SCTP Tail Loss Recovery Enhancementss

draft-nielsen-tsvwg-sctp-tlr-01.txt

Tsvwg, IETF 91, Haweii

Karen E E Nielsen, Ericsson

Status

- First draft, draft-nielsen-tsvwg-sctp-tlr-01.txt, available
- Implementation done in SCTP SW
- ”RFC6675” improvements active in SCTP SW in deployment
- Experiments (test environment) are ongoing
- Plans for collaboration with other SCTP groups/other SCTP SWs for ver02

Goal

- Improve SCTP Fast Recovery (SCTP FR) where dupthres (=3) mis indications/SACKs don't activate Fast Retransmission (not at all or late)
  - → Primary: Improve latency of loss recovery
  - → Secondary: Potential Throughput impr. (tbd)
- Presently supplements dupthres driven FR
  - Potential Future ? *Replace* dupthres driven FR with new timer driver approach more robust to packet re-ordering

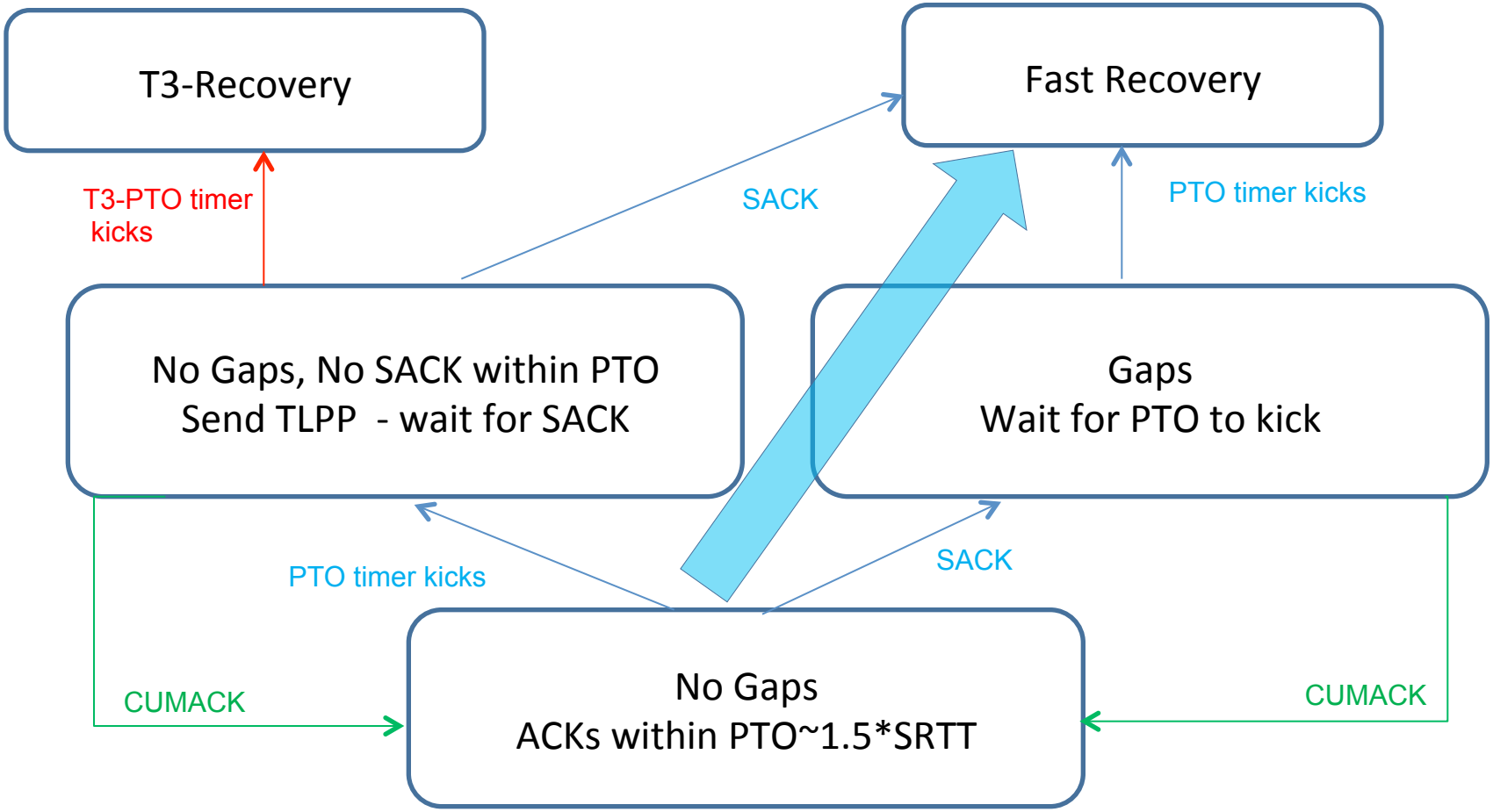
SCTP Loss Recovery Improvements

- ”RFC6675 improved” SCTP FR
  - Mis indication counts robust towards loss of SACKs
    - packet id variant and bytes variant (bytes variant not impl)
  - RFC6675 Last Resort features added to SCTP FR
- PTO Timer driven enter of FR (insp [1], [2])
  - SACK of higher TSN within PTO → Enter FR

    PTO1= 1.5 SRTT + MAX(RTTVAR, Delay_ack)

    PTO2= 1.5 SRTT + RTTVAR
- PTO Timer driven Tail Loss Probing (insp by [1])
  - TLPP sent if no SACK received within PTO
  - Rearm of timer as T3-timer of MAX(PTO, RTO-PTO)

Simplified View



More details see draft and/or http://www.ietf.org/proceedings/90/slides/slides-90-tsvwg-16.pdf

