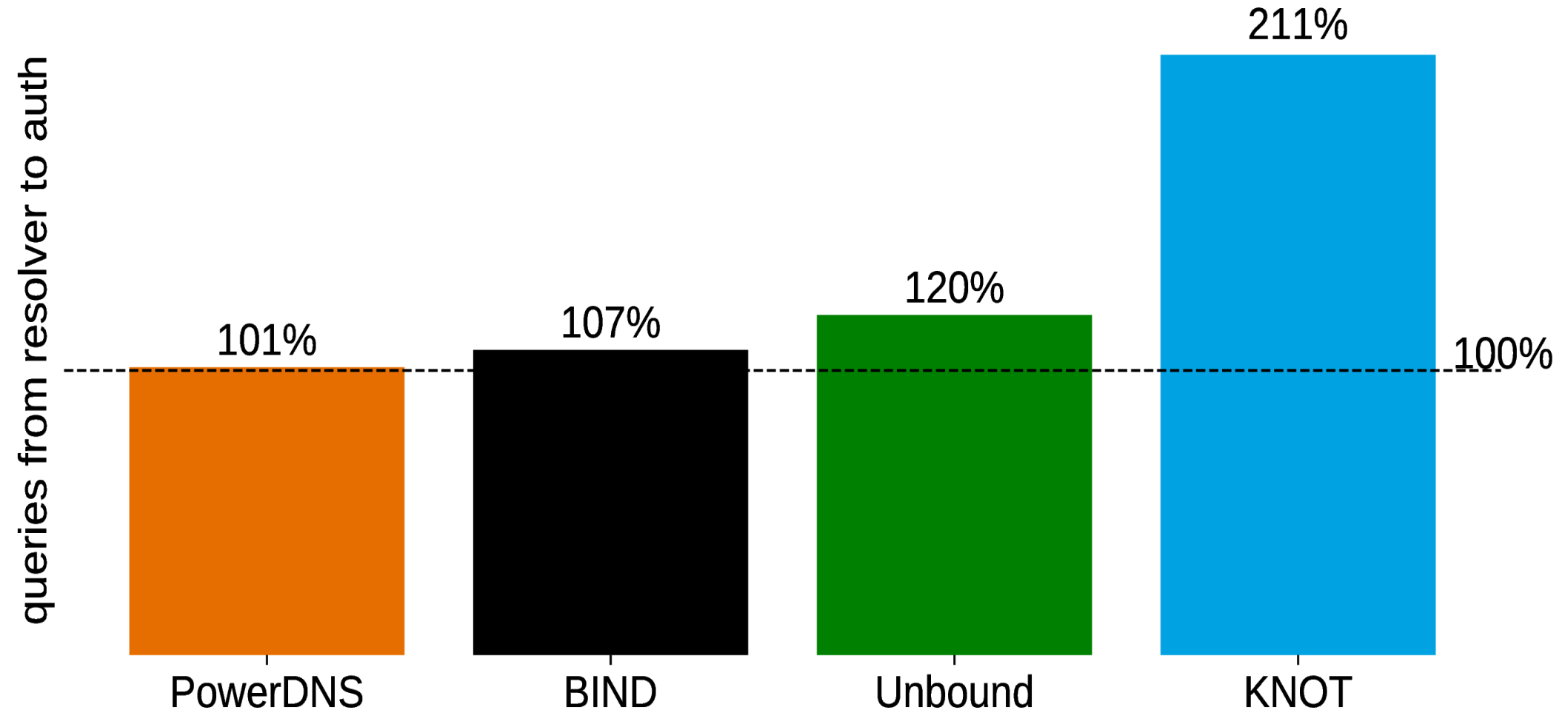
Tasmania: latency histogram



