Stateless Reverse Traceroute

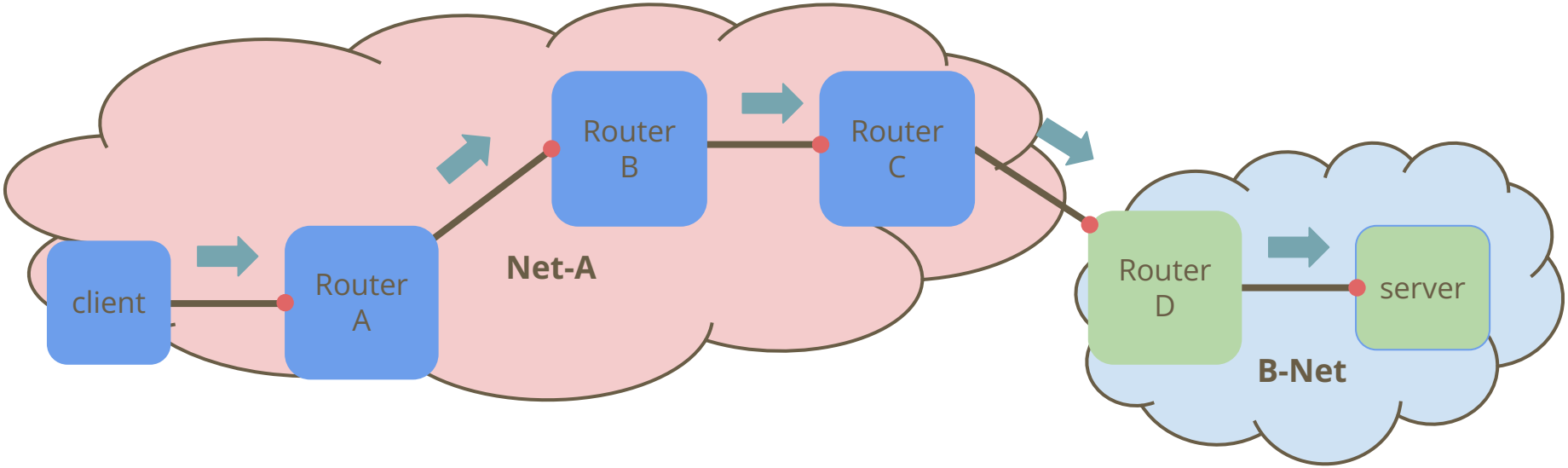
Rolf Winter

https://datatracker.ietf.org/doc/html/draft-heiwin-intarea-reverse-traceroute-stateless-02