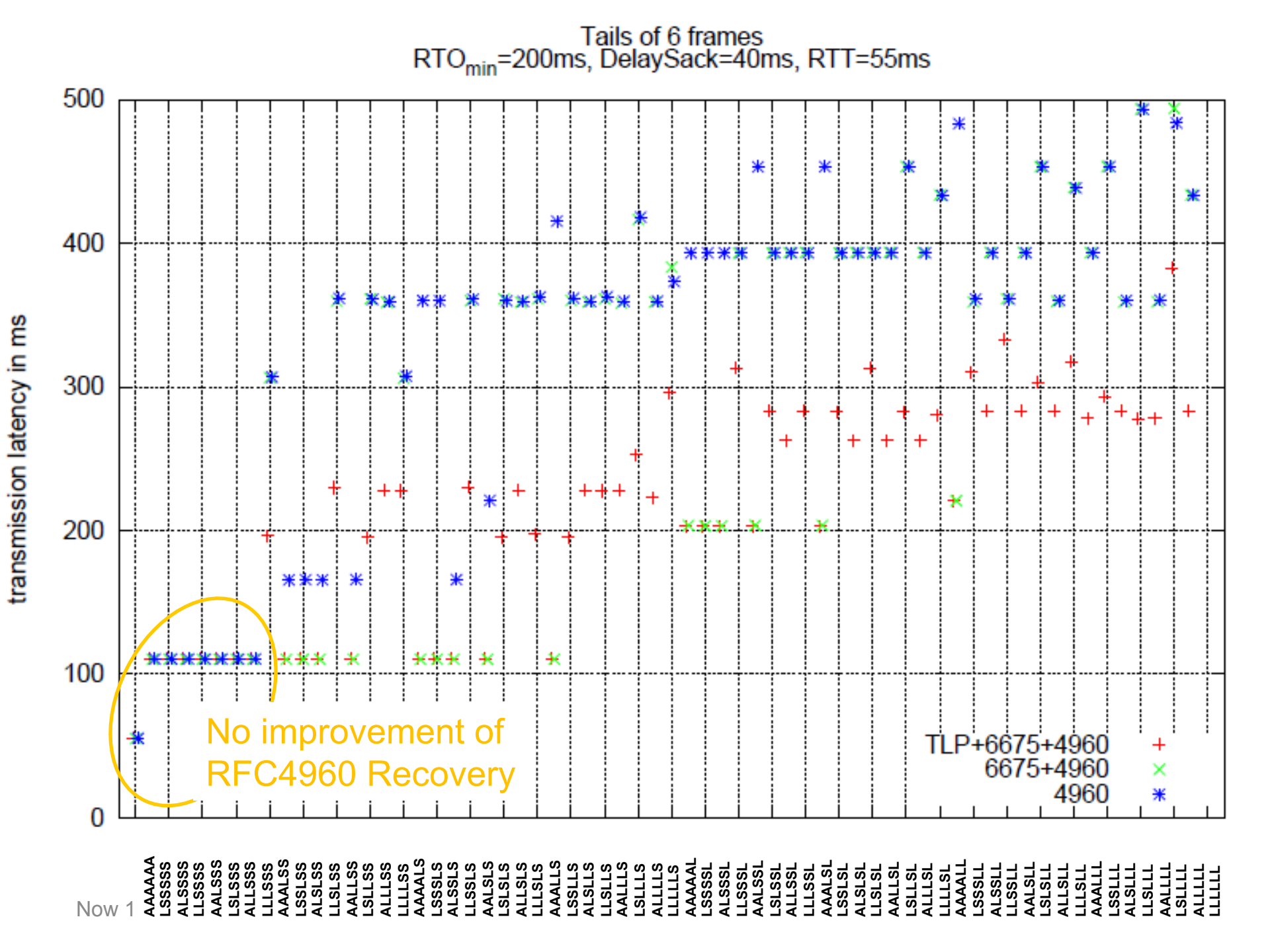
ISSUES

- TLPP Recovery Masking algorithm (almost as of [1])
  - may give spurious cwnd halving if TLPP is re-ordered with subsequently sent packets (as SACKed)
  - Possibly solution for SCTP is to introduce chunk id (Disscussed for other reasons as well, need protocol ext.)
- PTO timers:
  - PTO1, PTO2 for further consolidation
  - 2*PTO ~worst case~ 3*SRTT + 2*delay_ack (ideal not > RTO)
- PTO-Restart ”as of” [3] (not fully analysed yet)
  - Possibly implementation by SCTP Packet id embedding timestamp
  - PTO restart conjectured to be unconditionally
- CC Issues (not analysed, work from TCP may apply)

Some Preliminary RESULTS….

