

The RX Extended SACK Table Proposal - AuriStor RX Prototype Results

Wednesday, 27 October 2021 17:50 (25 minutes)

The RX protocol inherited from IBM AFS is incapable of filling a network pipe with a single RPC when the pipe's bandwidth delay product exceeds $44 \frac{1}{4}$ KB. On a 1 Gbit/sec pipe with a 1ms RTT, the maximum theoretical throughput is 360 Mbit/sec with a maximum window size of 44 KB. The RX ACK packet format provides for a theoretical maximum window of 65535 packets but the Selective Acknowledgment (SACK) Table is limited to 255 packets. With 255 packets the maximum window size is 351 KB or maximum throughput of 2.875 Gbit/sec with 1ms RTT. Increase the RTT to 8ms and the maximum throughput is once again reduced to 360 Mbit/sec. The RTT on cross-Atlantic commodity internet pipes often exceed 110ms which reduces the theoretical throughput to 28 Mbit/sec.

This presentation will describe AuriStor's proposed RX Extended SACK Table protocol extension, prior efforts at extending the maximum window size, and preliminary results on real-world networks using AuriStor's prototype supporting maximum windows up to 8192 packets or 11MB.

Desired slot length

20 minutes

Speaker release

Yes

Primary author: Mr ALTMAN, Jeffrey

Presenter: Mr ALTMAN, Jeffrey

Session Classification: Storage & File Systems

Track Classification: Storage & Filesystems