Quick Failover Algorithm in SCTP

draft-nishida-tsvwg-sctp-failover

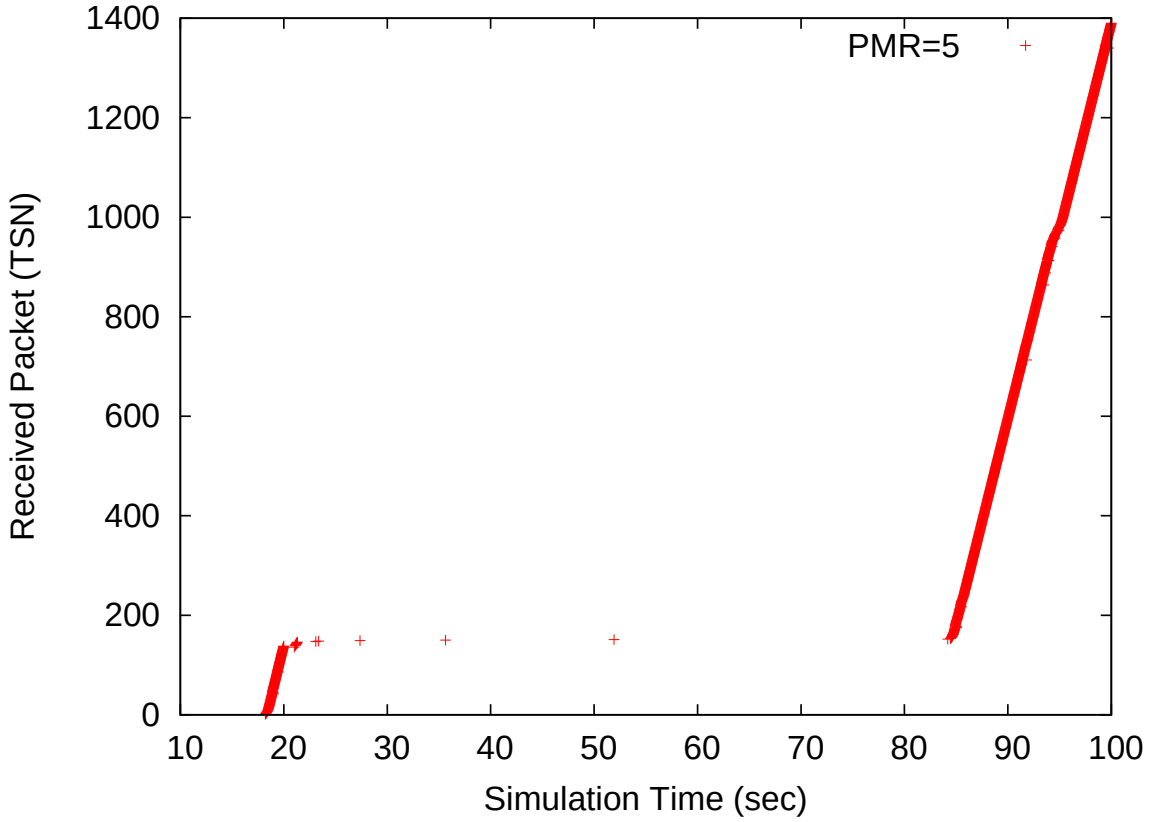
Yoshifumi Nishida, WIDE Project
Preethi Natarajan, CISCO systems

What is Quick Failover?

- A solution for failover issue in SCTP
  - SCTP needs 30-60 secs to failover in standard settings

Issues in SCTP Failover

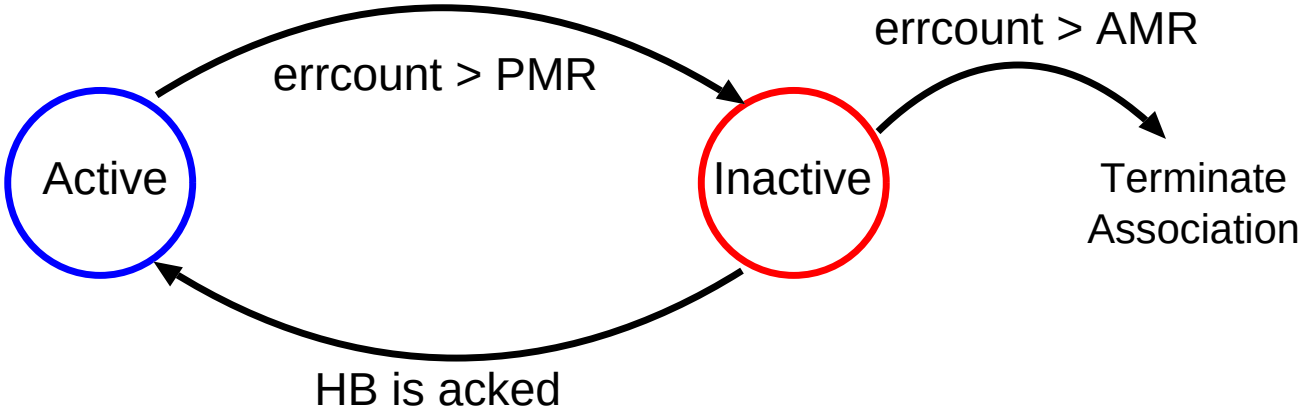
- SCTP needs 6 consecutive timeouts before failover
  - Path.Max.Retrans is recommended to be 5 in RFC4960



