2007 Day In The Life
DNS Root Server Analysis

Duane Wessels
The Measurement Factory/CAIDA

2007 DNS Ops Workshop
July 27, 2007

DITL 2007

- Day In The Life of The Internet. Okay, two days.
- 48 hour period: Jan 9 00:00:00 to Jan 10 23:59:59 UTC
- Primary focus is DNS and root servers, but other data was collected as well.
- We have data from C-, F-, K-, and M-roots, which is the subject of this presentation.
- Data is 740 GB compressed pcap files.
- 10,000,000,000 DNS queries.

Lessons

Problems Uploading

- These have already been covered by Keith
- OARC box ran out of disk space.
- Upload method does not preserve sender-side filename.
- Sender does not have shell access to fix mistakes.
- Receiving program did not handle partial uploads very well.
- Receiving program had bugs with the microsecond part of the first packet timestamp.
- No way to upload a "metadata" file.

Pcap file size and boundaries

- Inconsistent pcap file durations. Most pcap files are 1 hour long, except:
  - C-root: 5 minutes
  - K-root: random, other problems
- Inconsistent pcap file start times.
  - Some pcap files start at 1-hour boundaries.
  - C-, F-, and K-root start at random times.
- Consistent start time and lengths simplifies a number of tasks:
  - Selecting data for analysis
  - Merging pcap files together
  - Knowing if/when all data has been successfully uploaded

Clock Skew

- We sent queries with a timestamp-based query name to known anycast and unicast root servers.
- Found six nodes with skew greater than 3 seconds. One was off by 20 seconds and another by 17.
- Could affect anycast stability analysis?
- This technique is far from perfect.
  - Only hit 'a' nodes of F-root loadbalanced sites.
  - Did not have unicast addresses for C, K, M.
  - Hard to account for transmission delays.

Truncated Packets

- Most K-root instances have truncated packets (1500 vs 1514).
- b.orsn-servers.net has truncated packets (96 vs 1514).
- Probably only an issue for replies, rather than queries.

Unexpected Data

- Pcap files may contain packets for other servers, or even other protocols.
- For example, f-sfo2 pcap files also contain queries to ns-ext.isc.org and d.dns.br.
- We don't know the tcpdump command line and arguments used to capture the data.
- A simple analysis such as 'tcpdump -n -r - dst port domain — wc' may give incorrect results.
- pcap file may also contain queries sent *by* the server.
  - i.e., SOA queries for zone synchronization

Missing Data

- isc/f-dac1a: Missing about 6 hours
  - 2007-01-09 00:00 to 2007-01-09 06:00
  - although it includes an extra 6 hours of data after the end of the collection period. *cron* in wrong TZ?
- isc/f-muc1b: Missing about 22 hours
  - 2007-10-02:15 to the end
  - ISC had a hard time getting the pcap files from this node to the OARC server. Eventually they did upload 50 hours worth of data, but it seems to be shifted by 24 hours from the collection period.
- orsnb/b.root-servers.net: Missing 23 hours at the start, and 1 hour at the end.
- ripe/*: Much of RIPE's data is incomplete
  - Did not have enough local disk space for the capture files.
  - Some data given the wrong name ("poznan") when uploading. fixed?
  - DW accidentally deleted one of the ripe-brisbane files.
- wide/*: Most of the WIDE files are missing the final second or so of each hour from forgetting to call gzclose().

VLAN tags

- f-sfo2 is the only instance where packets are tagged with VLANs.
- A little bit annoying for people that write their own pcap readers.

Gzip Integrity

- Many uploaded pcap files fail a *gzip -t* test.
- Some software (e.g, Coral Reef) ignores the whole file if decompression fails.
- Should we keep the files as they are?
- Or re-compress them to remove the errors?
  - We re-compressed WIDE files

Pcap Integrity

- Some uploaded files encounter errors during pcap processing.
- Leave or fix?

Next Time?

- Intermediate storage sites to prevent data loss?
- Shell access for contributors?
- How much do we care about clock accuracy?
- *dnscap* will save us from truncated packets and other problems?
- Normalize pcap files after receipt by OARC?
  - start/end boundaries
  - remove pcap/gzip errors
  - remove VLAN tags
  - remove irrelevant packets

Results

Terminology

- Server: a collection of DNS nameservers operating under the same IP address.
  - c.root-servers.net is a server
- Instance: an anycast instance of a server.
  - k-milan is an instance of k.root-servers.net.

  Load-balanced nodes are combined into a single instance.
  - c-lax1a and c-lax1b are load-balanced members of the c-root LAX instance.
- Client: an IP address sending DNS queries.

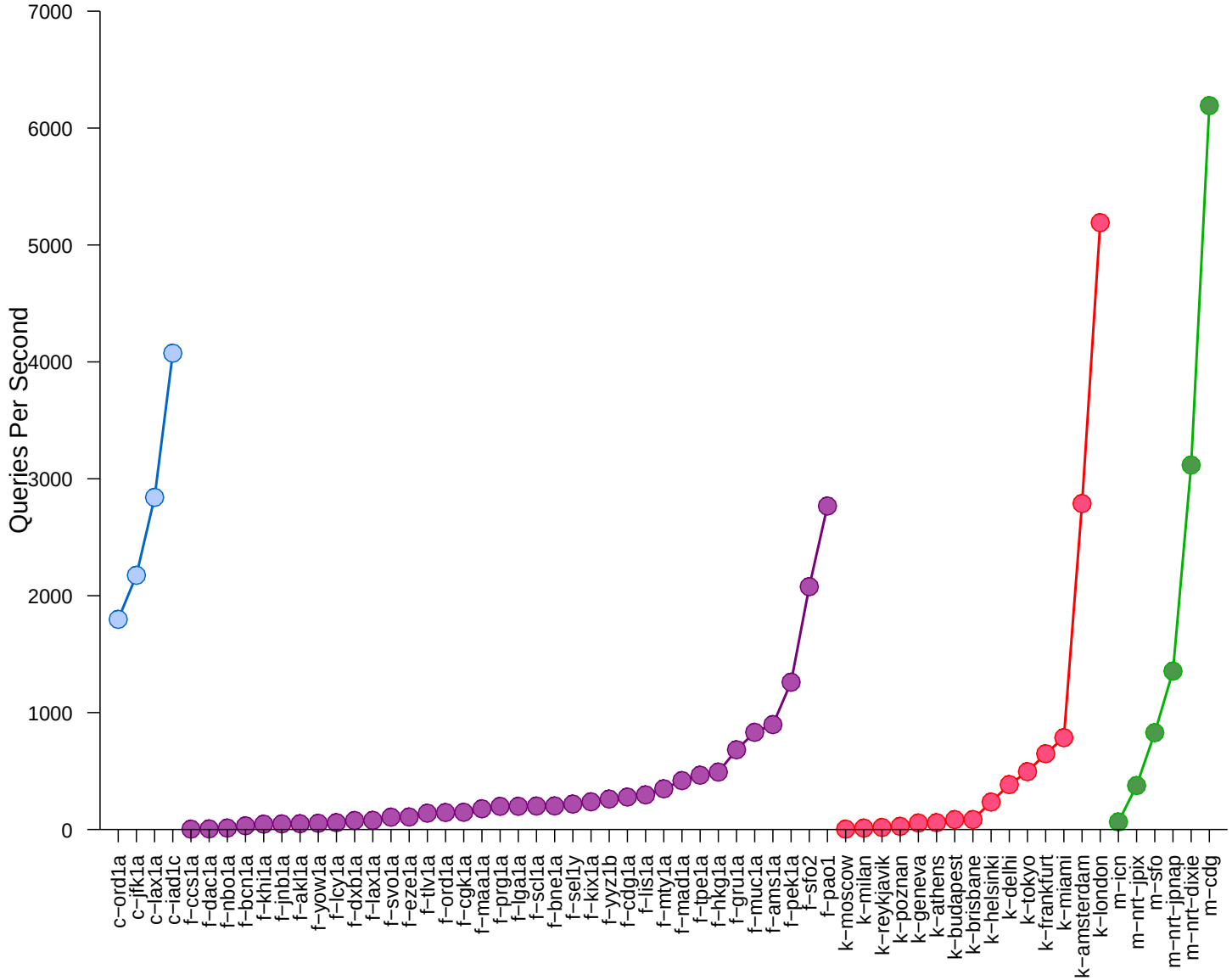
Merging Pcaps

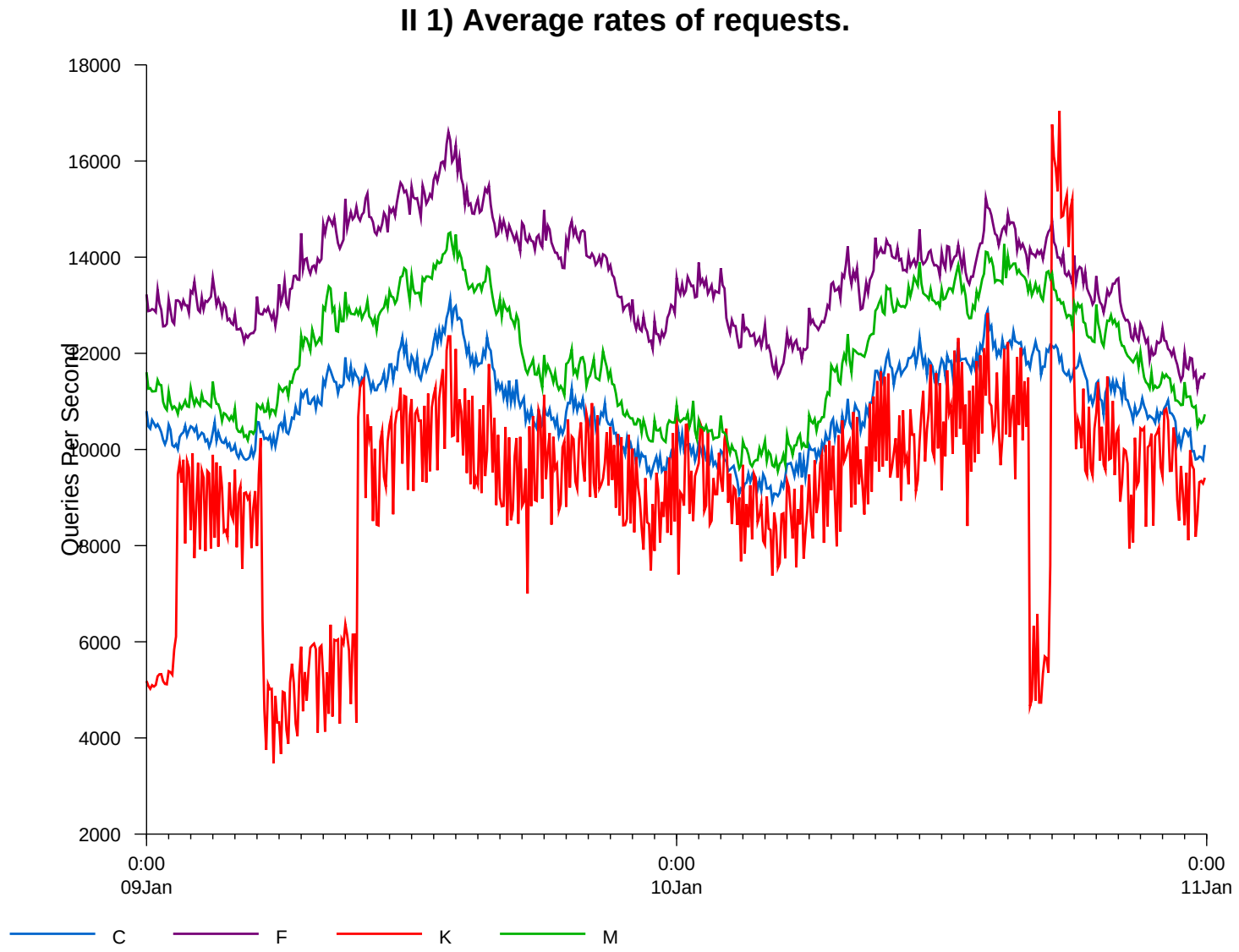
- First step was to create a single, "merged" pcap stream with all packets in chronological order.
- Created hour-long chunks for all instances, using *tcpdump-join* and *tcpdump-split*. Keep only data within the 48-hour DITL period. Queries only.
- Changed pcap timestamps for instances with known clock skew.
- Rewrote server IP addresses to encode server and instance.
  - e.g., 192.5.5.241 becomes 6.0.0.11 to represent the 11th instance of F-root.
- Merged all hour-long instance files into timestamp-sorted files with *mergecap*.

