Enhancing TCP's Loss Recovery Using Early Duplicate Acknowledgment Response

Mark Allman, Hari Balakrishnan, and Sally Floyd

draft-allman-tcp-lossrec-00.txt
June 2000

IETF, Transport Area Working Group
August 1, 2000

Proposing Two New TCP Mechanisms:

- Limited Transmit: Sending a new data segment in response to each of the first two duplicate acknowledgments that arrive at the sender.

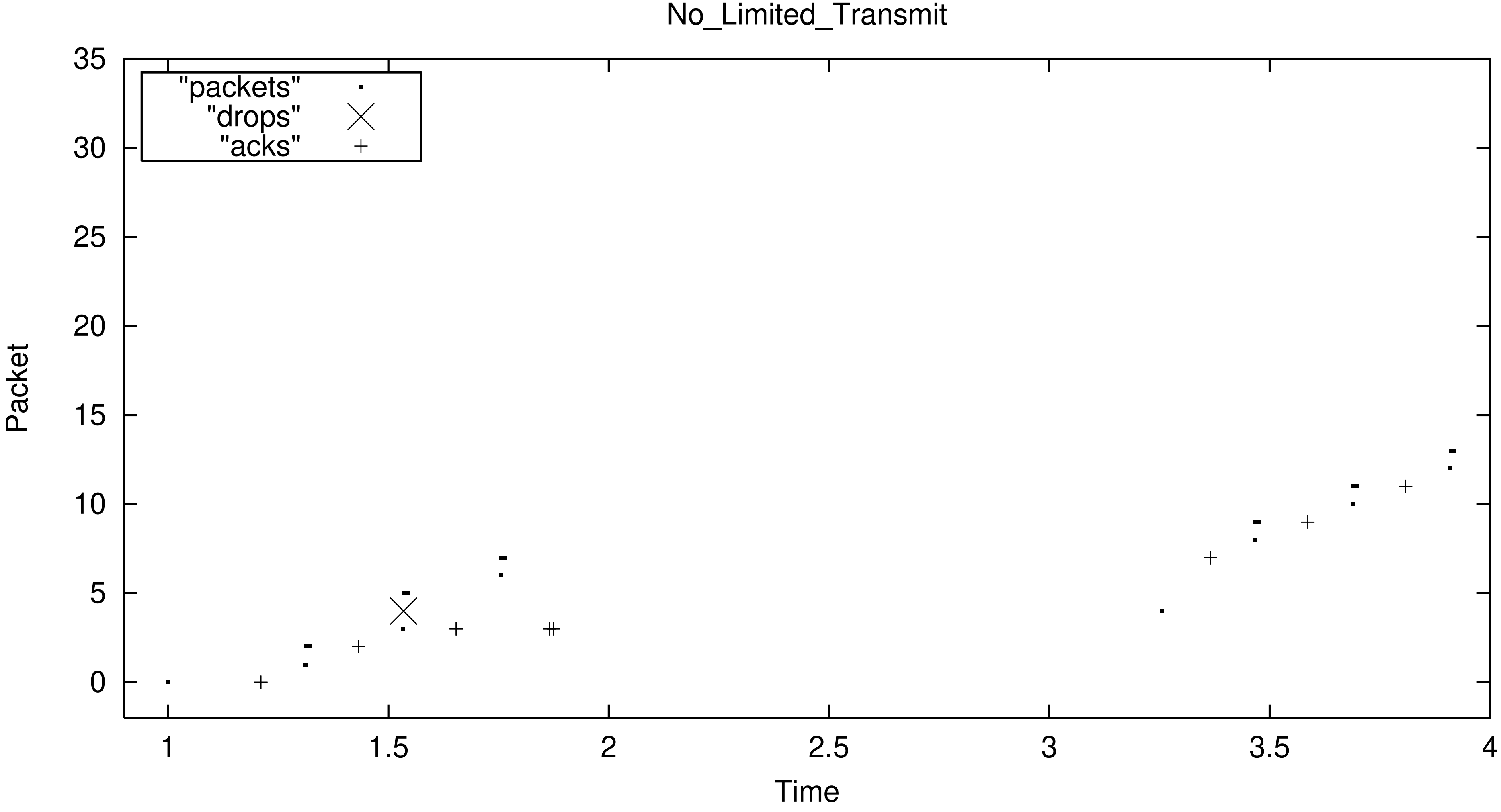
- Early Fast Retransmit: Reducing the dup ack threshold for triggering Fast Retransmit in certain special circumstances.

