SCTP based TML for ForCES protocol

Jamal Hadi Salim
ZNYX

SCTP vs TCP, UDP, DCCP

- Vs TCP:
  - provide ordered, reliable, connection oriented, flow controlled, congestion controlled data exchange
  - it does not provide byte streaming rather messages
- Vs UDP
  - provide message boundaries, unordered, unreliable data exchange
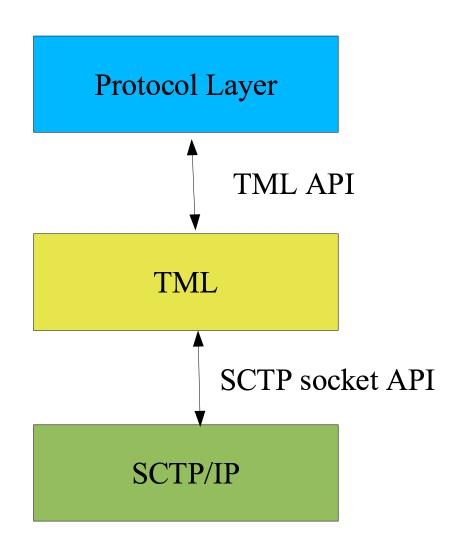 - does not provide multicast
- Vs DCCP
  - can provide unreliable, ordered, congestion controlled

Additional services unique to SCTP

- Multi-homing
- Runtime IP binding (via ADDIP)
- A range of reliability shades + congestion control
- Built-in heartbeating
- Multi-streaming
- Message boundaries + reliability
- Improved SYN DOS protection
- Simpler Transport events
- Simplified replicasting

So why SCTP?

- Mainly an all in one package
    - All other proposals require > 1 protocol
    - Allows for a much simpler programming
- Very mature (relative to DCCP for example)
    - Has been around for a few years
    - Widely deployed
- Provides more features with little effort
    - Example HA
    - Multiple streams for data vs control separation

Meeting TML requirements

Meeting TML requirements

- Reliability
  - It is possible to have reliable data exchange
- Congestion control
  - All data exchange is congestion controlled
- Timeliness
  - message can be time limited in PR-SCTP
    - If a message is not sent after timeout it is junked locally
      - a Forward-TSN message sent to remote to skip message
    - If a message is acknowledged after timeout, it is ignored

Meeting TML requirements

- Prioritization
  - Multiple streams can be made to be prioritized
    - The stream scheduler on Linux is incapable today
      - Some code is needed
- PL Addressing to peers
  - SCTP can be told to replicast a packet it receives (in the kernel) to several destinations
    - Not as good as UDP multicast, but saves local system memory bandwidth in multi-VM domain OSes (Unixes)
- Encapsulation
  - None needed by TML (if needed add new TLVs/chunks)

Meeting TML requirements

- HA
  - Multi-homing provides path diversty
    - When peer-IP is unreachable other can be accessed without TMLs intervention
  - Reachability fault detection
    - Built in HB on a per-peer IP address
    - Data transmission threshold on a per-peer IP address
  - Can coordinate migration of IP addresses from one node to another
    - ADDIP: allows adding IP of peers at runtime

What am I looking for?

- Make this a WG item