Decoupling TCP from IP with Multipath TCP

Olivier Bonaventure
http://inl.info.ucl.ac.be
http://perso.uclouvain.be/olivier.bonaventure

Thanks to Sébastien Barré, Christoph Paasch, Grégory Detal, Mark Handley, Costin Raiciu, Alan Ford, Michio Honda, Fabien Duchene and many others

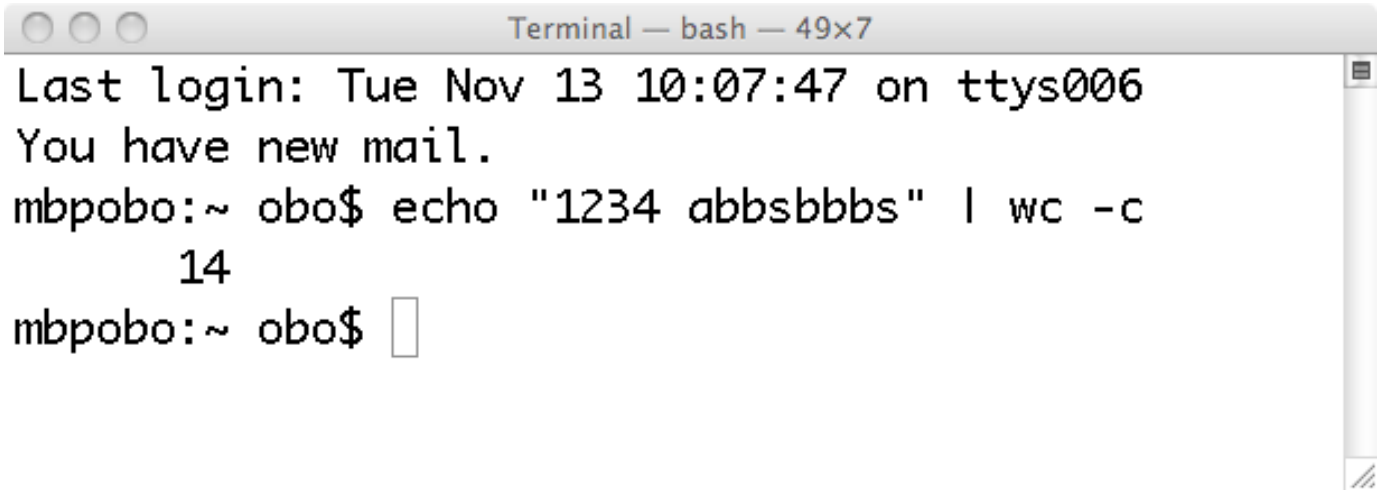
April 2013

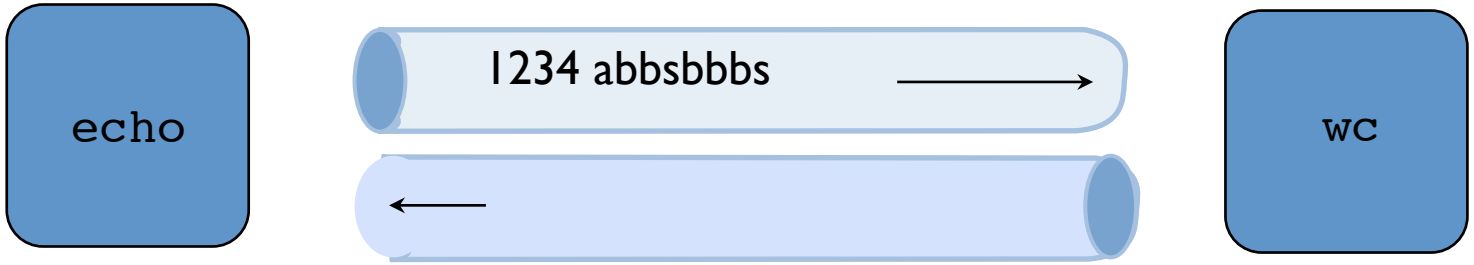
Agenda

The motivations for Multipath TCP

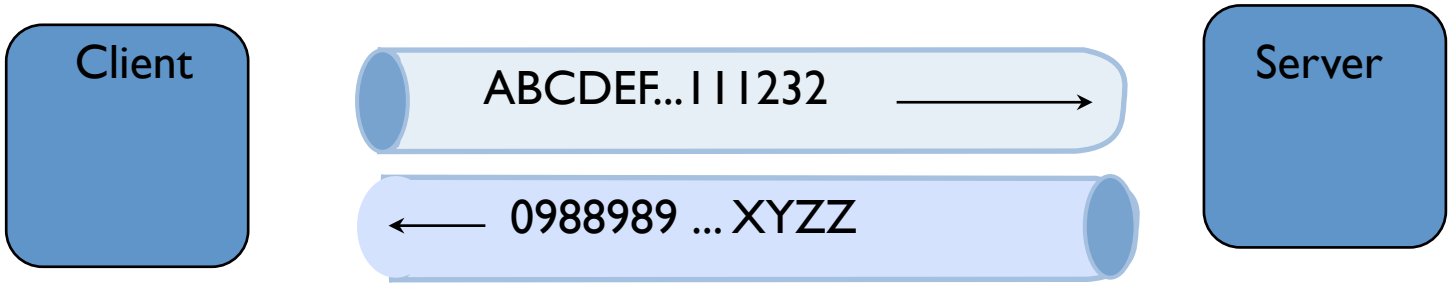
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases

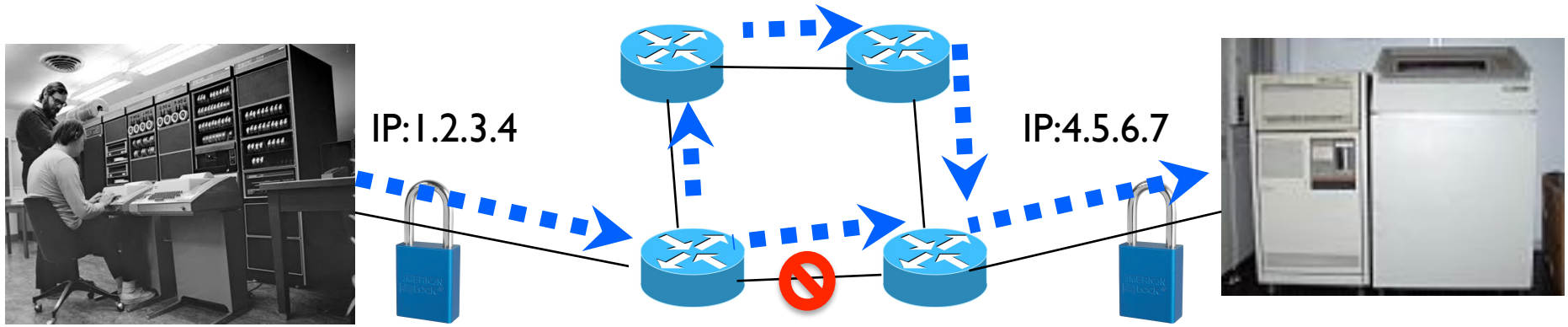
The Unix `pipe` model





The TCP bytestream model



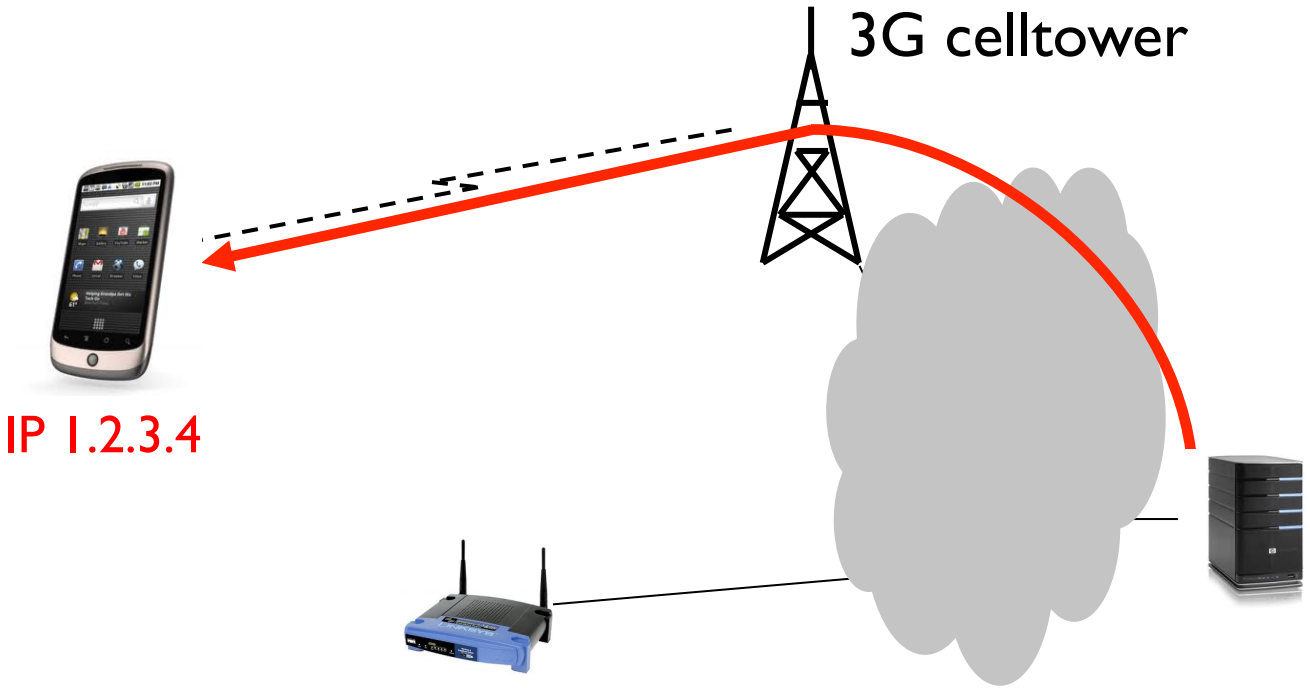


Endhosts have evolved

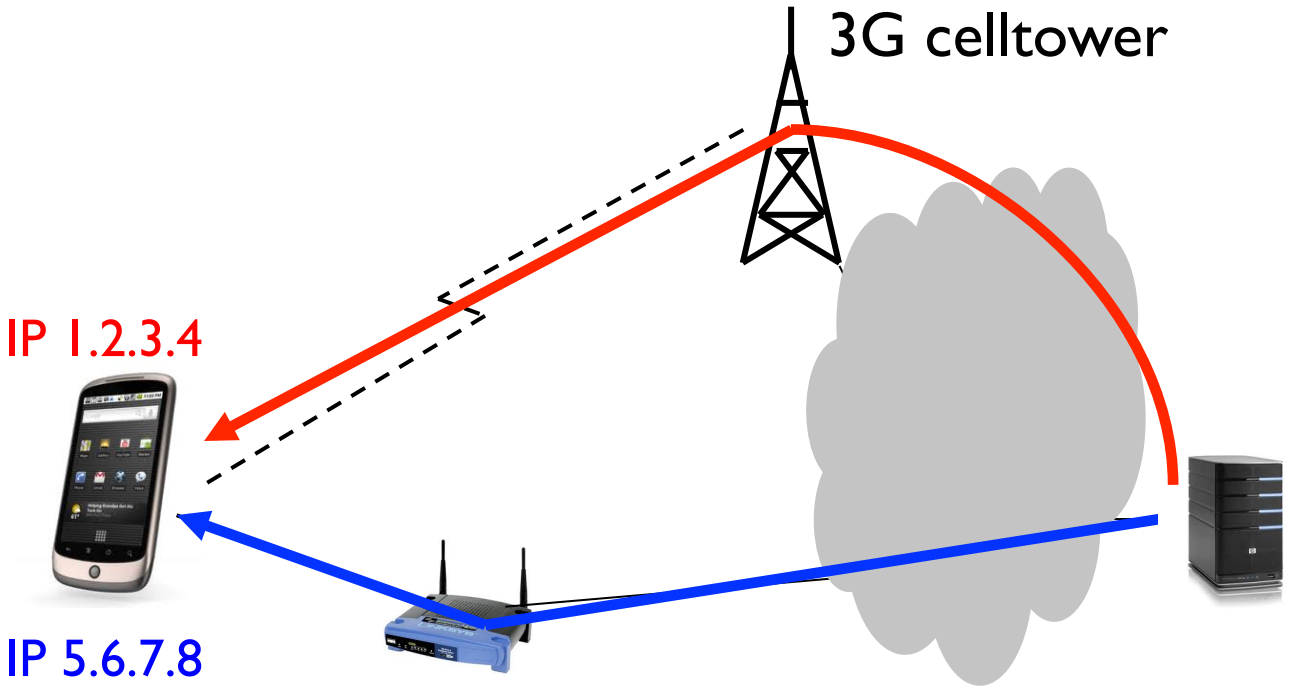
Mobile devices have multiple wireless interfaces

User expectations

What technology provides



What technology provides



When IP addresses change TCP connections have to be re-established !

Datacenters

Equal Cost Multipath

ECMP implementation

Packet arrival :

Hash(IP_{src}, IP_{dst}, Prot, $Port_{src}$, $Port_{dst}$) mod #oif

Packets from one TCP connection follow same path

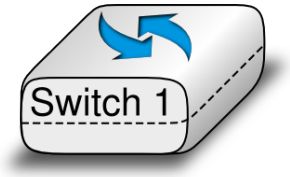
Different TCP connections follow different paths

Agenda

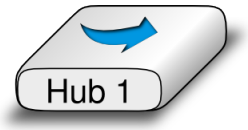
- The motivations for Multipath TCP
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases

The Internet architecture that we explain to our students

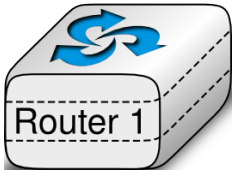
Application
Transport
Network
Datalink
Physical



Datalink
Physical



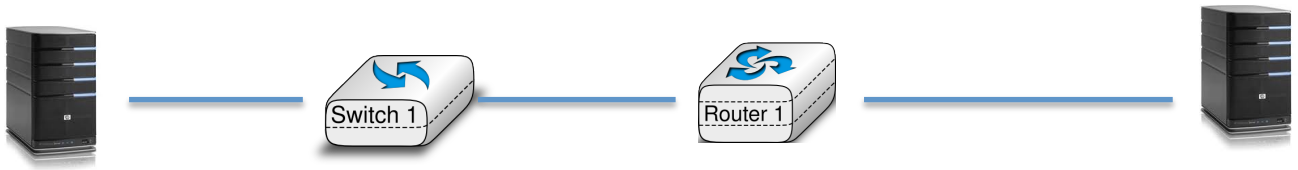
Physical

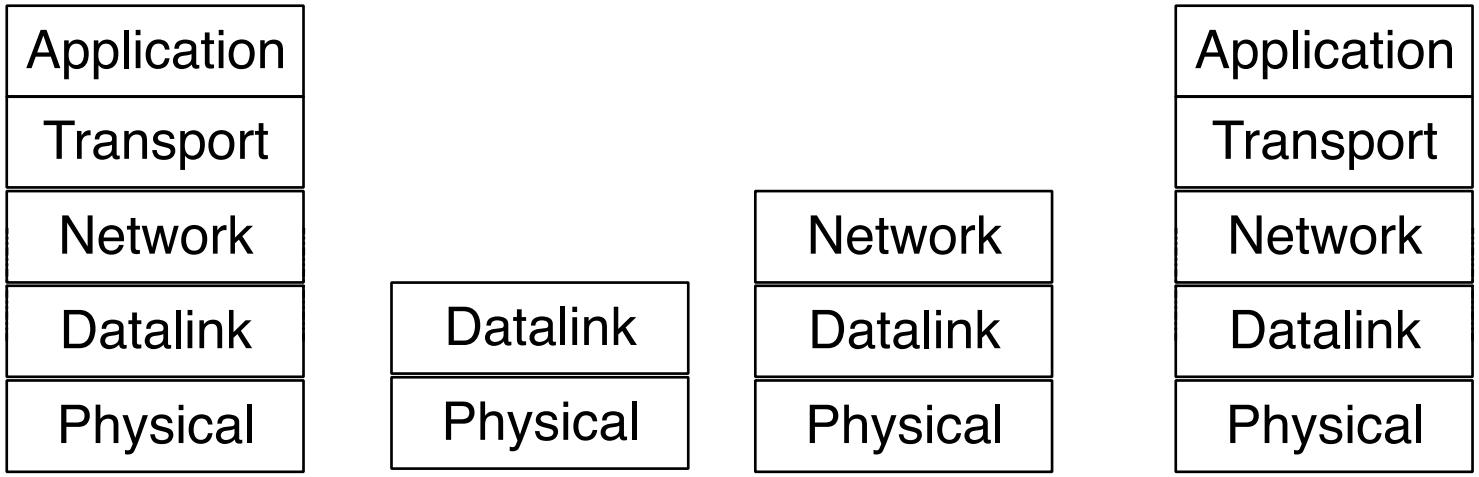


Network
Datalink
Physical

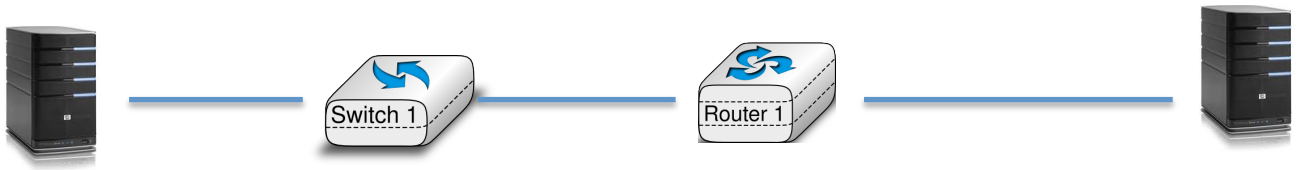
O. Bonaventure, Computer networking : Principles, Protocols and Practice, open ebook, http://inl.info.ucl.ac.be/cnp3

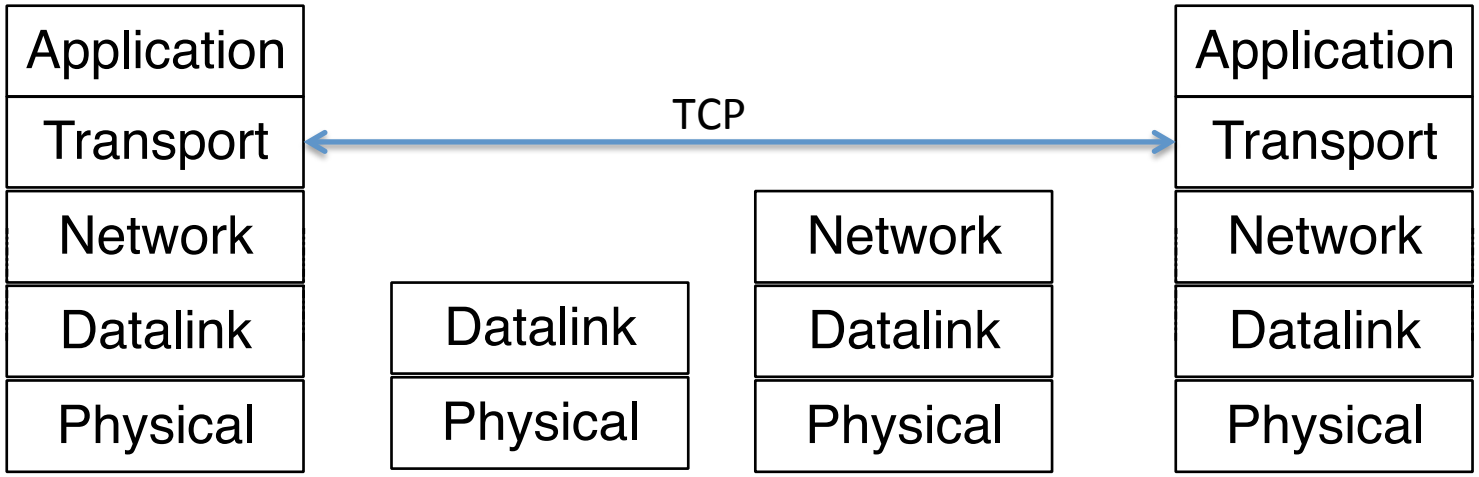
A typical "academic" network





The end-to-end principle





In reality

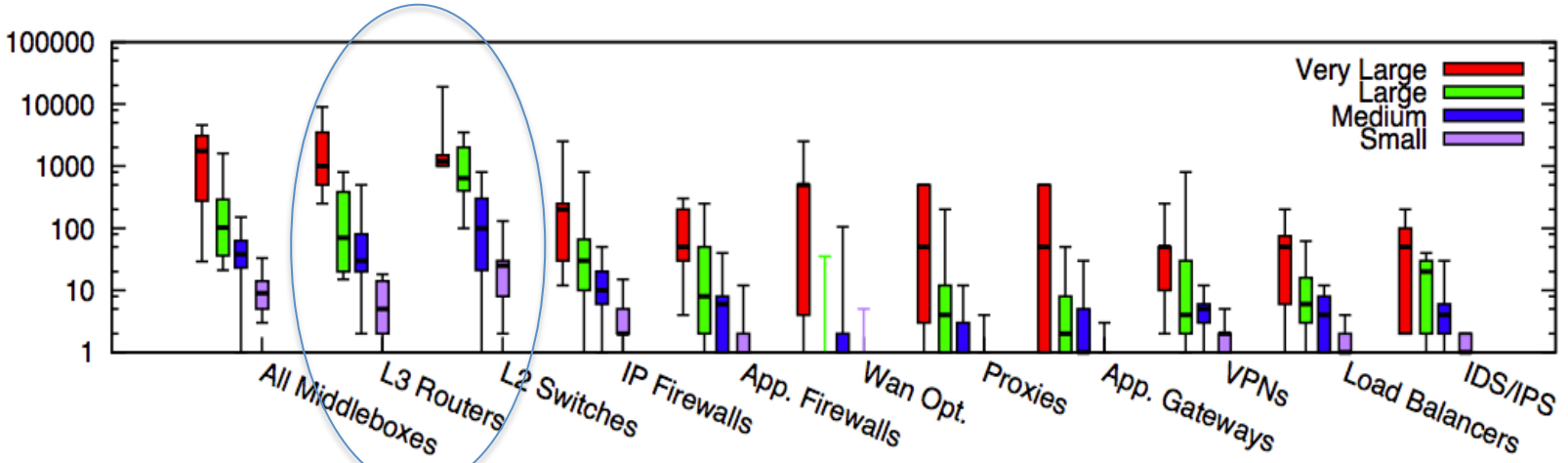


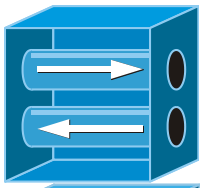
Figure 1: Box plot of middlebox deployments for small (fewer than 1k hosts), medium (1k-10k hosts), large (10k-100k hosts), and very large (more than 100k hosts) enterprise networks. Y-axis is in log scale.

