Dark Corners of Name Server Selection Algorithms

Yingdi Yu[1], Duane Wessels[2], Matt Larson[2], and Lixia Zhang[1]

[1]UCLA [2]VeriSign Lab

Server selection algorithms work well?

- In most cases: Yes!
- May not work well in some cases, e.g.,
  - Query rate is high
  - Servers become unresponsive
- Similar implementations may perform distinctively!

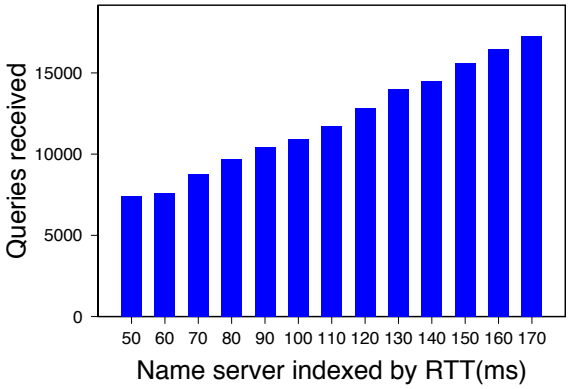
BIND 9.8 v.s. PowerDNS

- Both select the server with the least SRTT.
- SRTT decays in different ways:

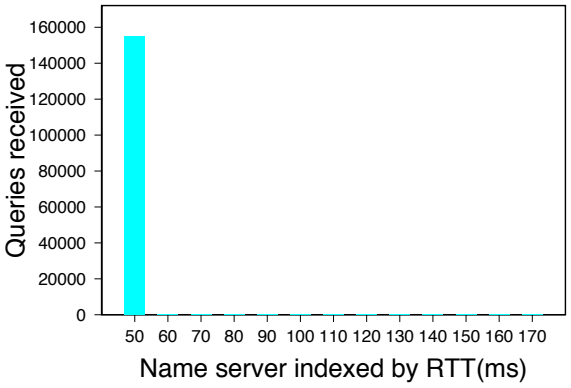
$$SRTT_{n+1} = \beta \cdot SRTT_n$$

  - BIND 9.8: constant β=0.98
  - PowerDNS: time dependent