Analysis Software

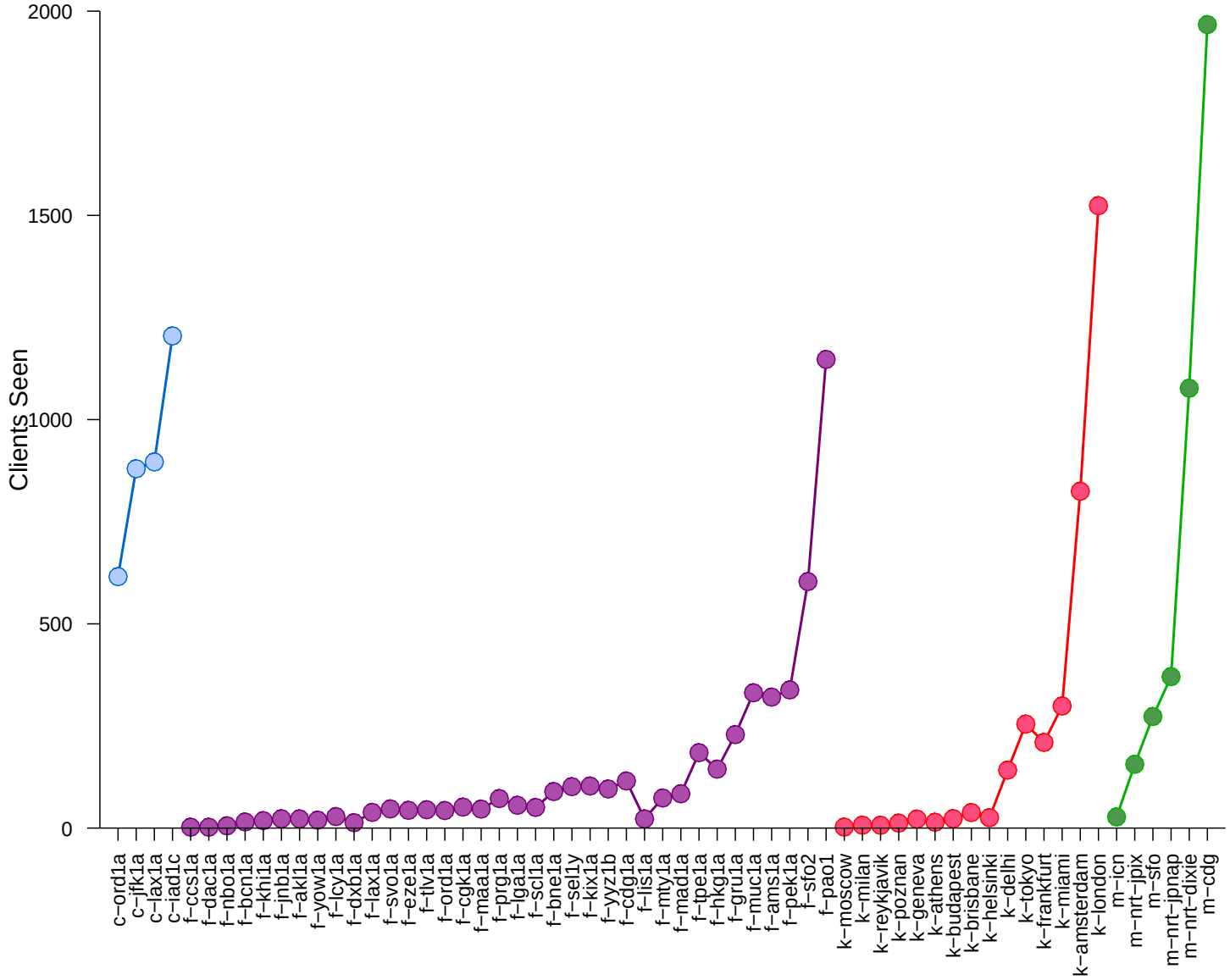
- C++ program reads pcap files and keeps various counters.
- Runs at about 40,000 packets/second, or about 80% the rate of "pcap time."
  - i.e., takes 60 hours to analyze 48 hours of data.
- Needs about 3GB RAM.
- Data goes into Postgres
- SQL SELECT statements and perl scripts produce data for ploting with *ploticus*.

II 1) Average rates of requests.