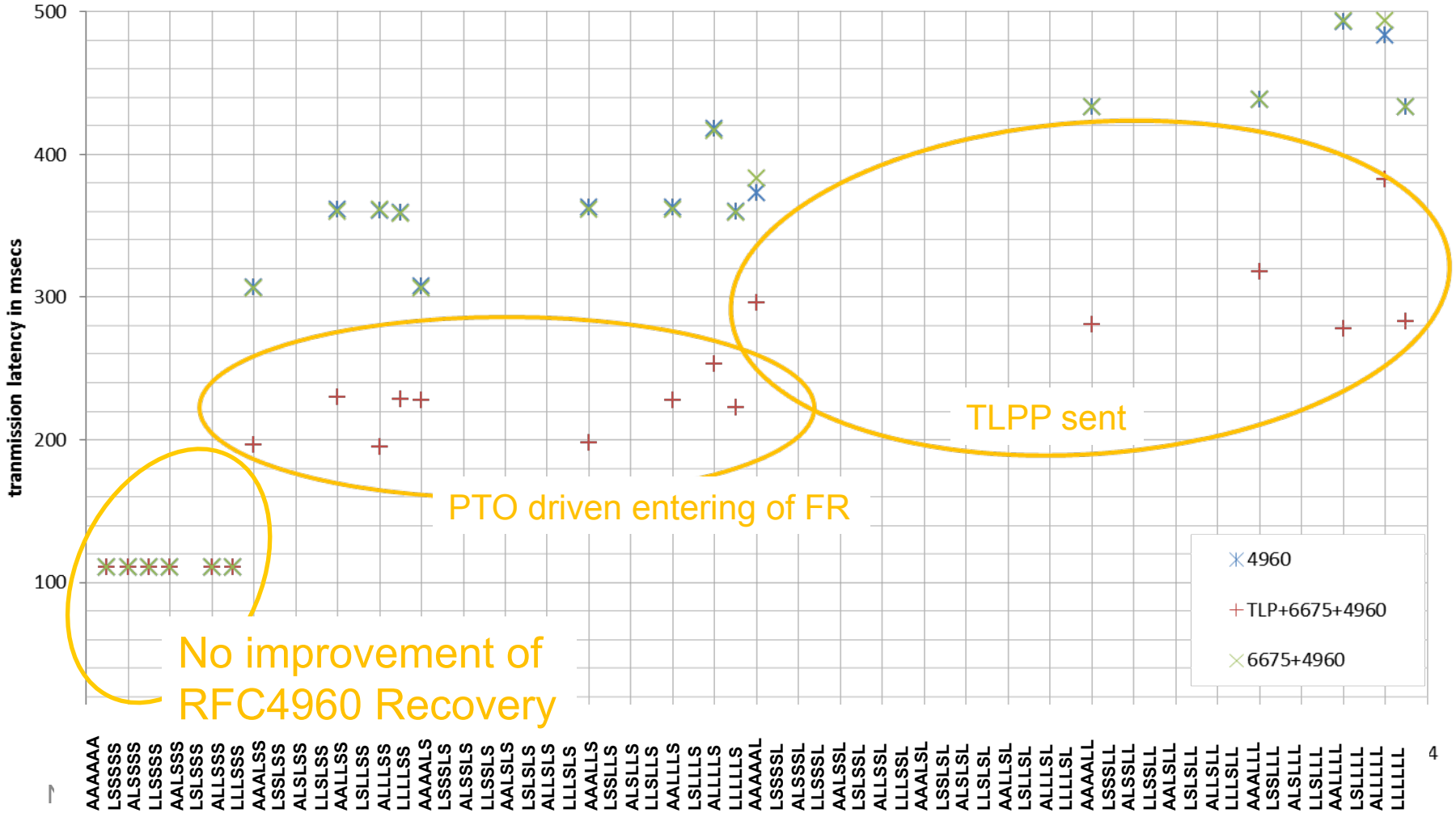
Functional tests

- Latency of Loss Recovery over a tail loss pattern over 6 SCTP packets transmitted instantly (IW of 4MTU)
- Latency = time from transmit (write) of packets to CUMACK all
- All combinations of tail loss patterns tested (drop filter in ETH egress).
- Results shows median value obtained over a number of tests. Variation only from timer-resolution issues. RTT, RTO_MIN, DELAY_ACK set to prevent algorithmic race conditions.
- No SACK loss, no RTT variation, no retransmission losses.


Tails of 6 frames
RTO$_{min}$=200ms, DelaySack=40ms, RTT=55ms
transmission latency in ms
500
400
300
200
100
0
No improvement of
RFC4960 Recovery
TLP+6675+4960
6675+4960
4960
AAAAAA
LSSSSS
ALSSSS
LLSSSS
AALSSS
LSLSSS
ALLSSS
LLLSSS
AAALSS
LSSLSS
ALSLSS
LLSLSS
AALLSS
LSLLSS
ALLLSS
LLLLSS
AAAALS
LSSSLS
ALSSLS
LLSSLS
AALSLS
LSLSLS
ALLSLS
LLLSLS
AAALLS
LSSLLS
ALSLLS
LLSLLS
AALLLS
LSLLLS
ALLLLS
LLLLLS
AAAAAL
LSSSSL
ALSSSL
LLSSSL
AALSSL
LSLSSL
ALLSSL
LLLSSL
AAALSL
LSSLSL
ALSLSL
LLSLSL
AALLSL
LSLLSL
ALLLSL
LLLLSL
AAAALL
LSSSLL
ALSSLL
LLSSLL
AALSLL
LSLSLL
ALLSLL
LLLSLL
AAALLL
LSSLLL
ALSLLL
LLSLLL
AALLLL
LSLLLL
ALLLLL
LLLLLL
Now 1