$$\beta(t_{n+1}) = e^{\frac{t_n - t_{n+1}}{C}}$$

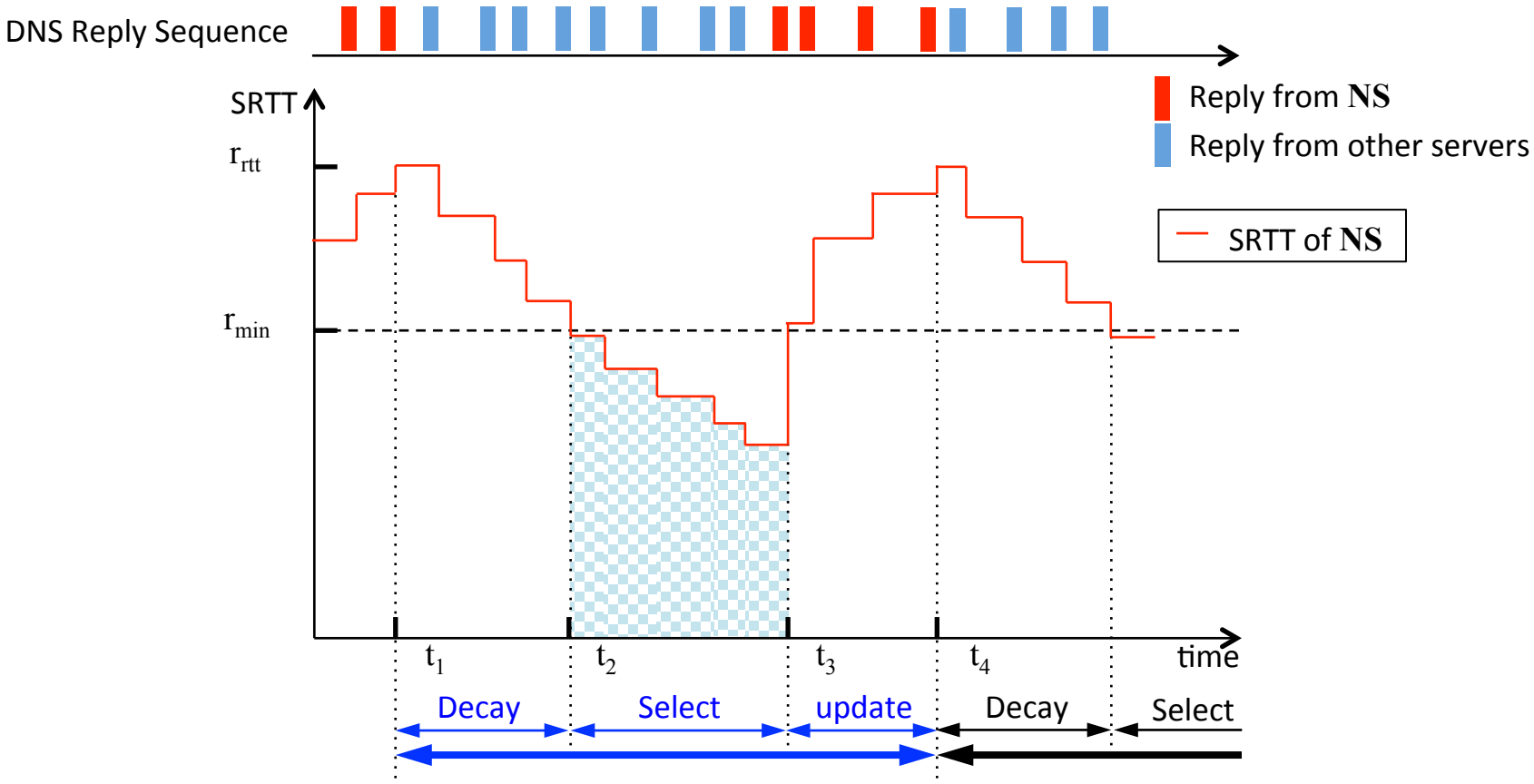


Bind 9.8 in Scenario 1



PowerDNS in Scenario 1

SRTT Variation



$$N_{query} \approx \frac{t_{select}}{t_{decay} + t_{select} + t_{increase}}$$

if t_{decay} is smaller, more queries are sent to slow servers.

When β is constant…

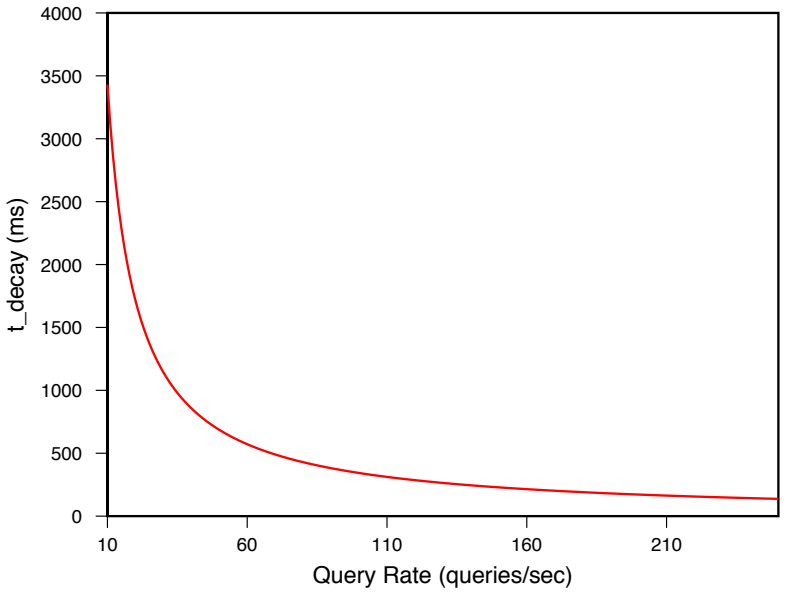
- $\beta = 0.98$ (BIND).

$$SRTT_{n+1} = \beta \cdot SRTT_n$$

- t_{decay} is actually determined by query rate Q .

$$t_{decay} = \frac{\log(r_{\min}/r_{rtt})}{\log \beta} \cdot \frac{1}{Q}$$

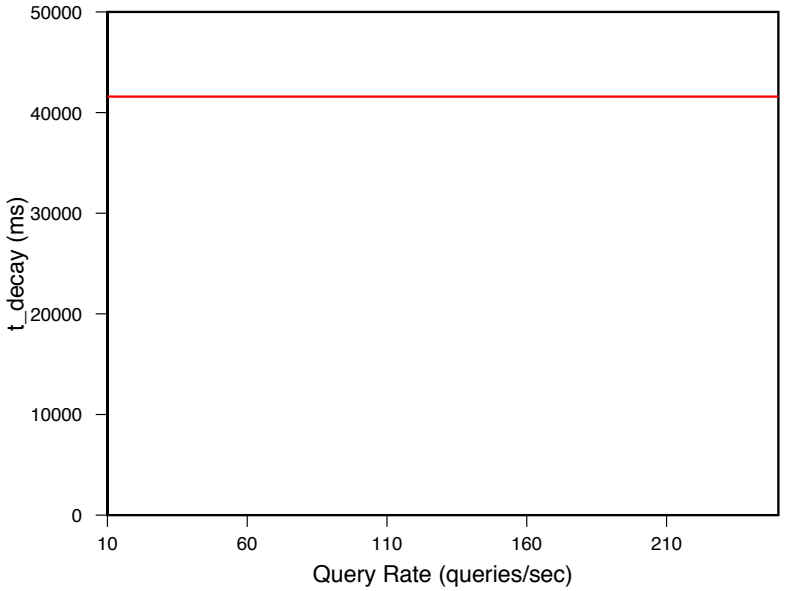
- Higher query rate, smaller t_{decay}!



