

IETFのTCP関連技術動向

Yoshifumi Nishida

nishida@sfc.wide.ad.jp

Who Am I?

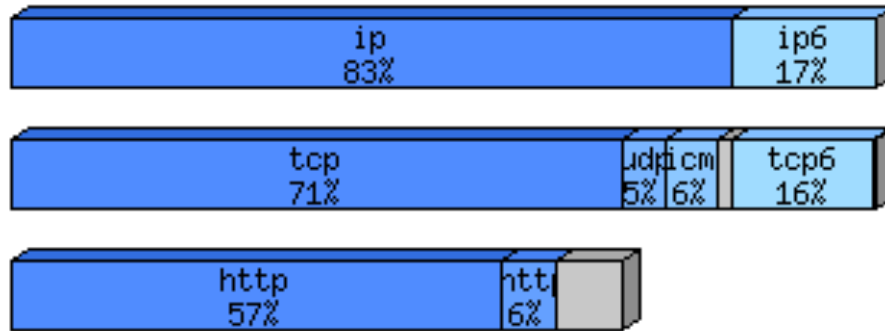
- Name: Yoshifumi Nishida
- Current Job: Senior Researcher at GE Global Research
- IETF Activities
 - Co-chair of TCPM Working Group
 - Co-chair of Multipath TCP Working Group
 - Transport Area Directorate

Agenda

- Why TCP?
- TCP Extensions in IETF
- Middlebox Issues for extending TCP
- TCP in the future

Why TCP? (1)

- It's still major traffic on the Internet



From: <http://mawi.wide.ad.jp/mawi/samplepoint-F/2013/201305121400.html>

– Why?

- Many apps still use TCP (e.g. web, e-mails)
 - UDP doesn't have much features (CC, reliability, etc)
- Many middleboxes (NAT) don't support new transport protocols such as SCTP, DCCP
 - New protocols cannot get through some parts of the Internet

Why TCP? (2)

- But, we have been using TCP for over 30 years. Do we still need to do something?
 - Short answer: Yes
 - More throughput
 - More quick response
 - More secure (encryption, authentication, DoS protection, etc)
 - More flexible transmission (multipath, mobility, etc)
- As long as Internet is growing and changing, We need to extend TCP

TCP extensions in IETF

- TCP Fast Open, IW10
- Multipath TCP
- TCPInc
- TCP Option space extension

TCP Fast Open, IW10

- Proposals from Google
 - Already implemented in Linux and used in google's servers
- TCP Fast Open (RFC7413)
 - Send data in SYN packet
 - Eliminate delay caused by 3 way handshake in TCP
- IW10 (RFC6928)
 - Increase initial window size of TCP up to 10 segments (previously 4 segments by RFC3390)
 - Mitigate delay caused by slow-start
 - Finish many connections in 1 RTT

TCPInc

- TCP Increased Security WG
 - Created 7/2014
- Provide security at TCP layer
 - Make all TCP traffic encrypted
 - Protect non-IPsec, non-SSL, non-TSL traffic
- Still under lots of discussions
 - Too much energies?

Multipath TCP

- Multipath TCP WG
 - Created 10/2009
- Allow TCP to use multiple IP addresses simultaneously
 - Utilize two links (e.g. 3G, Wifi) to increase throughput
 - Enhance error resiliency
 - Mobility supports
- Main spec is published
 - RFC6824
- Apple, Oracle starts supporting and using
 - Linux, FreeBSD code is also available

Extending TCP is not very easy

- TCP needs TCP options to use new features
 - Need to negotiate them before using
 - Put TCP options in TCP header in SYN packets
 - Sender TCP: “I want to use this feature, is it OK?”
 - Receiver TCP: “OK. I support the feature, so you can use it”
 - But, TCP’s Option space is limited to just 40 bytes!
 - TCP cannot carry many options!
 - When we discuss new features, option size is a typical topic
 - Small size is preferable, but need to carry many info for efficiency

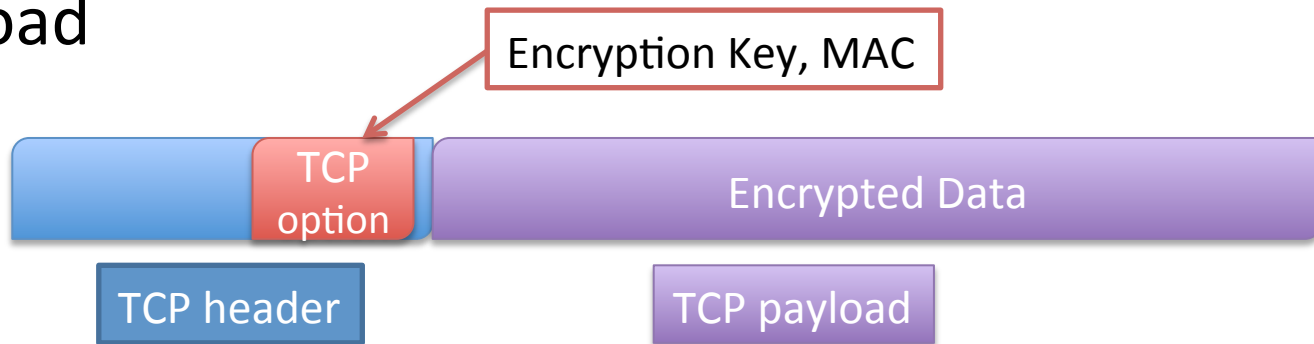


Middlebox Issue

- Middlebox causes issues in TCP's new features
 - Problematic TCP option handlings
 - Remove, duplicate, aggregate, shuffle options
 - Problematic TCP segment handlings
 - Coalesce, Split segments, Insert some data into segments
- Conventional TCP traffic (less options) can work, so the problem tends to be ignored

An example for middlebox issue

- Let's say, we have a new TCP option that encrypts payload



- If this segment is split into two segments..
 - Both packets cannot be decrypted



HOPS

- How Ossified is the Protocol Stack?
 - Research Group in IRTF
 - Measure and analyze middlebox behavior in the current Internet
 - Provide tangible data for protocol designs

Another issue: BufferBloat

- Identified by Jim Getty around 2009
- TCP Connections becomes really slow (e.g more than 10 secs RTT)
 - Caused by large amount of buffer at some devices (Cable modem, STB, AP router, etc)
 - Increasing buffer size sometime has adverse effects!

VJ's fountain analogy

- Don't use car analogy



- Use fountain for TCP's queuing model
 - Queue is water level of the pond in fountain
 - Water level has nothing to do with pump pressure, pump size, flow rate
 - Increasing queue length won't solve congestion at all
- New Queuing Scheme: CoDel

AQM

- Active Queue Management and Packet Scheduling WG
 - Created around 9/2013
 - Standardize CoDel and other Active Queuing Management schemes
 - Many AQM algorithms are proposed in academic, but they are not widely used. AQM WG will provide recommendations or guidelines to change this situation

Saving TCP's future

- Keep a TCP segment as it is when you forward it
- Copy TCP options as much as possible when you proxy a TCP connection
- When re-segmentation is required in middlebox, publish docs so that protocol designers can understand/predict its behavior
- Don't increase buffer size blindly in middleboxes even though memory cost is getting cheaper



Thank you so much!