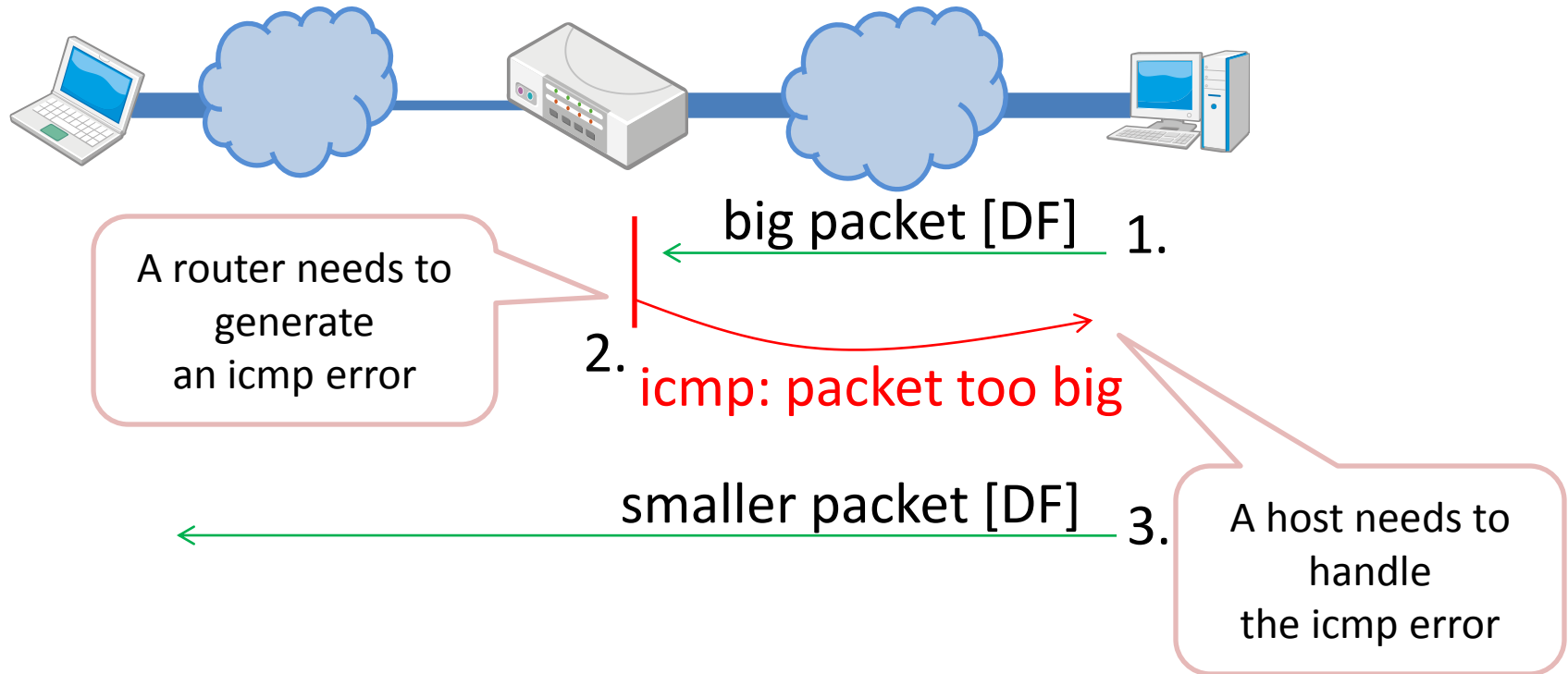# Path MTU Discovery
- failure cases -

Matsuzaki 'maz' Yoshinobu
<maz@iij.ad.jp>

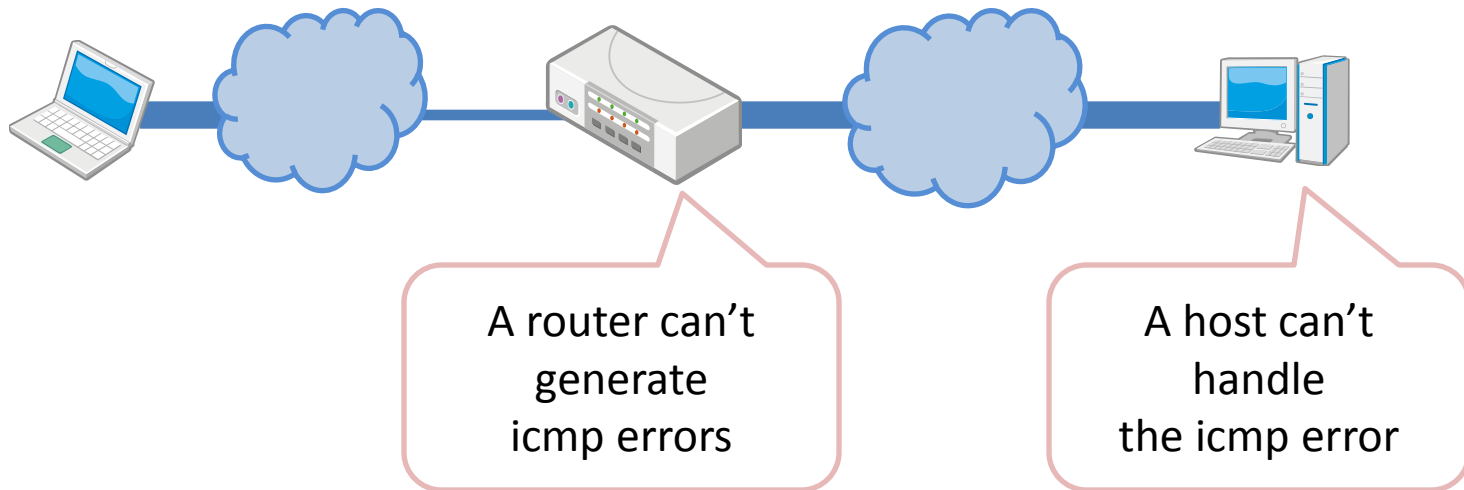# IPv6 and Path MTU Discovery

- Path MTU discovery for IPv6 [RFC1981]