- almost as many middleboxes as routers
- various types of middleboxes are deployed

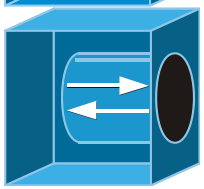
Sherry, Justine, et al. "*Making middleboxes someone else's problem: Network processing as a cloud service*." Proceedings of the ACM SIGCOMM 2012 conference. ACM, 2012.

A middlebox zoo

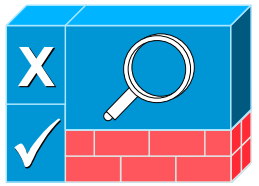
Web Security Appliance



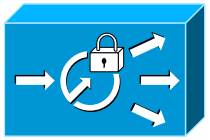
VPN Concentrator



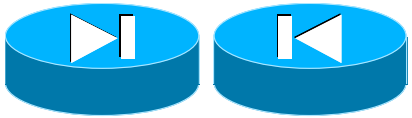
SSL Terminator



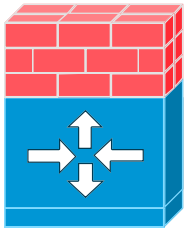
NAC Appliance



ACE XML Gateway

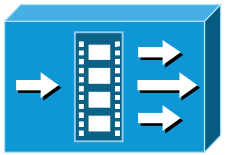


PIX Firewall Right and Left



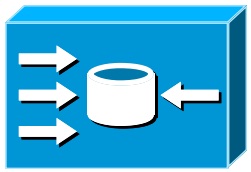
Cisco IOS Firewall

IP Telephony Router



Streamer

Voice Gateway



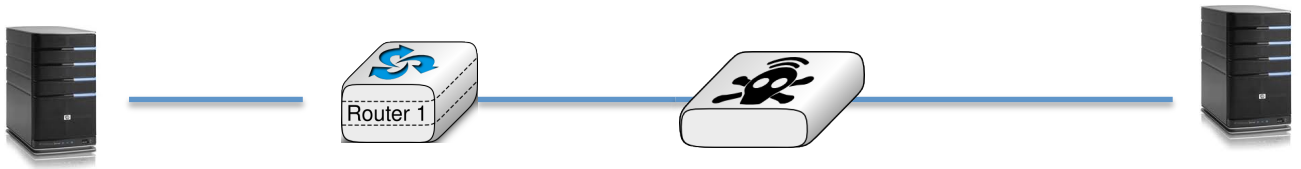
Content Engine

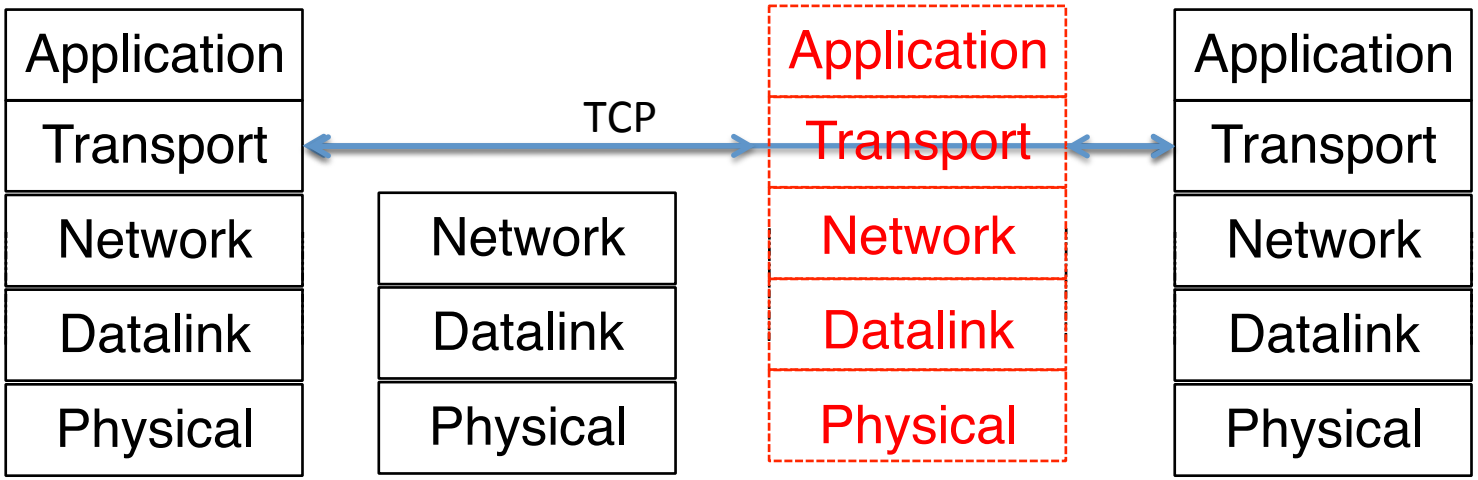
NAT

http://www.cisco.com/web/about/ac50/ac47/2.html

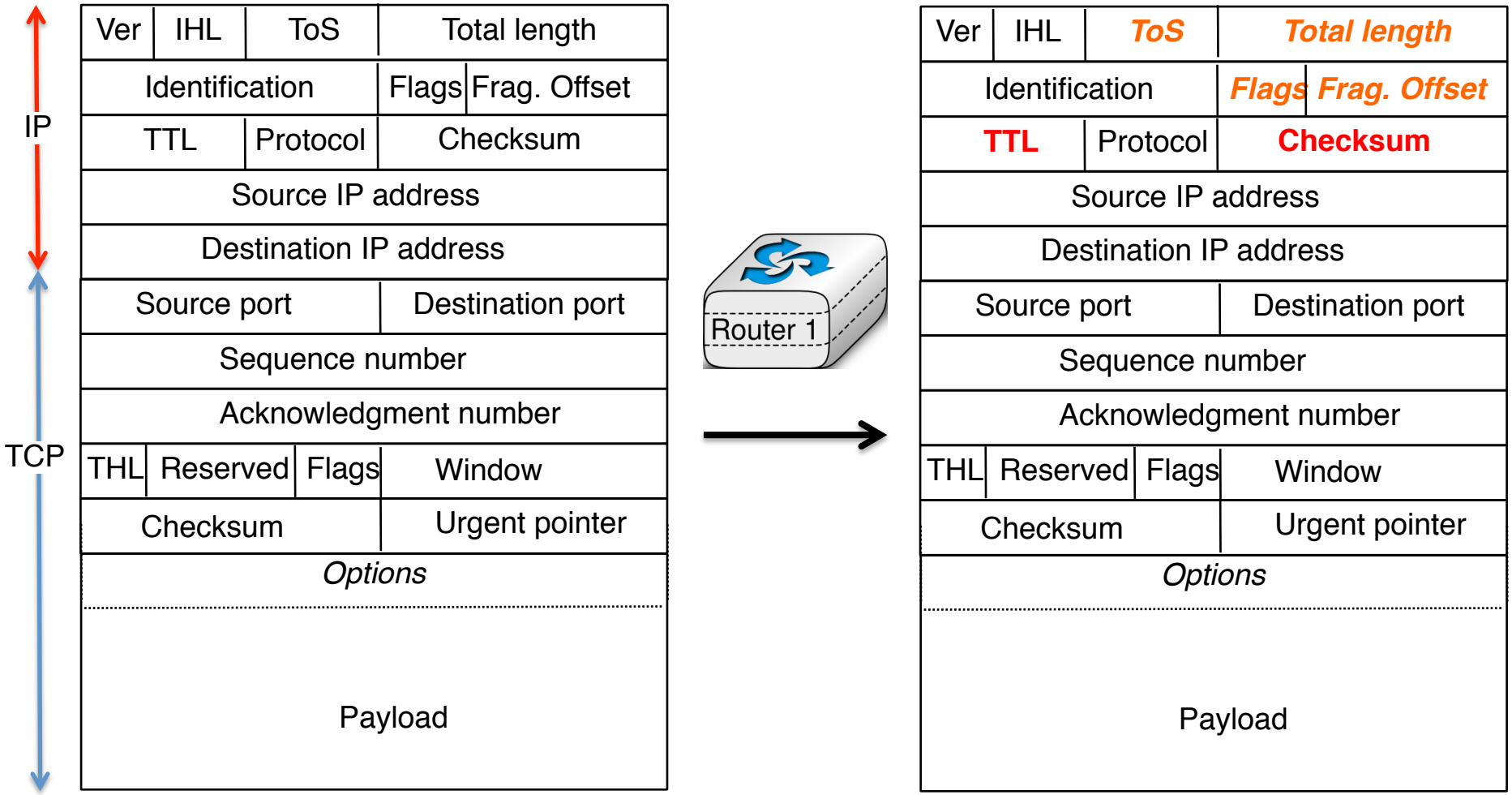
How to model those middleboxes ?

- In the official architecture, they do not exist
- In reality…

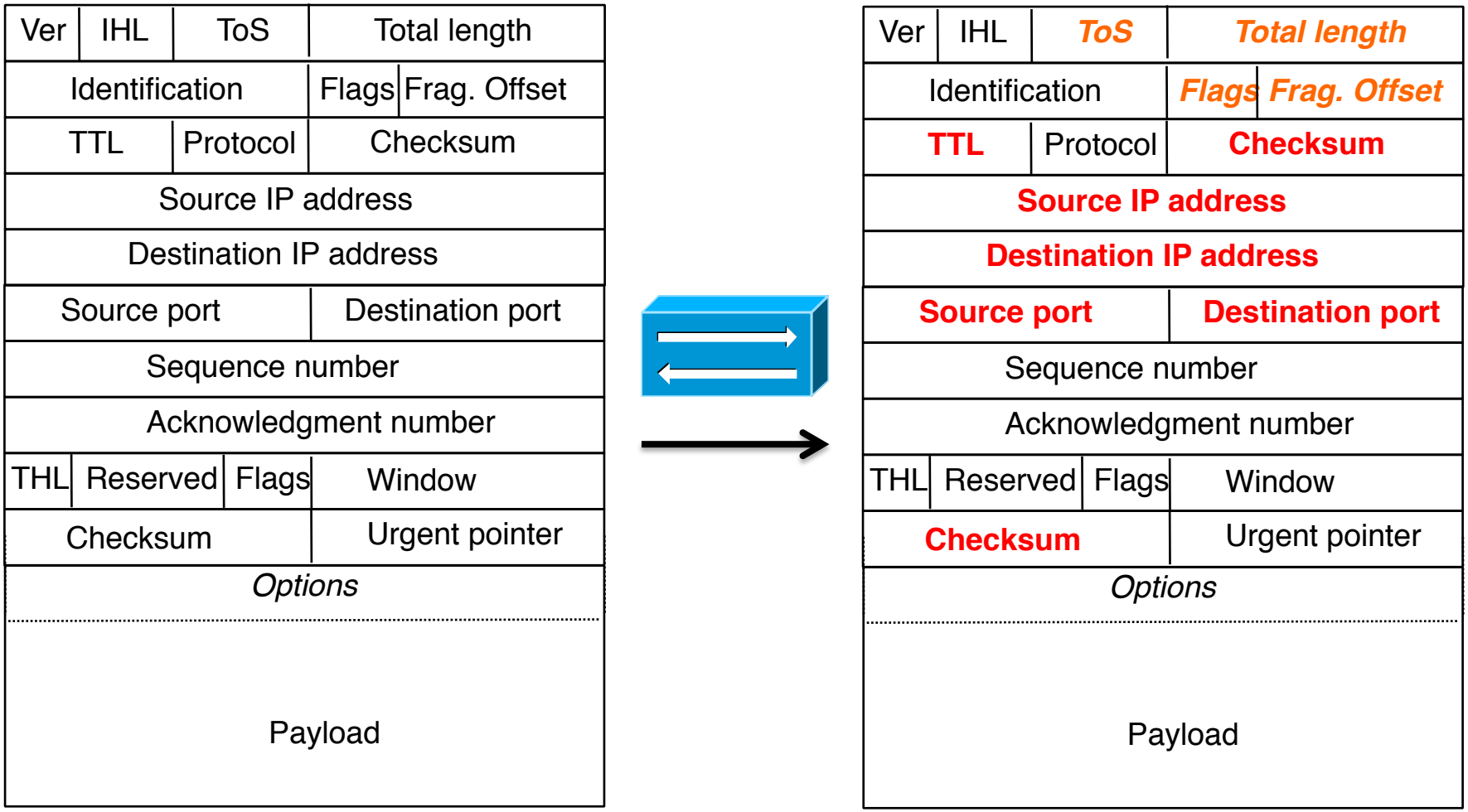




TCP segments processed by a router



TCP segments processed by a NAT



How transparent is the Internet ?

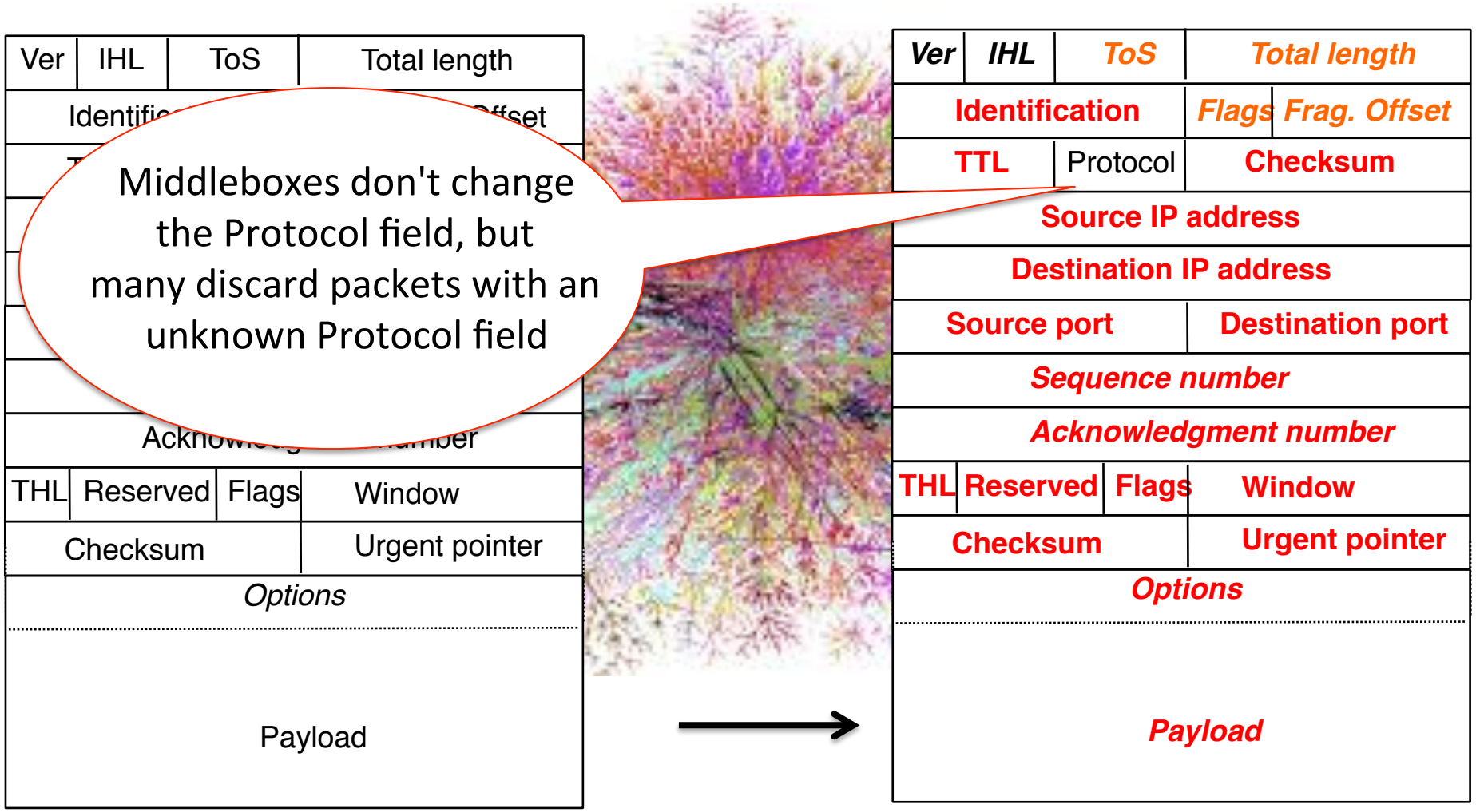
- 25th September 2010 to 30th April 2011
- 142 access networks
- 24 countries
- Sent specific TCP segments from client to a server in Japan

Table 2: Experiment Venues

Country	*Home*	*Hotspot*	*Cellular*	*Univ*	*Ent*	*Hosting*	Total
Australia	0	2	0	0	0	1	3
Austria	0	0	0	0	1	0	1
Belgium	4	0	0	1	0	0	5
Canada	1	0	1	0	1	0	3
Chile	0	0	0	0	1	0	1
China	0	7	0	0	0	0	7
Czech	0	2	0	0	0	0	2
Denmark	0	2	0	0	0	0	2
Finland	1	0	0	3	2	0	6
Germany	3	1	3	4	1	0	12
Greece	2	0	1	0	0	0	3
Indonesia	0	0	0	3	0	0	3
Ireland	0	0	0	0	0	1	1
Italy	1	0	0	0	1	0	2
Japan	19	10	7	3	2	0	41
Romania	1	0	0	0	0	0	1
Russia	0	1	0	0	0	0	1
Spain	0	1	0	1	0	0	2
Sweden	1	0	0	0	0	0	1
Switzerland	2	0	0	0	0	0	2
Thailand	0	0	0	0	2	0	2
U.K.	10	4	4	2	1	1	22
U.S.	3	4	4	0	4	2	17
Vietnam	1	0	0	0	1	0	2
Total	49	34	20	17	17	5	142

Honda, Michio, et al. "*Is it still possible to extend TCP?*" Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

End-to-end transparency today



Agenda

- The motivations for Multipath TCP
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases