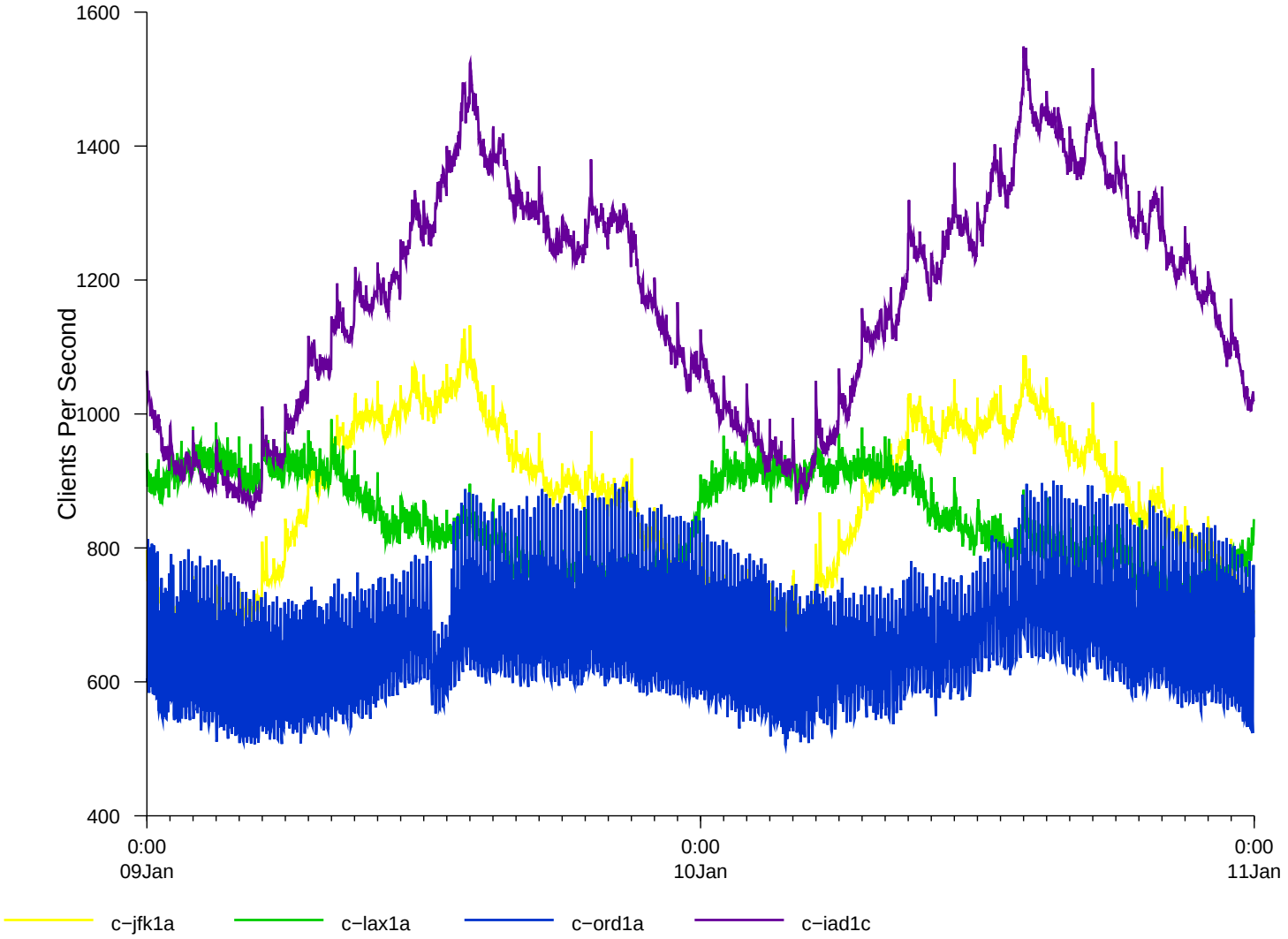


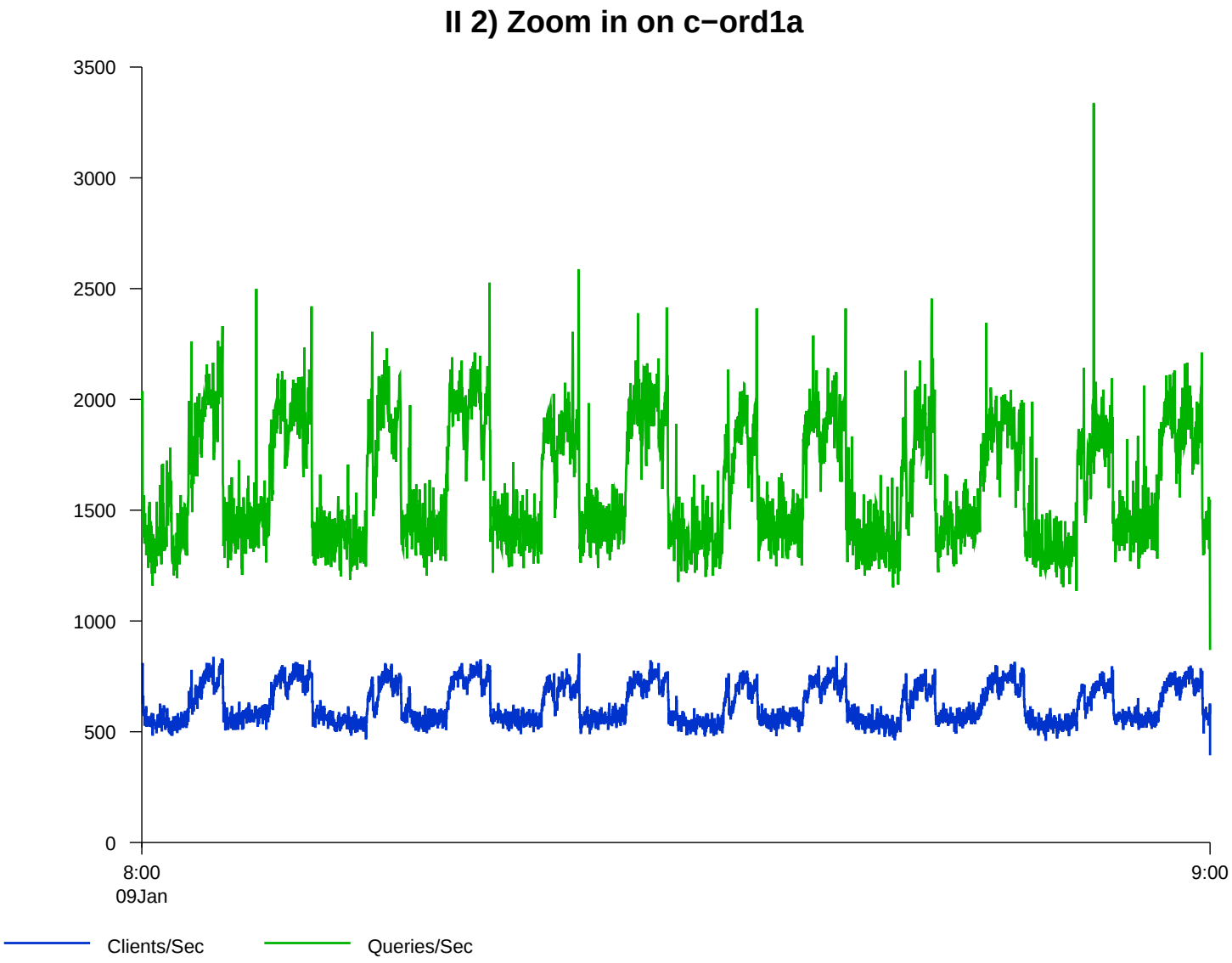

II 1) Average rates of requests.
Queries Per Second
18000
16000
14000
12000
10000
8000
6000
4000
2000
0:00
09Jan
0:00
10Jan
0:00
11Jan
C
F
K
M

II 2) The average number of clients per second seen at each instance.



II 2) The number of clients per second seen at each C−root instance.

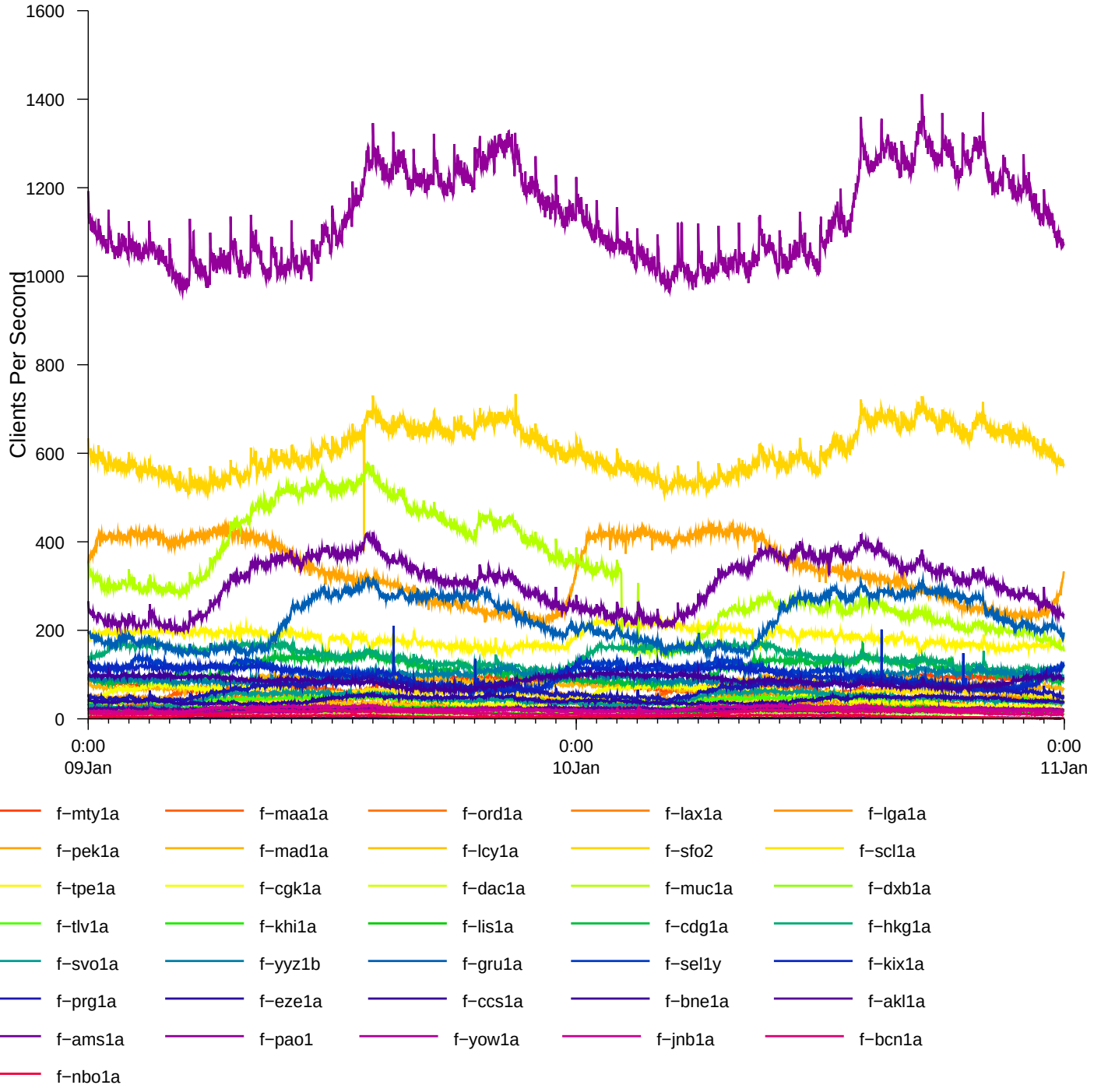



II 2) Zoom in on c−ord1a
3500
3000
2500
2000
1500
1000
500
0
8:00
09Jan
9:00
Clients/Sec
Queries/Sec

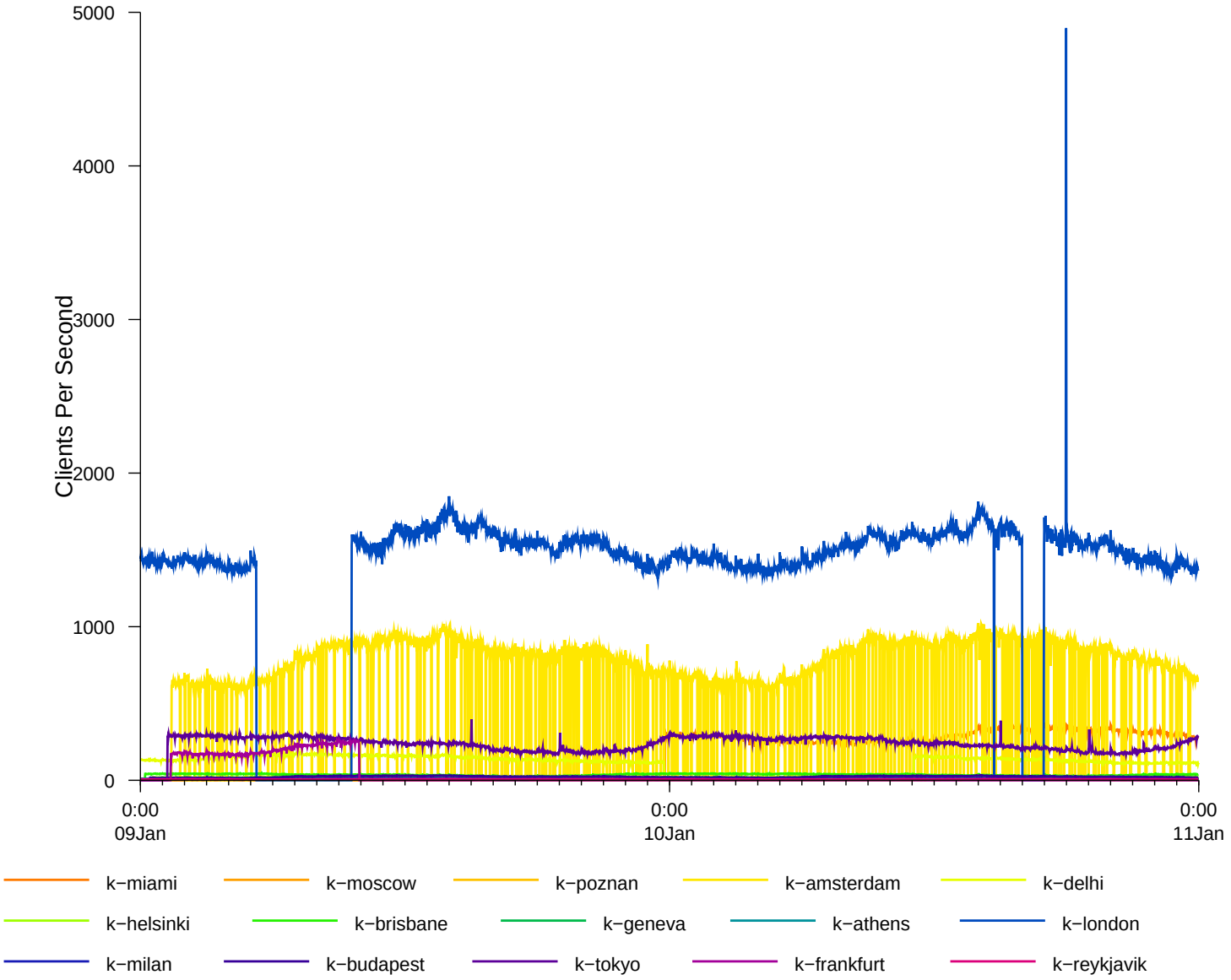
The cause??

```
Date: Thu, 11 Jan 2007 01:03:47 +0000
From: Paul Vixie <paul@vix.com>
To: wessels@Oarc.isc.org
Subject: oops

#ord1a.c:i386# jobs
[1]  + Running                       ./tcpdump -s 0 -n -w oarc.tcpd. -z gzip
-P 5 host c.root-servers.net
#ord1a.c:i386# kill %1

626638995 packets captured
667208048 packets received by filter
15549 packets dropped by kernel
[1]    Done                          ./tcpdump -s 0 -n -w oarc.tcpd. -z gzip
-P 5 host c.root-servers.net

i had two tcpdumps running on one of the c-root boxes...
```