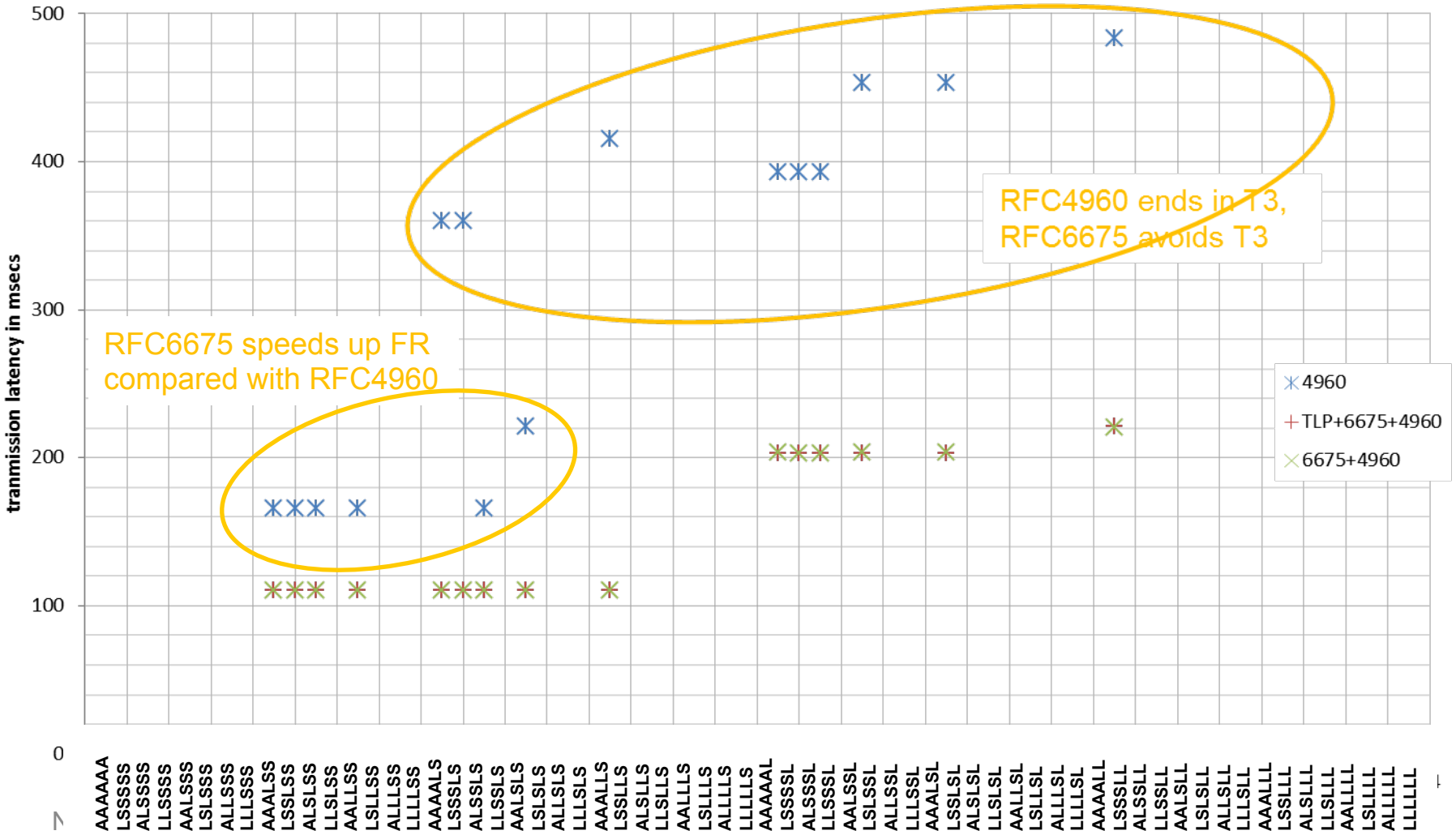
Evaluation of Results (1)