Technische Hochschule Augsburg

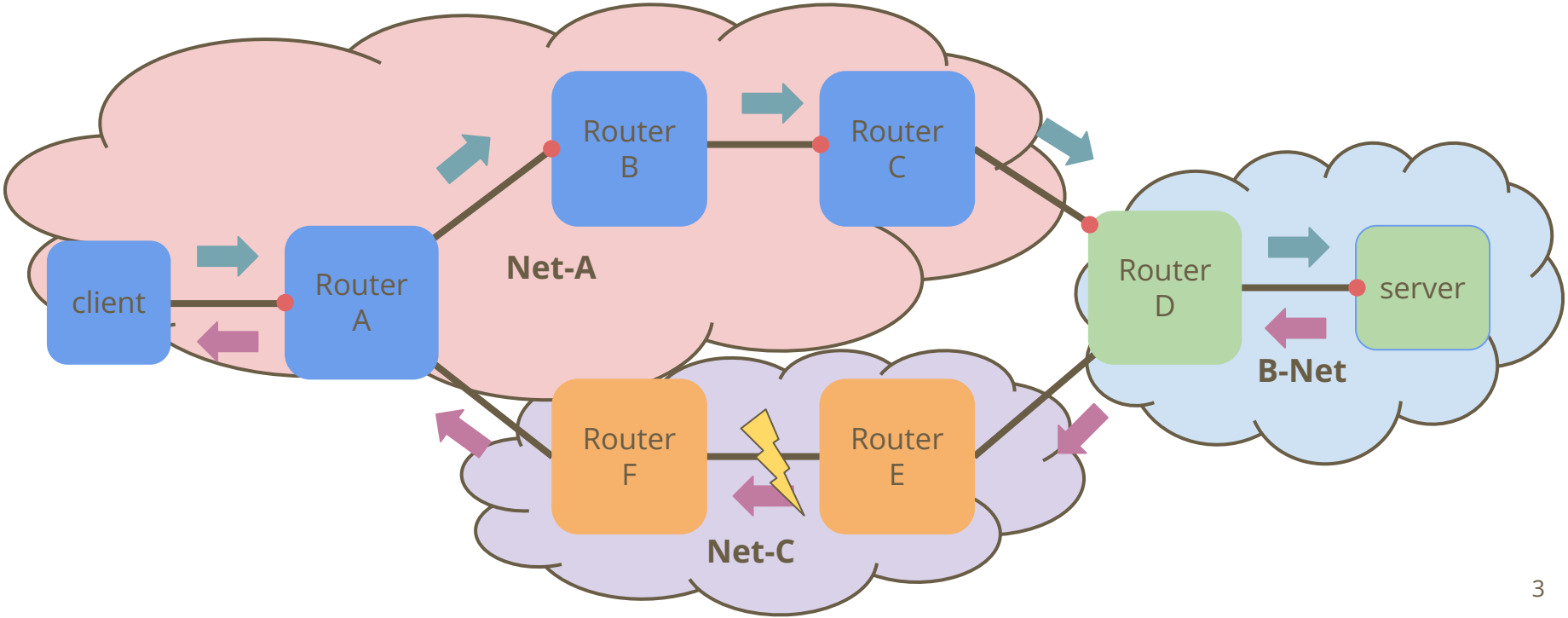
The problem

```
1    routerA.aug.net-a.com    (10.10.0.1)    1ms    2ms    1ms
2    routerB.muc.net-a.com    (10.10.0.2)    5ms    6ms    12ms
3    routerC.fra.net-a.com    (10.10.0.3)    11ms   21ms   14ms
4    routerD.fra.b-net.com    (20.20.0.1)    340ms  320ms  350ms
5    www.example.com          (20.20.0.2)    345ms  310ms  360ms
```



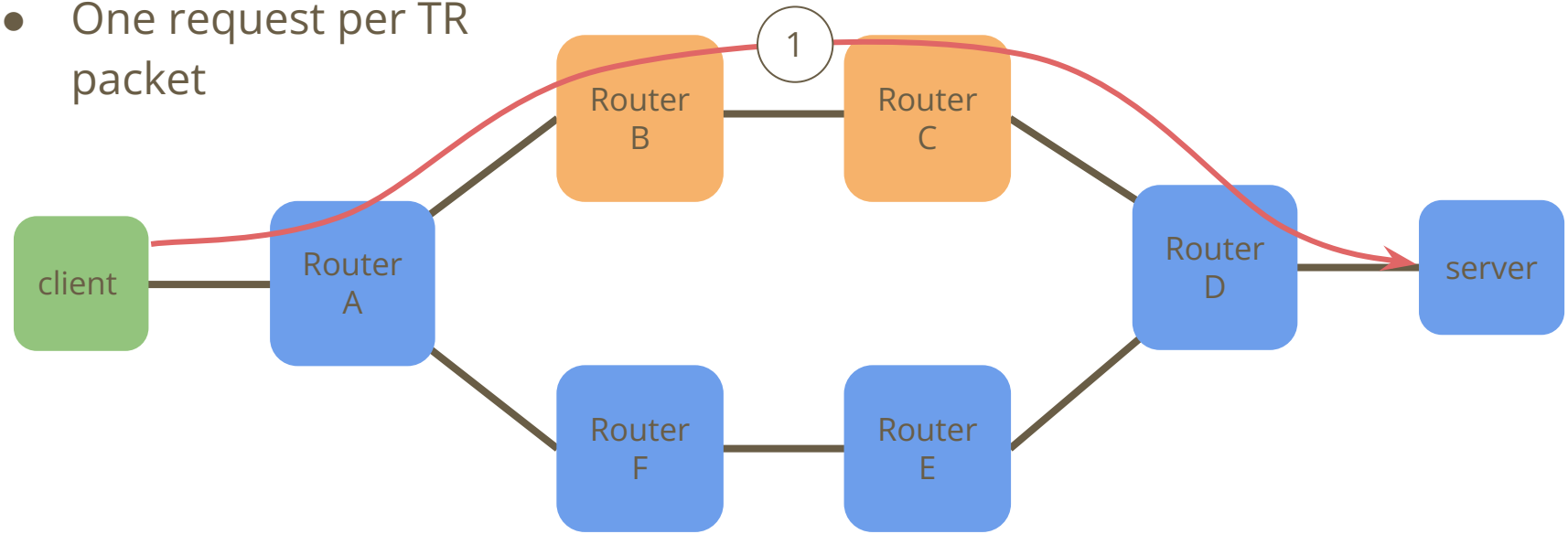
The problem

```
1    routerA.aug.net-a.com    (10.10.0.1)    1ms    2ms    1ms
2    routerB.muc.net-a.com    (10.10.0.2)    5ms    6ms    12ms
3    routerC.fra.net-a.com    (10.10.0.3)    11ms   21ms   14ms
4    routerD.fra.b-net.com    (20.20.0.1)    340ms  320ms  350ms
5    www.example.com          (20.20.0.2)    345ms  310ms  360ms
```