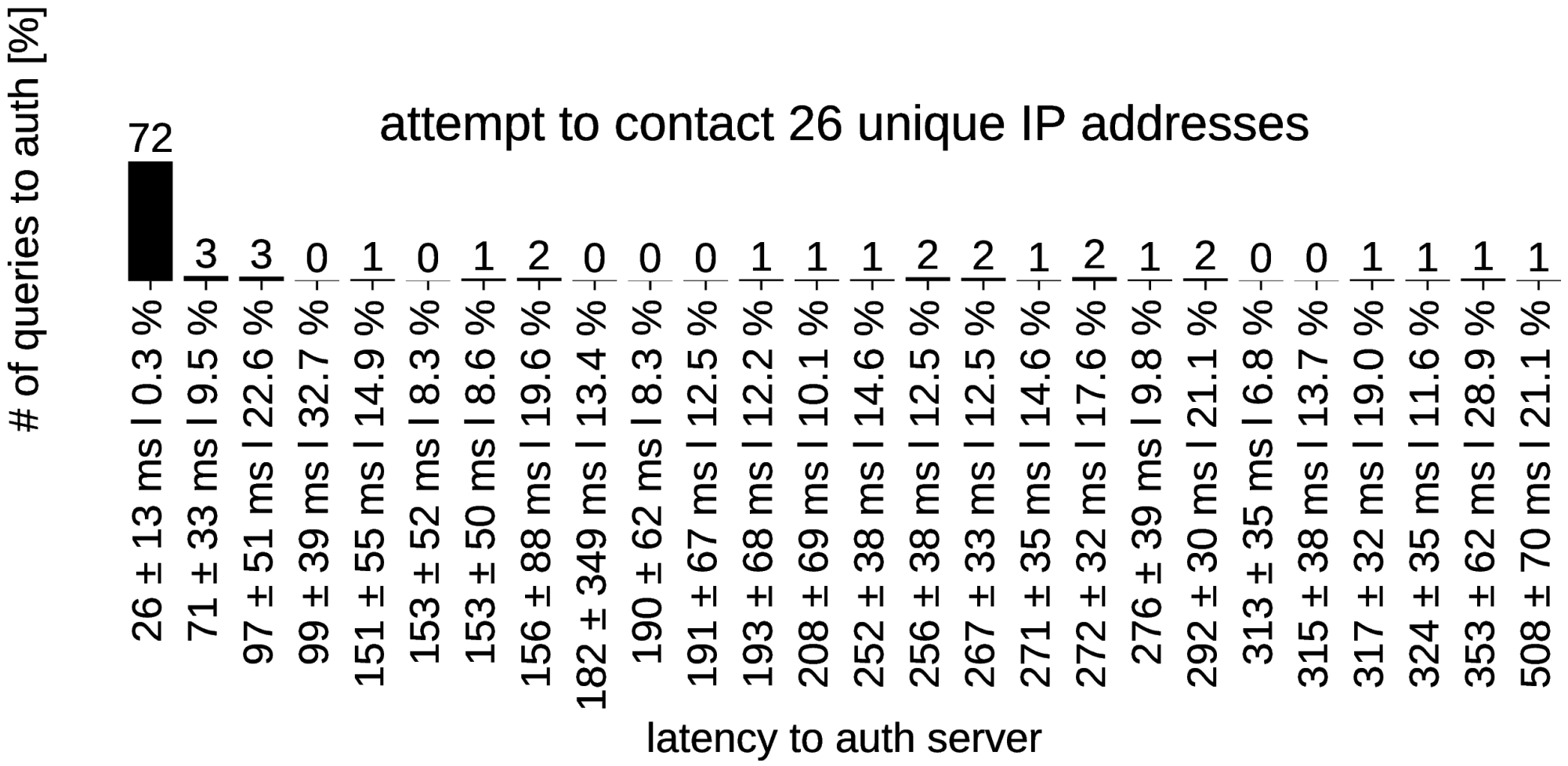
China Atlas #34903: auth servers