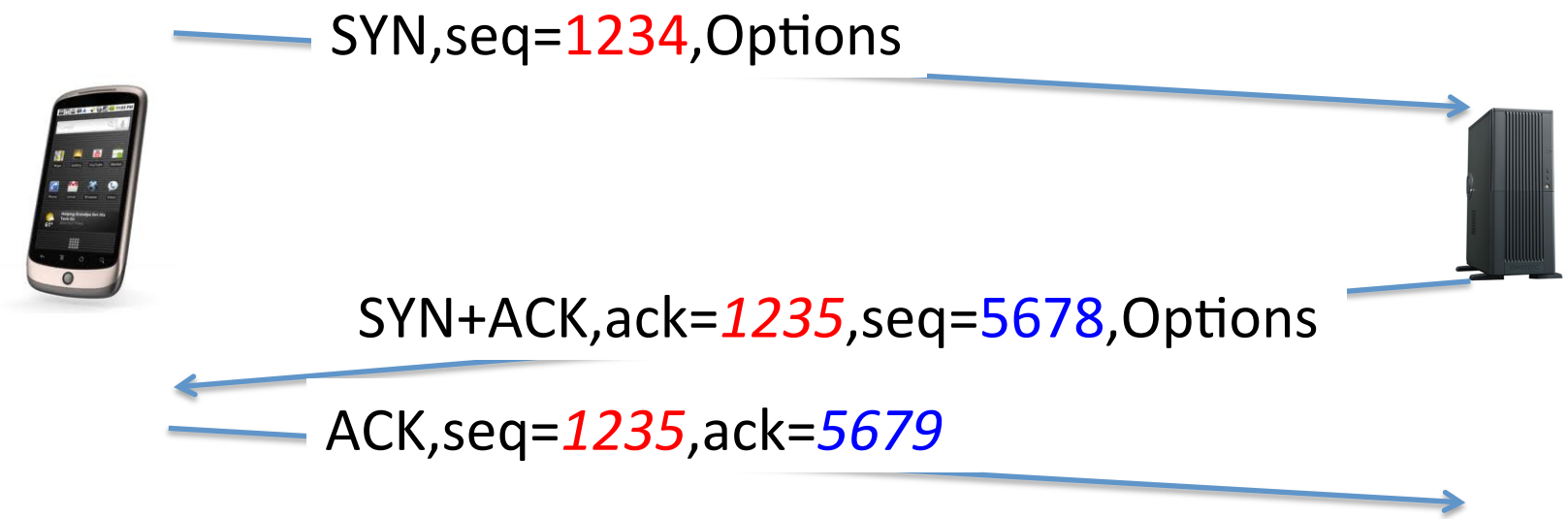
Design objectives

- Multipath TCP is an *evolution* of TCP

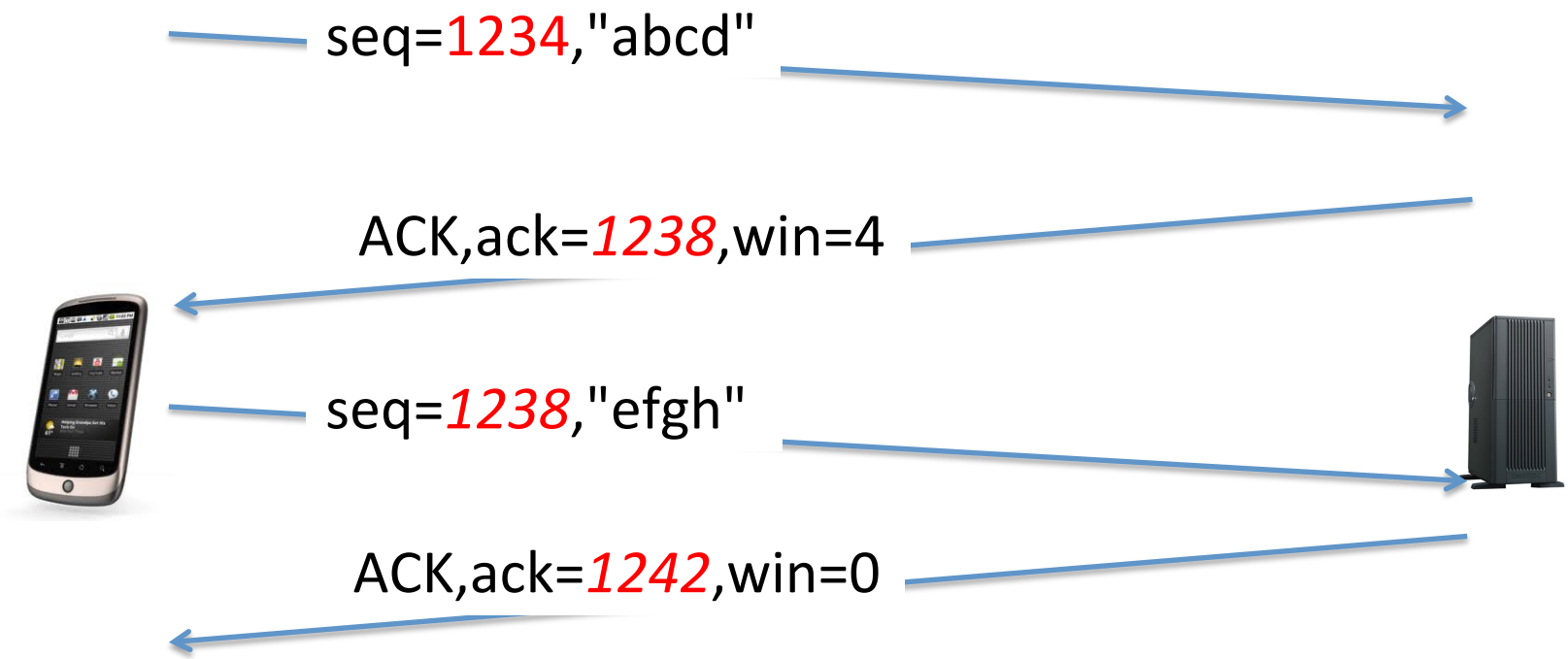
- Design objectives
  - Support unmodified applications
  - Work over today's networks (IPv4 and IPv6)
  - Works in all networks where regular TCP works

TCP Connection establishment

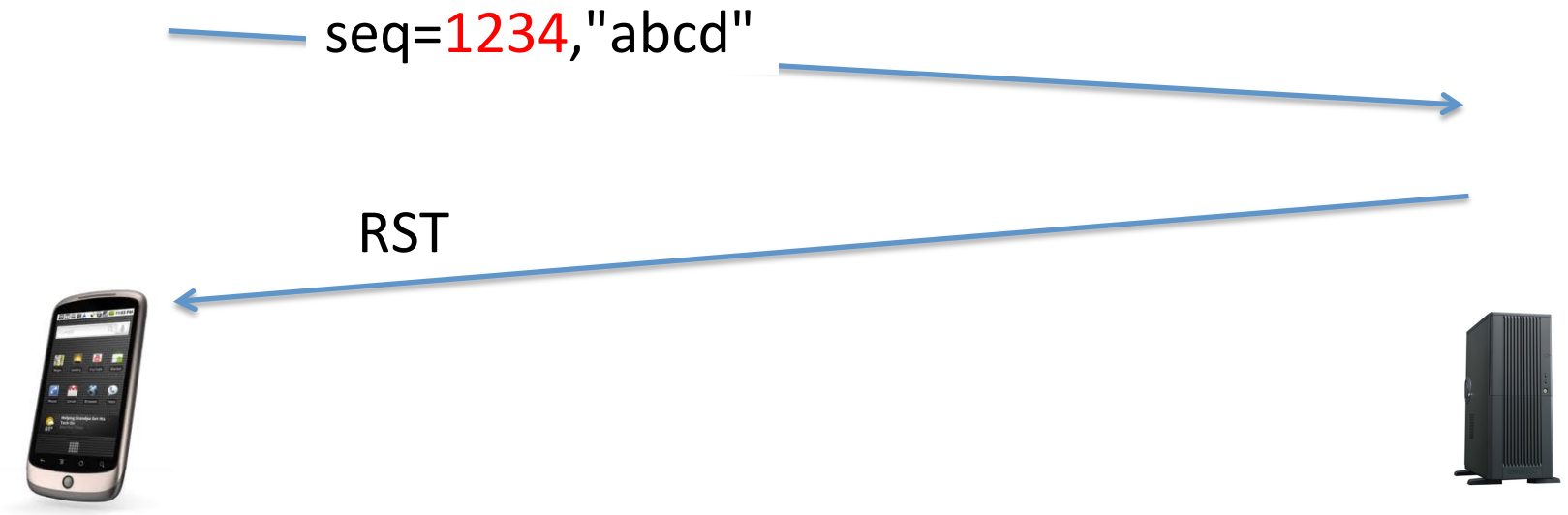
- Three-way handshake



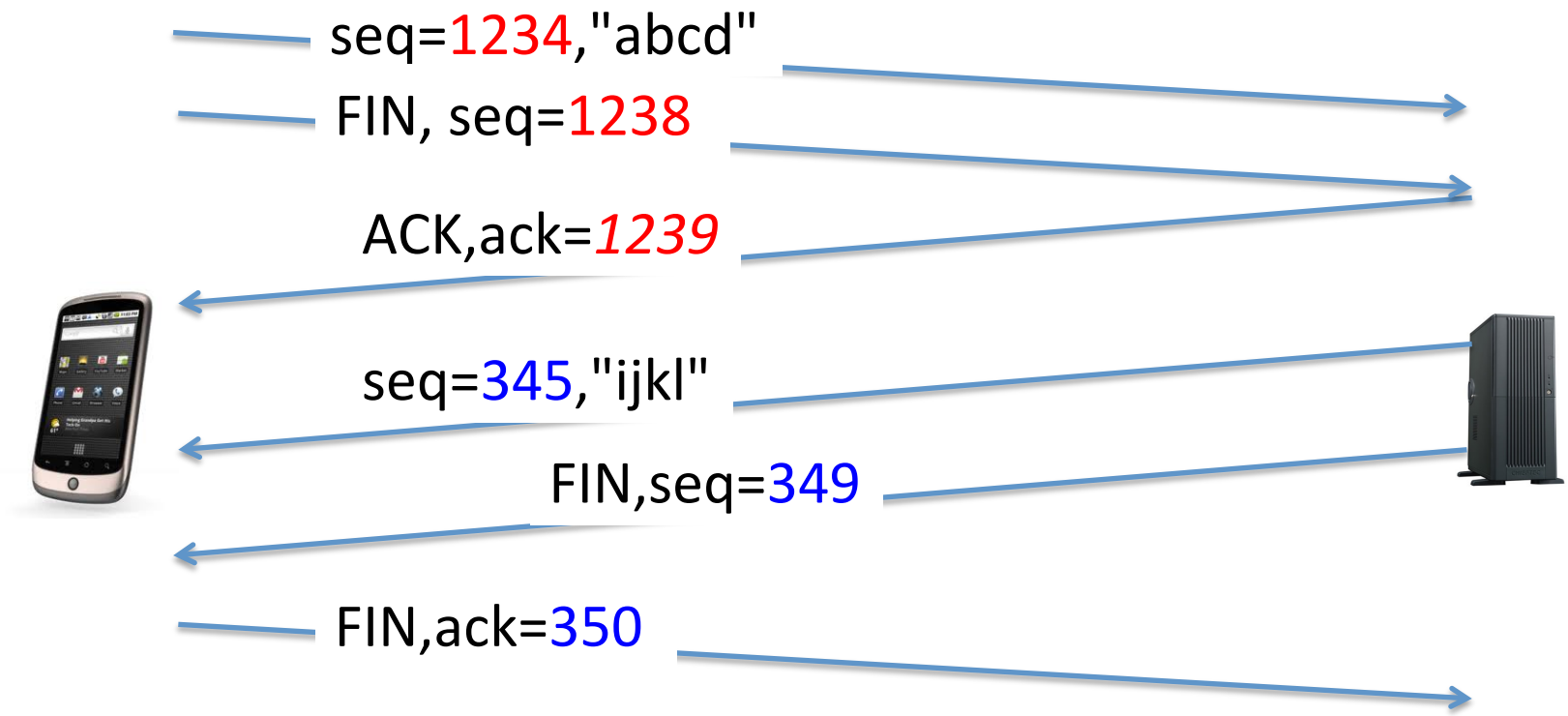
Data transfer



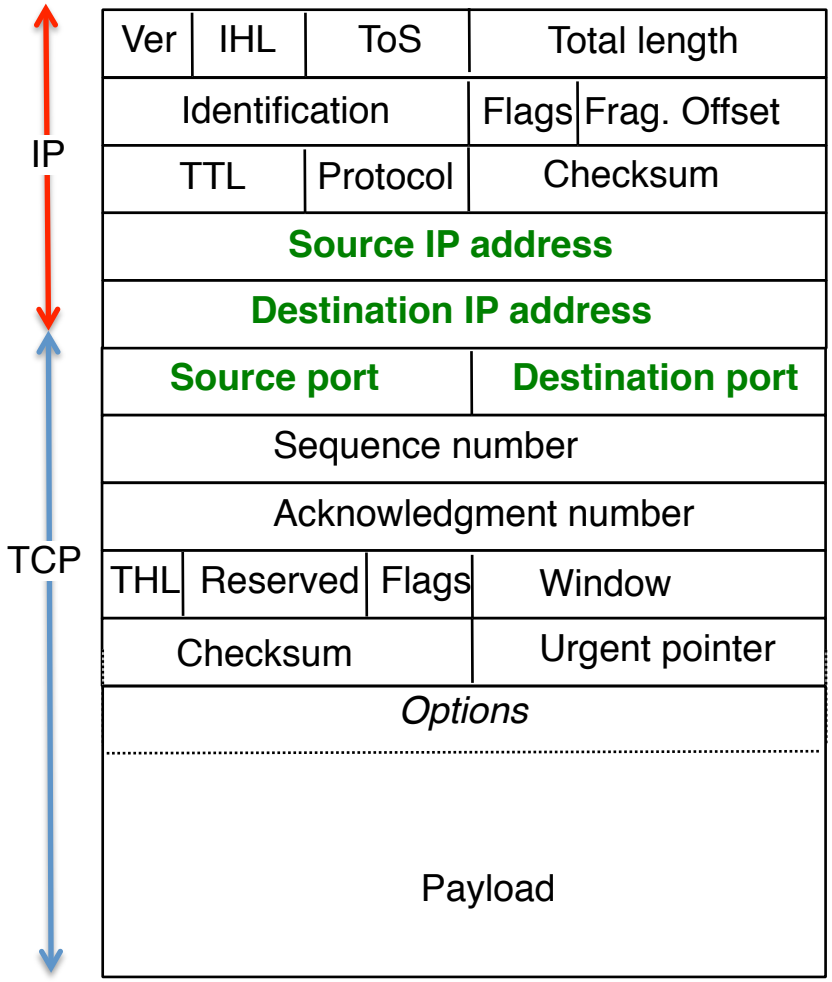
Connection release



Connection release



Identification of a TCP connection

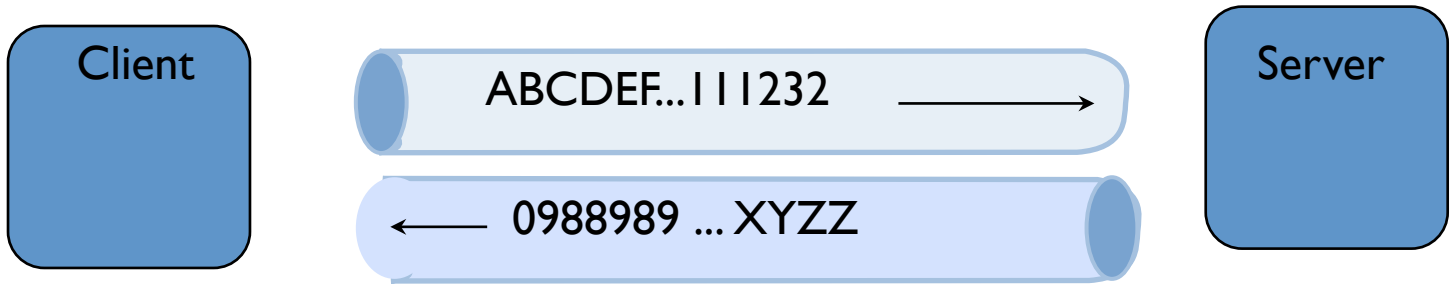


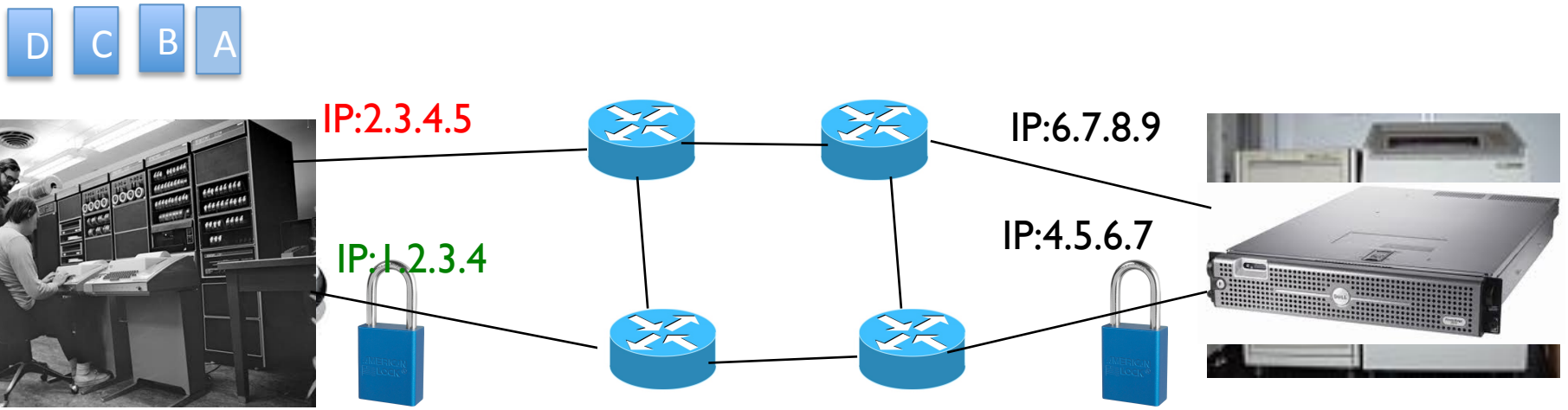
Four tuple

- IP_{source}
- IP_{dest}
- $Port_{source}$
- $Port_{dest}$

All TCP segments contain the four tuple

The *new* bytestream model

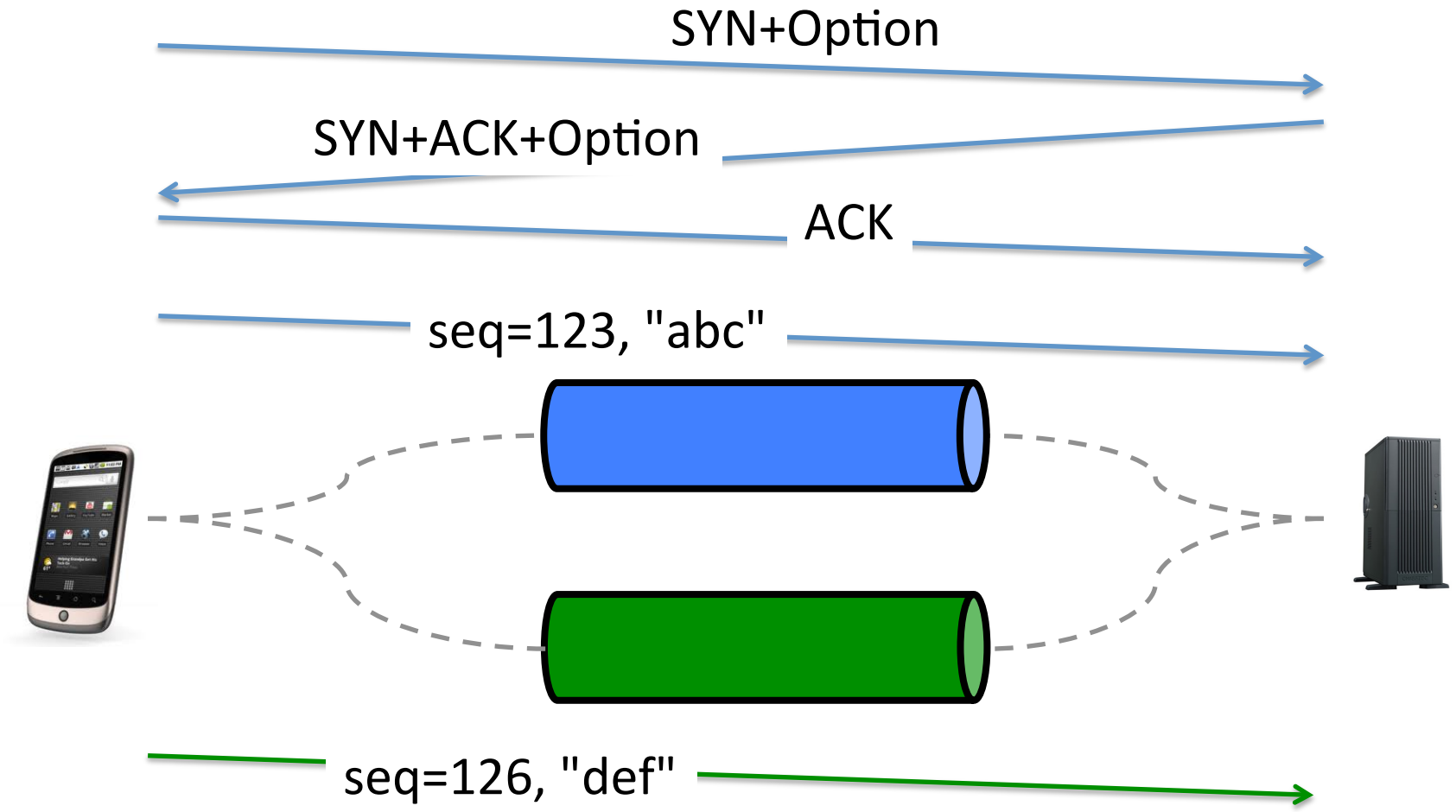




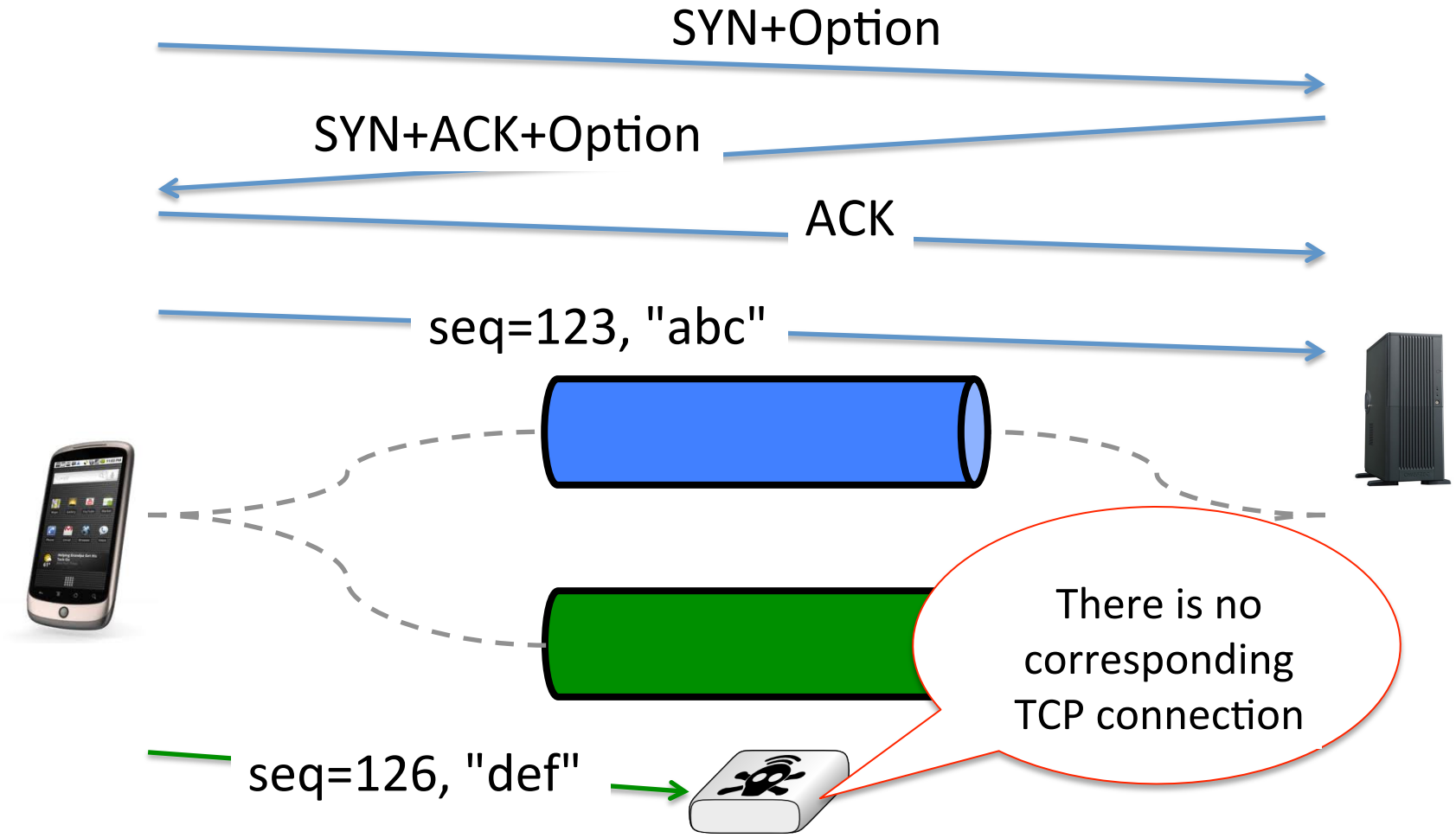
The Multipath TCP protocol

- Control plane
  - How to manage a Multipath TCP connection that uses several paths ?
- Data plane
  - How to transport data ?
- Congestion control
  - How to control congestion over multiple paths ?