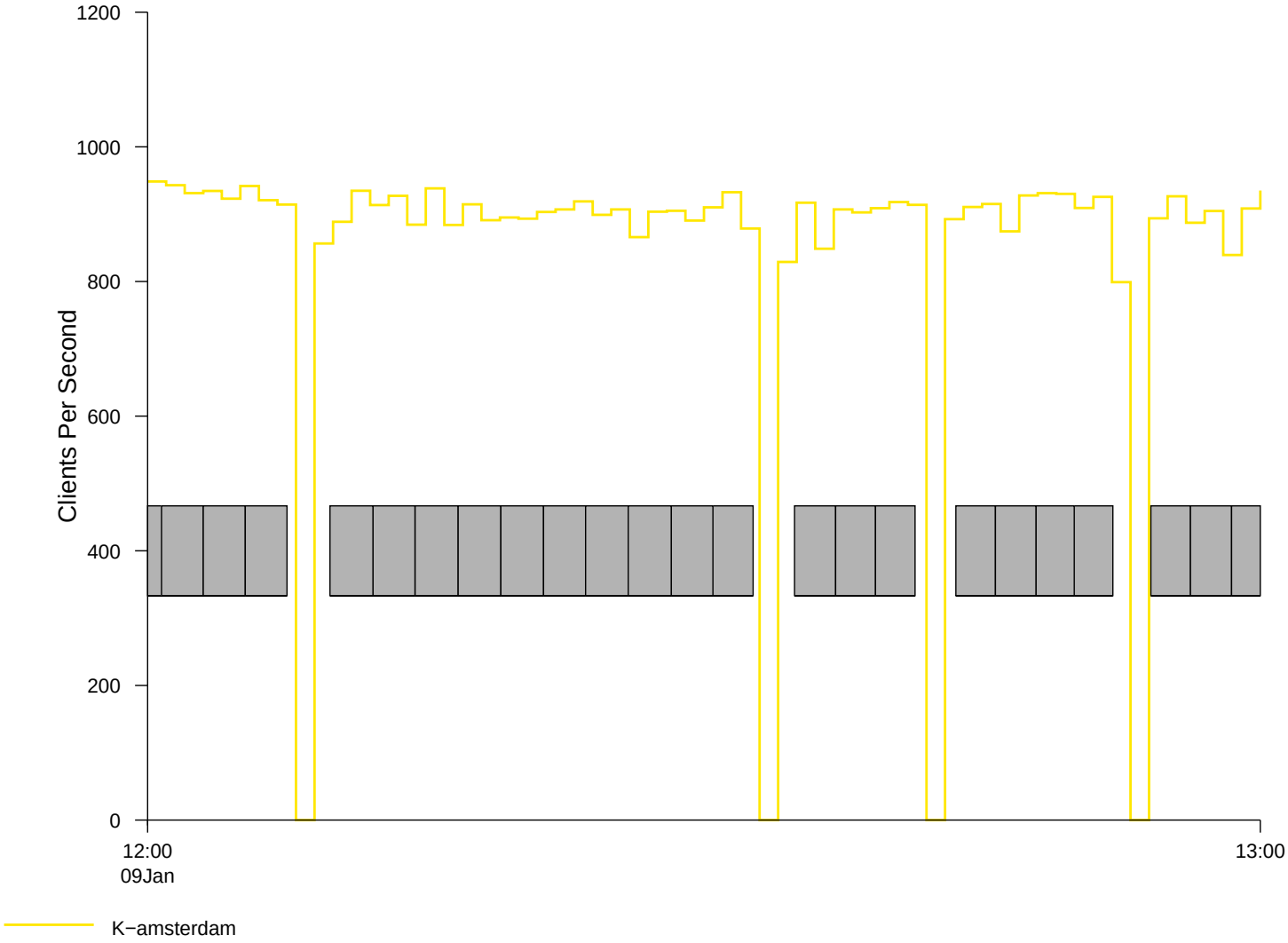
II 2) The number of clients per second seen at each F−root instance.



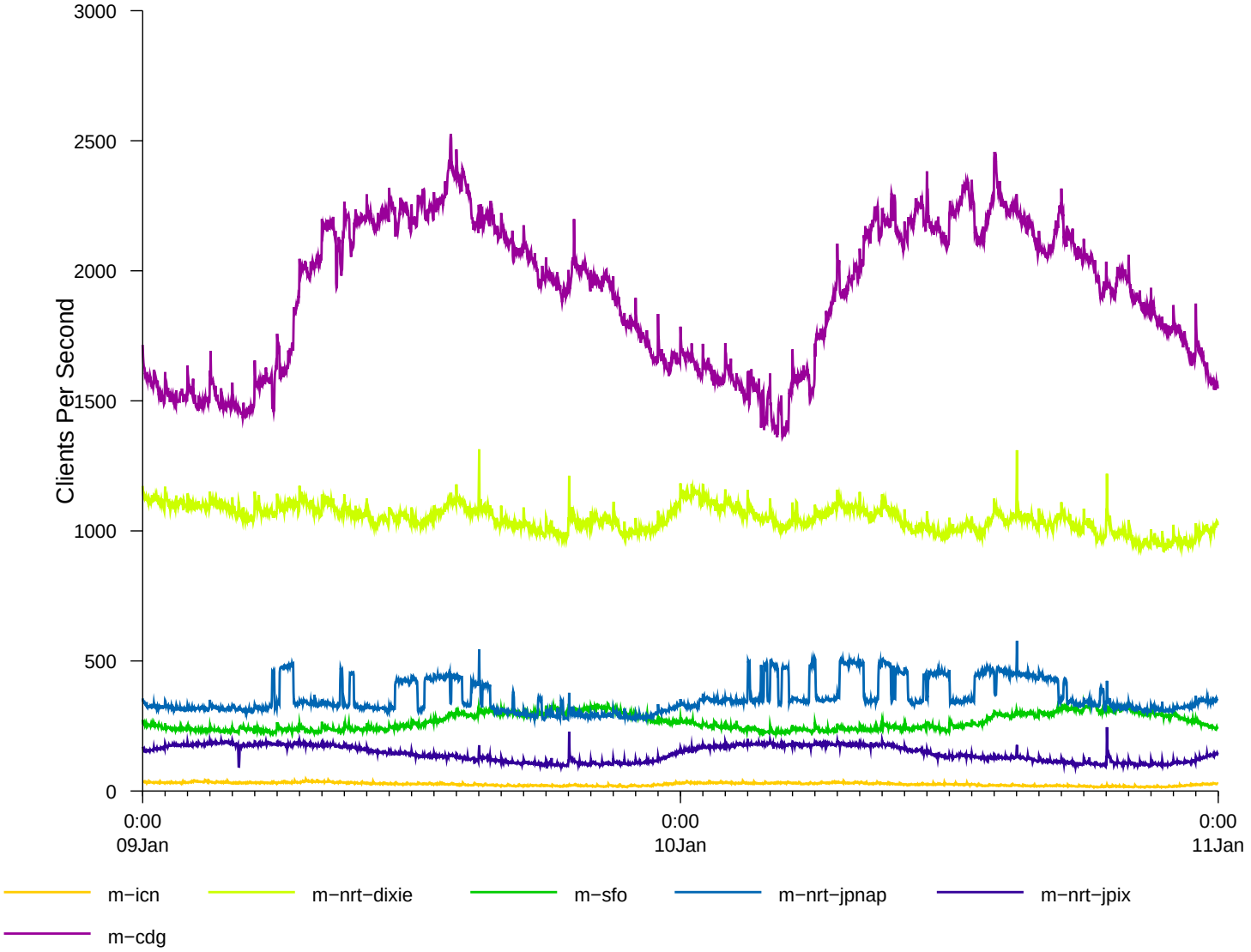
II 2) The number of clients per second seen at each K−root instance.