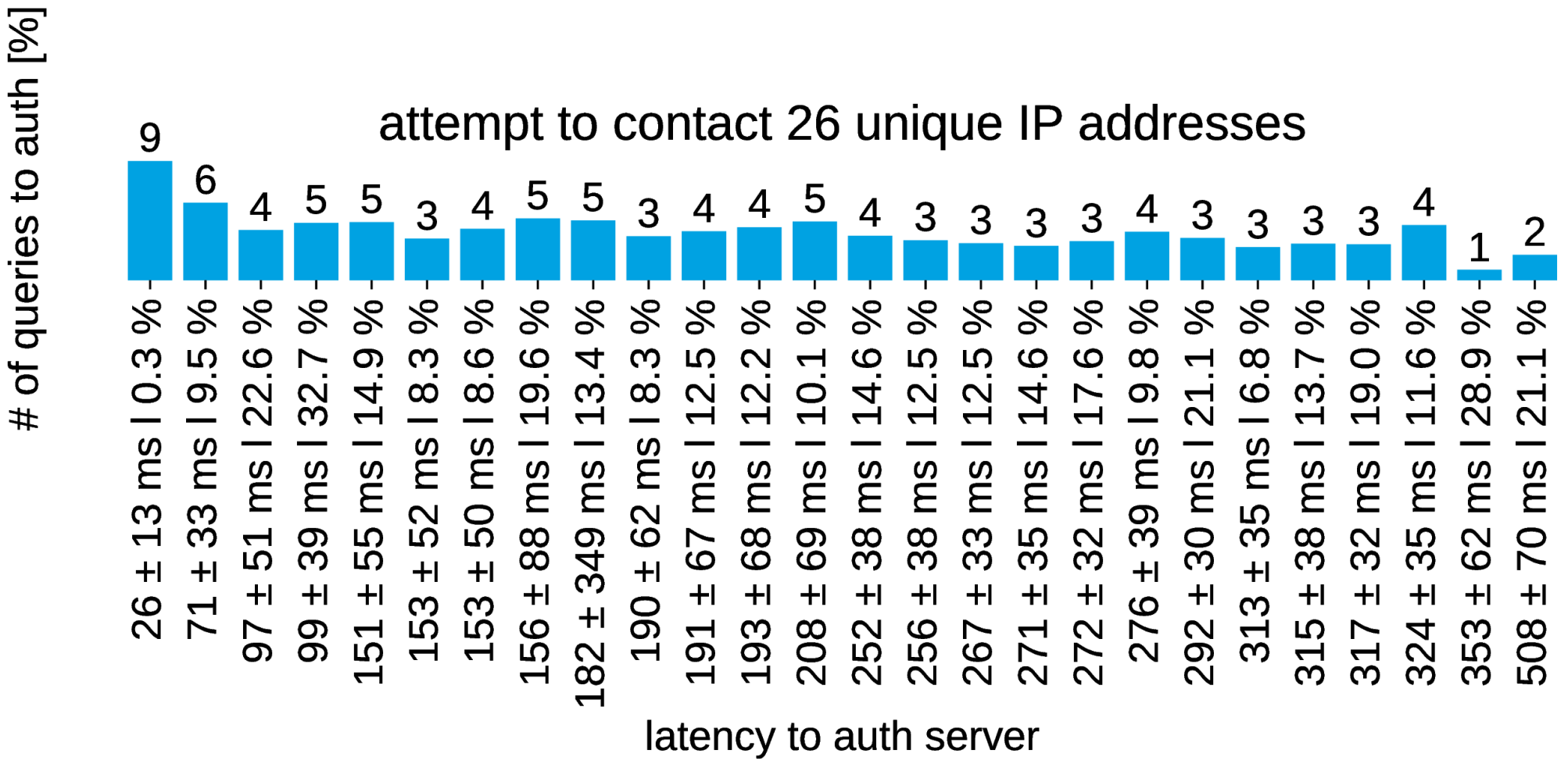
China Atlas #34903: # of queries to auth