What is the problem that needs to be fixed?

- TCP does not recover well from packet loss when the congestion window is less than four packets.
  - Fast Retransmit requires three dup ACKs.
  - The first and second dup ACKs are ignored.
  - The sender has to wait for a Retransmit Timeout.
  - The minimum RTO is one second.

- TCP does not recover well when too many packets are lost from a window of data.

The problem:



- To run: ns test-suite-LimTrans.tcl No_Limited_Transmit

Why does it matter?

- [BPS+98] shows that Retransmit Timeouts are common in traffic from a busy web server.

- [All00] shows that for one particular web server the median transfer size is less than four segments.

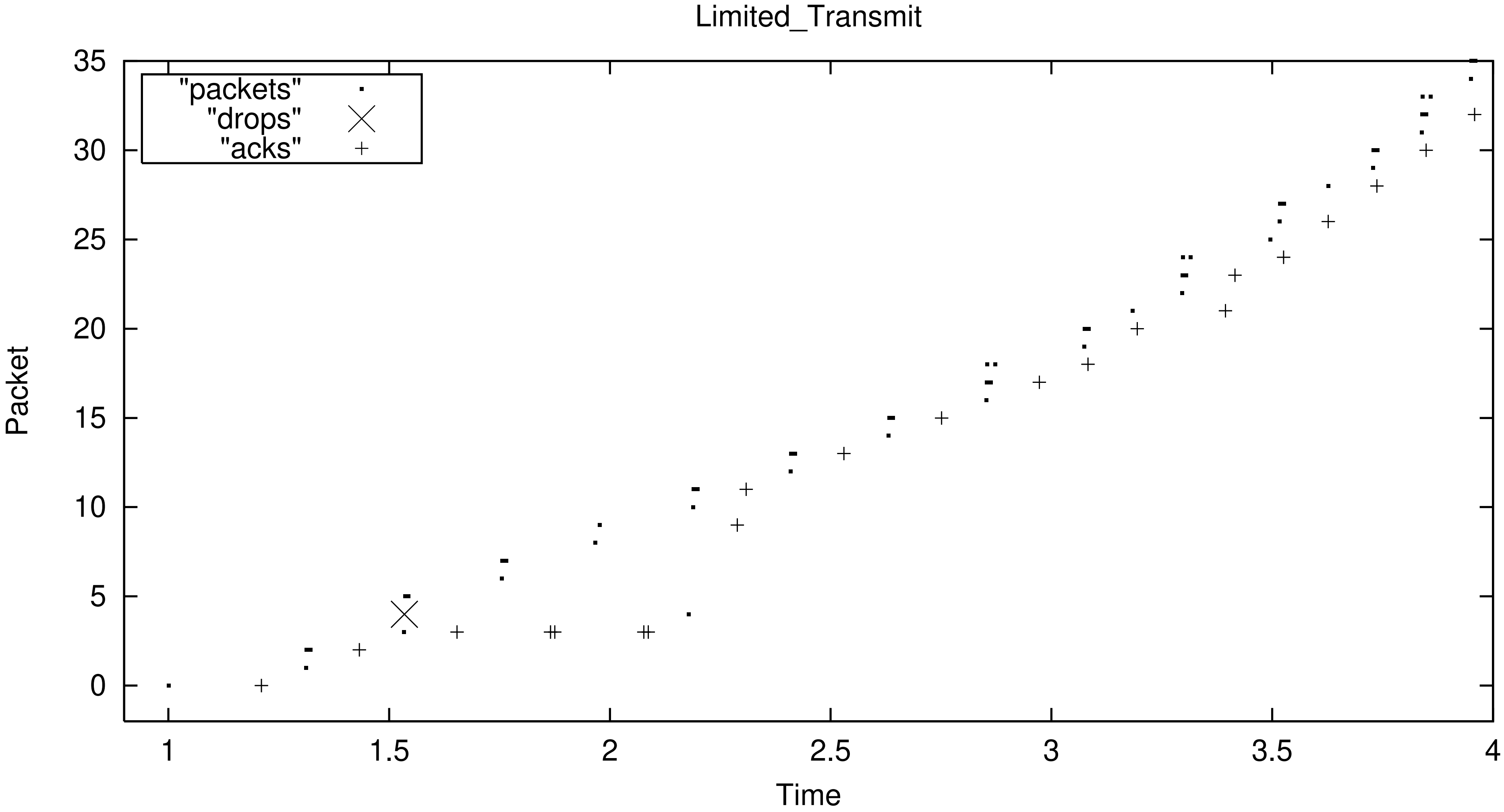
Limited Transmit:

- When the first or second dup ACKs arrive, the sender SHOULD send a new segment:
  - If allowed by the receiver's advertised window.
  - If the congestion window would be exceeded by at most two segments.

- The congestion window is not increased during Limited Transmit.

- For SACK TCP, Limited Transmit is used only if the dup ACK acknowledges new data.

The fix:



- To run: ns test-suite-LimTrans.tcl Limited_Transmit

The official language:

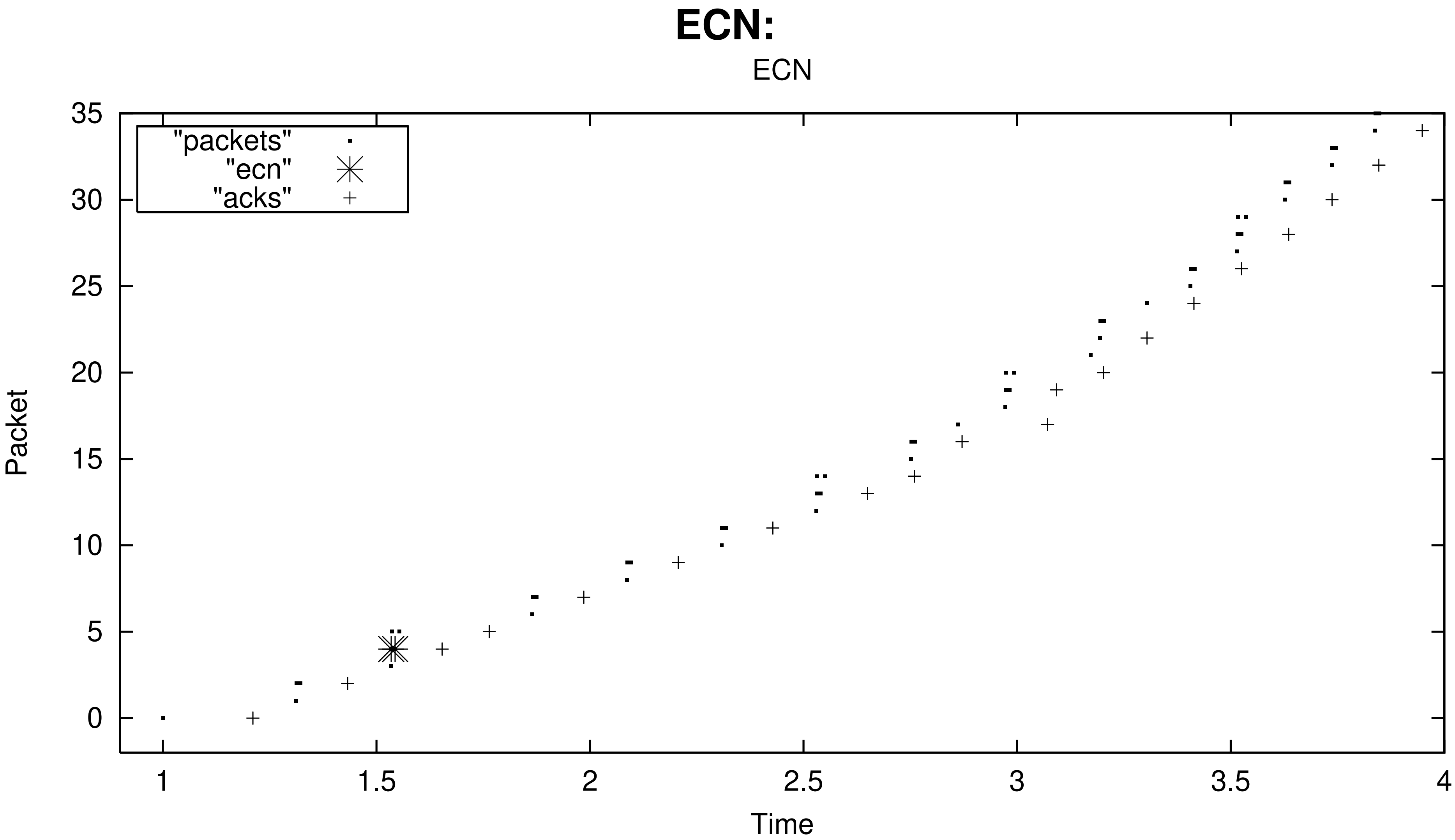
- “The amount of outstanding data would remain less than the congestion window plus the duplicate ACK threshold used to trigger Fast Retransmit. In other words, the sender can only send two segments beyond the congestion window (cwnd).”