$r_{\min}$=50ms, r_{rrt}=100ms, $\beta = 0.98$

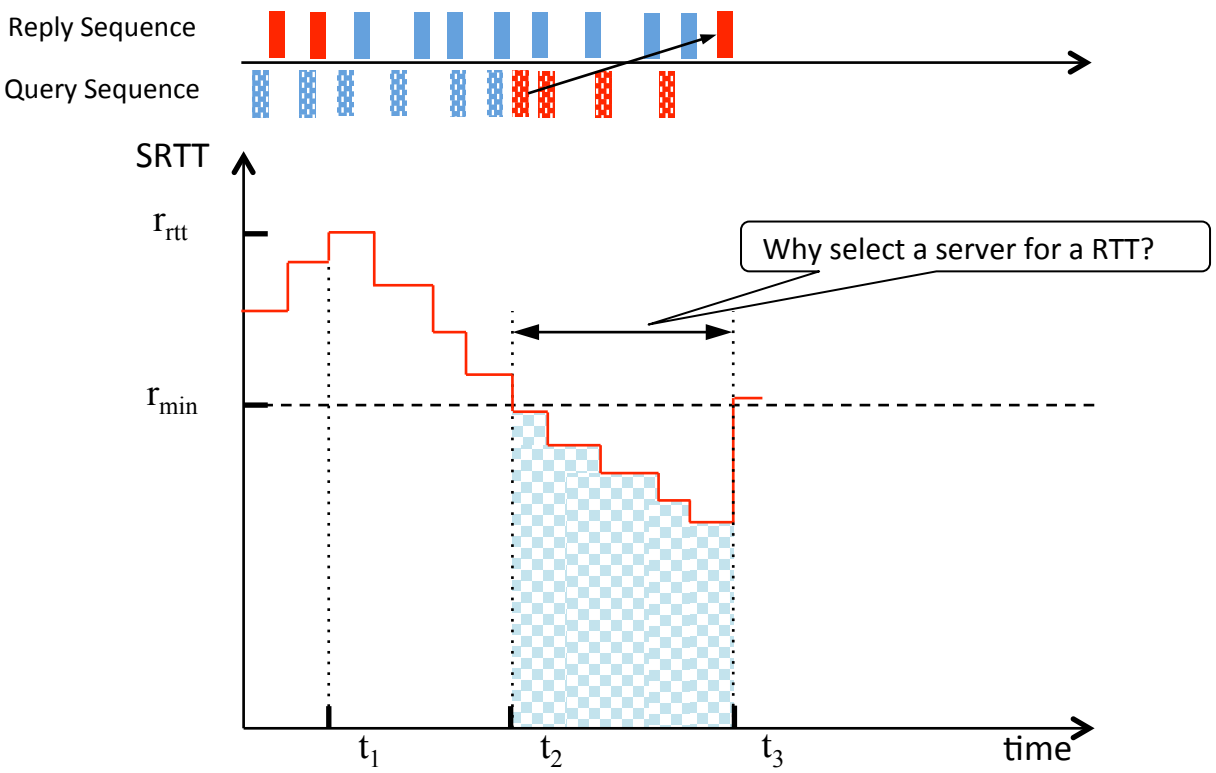
When β depends on time

- $\beta(t_2) = e^{\frac{t_1 - t_2}{C}}$ (PowerDNS)
- t_{decay} is independent from query rate.
- t_{decay} is determined by constant $r_{\min}/r_{\text{rtt}}$ and C.
  - Larger r_{rtt} , longer t_{decay}.
  - Larger C, longer t_{decay}.



$r_{\min}$=50ms, r_{rrt}=100ms, C=60

Revisit Selected phase



- Decaying is only for polling other server.
- Keep decaying leads to a long t_{select}
- Problem becomes more serious when a server is unresponsive.

$$N_{query} = \frac{t_{select}}{t_{decay} + t_{select} + t_{update}}$$

When an unresponsive server is selected

- No response will get back.
- Keep selecting the server until the timeout timer expires (t_{select} = timeout timer).
  - A lot of queries will be sent to the server.
  - DNS lookup will experience long latency.
- Do not have to hurry
  - Send the server a query at first.
  - Wait for response to decide whether to keep selecting the server.
  - (Unbound adopts this strategy)

What if the selected server is good?

- Do resolvers still need to wait for its response?
  - No, the performance would degrade.
- Responsive and unresponsive servers should be handled in different ways.

Conclusion

Detail matters

- Constant β couples decaying speed and query rate.
- Keeping decreasing SRTT of selected servers is unnecessary.
- Unresponsive server should not be treated as a server with a large SRTT.

Thanks!