A naïve Multipath TCP



A naïve Multipath TCP
In today's Internet ?

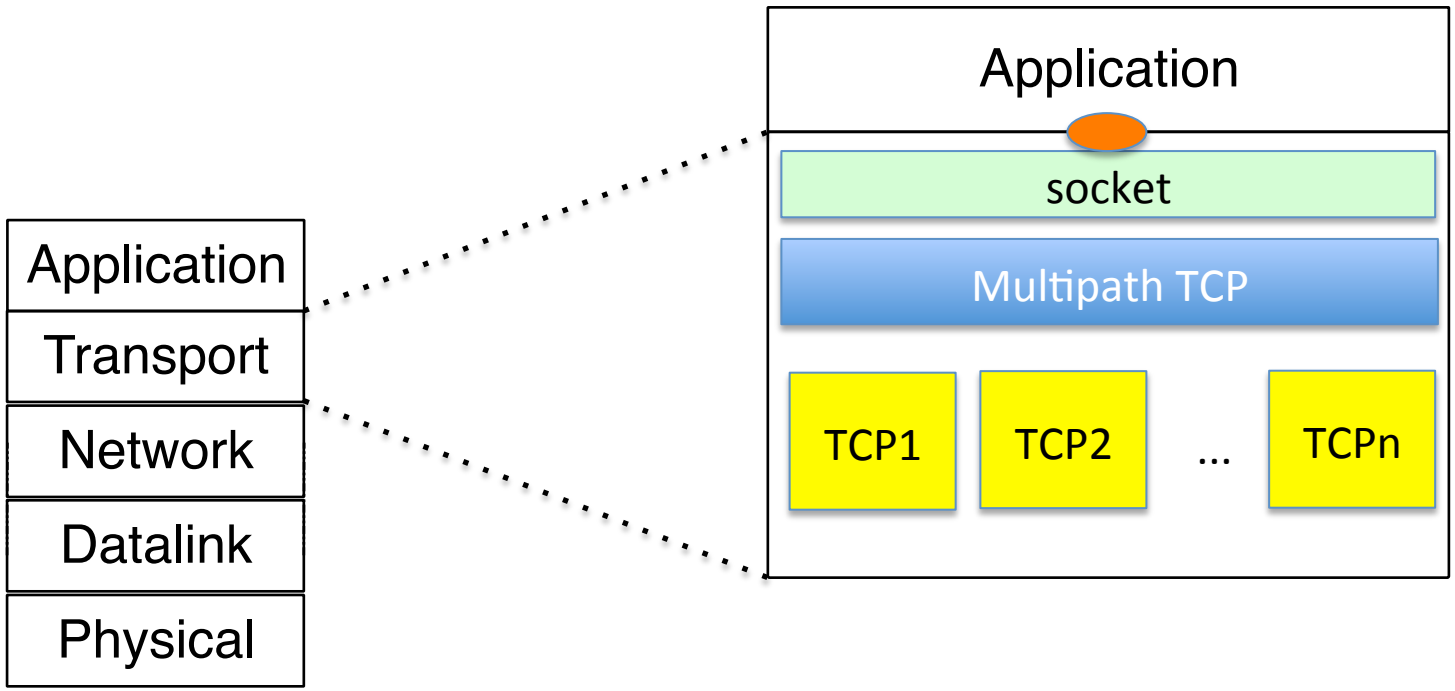


Design decision

– *A Multipath TCP connection is composed of one of more regular TCP subflows that are combined*

- Each host maintains state that glues the TCP subflows that compose a Multipath TCP connection together
- Each TCP subflow is sent over a single path and appears like a **regular TCP** connection along this path

Multipath TCP and the architecture

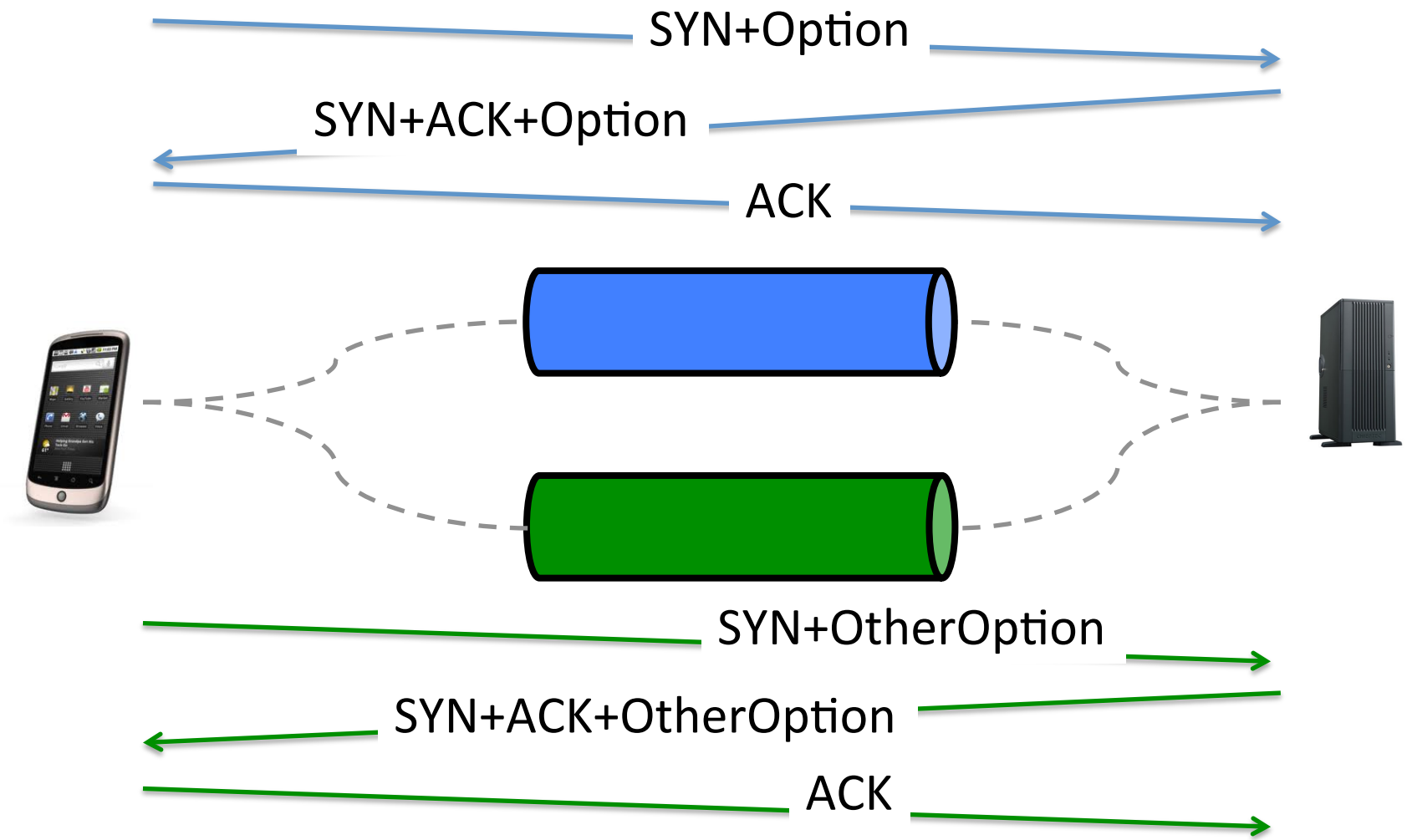


A. Ford, C. Raiciu, M. Handley, S. Barre, and J. Iyengar, “Architectural guidelines for multipath TCP development", RFC6182 2011.

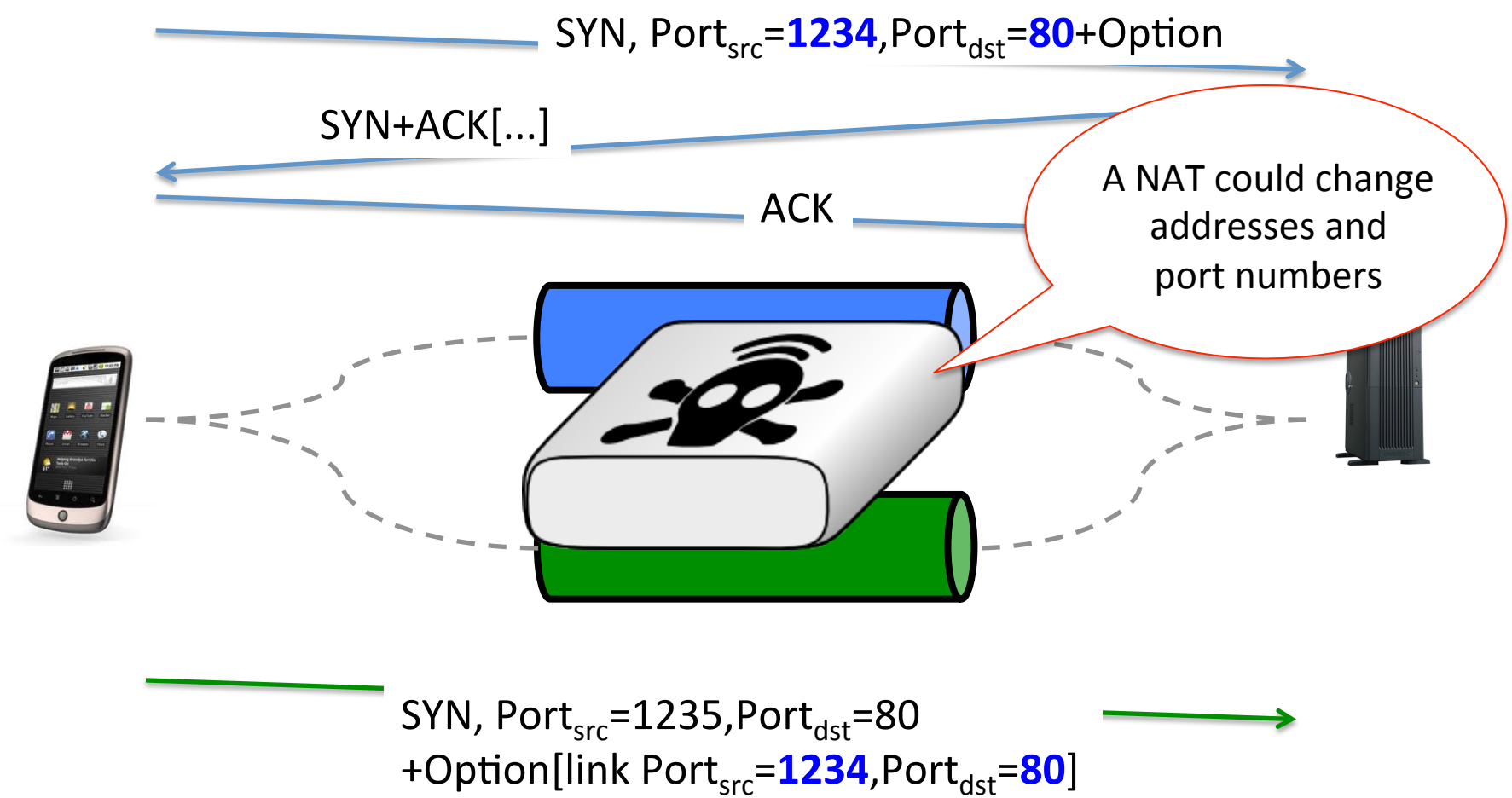
A *regular* TCP connection

- What is a *regular* TCP connection ?
  - It starts with a three-way handshake
    - SYN segments may contain special options
  - All data segments are sent in sequence
    - There is no gap in the sequence numbers
  - It is terminated by using FIN or RST

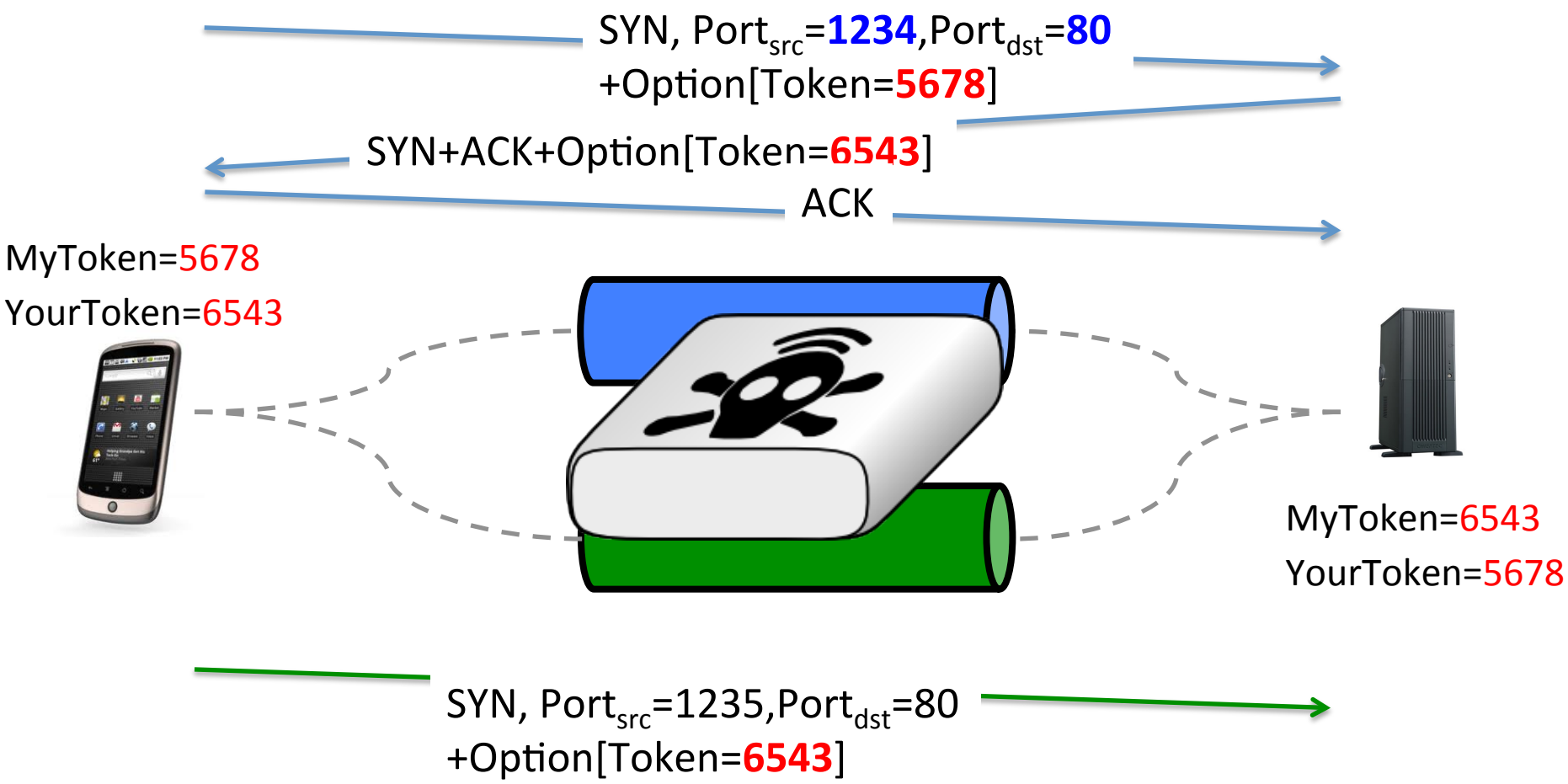
Multipath TCP



How to link TCP subflows ?

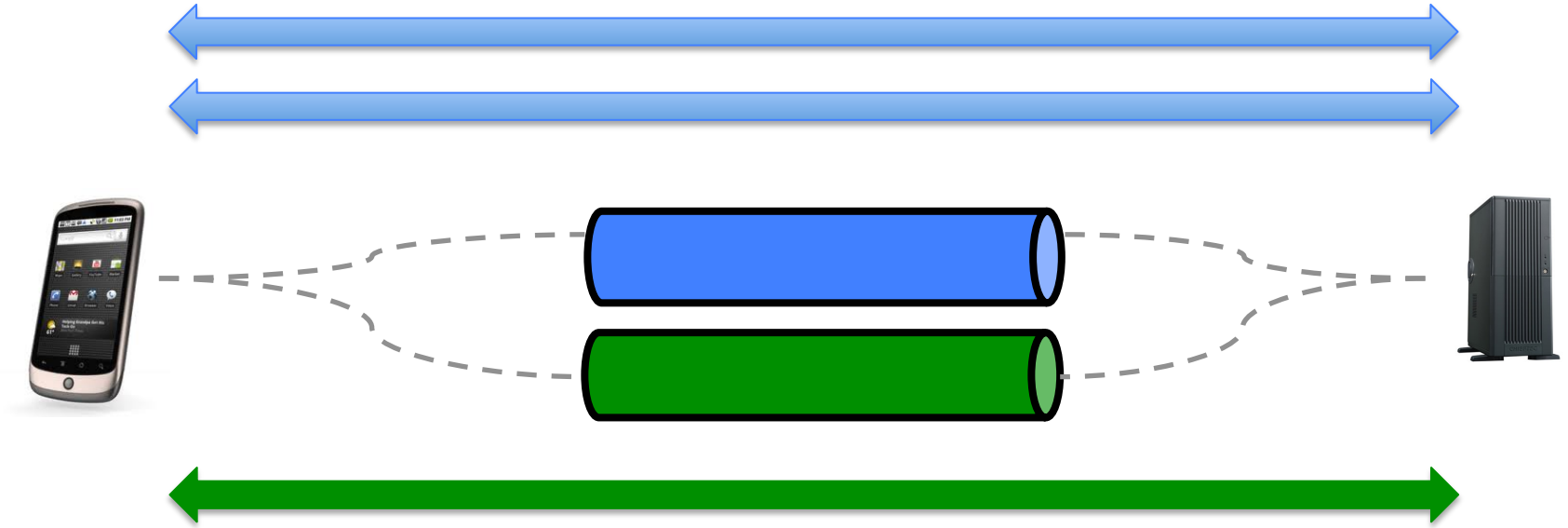


How to link TCP subflows ?