Properties of Limited Transmit:

- “Conservation of packets” is followed.
- Robust to reordered packets.
- Keeps the ACK clock going, ultimately allowing a Fast Retransmit.
  – Retransmit Timeouts are reduced.

Relationship to ECN:

- Limited Transmit gives non-ECN-enabled TCP performance closer to that of ECN-enabled TCP in an environment with ECN, for TCP with small congestion windows.



- To run: ns test-suite-LimTrans.tcl ECN

Fairness of Limited Transmit?

- "When ECN-enabled TCP traffic competes with non-ECN-enabled TCP traffic, ECN-enabled traffic can receive up to 30% higher goodput. For bulk transfers, the relative performance benefit of ECN is greatest when on average each flow has 3-4 outstanding packets during each roundtrip time." [ZQ00]

- A similar, slightly reduced benefit should occur for TCP with Limited Transmit, competing against non-ECN TCP without Limited Transmit.

Security Considerations of Limited Transmit:

- The problem: “False” dup ACKs. [SCWA99,RFC2581]

- Protection without SACK:
  – The sender exceeds the dup ACK threshold by at most two segments.

- Additional protection with SACK:
  – With SACK, the sender can verify that the dup ACK reports an acknowledgement for data that was actually sent.
  – With SACK, at most one valid acknowledgement recognized for a segment of data.

Early Fast Retransmit:

- The sender might not have new data to send when the first and second dup ACKs arrive.

- In this case, the sender could reduce the dup ACK threshold to "cwnd - 1".

- Example: When cwnd = 3, the dup ACK threshold would be 2.

- Problem: Less robust to reordered packets.
  - Fix: Early Fast Retransmit can be invoked only once in the life of a connection.
  - Fix: Early Fast Retransmit should only be invoked during an initial Slow-Start?

References:

- [All00] Mark Allman. A Web Server's View of the Transport Layer. May, 2000. In preparation.

- [BPS+98] Hari Balakrishnan, Venkata Padmanabhan, Srinivasan Seshan, Mark Stemm, and Randy Katz. TCP Behavior of a Busy Web Server: Analysis and Improvements. Proc. IEEE INFOCOM Conf., San Francisco, CA, March 1998.

- [ZQ00] Yin Zhang and Lili Qiu, Understanding the End-to-End Performance Impact of RED in a Heterogeneous Environment, Cornell CS Technical Report 2000-1802, July 2000.