Meet reverse traceroute

- Uses a new ICMP request to trigger a reverse traceroute
- One request per TR packet

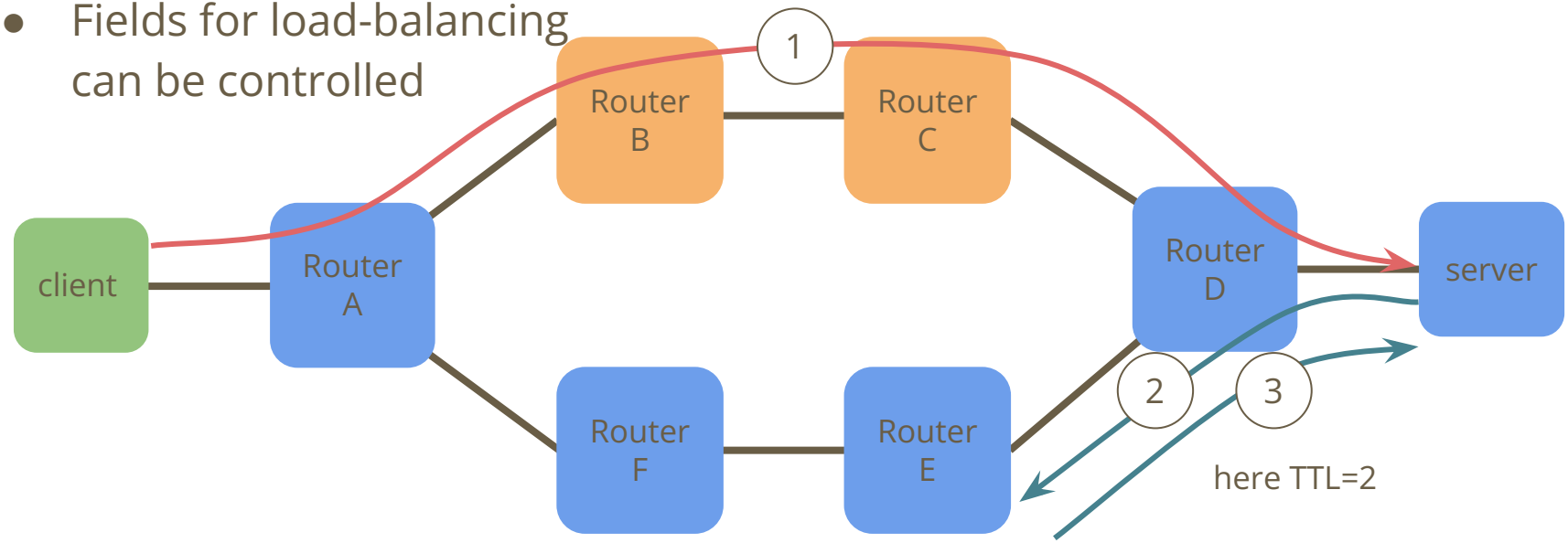


Routers reverse traceroute shows

Routers on the forward path

Meet reverse traceroute

- A regular TR packet is sent (UDP, ICMP or TCP)
- Fields for load-balancing can be controlled