Effects for One Tail Loss Event types



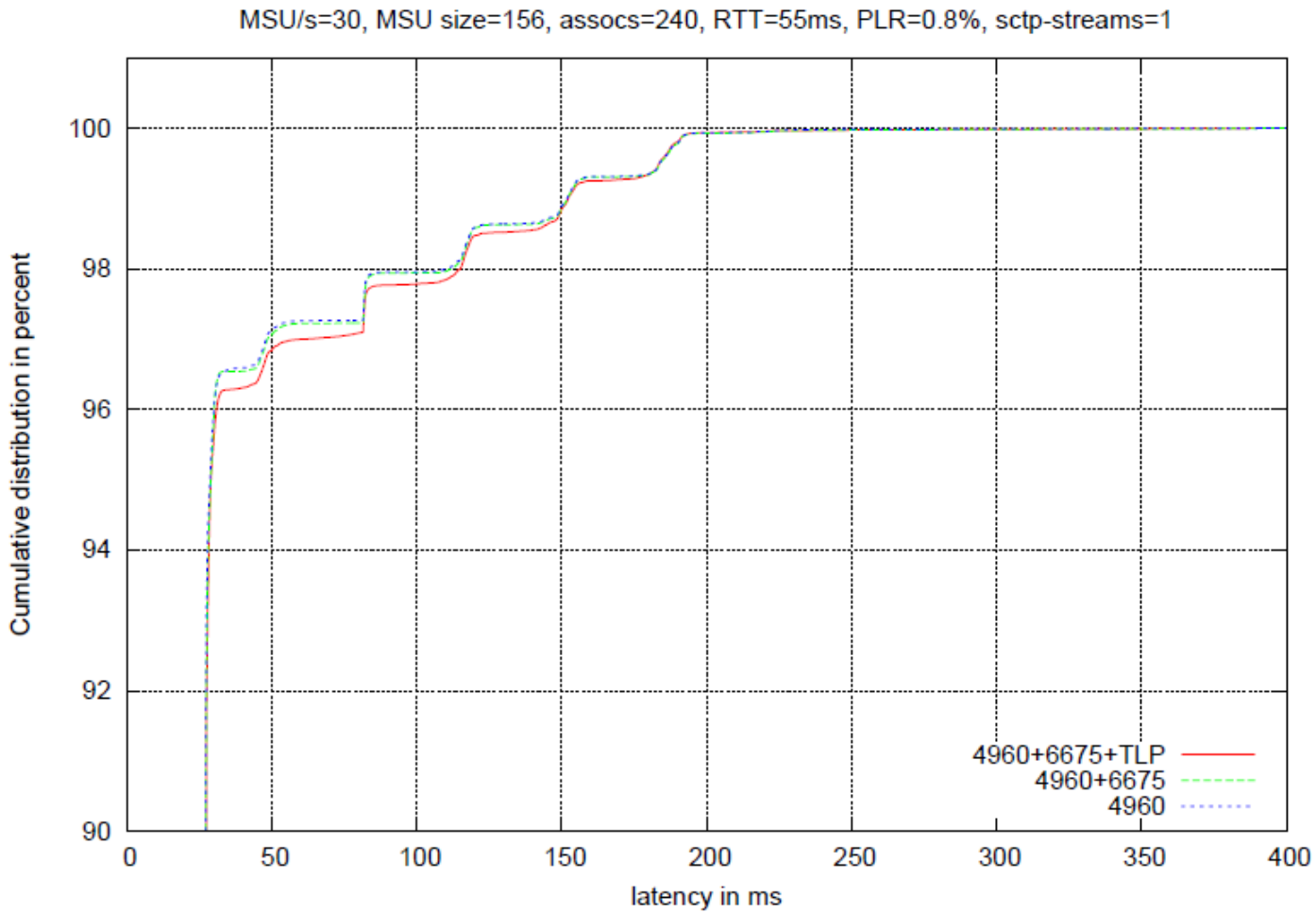
Evaluation of Results (2)

