Achieving Practical Reliable Multicast with TCP-SMO

Michael Chan, Sam Liang and David Cheriton

mcfchan@stanford.edu, sliang@dsg.stanford.edu, cheriton@stanford.edu

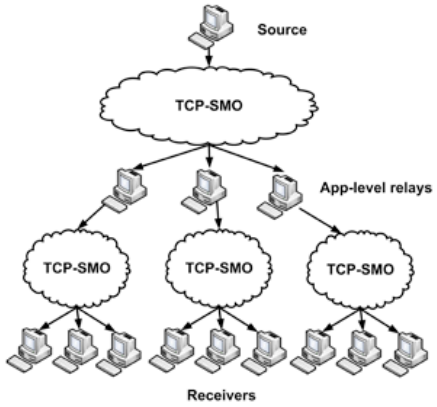
Motivation

Applications involving reliable data transfer to a large number of hosts in enterprise and datacenter networks, e.g. software updates, content distribution in CDNs, publish-subscribe systems, and live lecture broadcasts on campus, can be significantly improved with multicast to achieve better bandwidth utilization and lower latency.

Issues with existing solutions

- Application-level multicast – inefficient bandwidth usage as amount of data sent grows linearly with # receivers
- Reliable transport protocols – Many solutions designed for large-scale operation, some require router support, but none are widely accepted and deployed

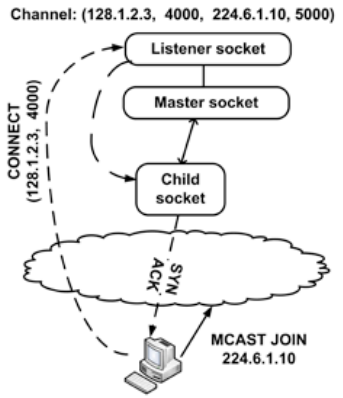
Approach with TCP-SMO



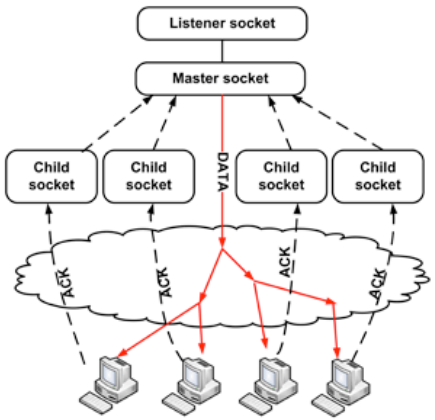
- Adds reliable single-source multicast optimization (SMO) to TCP, the most widely used transport
- Changes only end hosts – easier deployment
- Provides a familiar API – only need to set one extra socket option
- Embraces network-layer multicast for significant bandwidth savings over multiple unicast transfers
- Operates at medium scale (up to 1000 receivers) to maintain receiver homogeneity – mitigates "crying baby" problem, where one slow receiver slows down entire group
- Achieves large scale with application-level relays – hierarchy of multiple TCP-SMO sessions
- Supports multiple senders with session relay – data from receiver relayed by source

TCP-SMO illustration

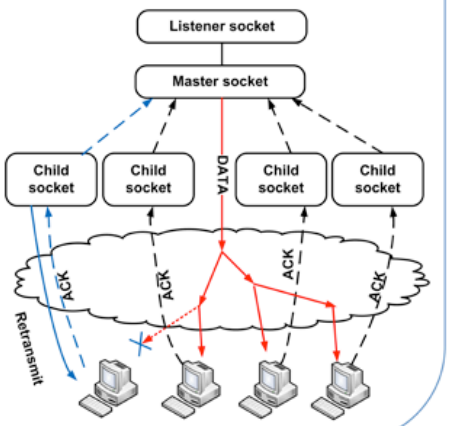
Receiver-initiated connections



Data multicast by master socket



Retransmissions by child sockets



Early evaluation

- 2% loss rate at 200 receivers
- Source: PIII 930MHz CPU, 0.1% utilization