



Contribution ID: 103

Type: **not specified**

Tuning Linux TCP for data-center networks

Monday, September 12, 2022 12:00 PM (30 minutes)

For better or worse, TCP remains the main transport of many hyperscale data-center networks. Optimizing TCP has been a hot topic in both academic research and industry R&D. However individual research paper often focuses on solving a specific problem (e.g. congestion control for data-center incast) and the industry solutions are often not public or generically applicable. Since Linux TCP default configurations are more or less tuned for wide-area Internet settings, it's not easy to tune Linux TCP for low-latency data-center environments. For example, simply changing the congestion control to the well-known "dctcp" congestion control may not fully deliver all the benefits Linux TCP can provide.

In this talk, we'd like to share our knowledge and best practices after a decade-long experience of tuning TCP for data-center networks and applications, covering congestion control, protocol and IO enhancements. We will discuss the trade-offs among latency, utilization, CPU, memory, and complexity. In addition we'll present the inexpensive instrumentation to trace application frame-aware latency beyond general flow level statistics. It's worth emphasizing that the goal is not to promote the authors' own works but to help promote interesting discussions with other data center networking developers and guide new comers. After the meeting we hope to synthesize our recommendations into Documentation/networking/tcp_datacenter.txt

I agree to abide by the anti-harassment policy

Yes

Primary authors: CHENG, Yuchung (Google); CARDWELL, Neal (Google)

Presenter: CHENG, Yuchung (Google)

Session Classification: eBPF & Networking

Track Classification: eBPF & Networking Track