  - IPv6 nodes SHOULD implement Path MTU Discovery in order to discover and take advantage of paths with PMTU greater than the IPv6 minimum link MTU [IPv6-SPEC].
- IPv6 minimum link MTU [IPv6-SPEC] == 1280
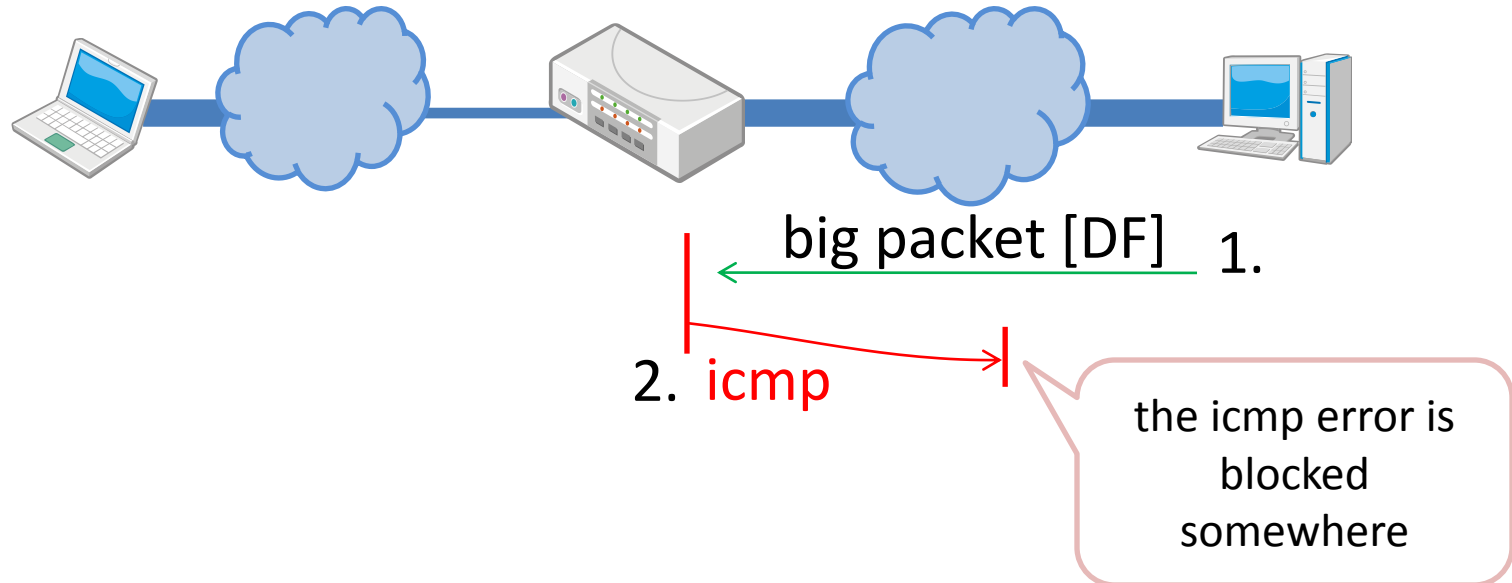
path MTU discovery scenario