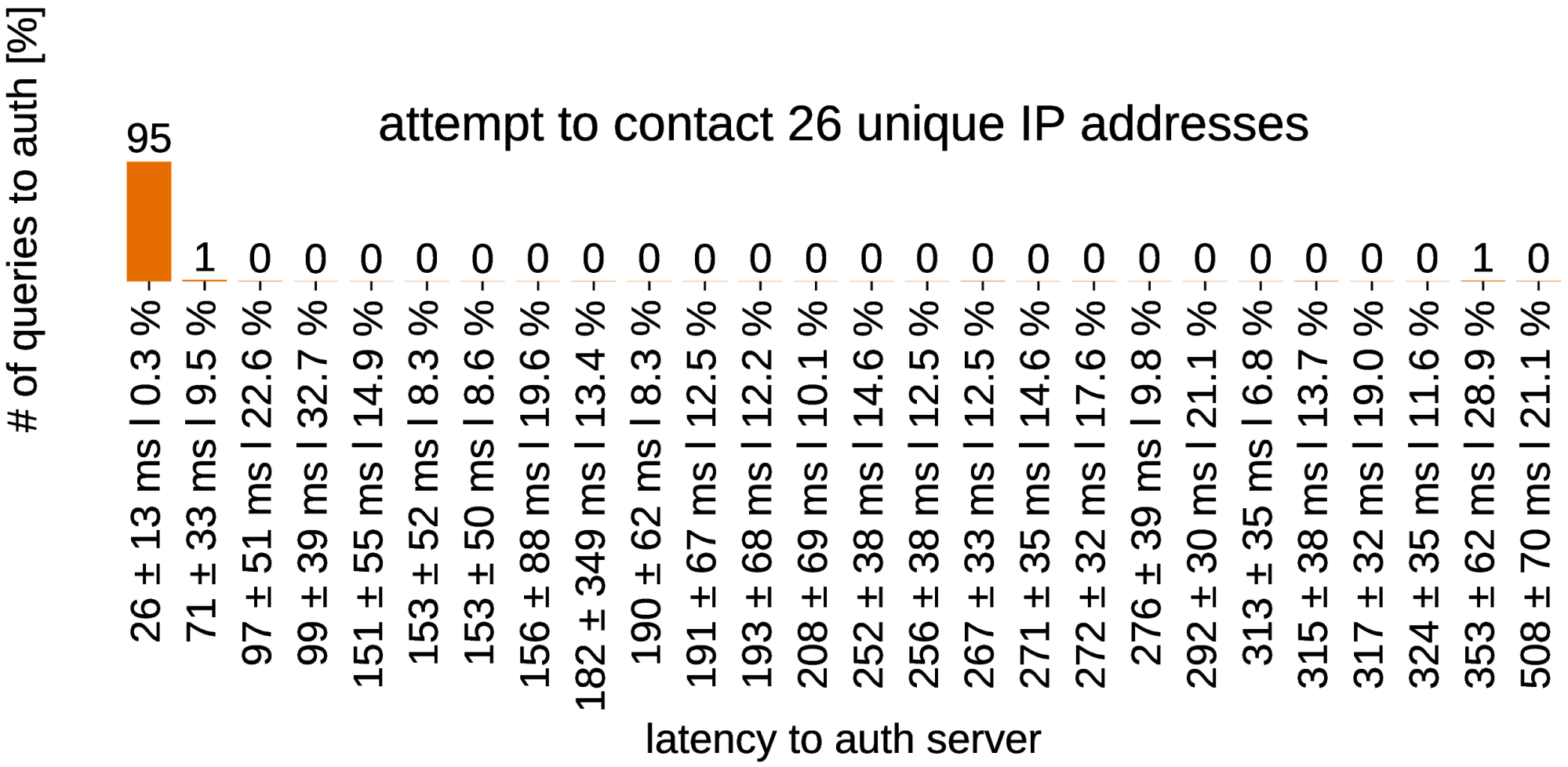
China Atlas #34903: BIND packets to auth