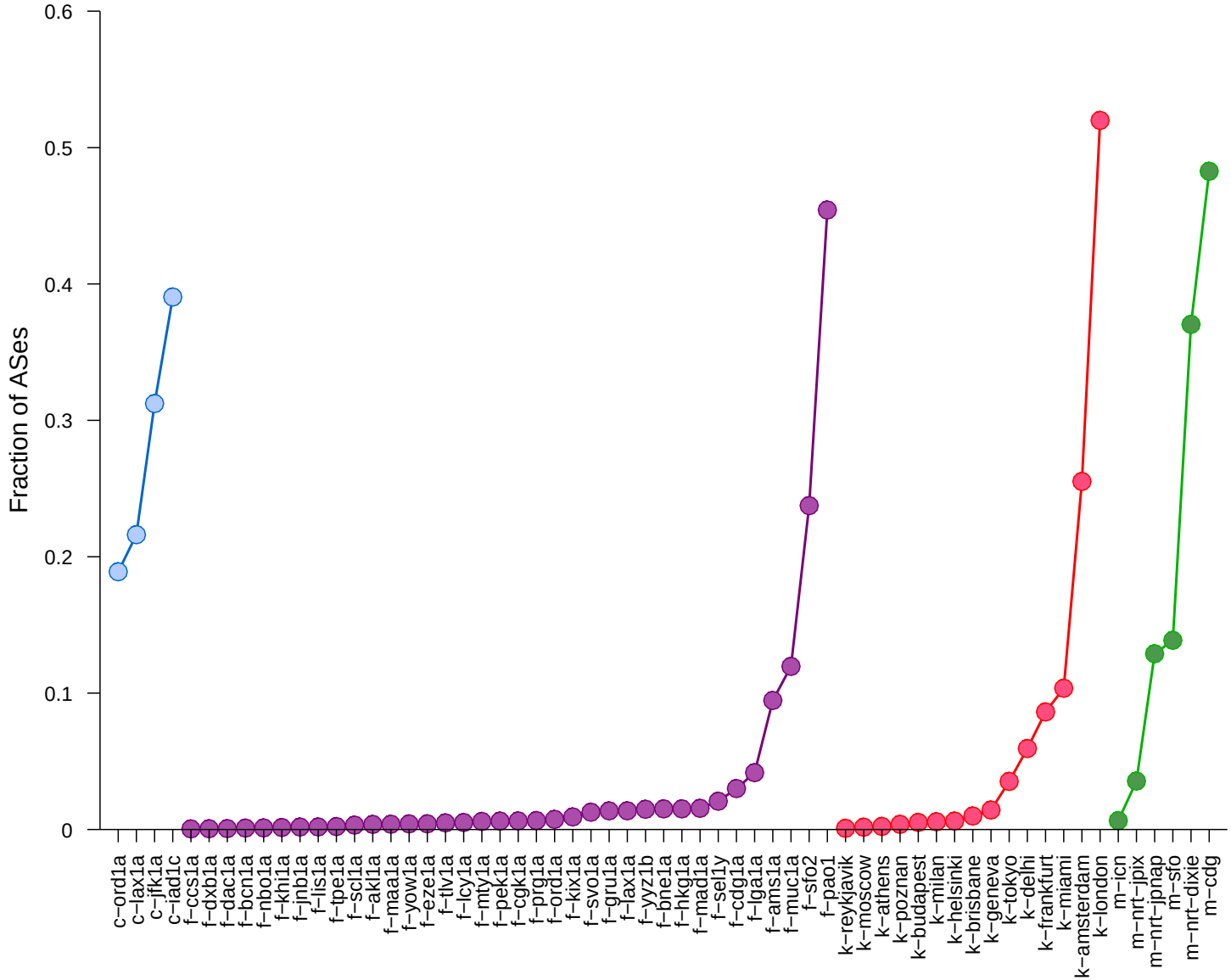
II. 2) Zoom in on K−amsterdam node



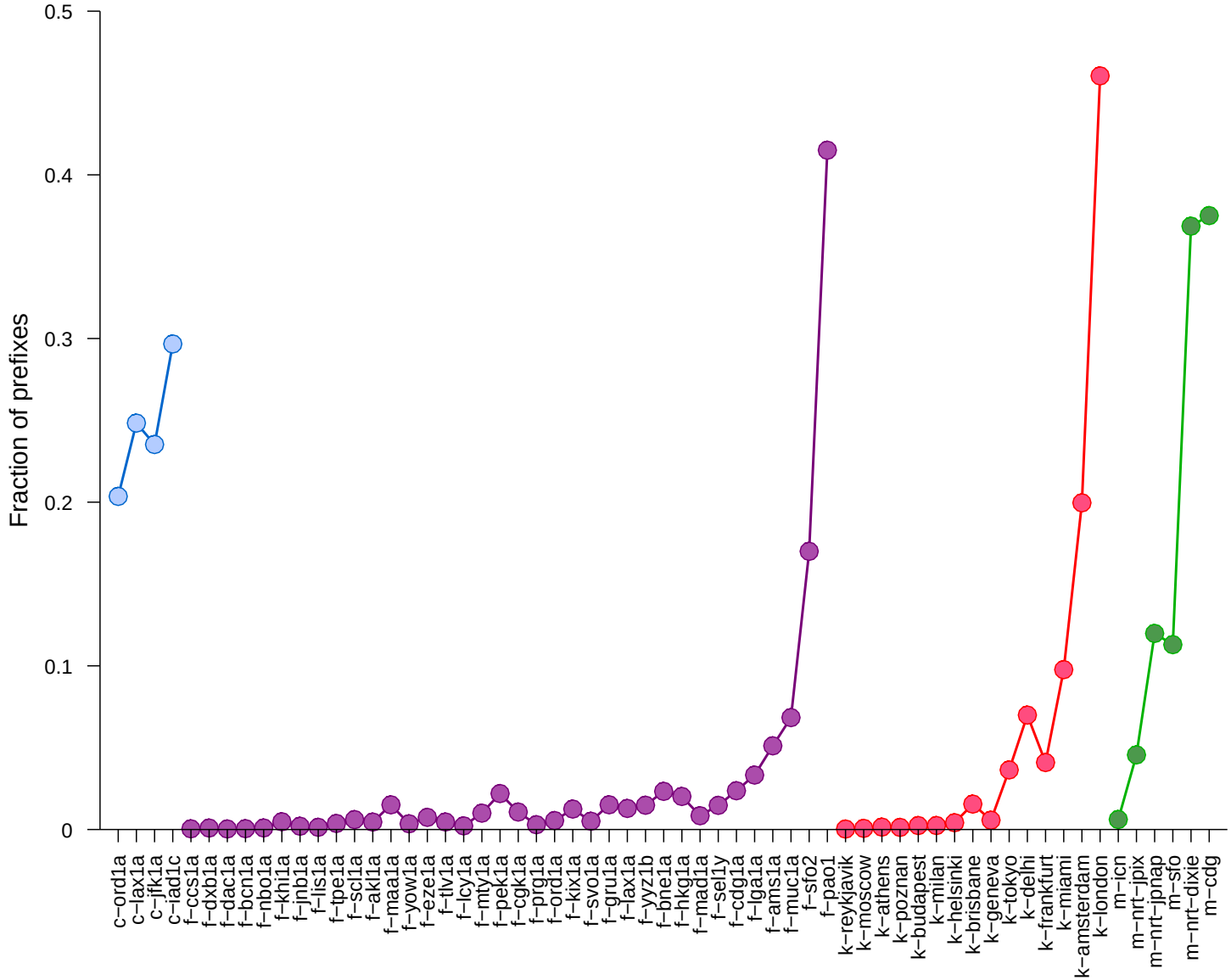
II 2) The number of clients per second seen at each M−root instance.



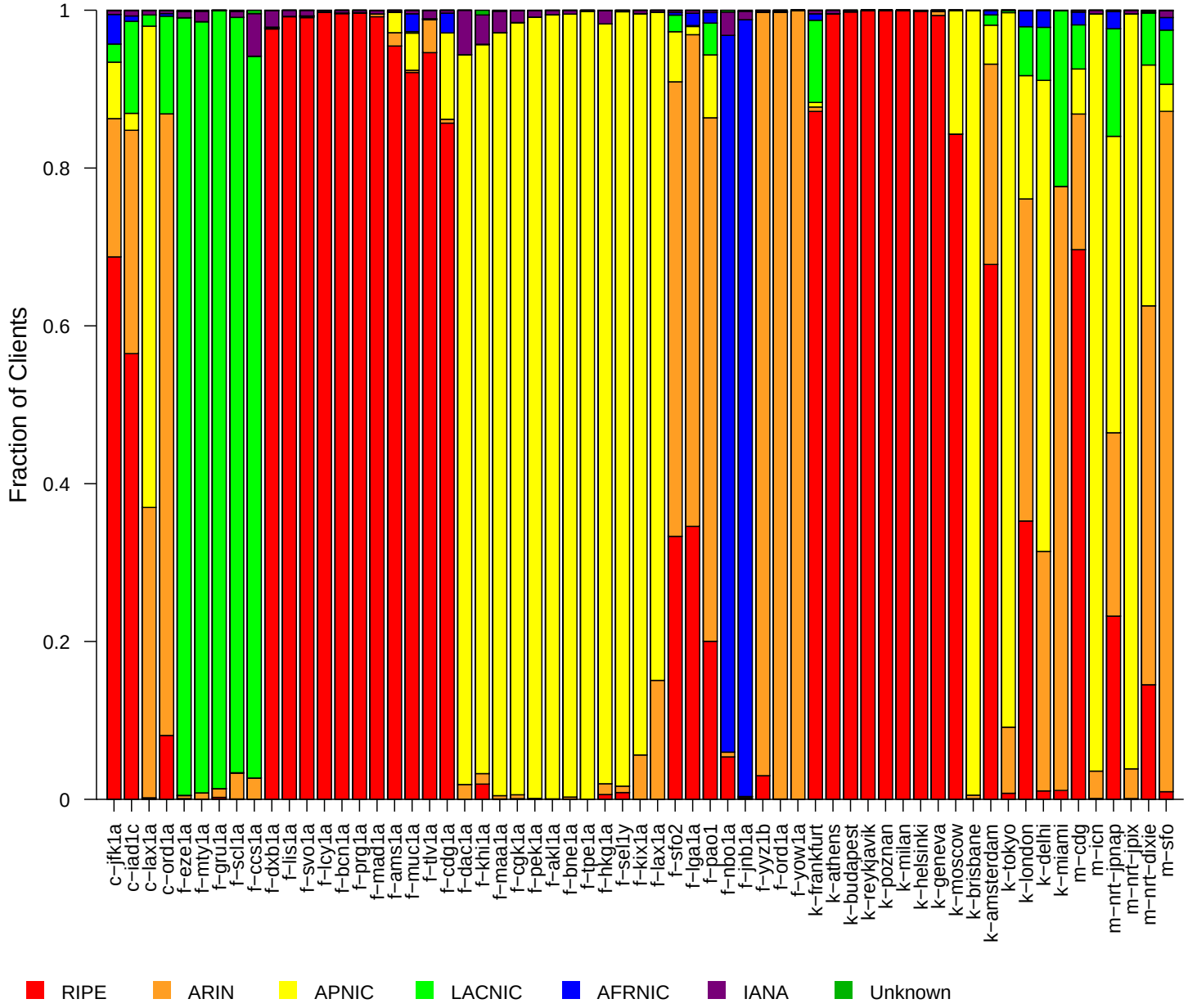
II 3) Topological coverage by ASes.