A is sending data to B and B has two address B1, B2 (B1 is primary) when primary becomes unavailable at 20 sec, it takes 60 secs to restart data transmission. (Path.Max.Retrans = 5)

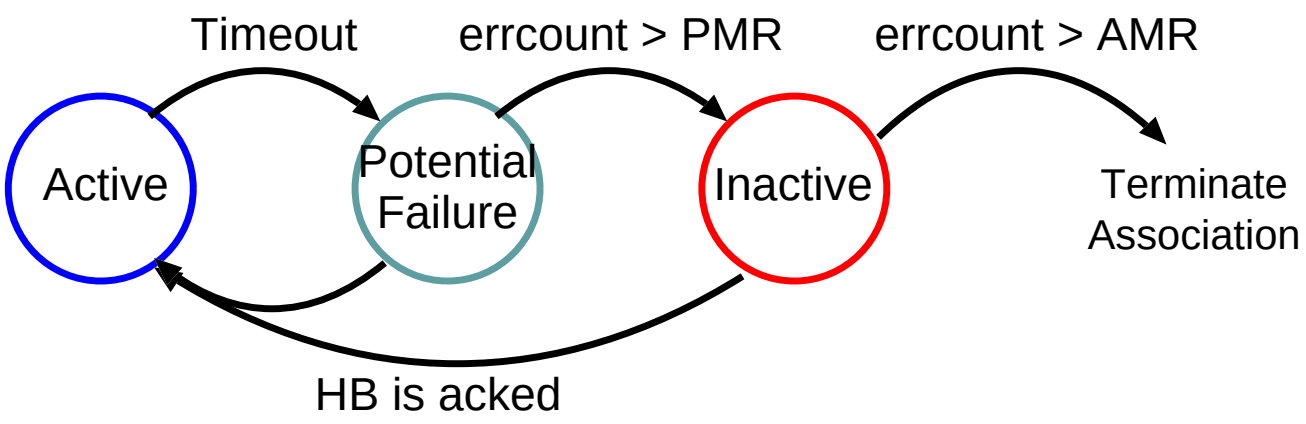
SCTP Path Management

- SCTP marks path inactive when errcount > PMR
  - Failover happens after path is marked as inactive
- SCTP terminate association when errcount > AMR



Quick Failover

- Introduce an intermediate state
  - When path is in PF, SCTP can utlize secondary path
    - Send HB to the primary and if HB ack returns, it quickly fallback to active



Quick Failover Summary

- Use secondary path quickly in case of path failure
- Simple and sender only logic
- Research results indiacate it’s useful and harmless
- No need to change current PMR, AMR, HB.Interval
  - No need to change applications or OS’s settings
- It can be applied to both RFC4960 and CMT proposal
  - draft-tuexen-tsvwg-sctp-multipath
- Behavior is configurable
  - Apps can preserve original behavior if they want

Do We Really Need This?

- We have several choices
  - Do nothing. 30-60 secs delay can be acceptable
  - Expect developers and sysadmins to solve this
    - Tuning several parameters will work in some situations
  - Update the spec to support PF
    - No requriment for app developers or sysadmins
    - Give consistent behavior to users

PMR = 0 Solution

- Setting PMR=0 can also be a solution for this
- But,
  - We'll still need to update RFC4960 to some extent
    - Recommended value for PMR
    - Behavior in dormant state
    - Relationship between PMR and AMR
      - RFC4960 states 'users should avoid having the value of Association.Max.Retrans larger than the summation of the 'Path.Max.Retrans'
    - May need to add special logic for sending HB