Subflow agility

- Multipath TCP supports
  - addition of subflows
  - removal of subflows



TCP subflows

- Which subflows can be associated to a Multipath TCP connection ?
  - At least one of the elements of the four-tuple needs to differ between two subflows
    - Local IP address
    - Remote IP address
    - Local port
    - Remote port

The Multipath TCP protocol

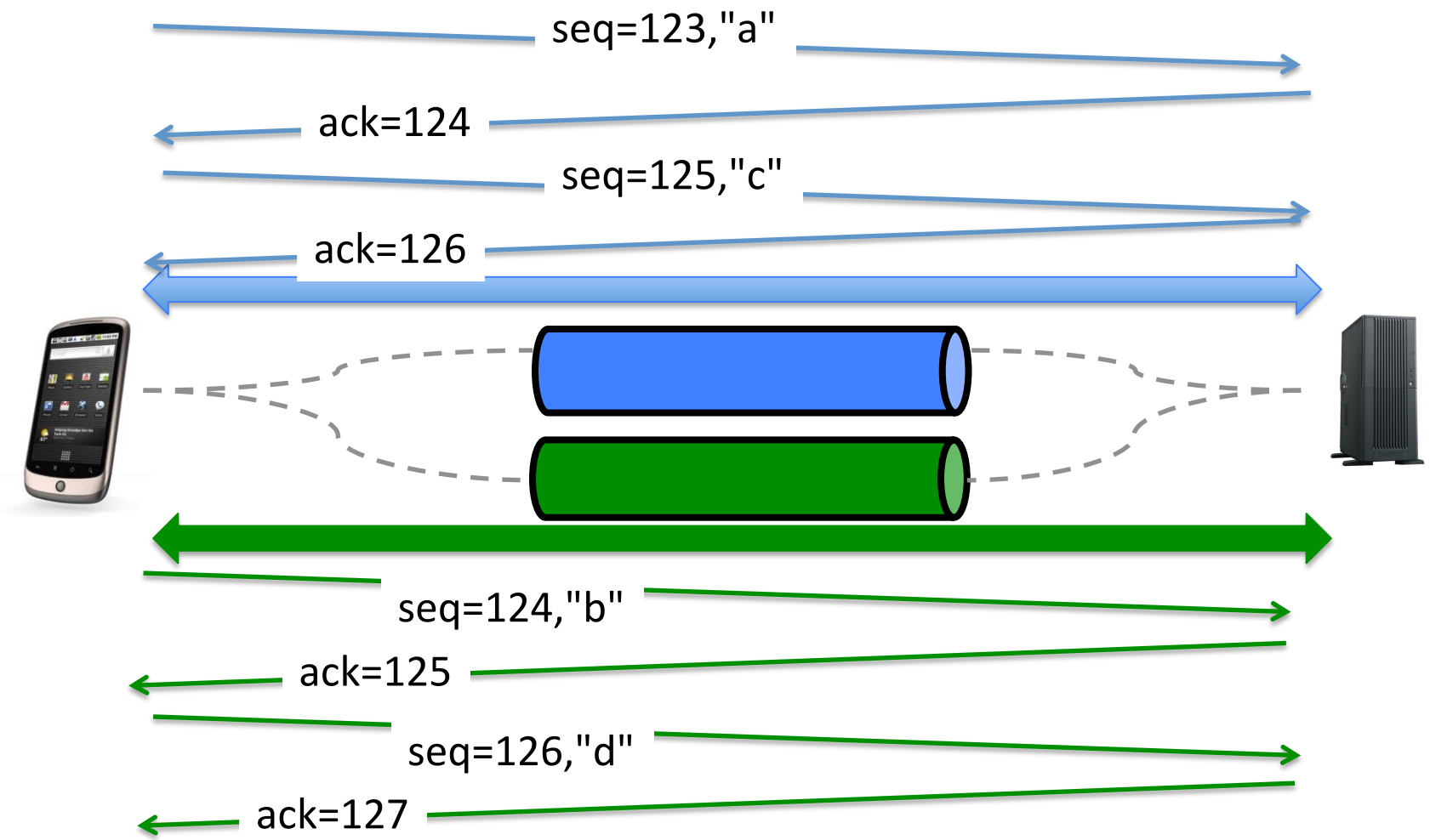
- Control plane
  - How to manage a Multipath TCP connection that uses several paths ?

Data plane

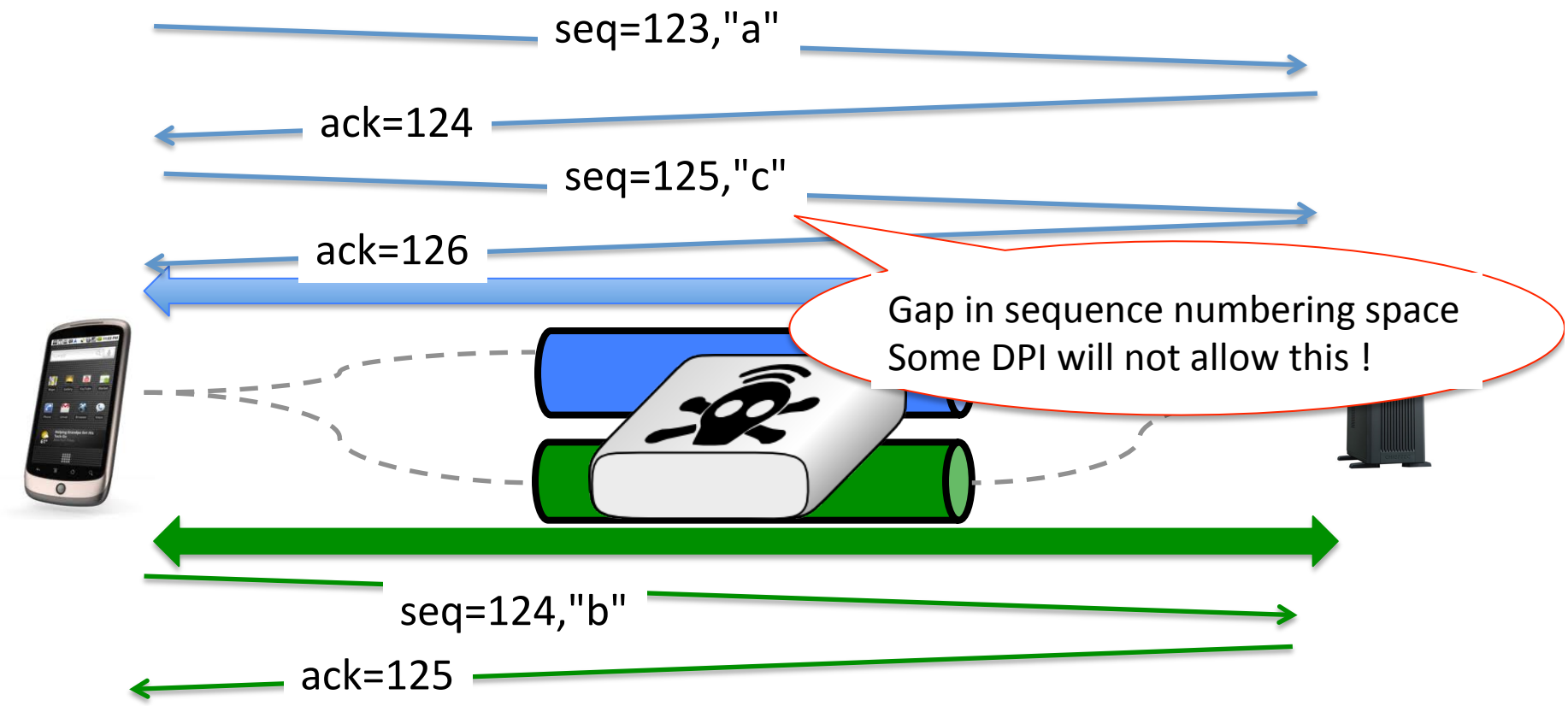
  - How to transport data ?

- Congestion control
  - How to control congestion over multiple paths ?

How to transfer data ?

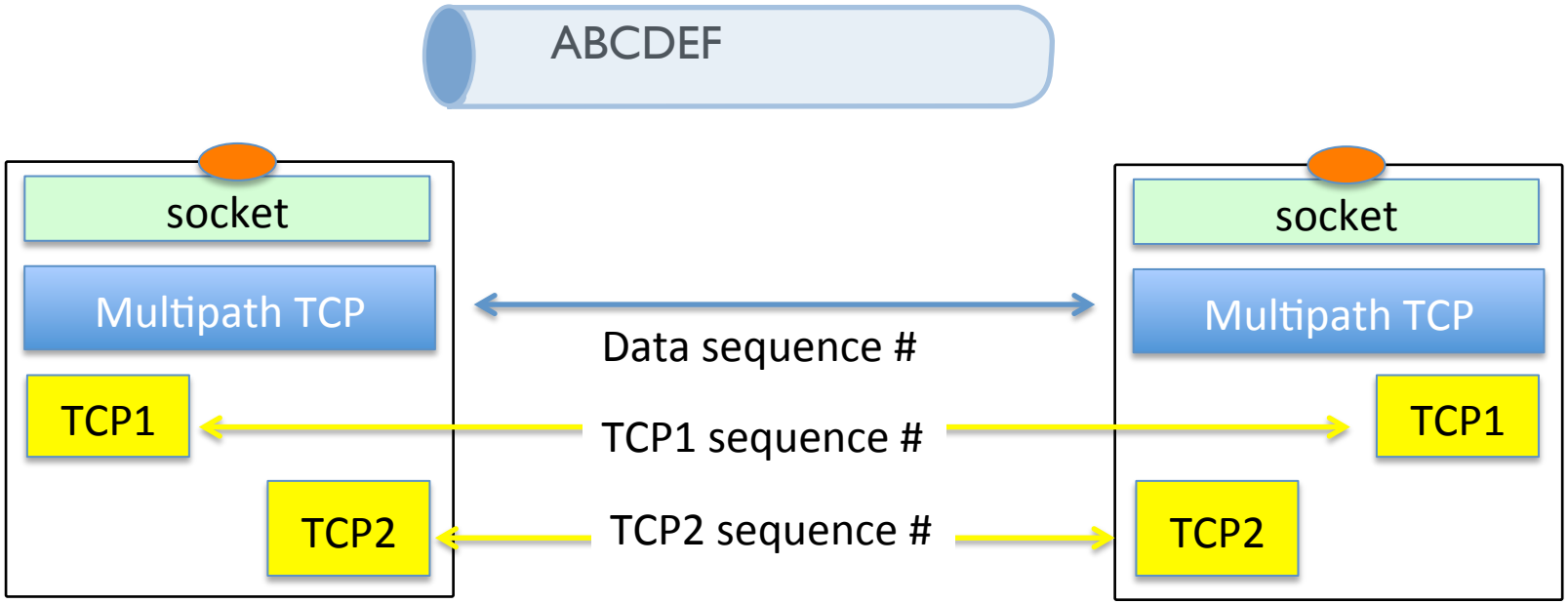


How to transfer data in today's Internet ?

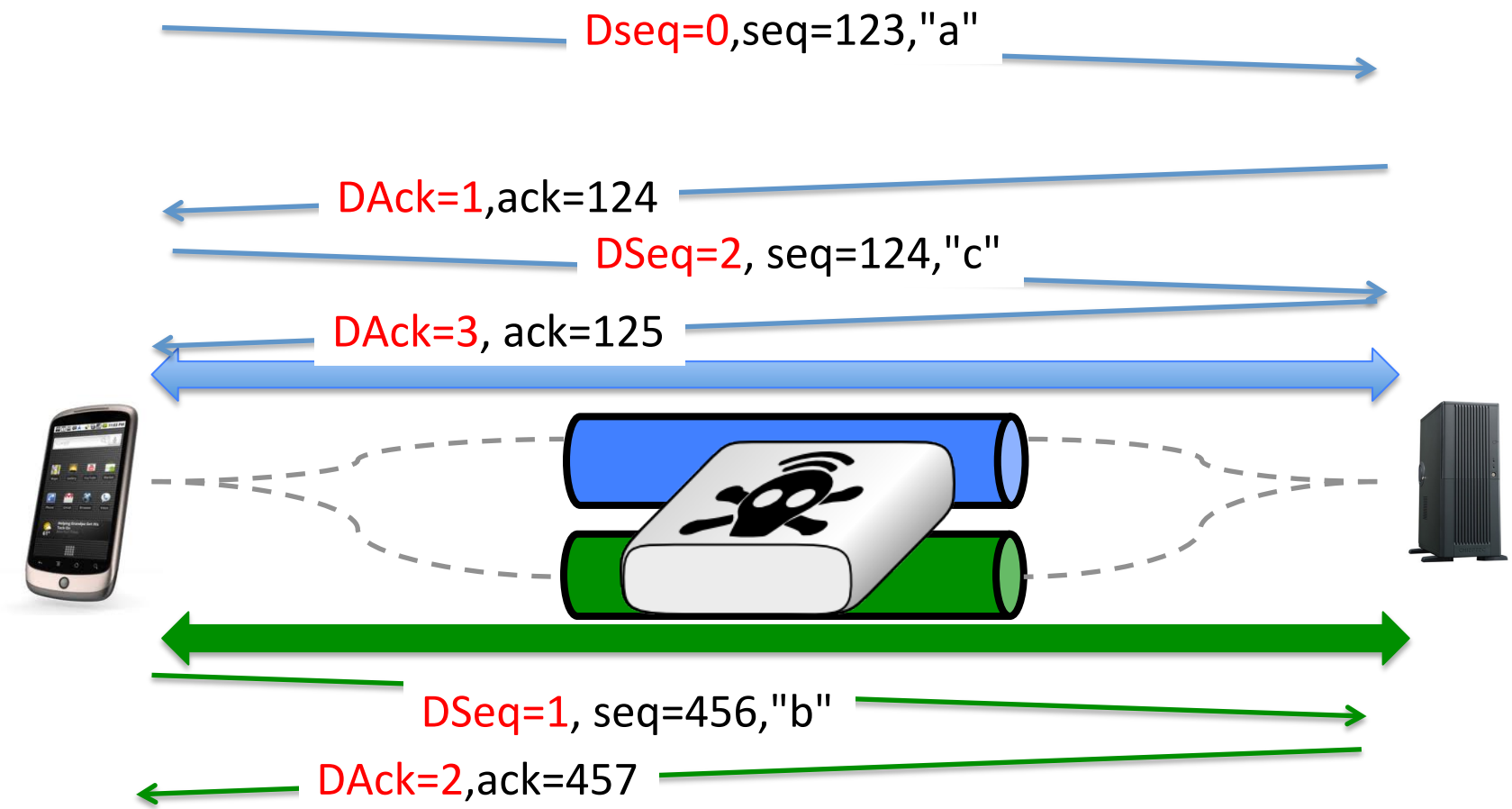


Multipath TCP Data transfer

- Two levels of sequence numbers

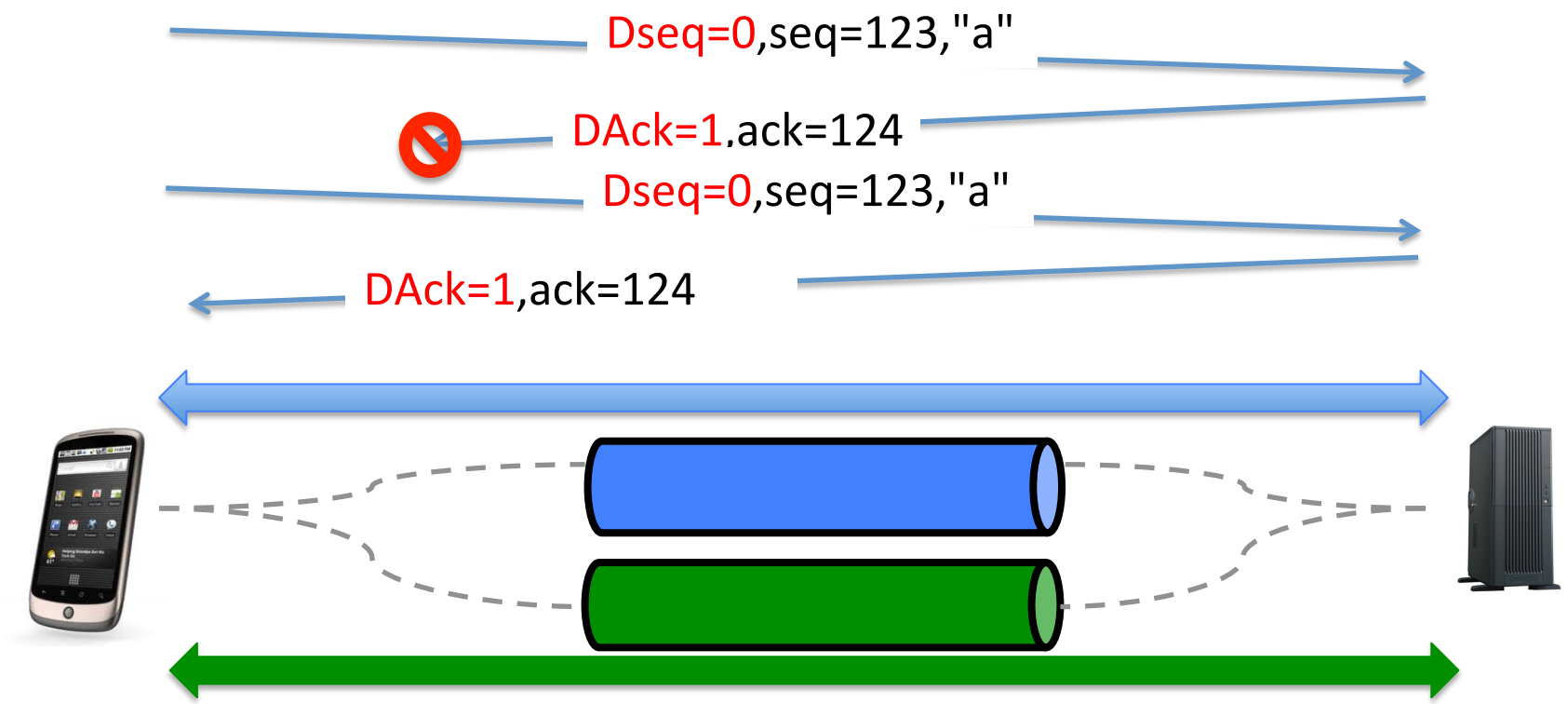


Multipath TCP
Data transfer



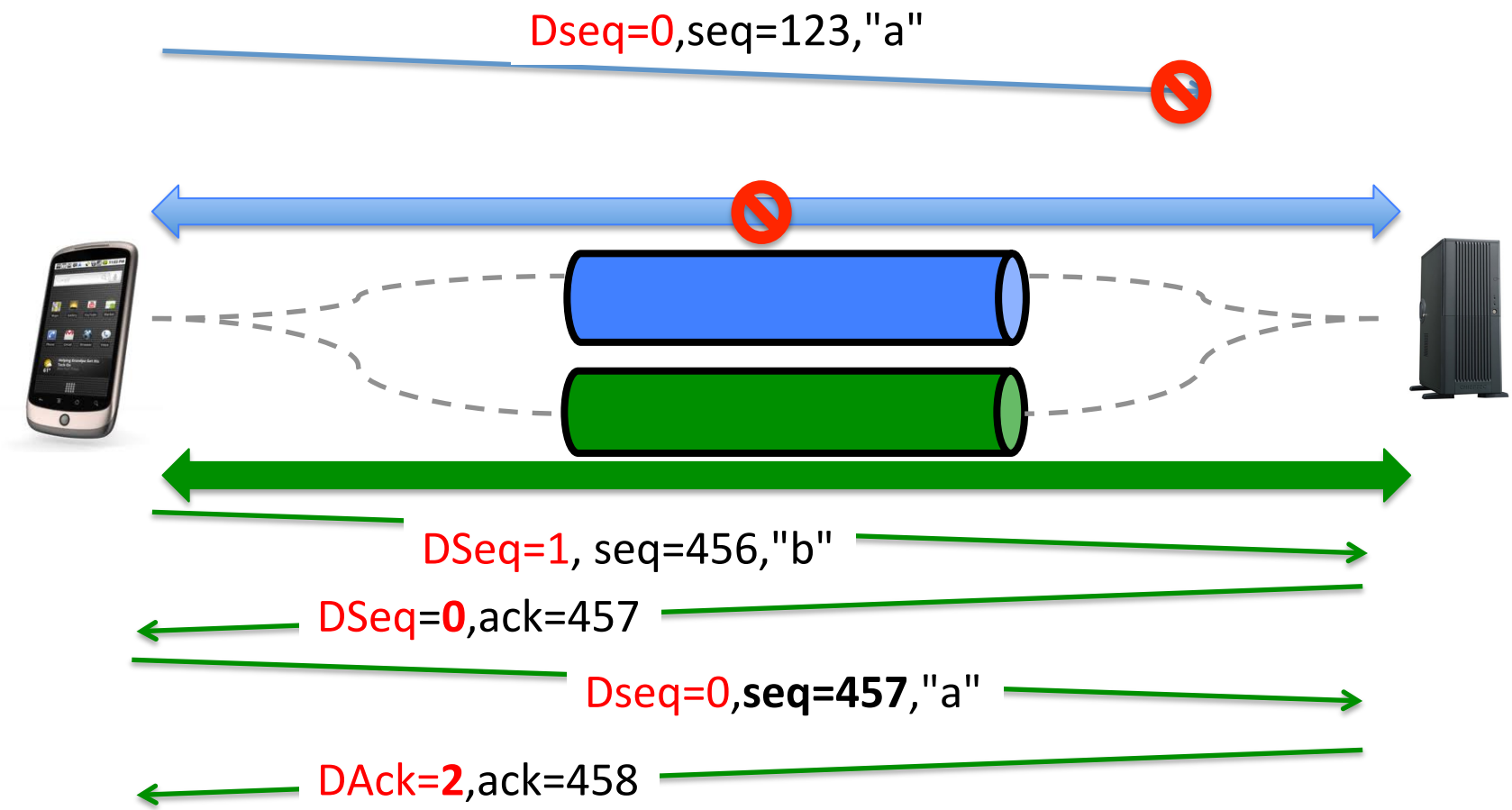
Multipath TCP
How to deal with losses ?

- Data losses over one TCP subflow
  - Fast retransmit and timeout as in regular TCP



Multipath TCP

- What happens when a TCP subflow fails ?