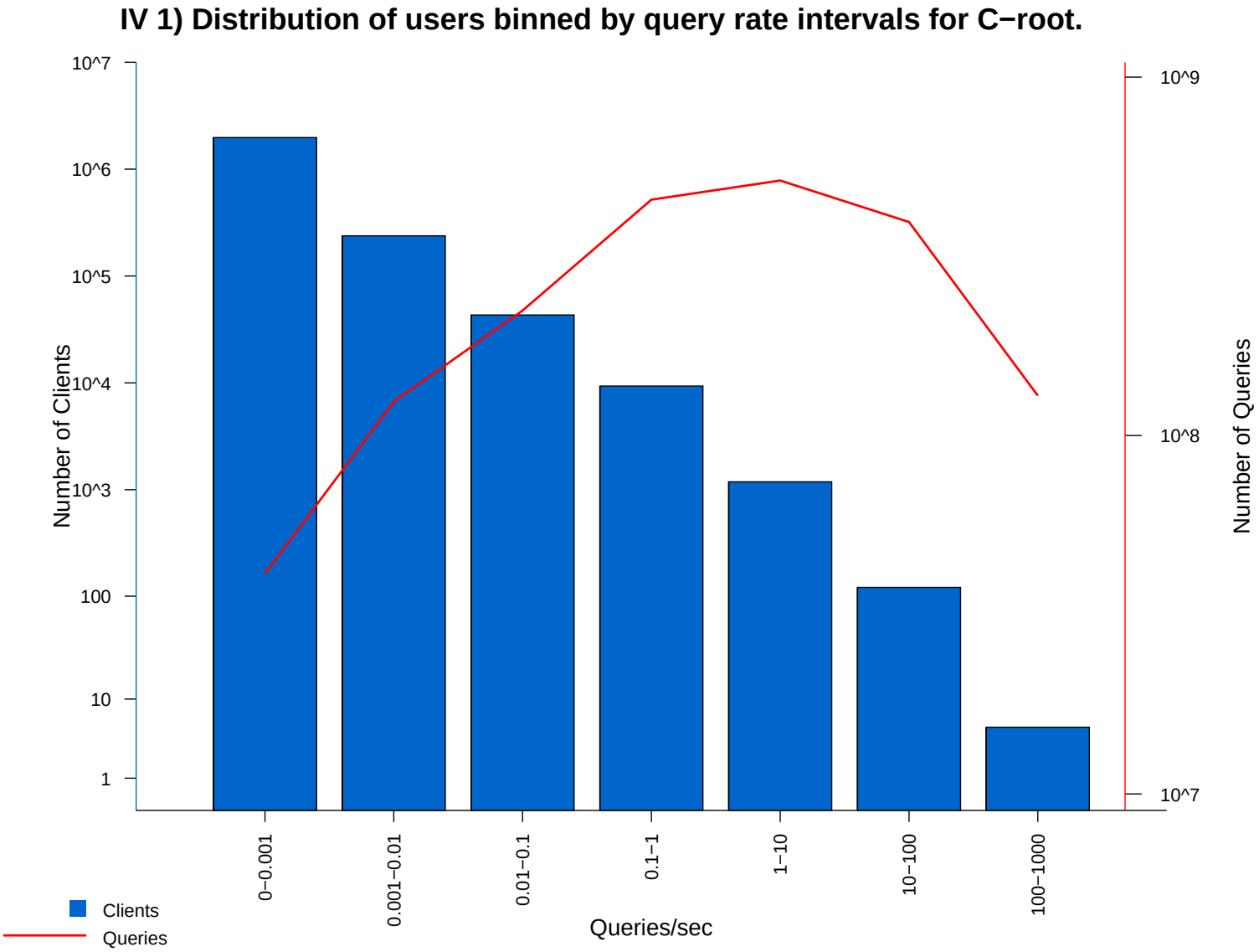


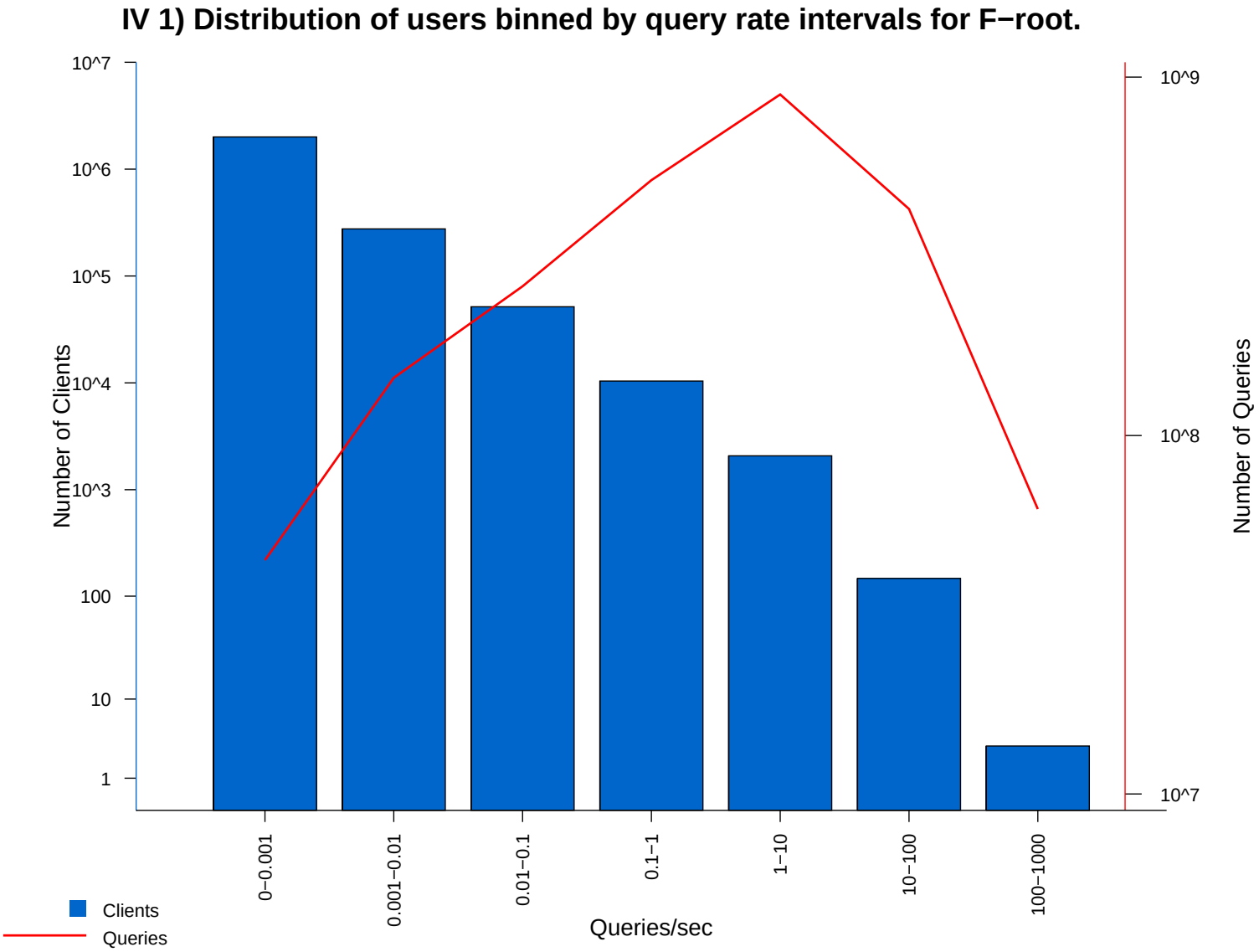
II 4) Topological coverage by prefixes.

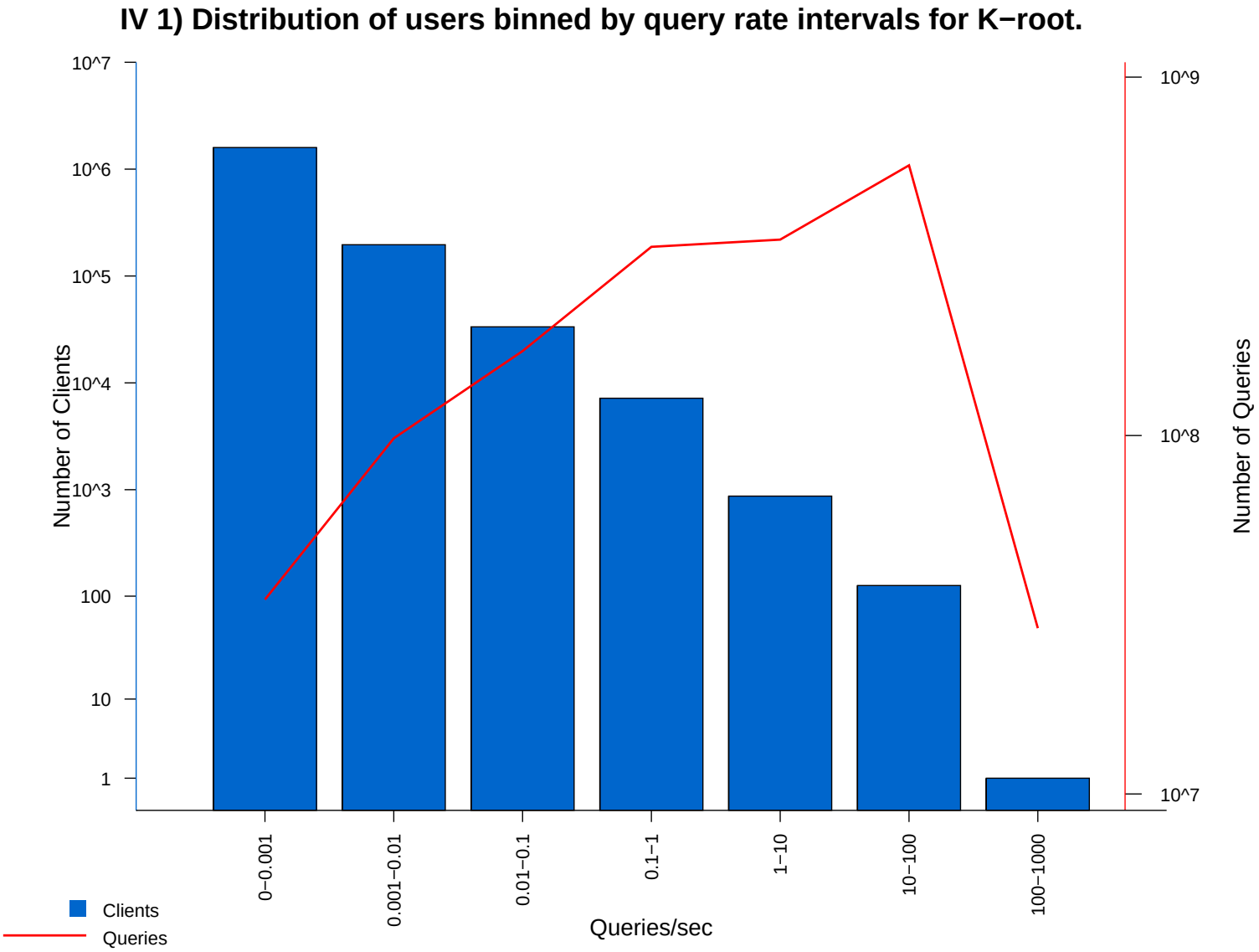


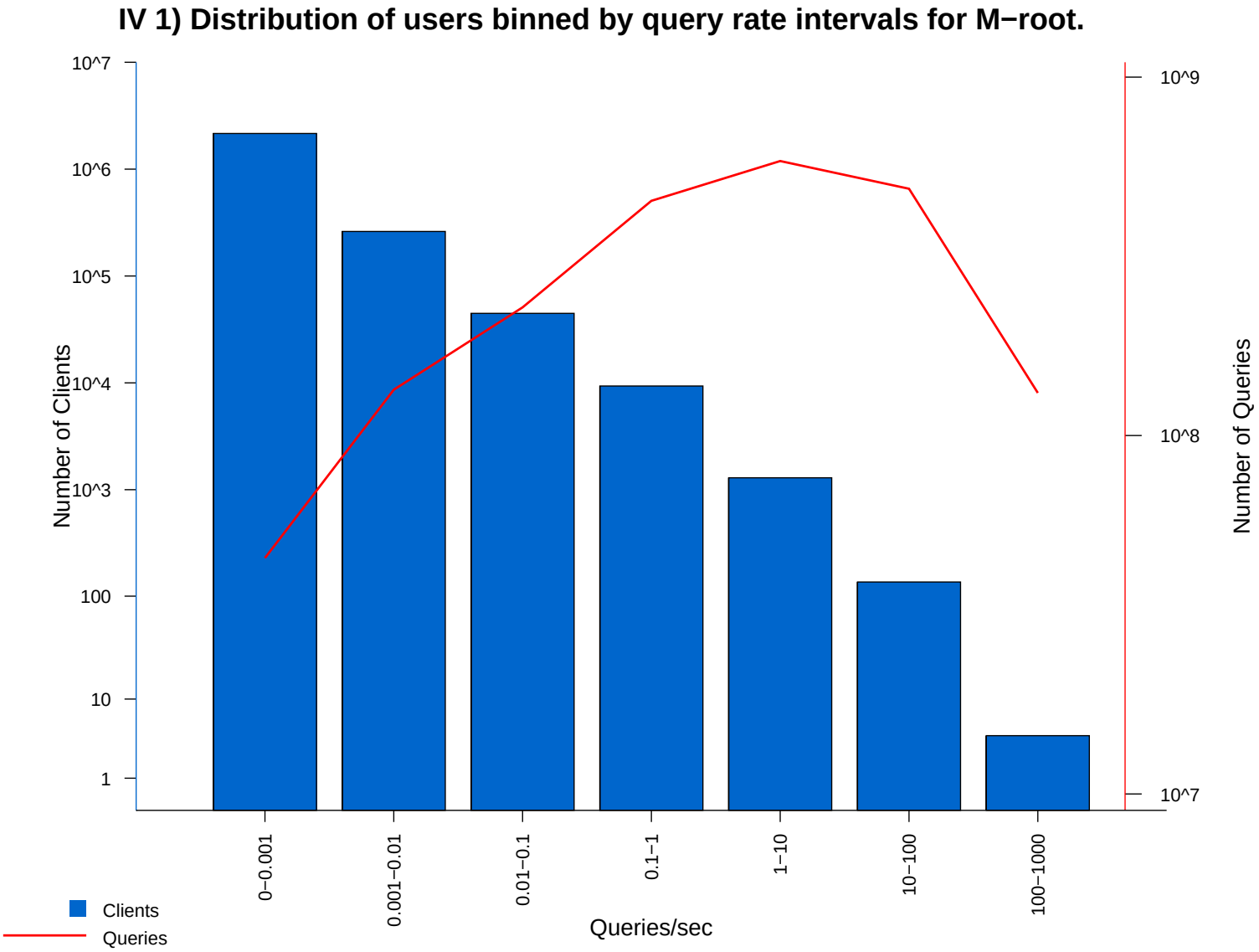
III 1) Clients distribution by RIR for each instance