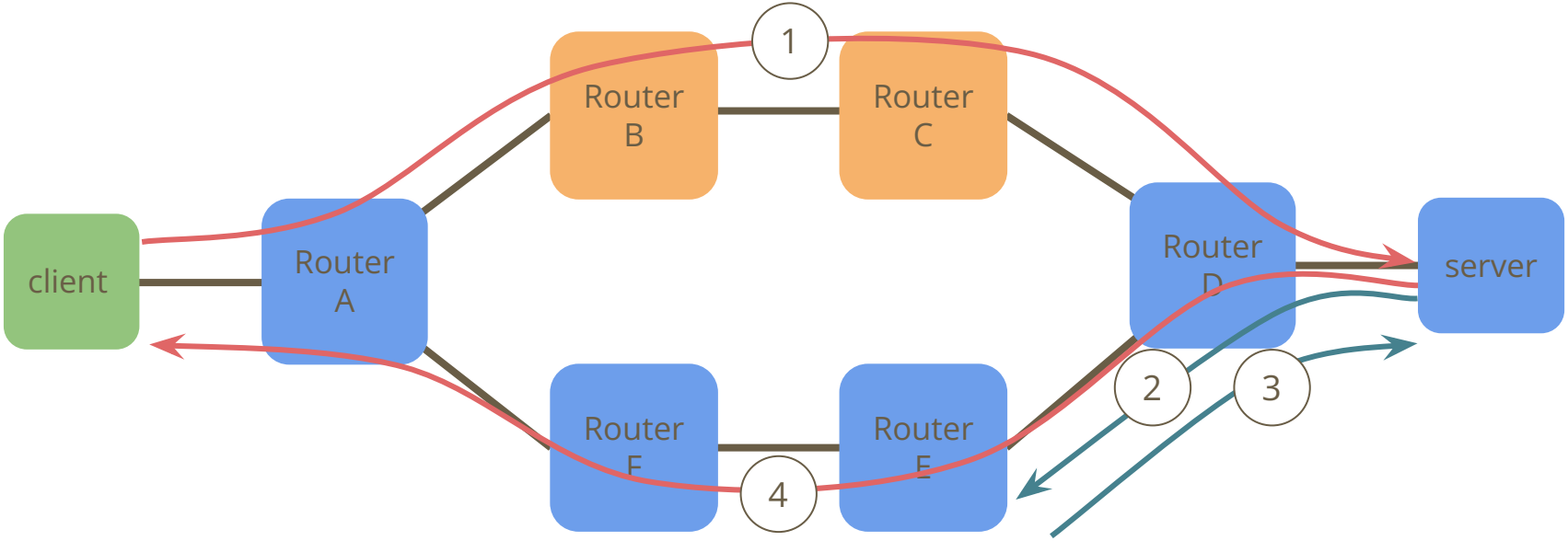


Routers reverse traceroute shows

Routers on the forward path

Meet reverse traceroute

- For that single probe, an ICMP response is sent back

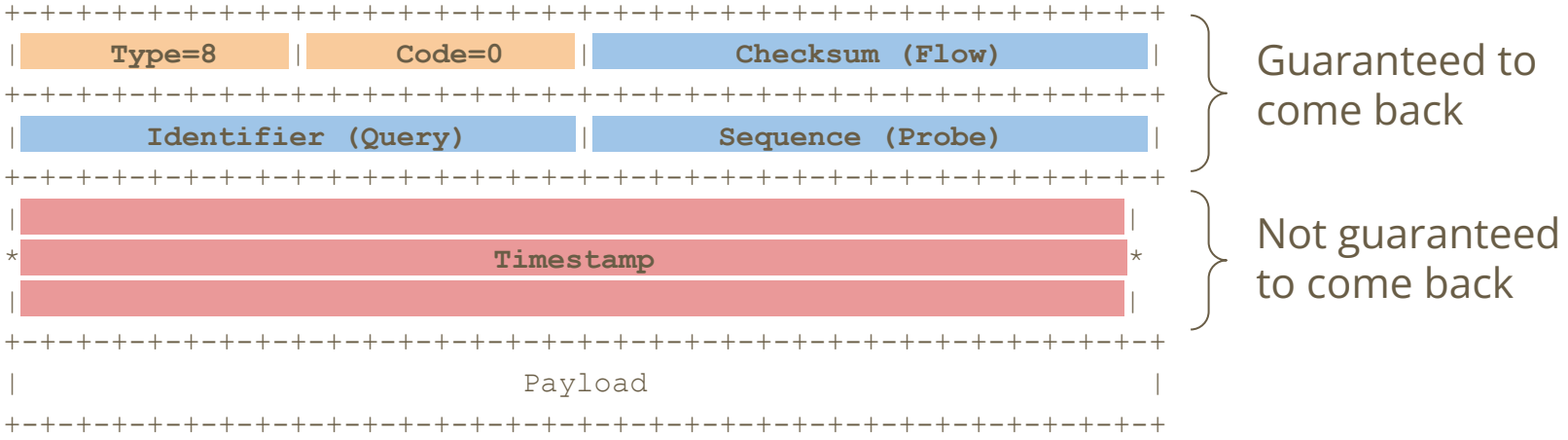


Routers reverse traceroute shows

Routers on the forward path

Stateless - how?

- Stateless: all the information needed has to be put into the probe packets
  - For IPv4, the ICMP Time Exceeded packet only guarantees to contain the original IP header plus the 64 bits following it
- ICMP probes (as an example, applies similarly to UDP and TCP*):



- RTT estimations are not guaranteed to work

*well, TCP needs 28 bits - instead of 16

And on today's internet?

- RIPE Atlas, publishes archives of their measurement data, or rather metadata about the actual measurements
  - We filtered that data to only contain IPv4 ICMP Time Exceeded packets without extensions to reliably infer the true return size
  - Removed duplicates with the same (payload size, IP address) tuple
  - Data isn't perfect: e.g. sometimes it seems that answers are bigger than the actual probe → filter data that seems inconsistent away, too