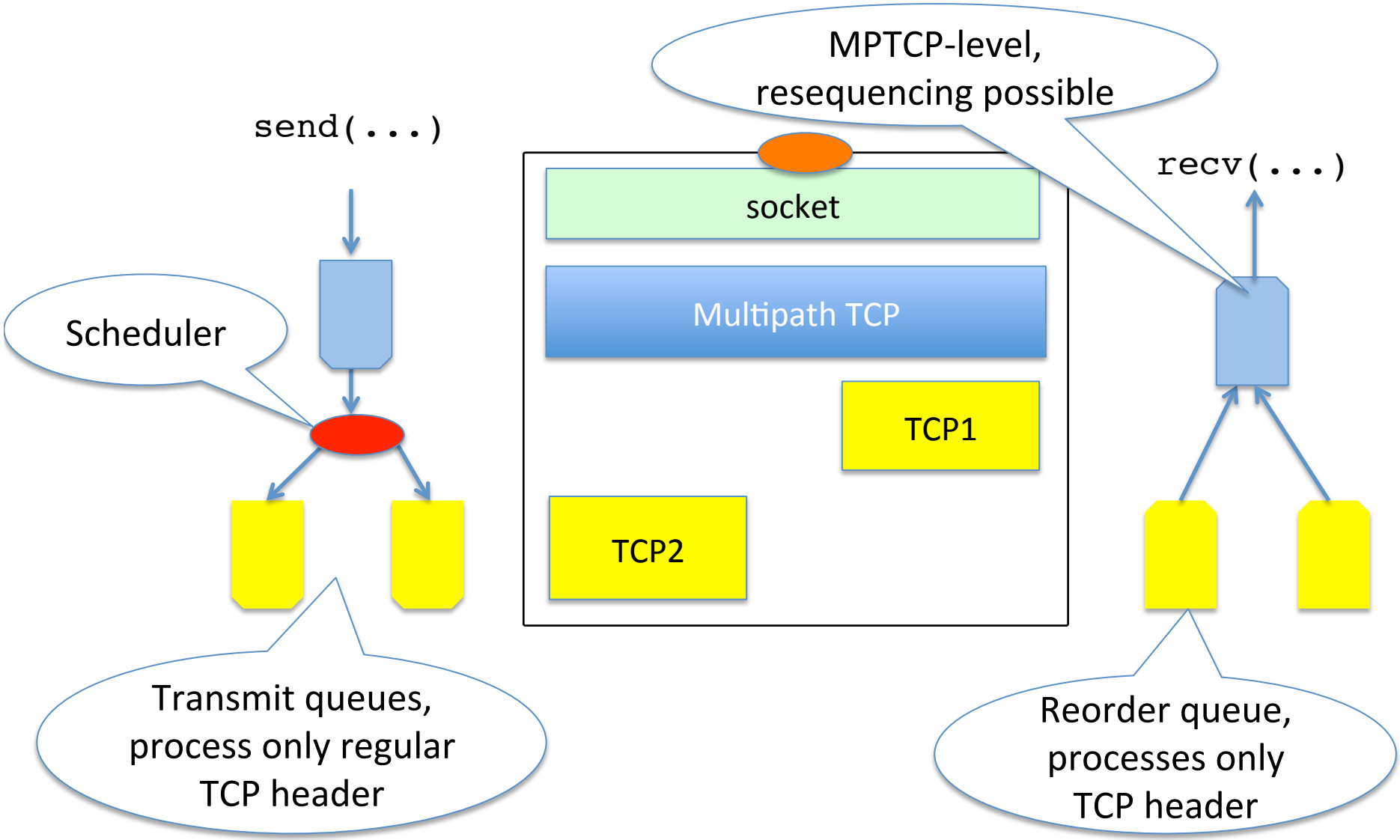
Retransmission heuristics

- Heuristics used by current Linux implementation
  - Fast retransmit is performed on the same subflow as the original transmission
  - Upon timeout expiration, reevaluate whether the segment could be retransmitted over another subflow
  - Upon loss of a subflow, all the unacknowledged data are retransmitted on other subflows

Multipath TCP Windows

- Multipath TCP maintains one window per Multipath TCP connection
  - Window is relative to the last acked data (Data Ack)
  - Window is shared among all subflows
    - It's up to the implementation to decide how the window is shared
  - Window is transmitted inside the `window` field of the regular TCP header
  - If middleboxes change `window` field,
    - use largest `window` received at MPTCP-level
    - use received `window` over each subflow to cope with the flow control imposed by the middlebox

Multipath TCP buffers



Sending Multipath TCP information

- How to exchange the Multipath TCP specific information between two hosts ?

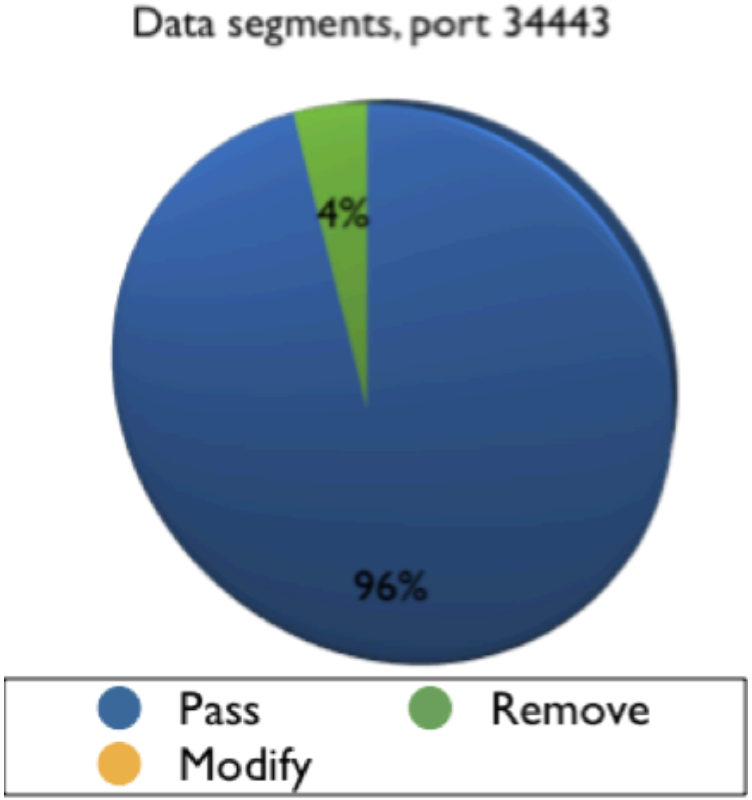
- Option 1
  - Use TLVs to encode data and control information inside payload of subflows

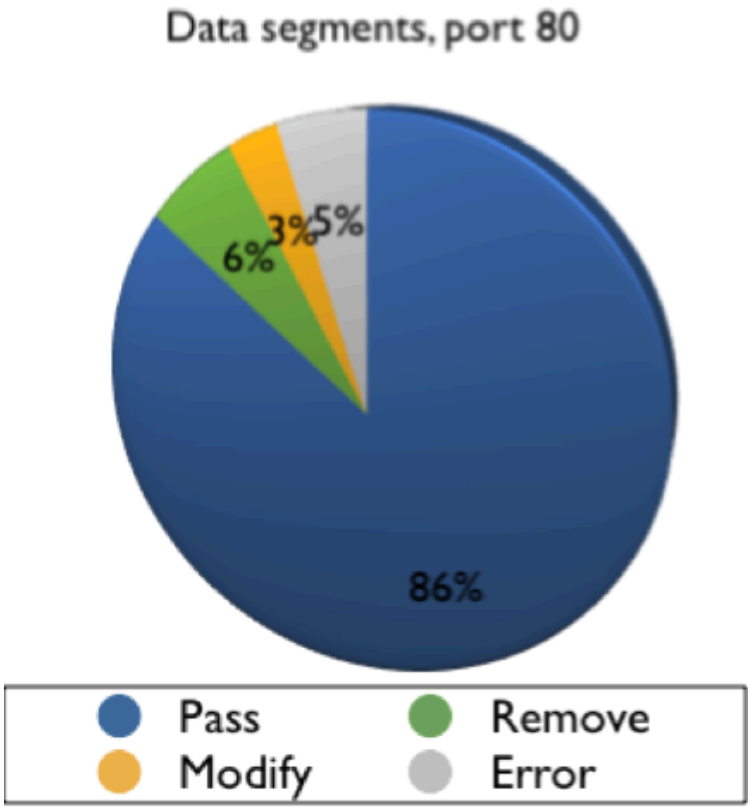
- **Option 2**
  - Use TCP options to encode all Multipath TCP information

Option 1 : Michael Scharf, Thomas-Rolf Banniza , *MCTCP: A Multipath Transport Shim Layer*, GLOBECOM 2011

Is it safe to use TCP options ?

- Known option (TS) in Data segments

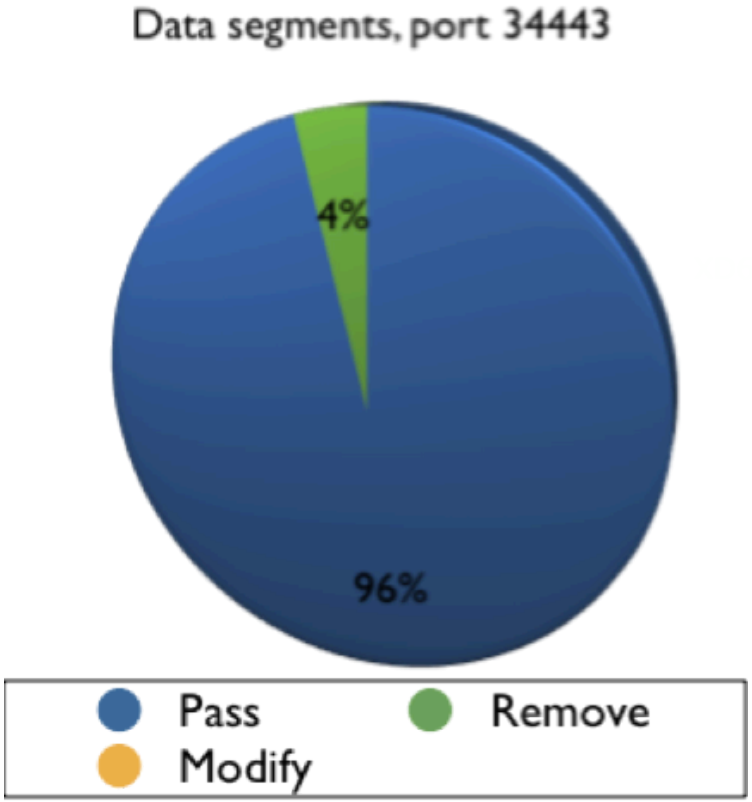


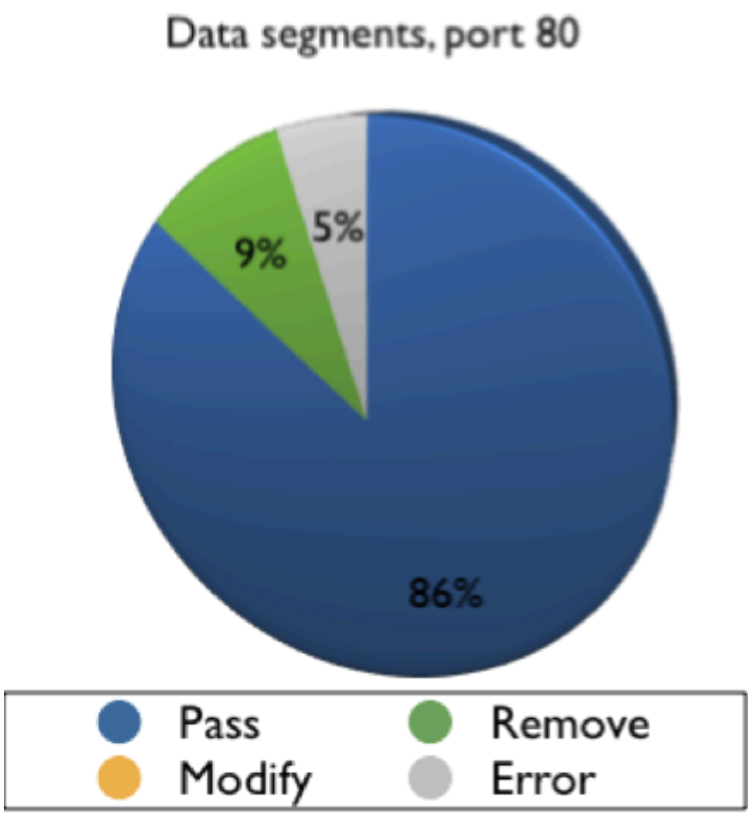


Honda, Michio, et al. "Is it still possible to extend TCP?." Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

Is it safe to use TCP options ?

- Unknown option in Data segments

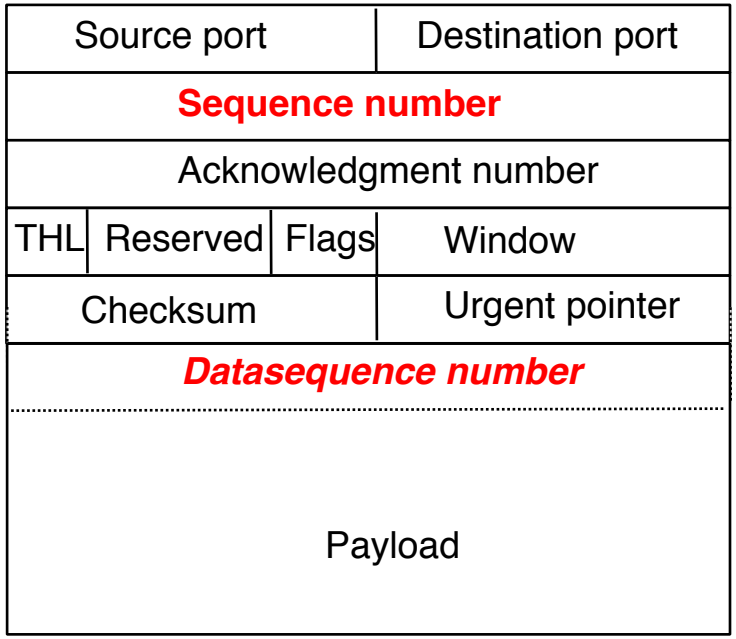




Honda, Michio, et al. "Is it still possible to extend TCP?." Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

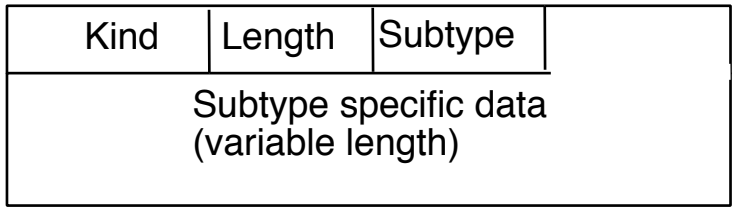
Data sequence numbers and TCP segments

- How to transport Data sequence numbers ?
  - Same solution as for TCP
    - Data sequence number in TCP option is the Data sequence number of the first byte of the segment

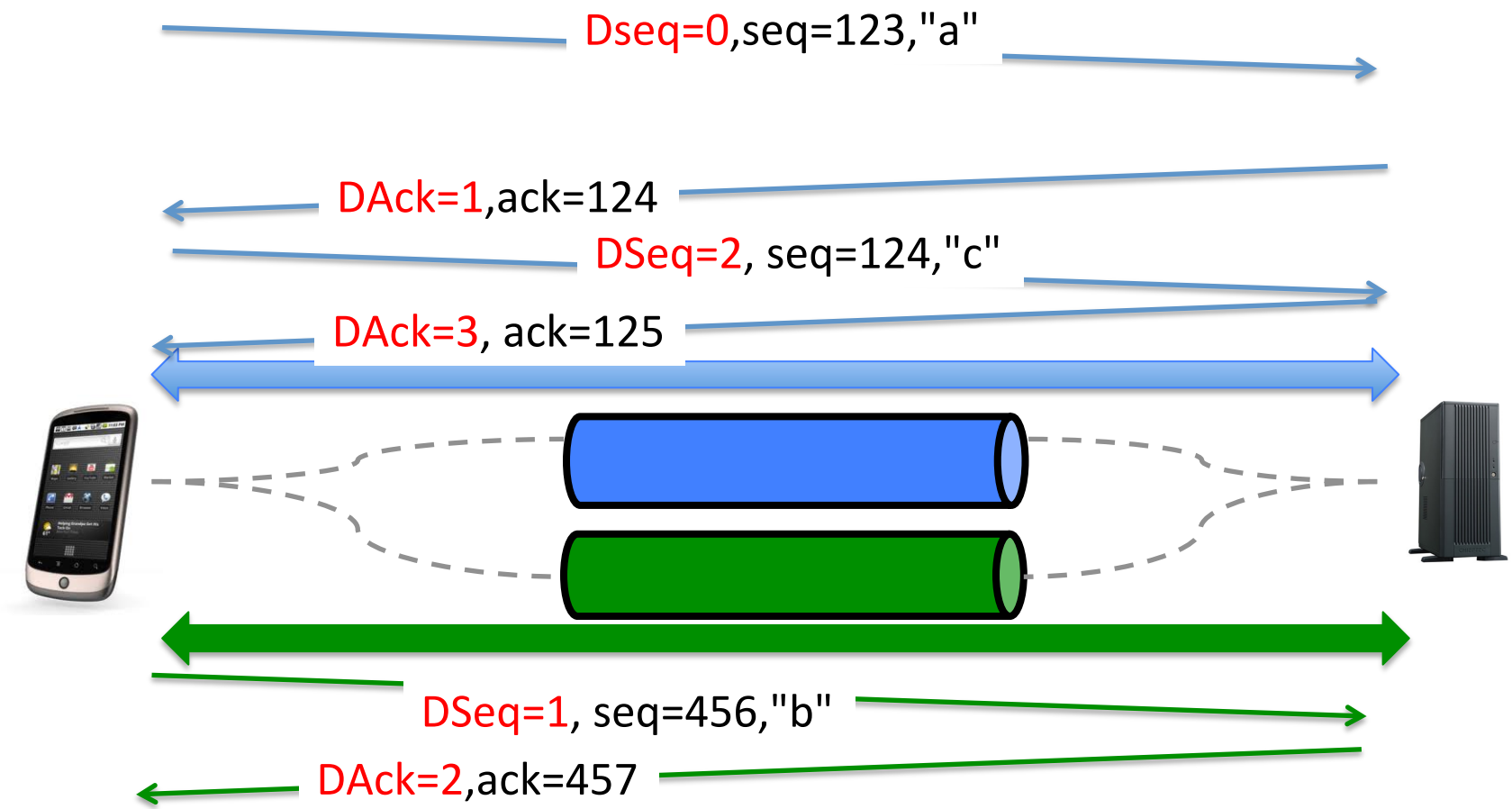


Multipath TCP option

- A single option type
  - to minimise the risk of having one option accepted by middleboxes in SYN segments and rejected in segments carrying data