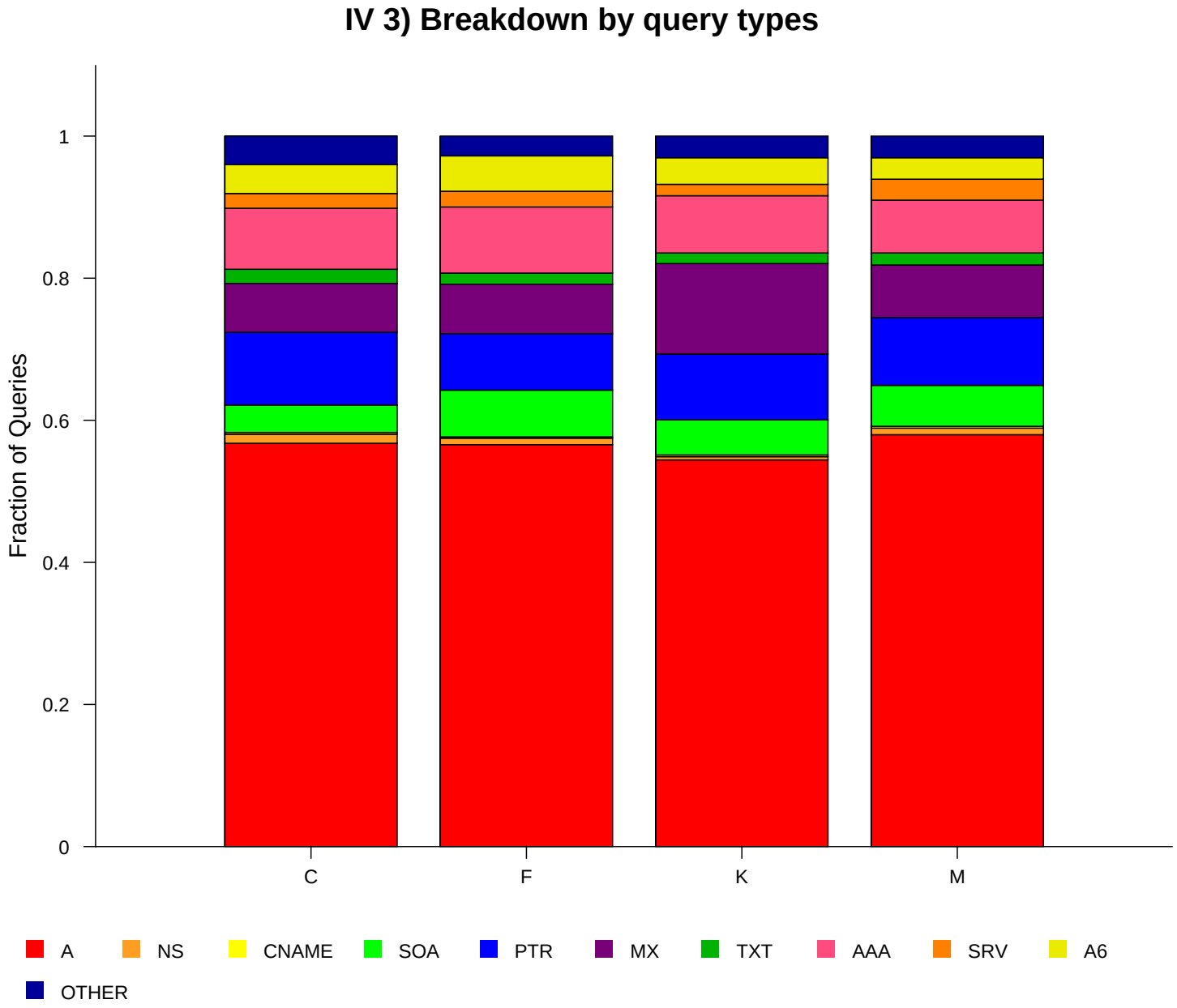


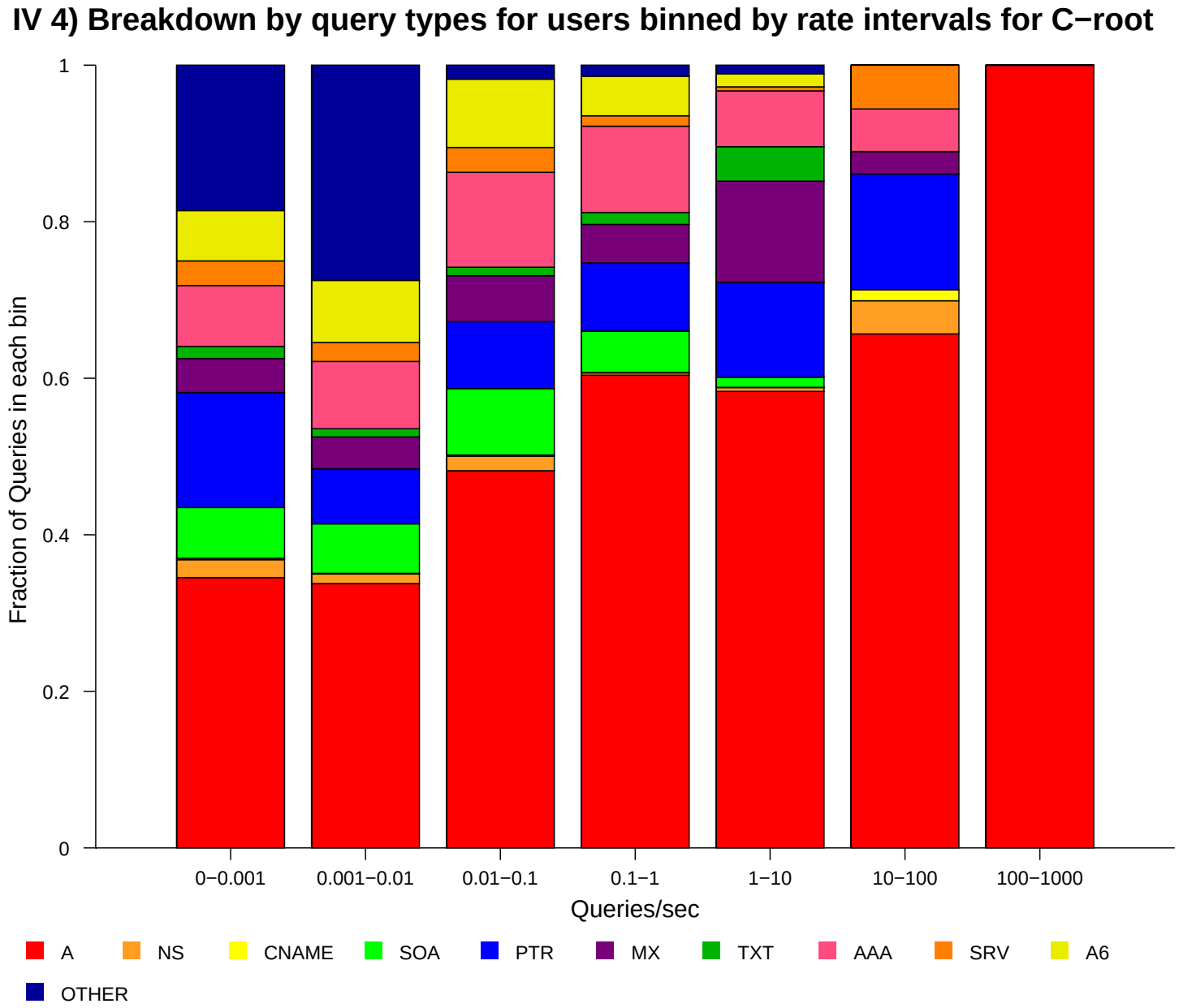

IV 1) Distribution of users binned by query rate intervals for C−root.
10^7
10^6
10^5
10^4
10^3
100
10
1
Number of Clients
10^9
10^8
10^7
Number of Queries
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
Clients
Queries