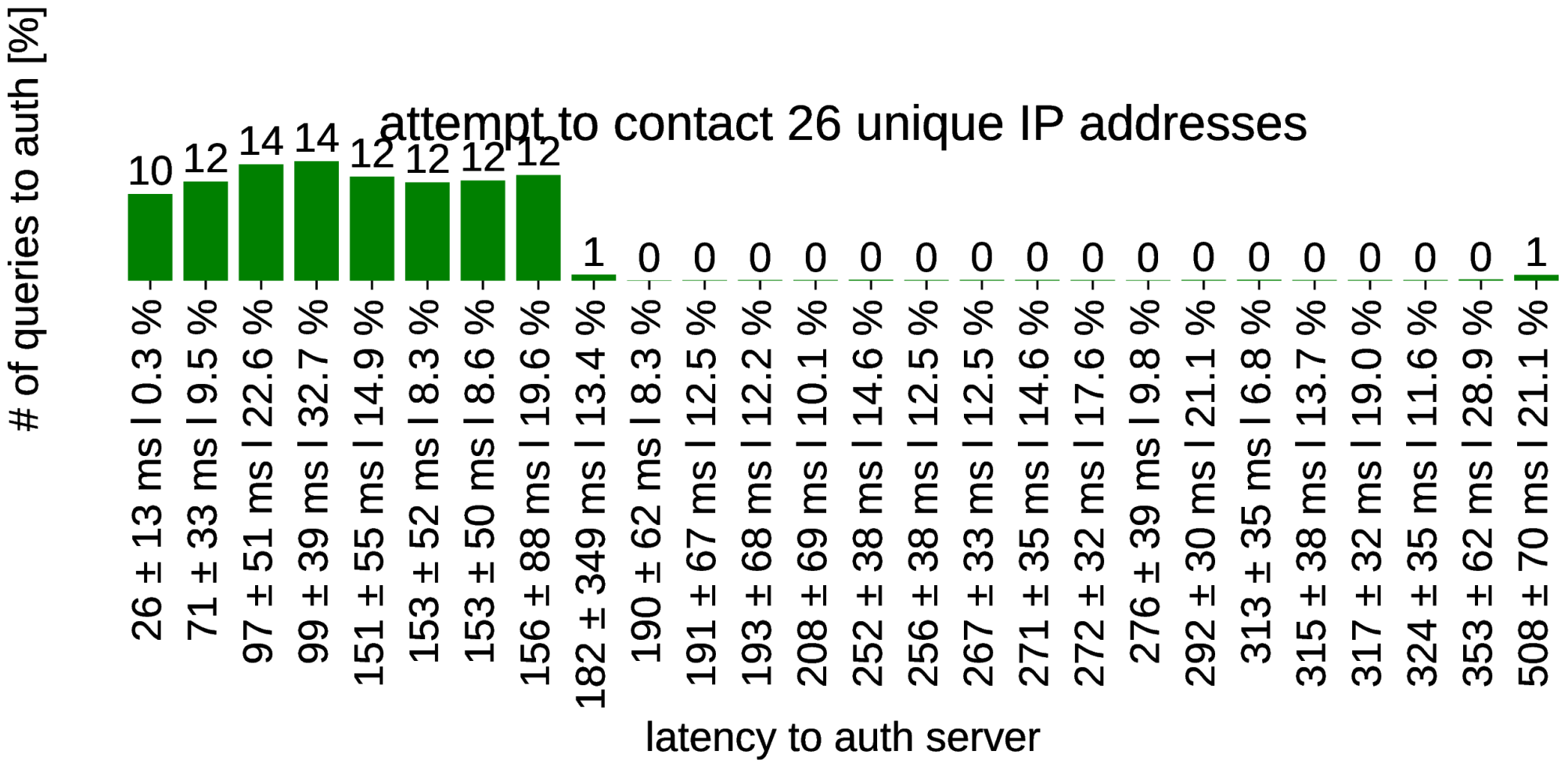
China Atlas #34903: KNOT packets to auth