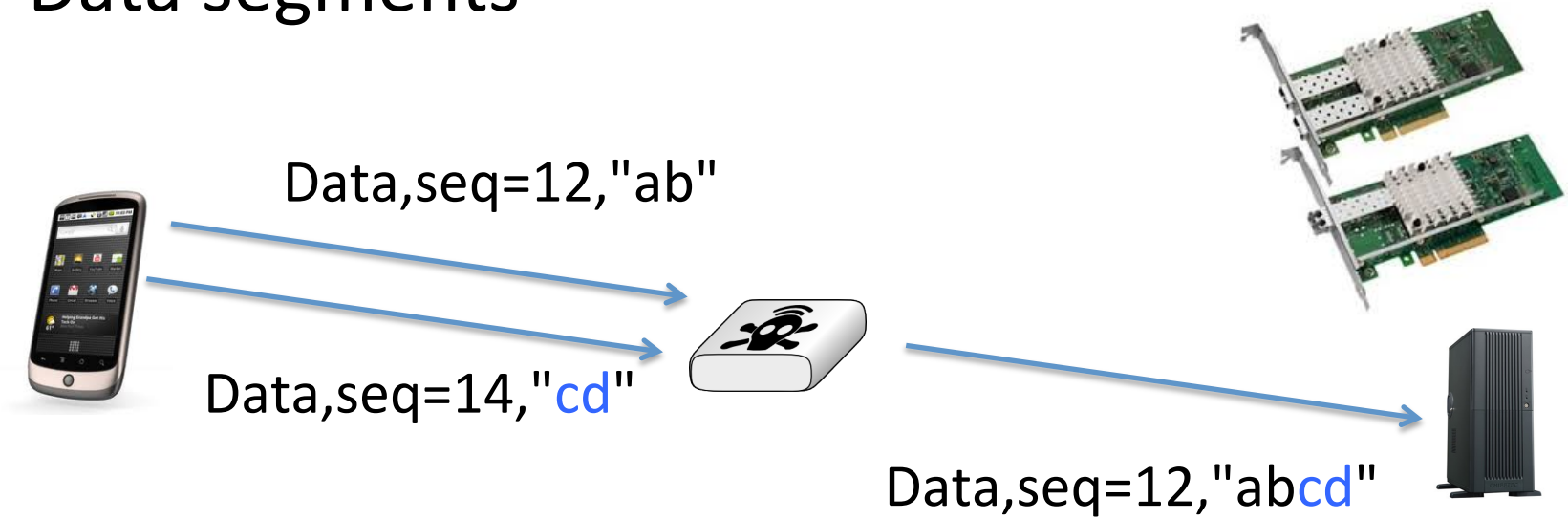


Multipath TCP
Data transfer



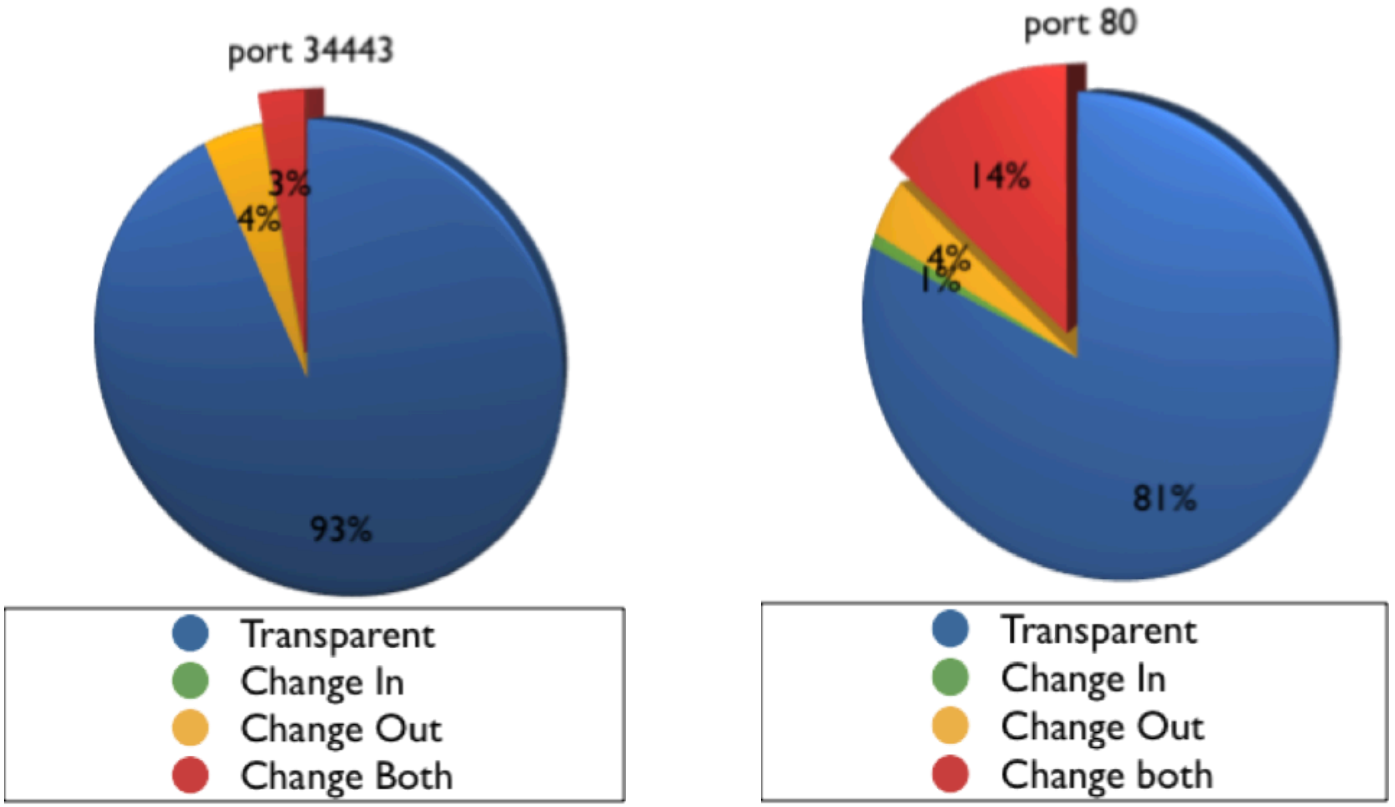
Other types of middlebox interference

- Data segments



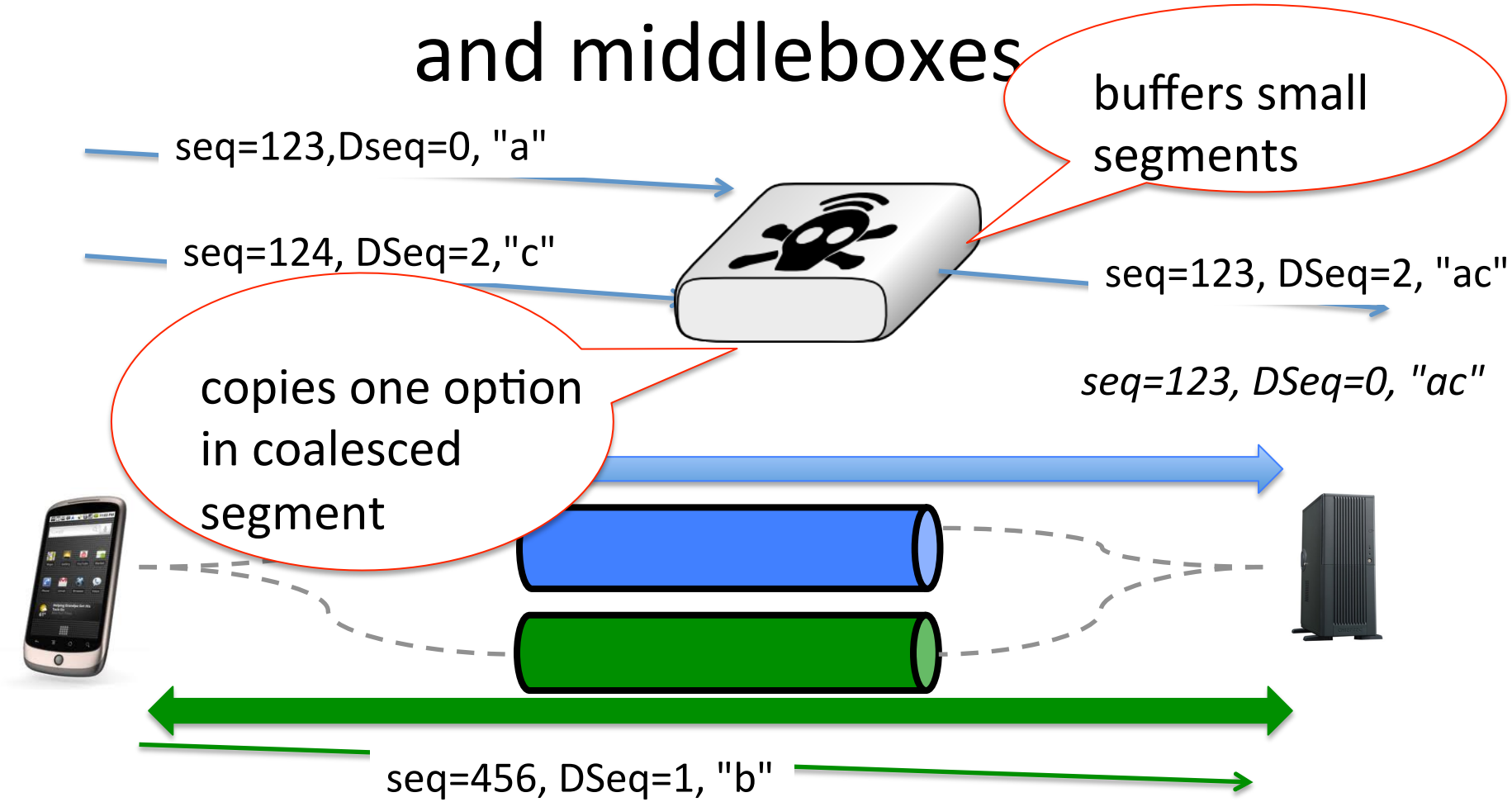
Such a middlebox could also be the network adapter of the server that uses LRO to improve performance.

Segment coalescing

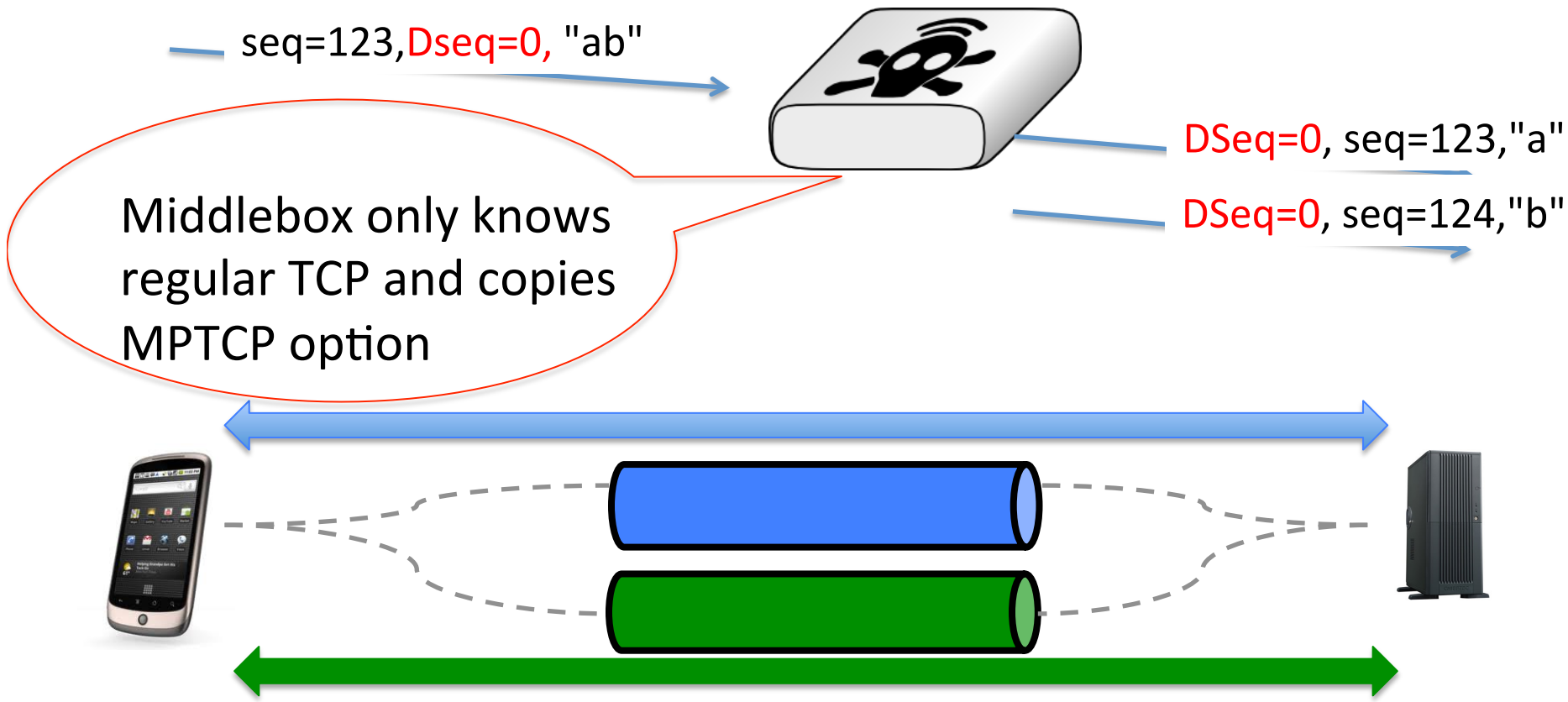


Honda, Michio, et al. "Is it still possible to extend TCP?." Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

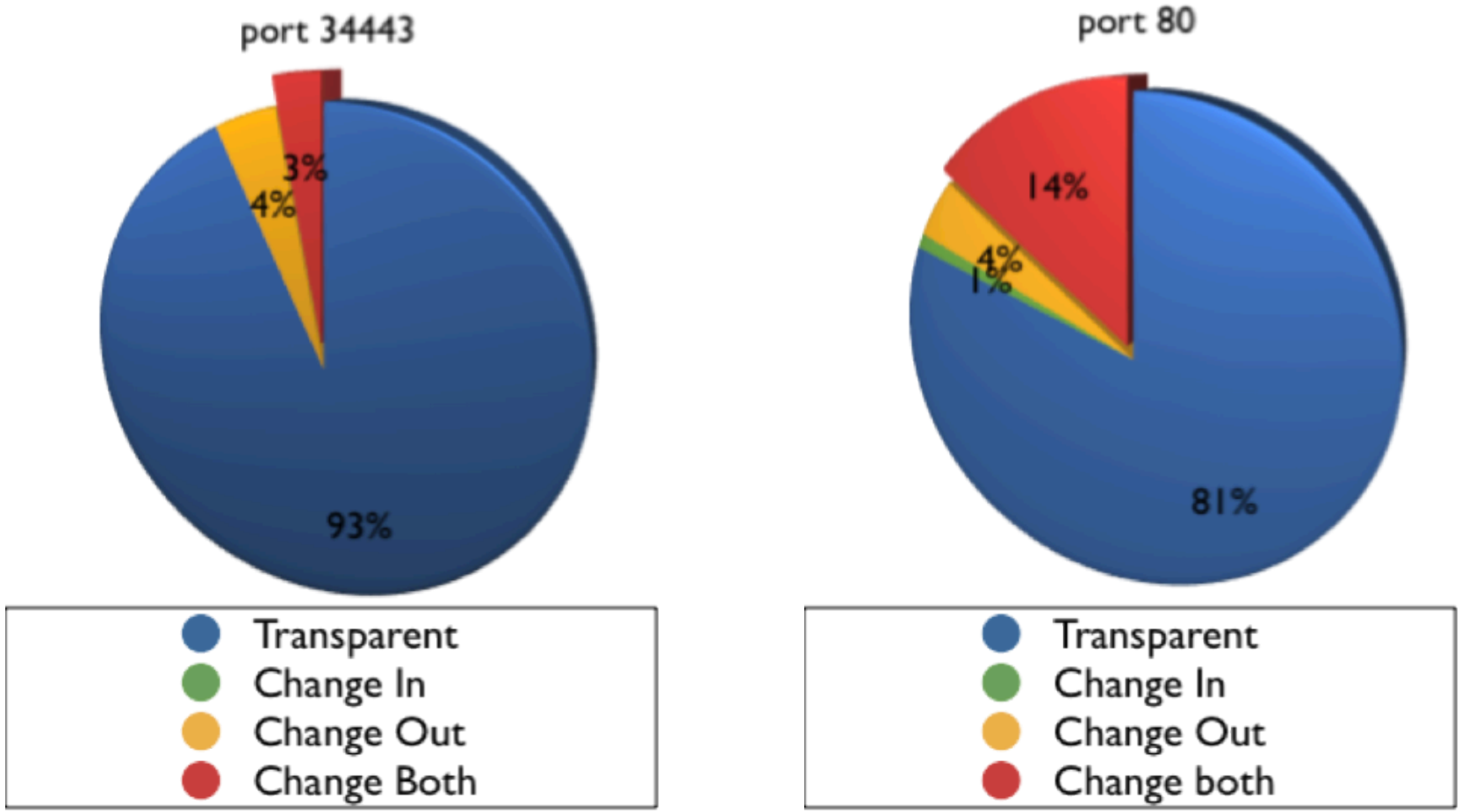
Data sequence numbers and middleboxes



Data sequence numbers and middleboxes



TCP sequence number and middleboxes

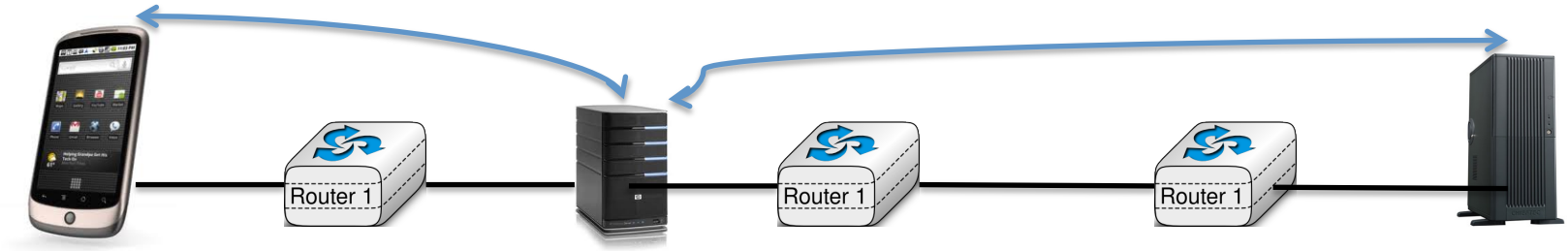


Honda, Michio, et al. "Is it still possible to extend TCP?." Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

Which middleboxes change TCP sequence numbers ?

- Some firewalls change TCP sequence numbers in SYN segments to ensure randomness
  - fix for old windows95 bug

- Transparent proxies terminate TCP connections

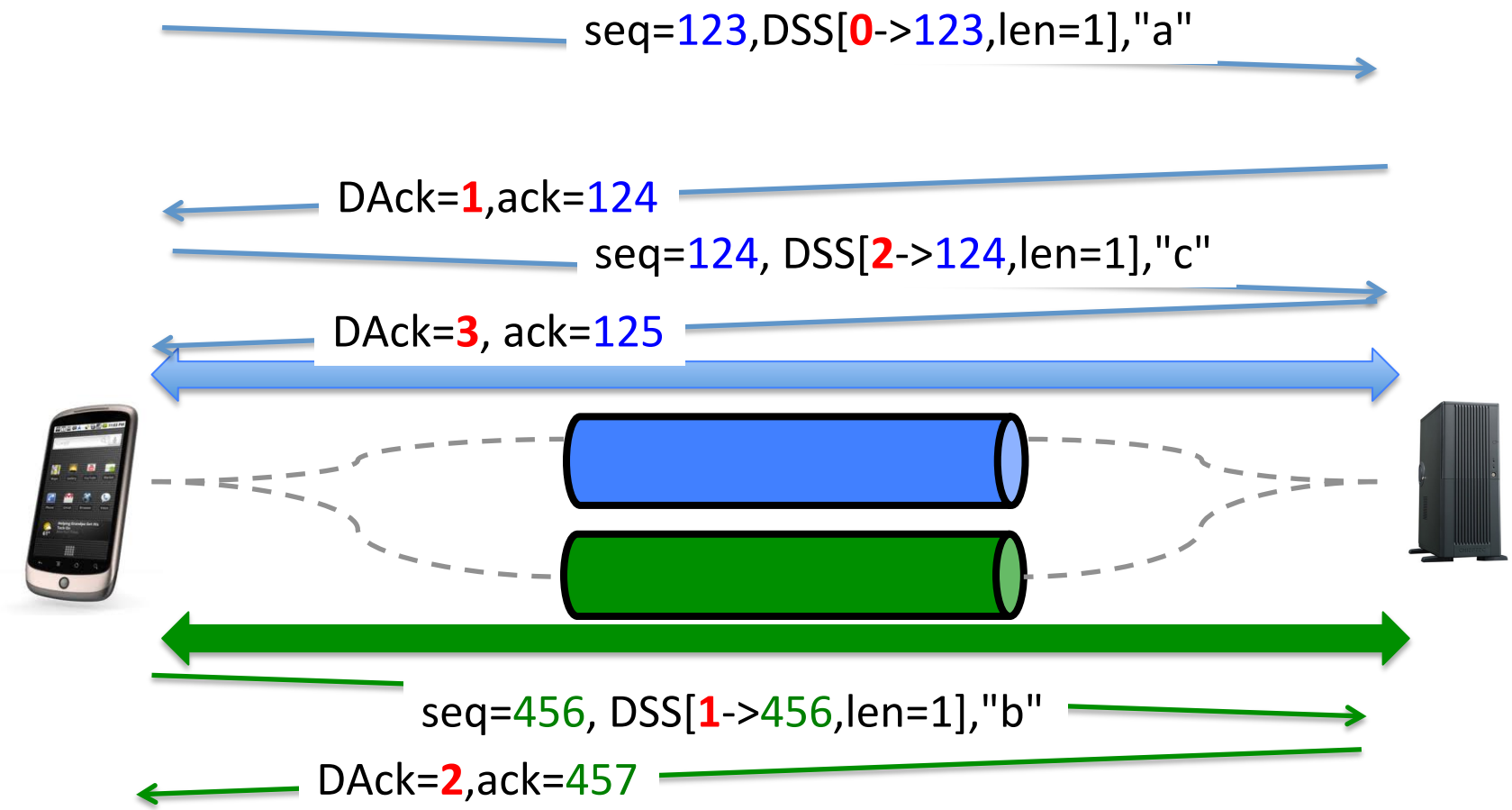


Data sequence numbers and middleboxes

- How to avoid desynchronisation between the bytestream and data sequence numbers ?

