Towards a Composable Transport Protocol: TCP without Congestion Control

Roland Kempter+, Xin Bin†, Sneha Kumar Kasera‡, Christian Schlegel×

SIGCOMM 2004, Portland/Oregon/USA

Introduction
- TCP features: error control, congestion control, in-order delivery, flow control
- App. diversity: not all features needed for diverse applications, networks
- Performance issue: including unnecessary features could reduce performance
- Our goal: make TCP features composable
  - Allow applications to use only the required features
  - Study the interactions between features
  - Support dynamic protocol negotiation and upgrading
  - Investigate the impact of congestion control on the performance of TCP over lossy links
- Other benefits of modularity and compositability

TCP w/o Congestion Control
- Our focus in this poster: separating congestion control feature and make it composable
- Motivation: where can this be useful?
  - Packet losses due to reasons other than network congestion
  - Fair access to network resources achieved through other means
    - Example: wireless networks
    - Losses due to random channel errors
    - Fairness handled at the link layer

System Emulation
- Measurement approach: emulation instead of simulation
- Measurement Setup:
  - Use Emulab environment [Emu]
  - Consists of four hosts, two routers, and a shared link
  - Loss rates vary from p=0 to p=0.2, uniformly distributed
  - Four variants of TCP compiled in kernel and loaded onto hosts
  - Data collected using tcpdump, analyzed with Ethereal [Eth]

Experiment
- Modification based on TCP New Reno in FreeBSD 4.7
  - Real system, instead of specialized frameworks, APIs
  - Around 10K LOC with comments
  - Congestion Control separated using C preprocessors
  - Challenge: code spread out in several files
- In the modified TCP: amounts to 250 LOC
- Also investigate the use of SACK: 1K LOC

Results
- Four different flavors of TCP were composed:
  1. TCP New Reno without SACK, BASE
  2. TCP New Reno with SACK, SACKBASE
  3. TCP New Reno with neither of congestion control nor SACK, EXP
  4. TCP New Reno with SACK but without Congestion Control, SACKEXP

Table 1: Summary of emulation results. Goodput of the experimental TCP compared to TCP New Reno. Efficiency η = goodput/throughput.

<table>
<thead>
<tr>
<th>Flavors</th>
<th>PLR=10^-3</th>
<th>PLR=10^-5</th>
<th>Efficiency</th>
<th>At 10% loss rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SACKBASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SACKEXP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References/Related Work
- [Hutchinson98] K. Kerner-Manor et al., protocol layers
- [Lan95] Ensemble, Li et al., routing protocol composed from micro-layers
- [Kerner99] TCP in FreeBSD: Kehler et al., a modular TCP in a specialized programming language
- [Saltzer84] Jerome H. Saltzer, David P. Reed, and David D. Clark, End-to-End arguments for a reliable protocols, 1984
- [RamFloyd99] Shawn Ostermann. tcptrace tool.
- [Tcptrace] Shawn Ostermann. tcptrace tool.
- [RP-CDMA] Kempter, R., PhD Dissertation Proposal, available upon request

Conclusion
- W/o congestion control and w/o SACK: compare EXP to BASE
  - Goodput difference (111%-280%): bigger sender window
  - Efficiency ratio (~<125%): go-back-N wastes bandwidth
- W/o congestion control but w/ SACK: compare SACKEXP to SACKBASE
  - Goodput difference (184%-728%): much higher ratio than the above case, i.e., w/o SACK
  - Efficiency ratio (65%-99%): Compare SACKEXP to EXP, SACK greatly improves efficiency for TCP w/o congestion control
- At PLR=10^-2, SACKEXP achieves 91% of lossless link, SACKBASE drops to 60%.
  - PLR=10^-2 is achievable with Random Channel Access up to high loads (BP-COMA)
- We suggest to apply SACKEXP to wireless environments

Future Work
- Error control feature: our initial efforts indicate that separating this feature is likely to be much harder (c.f. Kehler02)
- Congestion Control w/ ECN: investigate the performance of SACKEXP w/ congestion control based on ECN bit [RamFloyd99]
- Comparison of performance of SACKEXP and TCP that can resort to a Link Layer retransmission scheme

Author and Contact Information
- Student, Dept. of Electrical and Computer Engineering, kempter@eng.utah.edu
- Student, School of Computing, xinb@cs.utah.edu
- Assistant Professor, School of Computing, kasera@cs.utah.edu
- Professor, Core Chair, Director HCDC Laboratory
  - Electrical & Computer Engineering Research Facility, Rm ECEB 2W-124 University of Alberta, Edmonton, Canada T6G 2V4