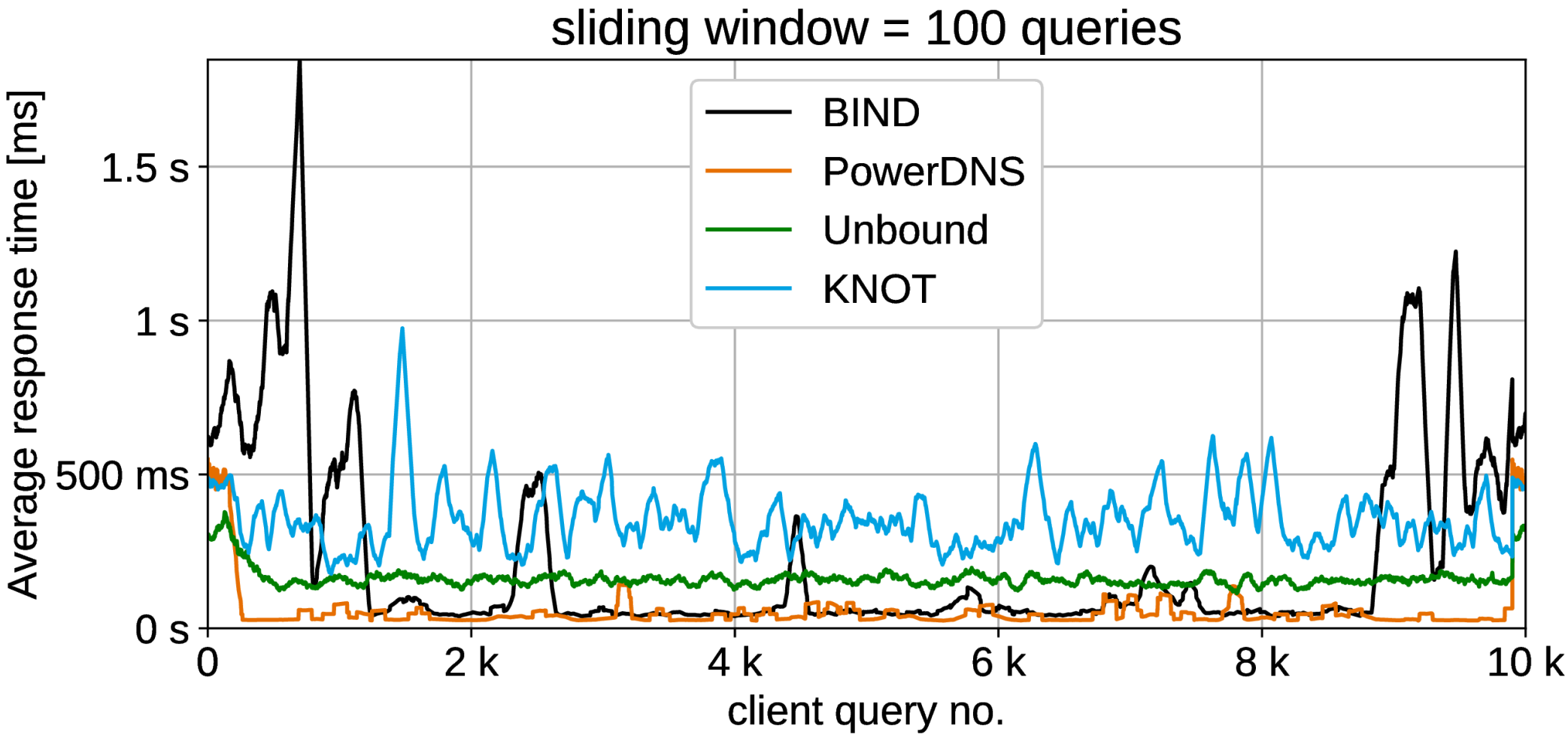
China Atlas #34903: PowerDNS packets to auth