A router needs to generate an icmp error

big packet [DF]   1.

2.
icmp: packet too big

smaller packet [DF]   3.

A host needs to handle the icmp error

failure case #1: incapable

- pMTUd blackhole router
- lack/mis-implementation of icmp handling

A router can't
generate
icmp errors

A host can't
handle
the icmp error

failure case #2: filtered

- careless packet filter
- clueless security policy

big packet [DF]   1.

2. icmp

the icmp error is blocked somewhere

failure case #3: limited

- how often can a router generate icmp errors?
- how many networks put rate-limit for icmp?
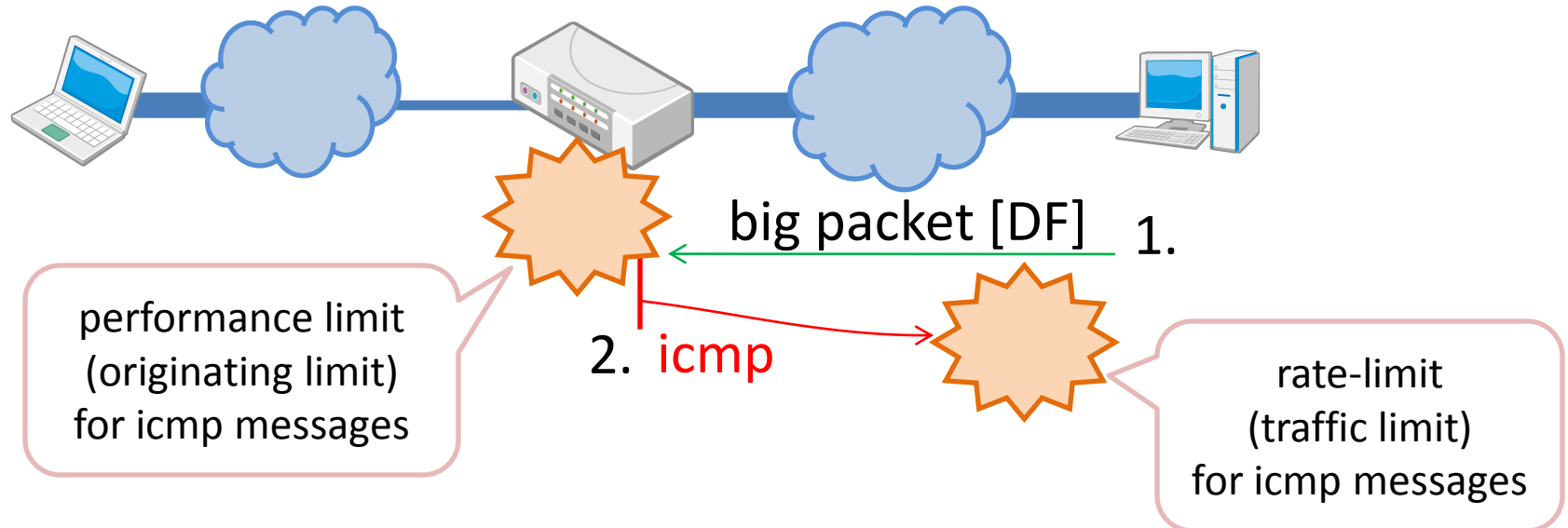
icmp originating-limit

- cisco ios
  - ip icmp rate-limit unreachable 500
    - means icmp errors are limited to one every 500msec
  - ipv6 icmp error-interval 100
    - means icmp errors are limited to one every 100msec
- juniper junos
  - icmpv4-rate-limit {packet-rate 1000;};
    - means max 1000pps for icmp to/from RE
  - icmpv6-rate-limit {packet-rate 1000;};
    - means max 1000pps for icmp to/from RE

summary of failures

- The Path MTU Discovery could fail even if all of given devices support it
  - performance issue
  - icmp message could be discarded
- The Path MTU Discovery is like an "exceptional handling"
  - network ops are usually focused on 'forwarding performance' of routers.