RFC6675 Effects



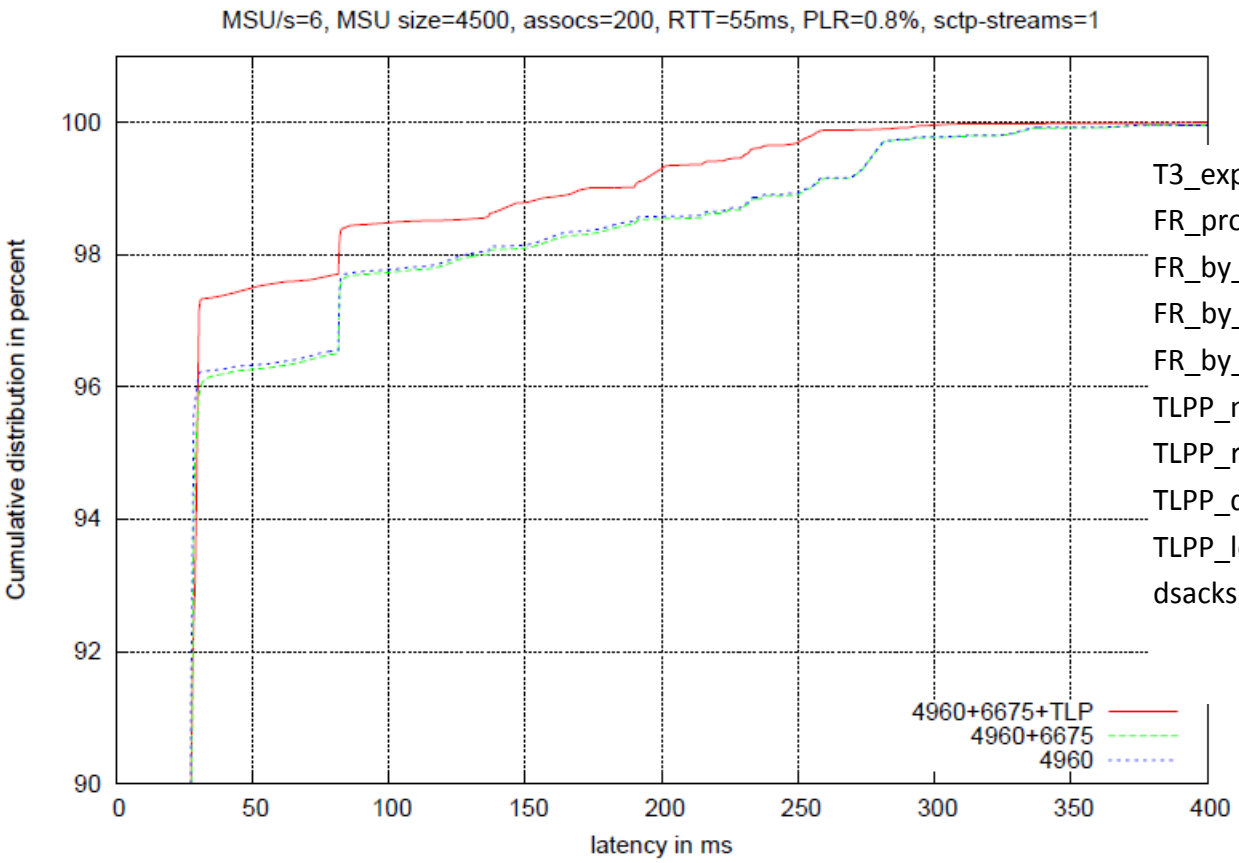
Synthetic Traffic (1)

- Bottleneck simulated by Gilbert-Elliott packet losses



Synthetic Traffic (2)

- Bottleneck simulated by Gilbert-Elliott packet losses



Statistics

	RFC4960	RFC4960+ RFC6675	TLR
T3_expireds	6651	6758*	227
FR_prolong	2	0	10
FR_by_3miss	8006	8241	2142
FR_by_probe_wait	0	0	2425
FR_by_delay_wait	0	0	9230
TLPP_new	0	0	759
TLPP_rtx	0	0	1827
TLPP_denied	0	0	0
TLPP_loss_mask	0	0	158
dsacks	42	29	753

- Under analysis.
  Double loss or 3) block in CC

References

[1] Dukkipati et al., Tail Loss Probe (TLP): An Algorithm for Fast Recovery of Tail Losses, Expired work.
http://tools.ietf.org/html/draft-dukkipati-tcpm-tcp-loss-probe-01

[2] Dukkipati et al, "Proportional Rate Reduction for TCP", Proceedings of the 11th ACM SIGCOMM Conference on Internet Measurement 2011, Berlin, Germany, November 2011.

[3] Hurtig et al, TCP and SCTP RTO Restart, draft-ietf-tcpm-rtorestart-04, Work In Progress