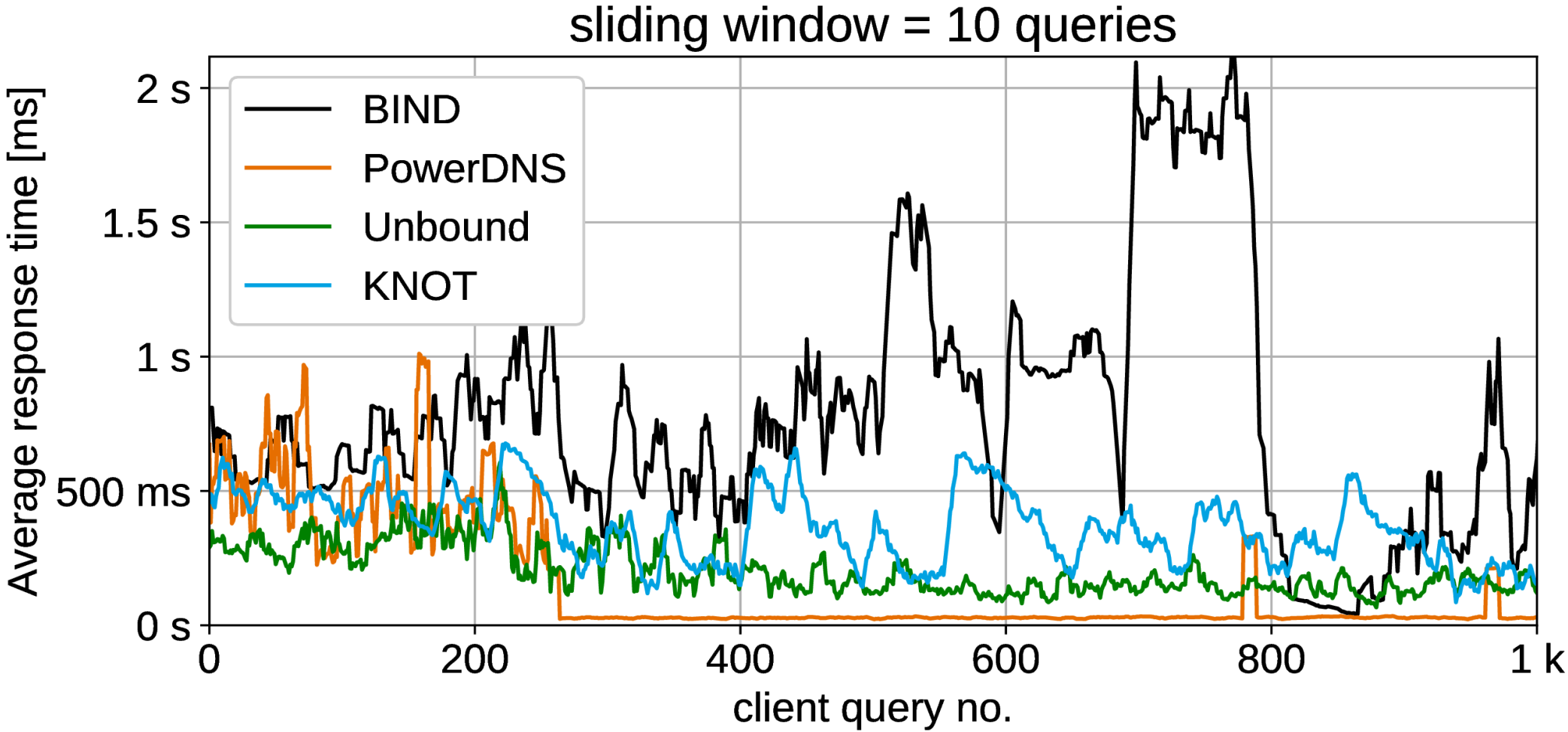
China Atlas #34903: Unbound packets to auth



