How to write UDP/IPv6 applications that care about path MTU

JINMEI, Tatuya
Toshiba Corporation/The KAME Project
jinmei@{isl.rdc.toshiba.co.jp, kame.net}

Background

- Path MTU (PMTU) discovery is mandatory in IPv6.
  - Routers always discard too big packets.
- rfc2292bis(-02) socket options and ancillary data items
  - IPV6_USE_MIN_MTU
    - requires the kernel to send packets at the minimum MTU.
    - for applications that do not want packet drop during the PMTU discovery phase.
    - e.g. DNS, IKE, DHCP(?)
  - IPV6_RECVPATHMTU
    - requires the kernel to tell applications an appropriate MTU.
    - for applications that can divide a single transaction into multiple packets.
    - in order to avoid fragmentation.

IPV6_USE_MTU(new ancillary data item)

- Originally proposed by Robert Elz in April 2000.
- Natural generalization of IPV6_USE_MIN_MTU.
  - takes an integer.
  - requires the kernel to send packets at a specified MTU.
  - would obsolete IPV6_USE_MIN_MTU.
    - IPV6_USE_MTU(1280) == IPV6_USE_MIN_MTU
- Useful when
  - an application can't divide a single transaction,
  - it would rather perform PMTU discovery than just sending at the minimum MTU, and
  - we can't assume the kernel's behavior about PMTU discovery.
    - e.g. introduce a strict condition to reject forged TOO BIG erros
- But
  - At least every BSD kernel always performs PMTU discovery.
  - not necessary if we can always rely on the kernel's behavior.
- So
  - would like to hear opinions.