- In sum, we looked at 200 GB (compressed) archives

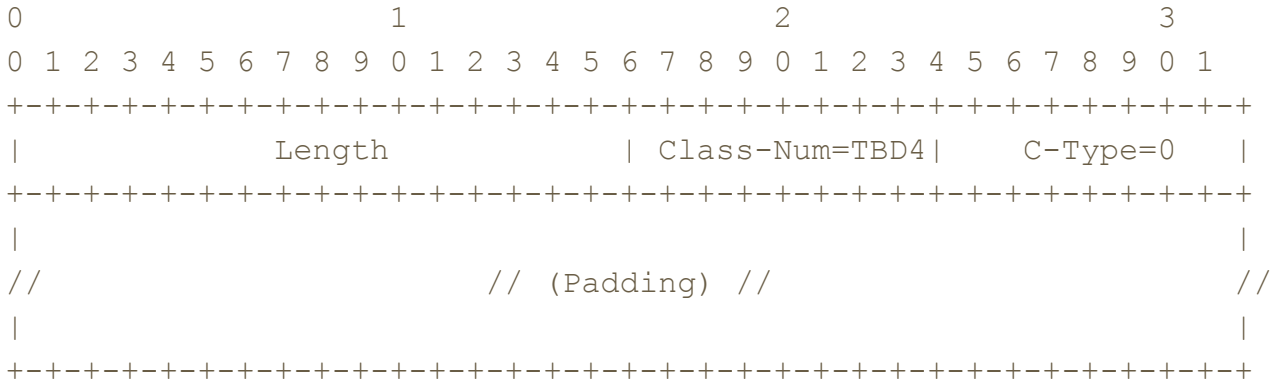
Response size [bytes]	8	12	20 - 28	32 - 39	44	>=48
% of responses	48.6%	0.1%	0.4%	0.5%	0.4%	50%

Stateless vs. Stateful

Stateless	Stateful
Identifies all routers (that send ICMP time exceeded packets today)	Identifies all routers (that send ICMP time exceeded packets today)
Can estimate the RTT in about 50% of all cases	Can always estimate the RTT
Is restricted to trace towards the requestor	Can potentially trace to some other host on the internet

Extensions

- There is some amplification: a client request triggers a single traceroute probe being sent AND also a reverse traceroute response
- Solution: make the client send more bytes using a padding extension



Next steps

- Ask for WG adoption
  - We did the measurements
  - We did the implementation
  - The document is in good shape (for an individual document)
- If the document gets adopted
  - Fold the stateful document back into this one and just name it “Reverse Traceroute”
  - The party offering the reverse traceroute function should be able to pick stateful over stateless (this is where the state is kept)