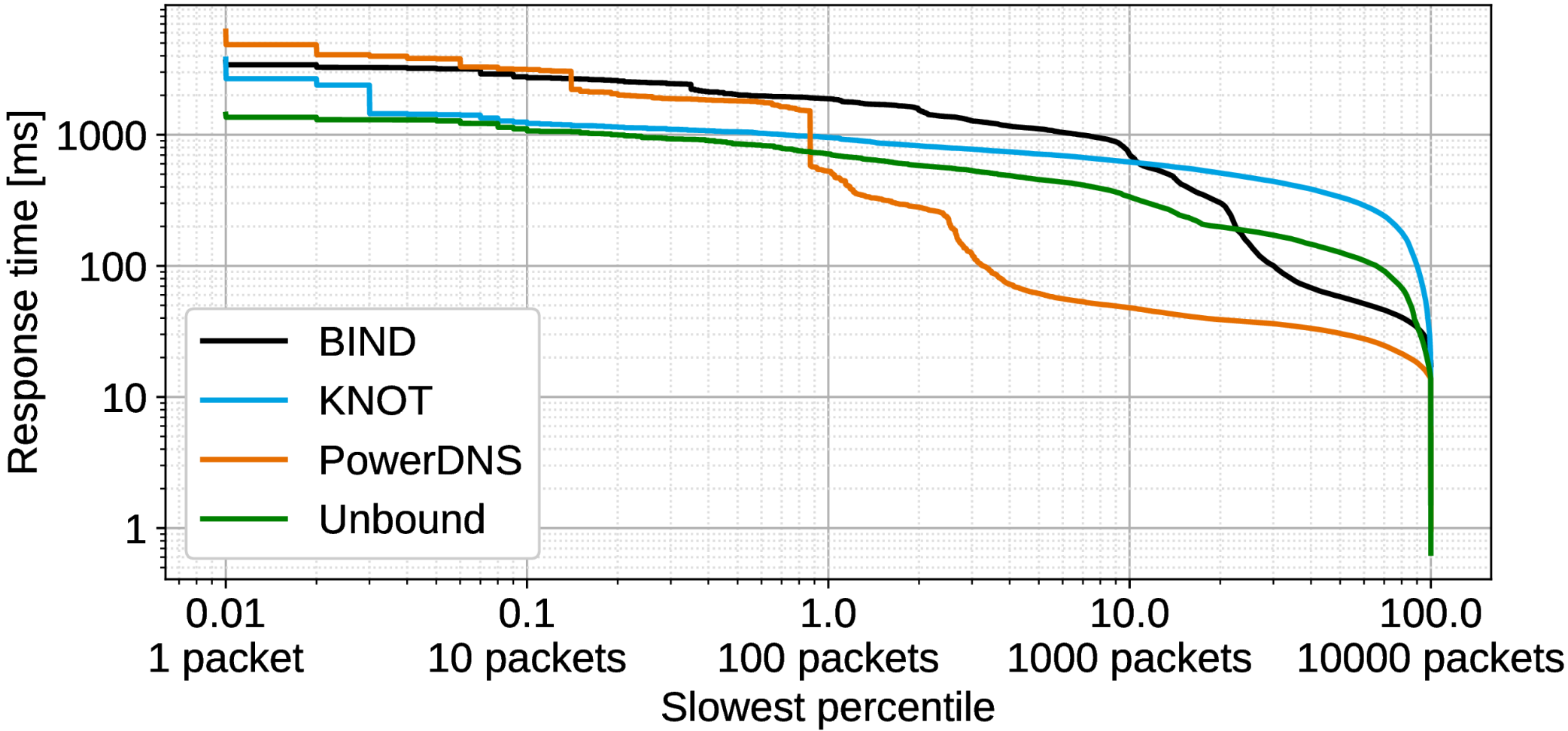
China Atlas #34903: latency over time



China Atlas #34903: latency over time (zoom)



China Atlas #34903: latency histogram



DNS Maze: limitations

- Black-box approach
  - No insight "why"
  - Client-visible latency
    - Not only server selection ...

Further work

- Automate evaluation
- Re-implement algorithms in Teachable DNS resolver
  - Isolate server selection algorithm
- Simulate dynamic changes
- Rewrite server selection in Knot Resolver :-)

Conclusion: Latency != throughput

- QPS throughput practically **NOT** limited by server selection
- Server selection quality
  - this talk
- Throughput
  - talk "Benchmarking and Optimizing DNS Resolvers on the ISP level"
  - http://ripe79.ripe.net/programme/meeting-plan/dns-wg/
- Test your code! :-)