IPv4 TCP SYN rate
at a consumer aggregation router
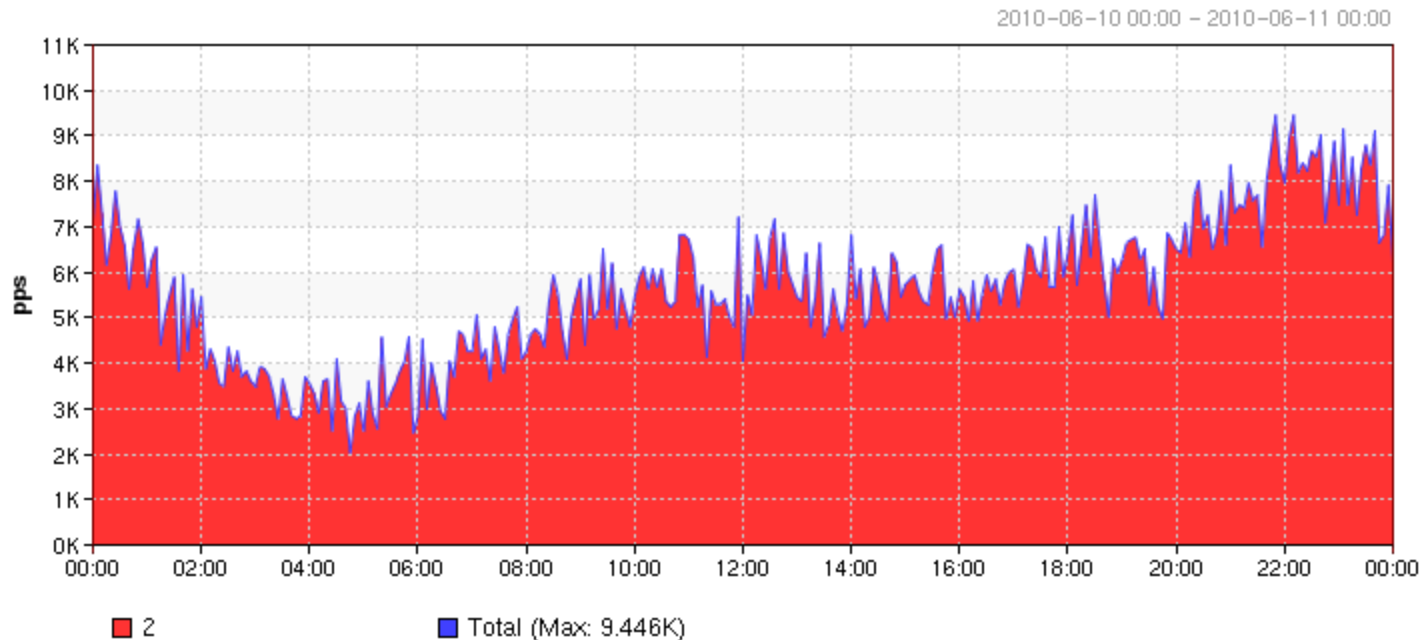
learning from IPv4

- Almost of all broadband routers have a TCP MSS hack capability

- It chokes TCP MSS on a tunnel link
  - PPPoE, or whatever the link MTU is less than 1500
  - to avoid unnecessary fallbacks

- The TCP MSS hack works fine
  - No complaint from customers

options for IPv6

- MTU notification by RA
  - think about 1000base-T at home
    - people tend to use jumbo frame at home

- TCP MSS hack by broadband router
  - though this works for TCP only

suggestion

- ask broadband venders to implement the TCP MSS hack for IPv6 as well
  - TCP MSS hack is a dirty hack, but still it works

- improve the support of path MTU discovery
  - revise filters/rate-limits
  - care about MTU steps