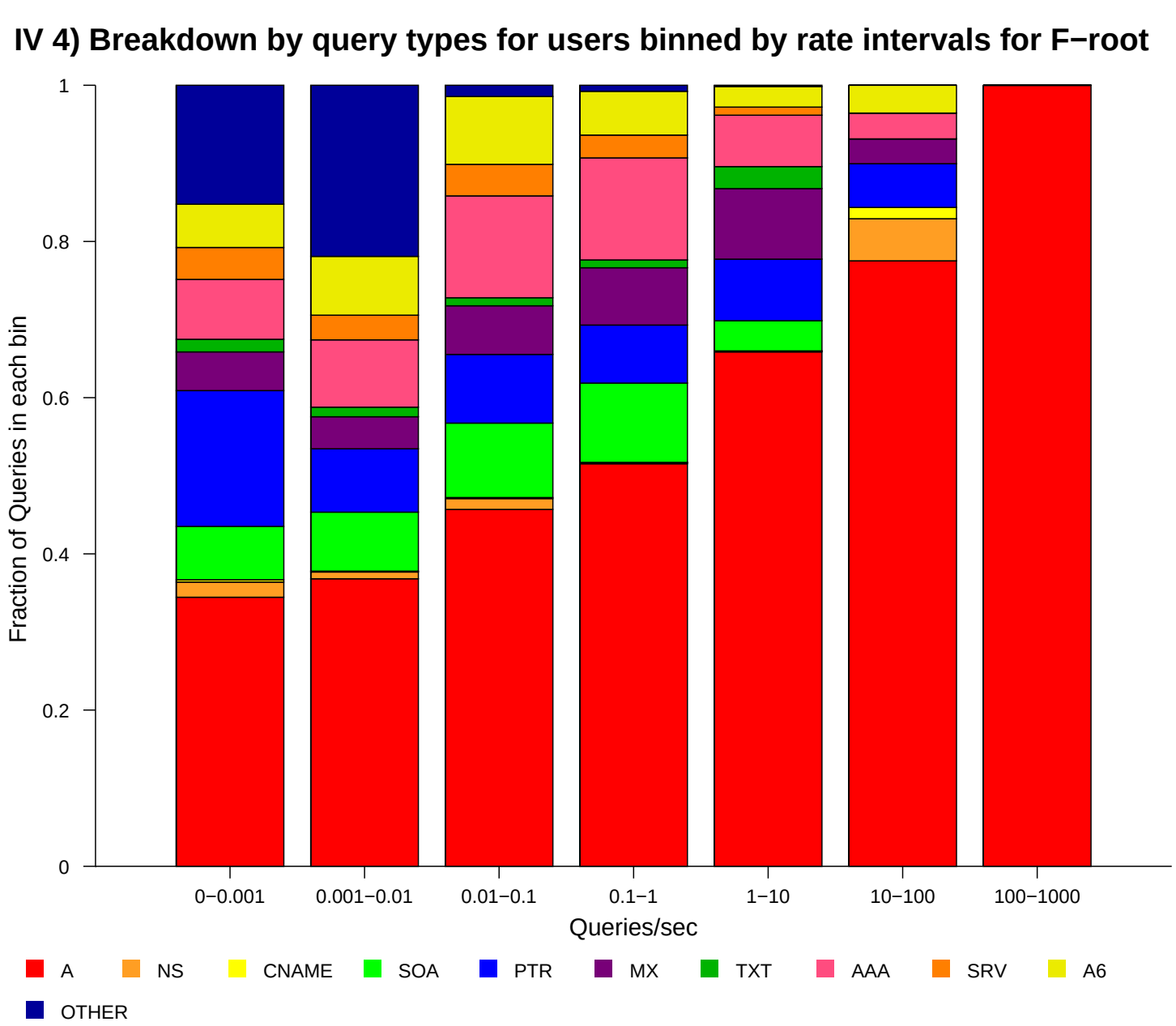

IV 1) Distribution of users binned by query rate intervals for F−root.
10^7
10^6
10^5
10^4
10^3
100
10
1
Number of Clients
10^9
10^8
10^7
Number of Queries
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
Clients
Queries

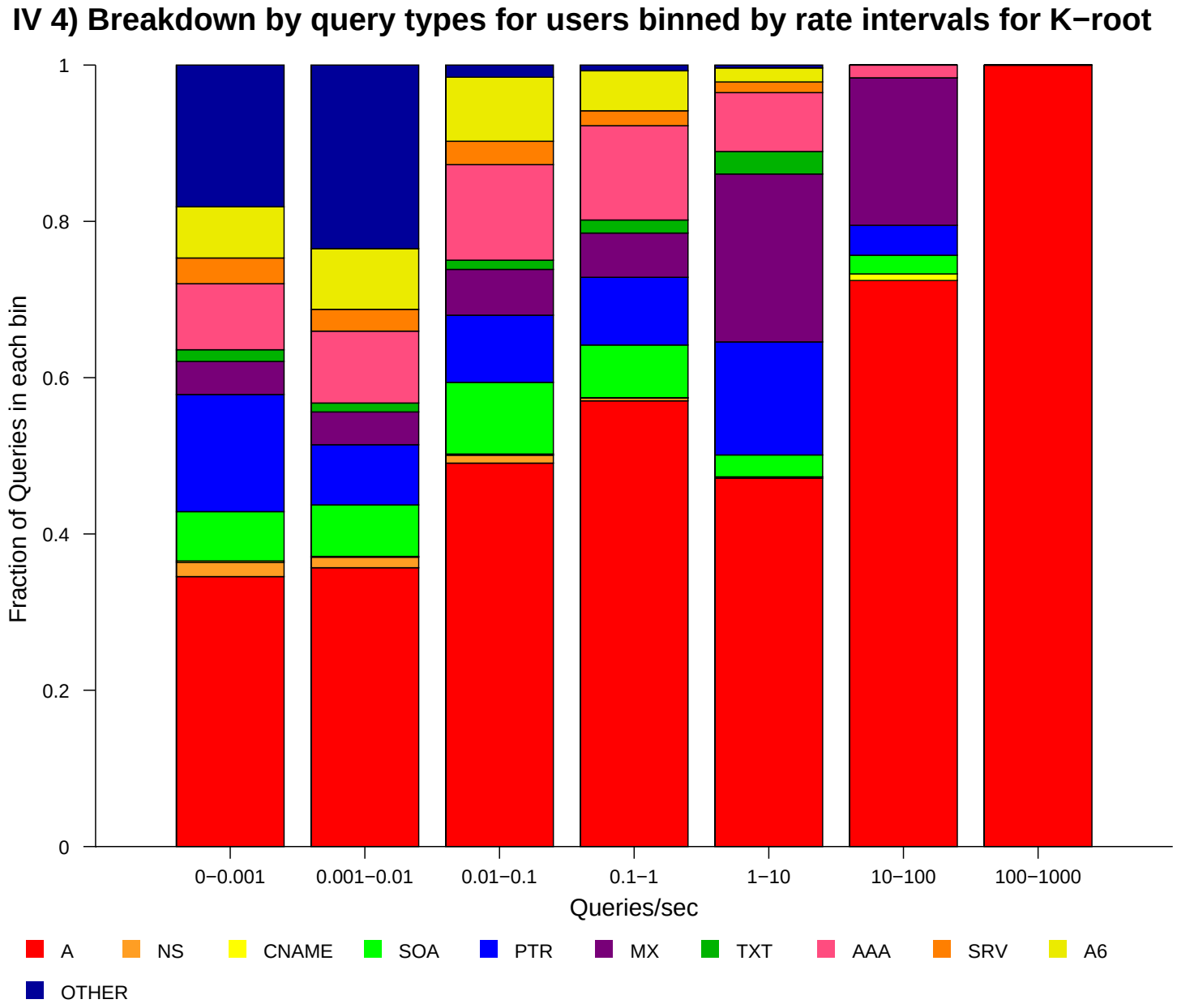

IV 1) Distribution of users binned by query rate intervals for K−root.
10^7
10^6
10^5
10^4
10^3
100
10
1
Number of Clients
10^9
10^8
10^7
Number of Queries
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
Clients
Queries

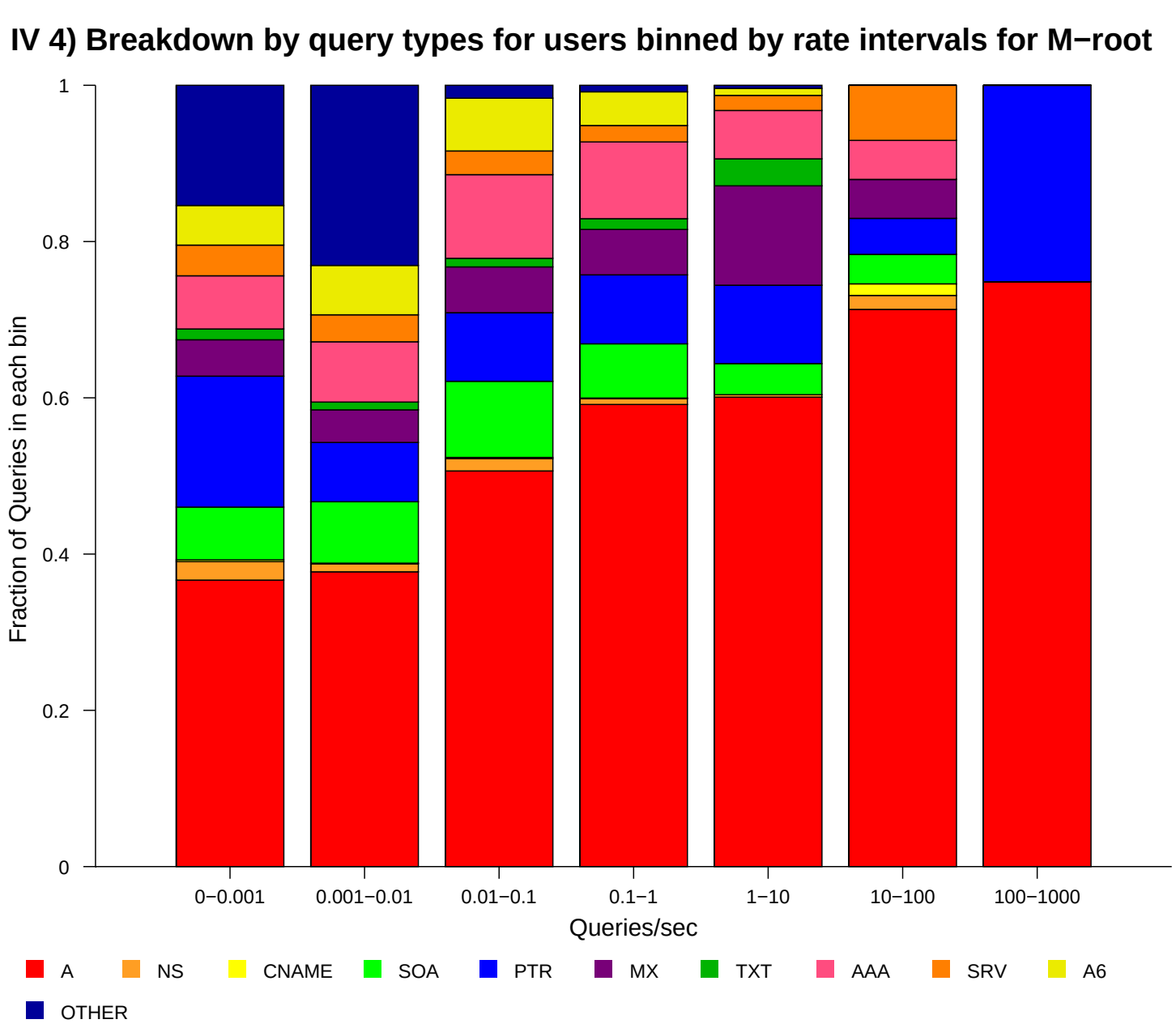

IV 1) Distribution of users binned by query rate intervals for M−root.
10^7
10^6
10^5
10^4
10^3
100
10
1
Number of Clients
10^9
10^8
10^7
Number of Queries
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
Clients
Queries


IV 3) Breakdown by query types
Fraction of Queries
1
0.8
0.6
0.4
0.2
0
C
F
K
M
A
NS
CNAME
SOA
PTR
MX
TXT
AAA
SRV
A6
OTHER


IV 4) Breakdown by query types for users binned by rate intervals for C−root
Fraction of Queries in each bin
1
0.8
0.6
0.4
0.2
0
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
A
NS
CNAME
SOA
PTR
MX
TXT
AAA
SRV
A6
OTHER


IV 4) Breakdown by query types for users binned by rate intervals for F−root
Fraction of Queries in each bin
1
0.8
0.6
0.4
0.2
0
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
A
NS
CNAME
SOA
PTR
MX
TXT
AAA
SRV
A6
OTHER


IV 4) Breakdown by query types for users binned by rate intervals for K−root
Fraction of Queries in each bin
1
0.8
0.6
0.4
0.2
0
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
A
NS
CNAME
SOA
PTR
MX
TXT
AAA
SRV
A6
OTHER


IV 4) Breakdown by query types for users binned by rate intervals for M−root
Fraction of Queries in each bin
1
0.8
0.6
0.4
0.2
0
0−0.001
0.001−0.01
0.01−0.1
0.1−1
1−10
10−100
100−1000
Queries/sec
A
NS
CNAME
SOA
PTR
MX
TXT
AAA
SRV
A6
OTHER

The End