- Solution
  - Multipath TCP option carries **mapping** between Data sequence numbers and *(difference between initial and current)* subflow sequence numbers
    - mapping covers a part of the bytestream (length)

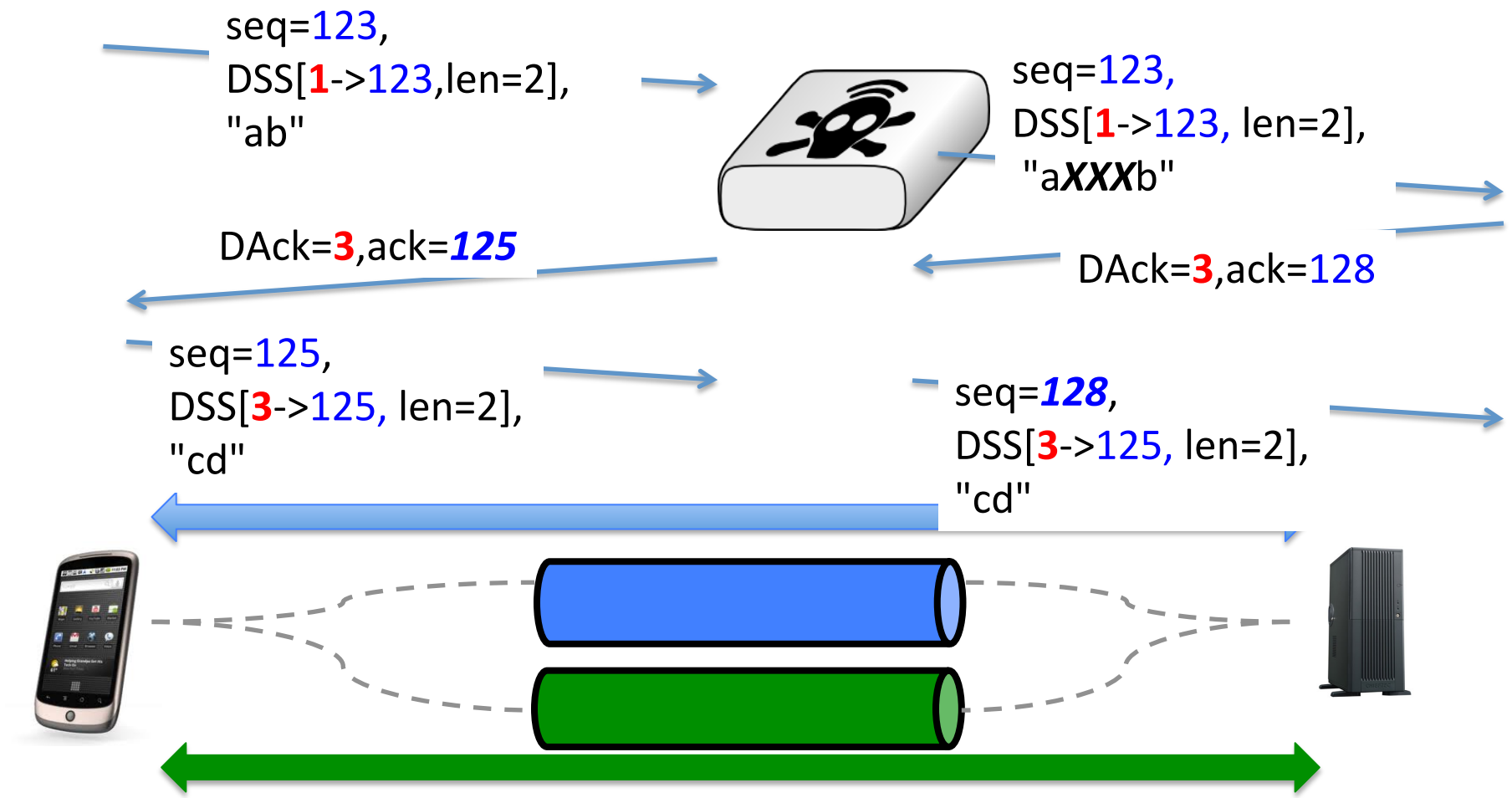
Multipath TCP
Data transfer



Multipath TCP and middleboxes

- With the DSS mapping, Multipath TCP can cope with middleboxes that
  - combine segments
  - split segments
- Are they the most annoying middleboxes for Multipath TCP ?
  - Unfortunately not

The worst middlebox



- Is this an academic exercise or reality ?

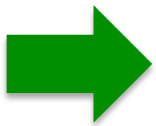
The worst middlebox

- Is unfortunately very old and widely used...
  - Any ALG for a NAT

```
220 ProFTPD 1.3.3d Server (BELNET FTPD Server) [193.190.67.15]
ftp_login: user `<null>' pass `<null>' host `ftp.belnet.be'
Name (ftp.belnet.be:obo): anonymous
---> USER anonymous
331 Anonymous login ok, send your complete email address as your password
Password:
---> PASS XXXX
---> PORT 192,168,0,7,195,120
200 PORT command successful
---> LIST
150 Opening ASCII mode data connection for file list
lrw-r--r--   1 ftp      ftp            6 Jun  1  2011 pub -> mirror
226 Transfer complete
```

Coping with the worst middlebox

- What should Multipath TCP do in the presence of such a worst middlebox ?
  - Do nothing and ignore the middlebox
    - but then the bytestream and the application would be broken and this problem will be difficult to debug by network administrators



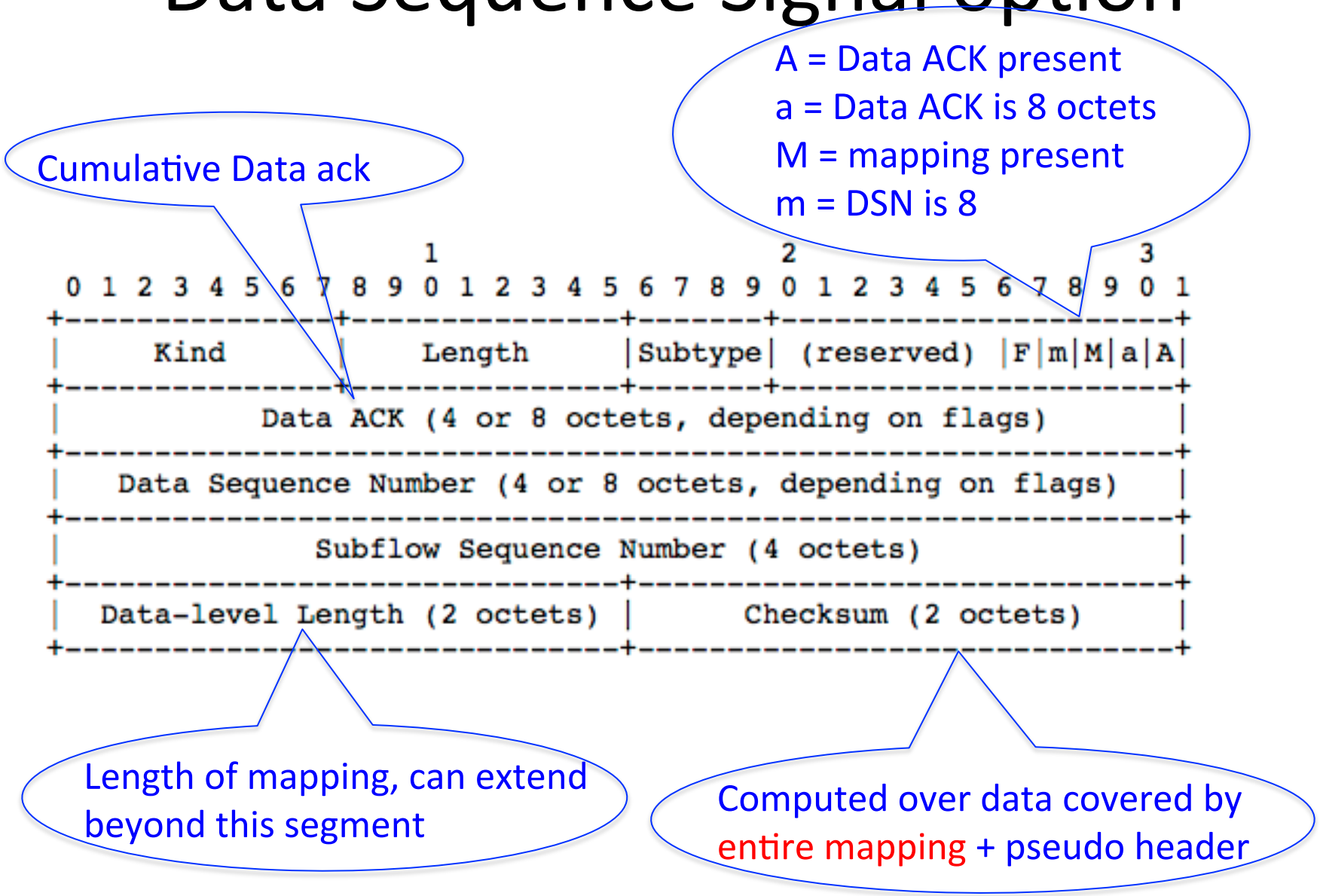
  - Detect the presence of the middlebox
    - and fallback to regular TCP (i.e. use a single path and nothing fancy)

Multipath TCP **MUST** work in all networks where regular TCP works.

Detecting the worst middlebox ?

- How can Multipath TCP detect a middlebox that modifies the bytestream and inserts/ removes bytes ?
  - Various solutions were explored
  - In the end, Multipath TCP chose to include its own checksum to detect insertion/deletion of bytes

Data Sequence Signal option



The Multipath TCP protocol

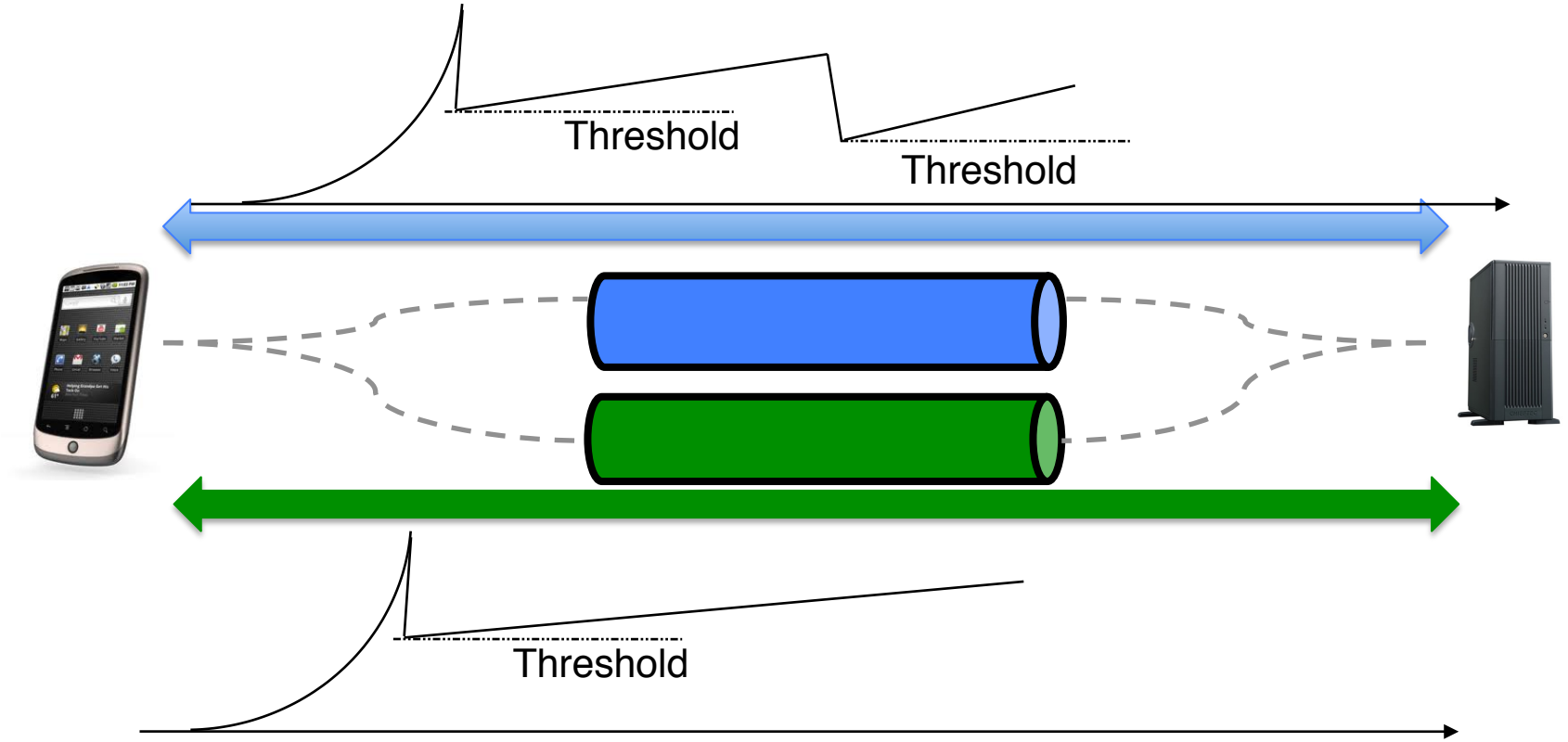
- Control plane
  - How to manage a Multipath TCP connection that uses several paths ?
- Data plane
  - How to transport data ?

Congestion control

  - How to control congestion over multiple paths ?

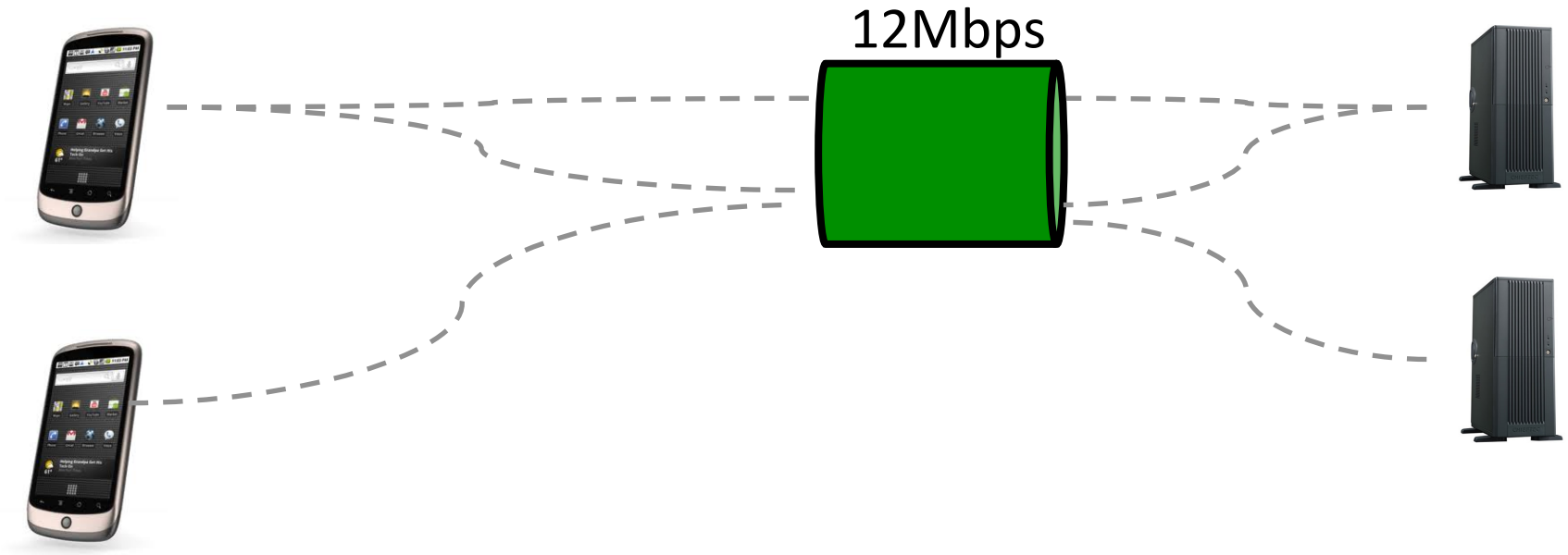
Congestion control for Multipath TCP

- Simple approach
  - independant congestion windows



Independant congestion windows

- Problem



Multipath TCP congestion control

- Goals
  - Improve throughput
    - MPTCP flow should get at least as much as single flow on the best path
  - Do no harm
    - fairness with regular TCP flows
  - Balance congestion
    - Multipath TCP should move as much traffic as possible out of its most congested paths while meeting the above goals

The Multipath TCP protocol

Control plane

- How to manage a Multipath TCP connection that uses several paths ?

- Data plane
  - How to transport data ?

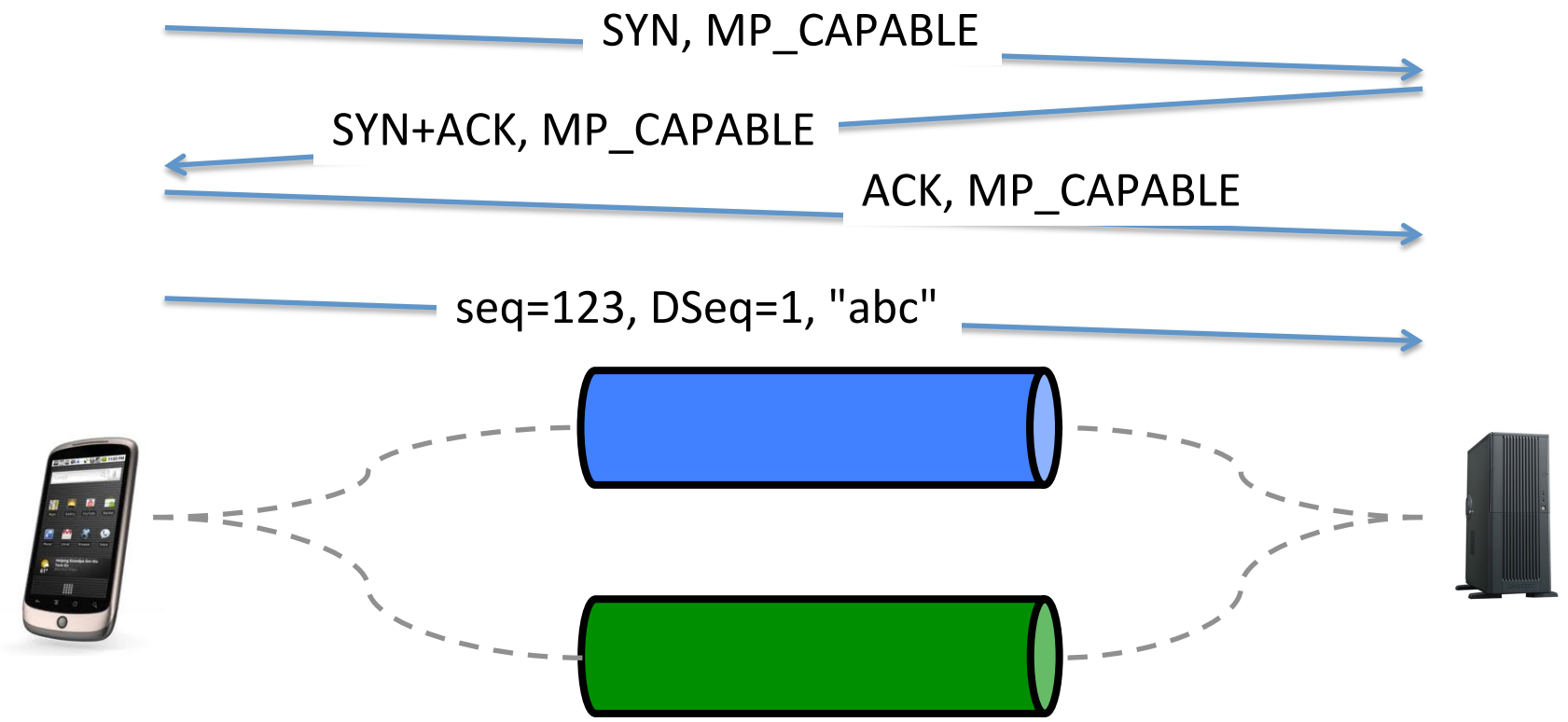
- Congestion control
  - How to control congestion over multiple paths ?

The Multipath TCP control plane

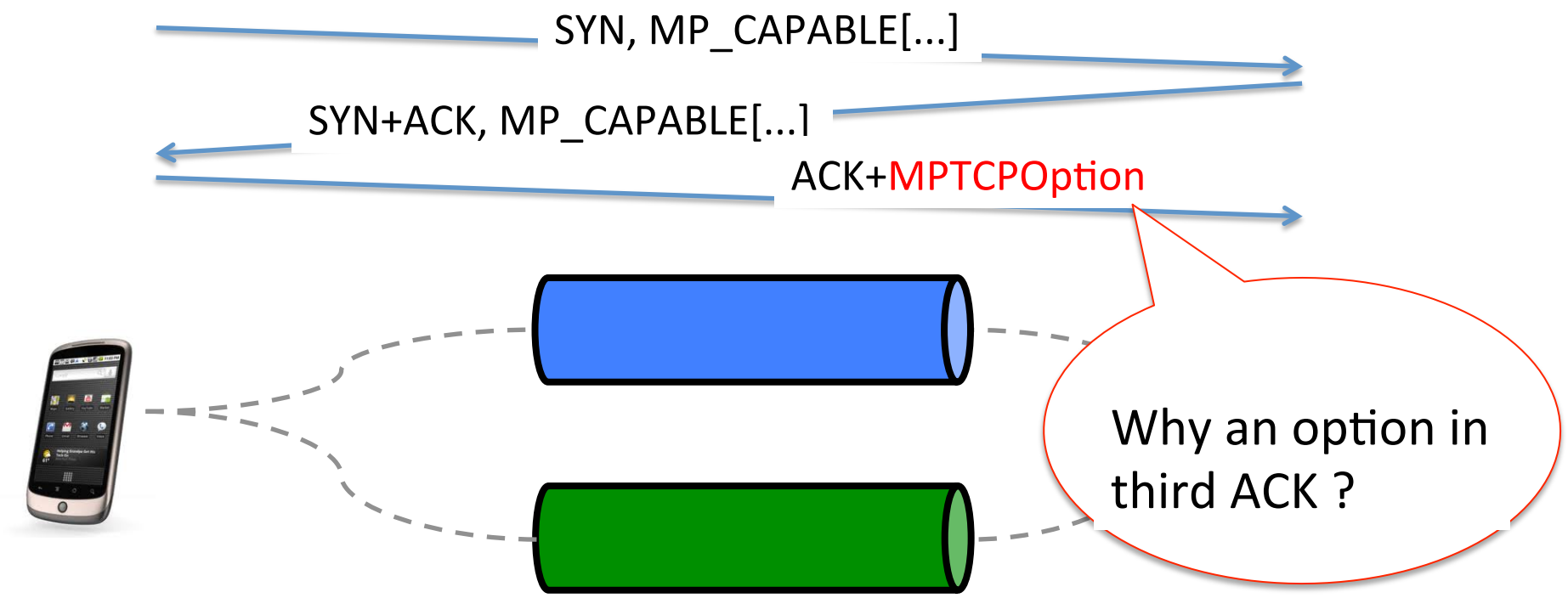
- Connection establishment
  - Beware of middleboxes that remove TCP options
  - Limited space inside TCP option in SYN
- Closing a Multipath TCP connection
  - Decouple closing the Multipath TCP connection from closing the subflows
- Address dynamics

Multipath TCP
Connection establishment

- Principle

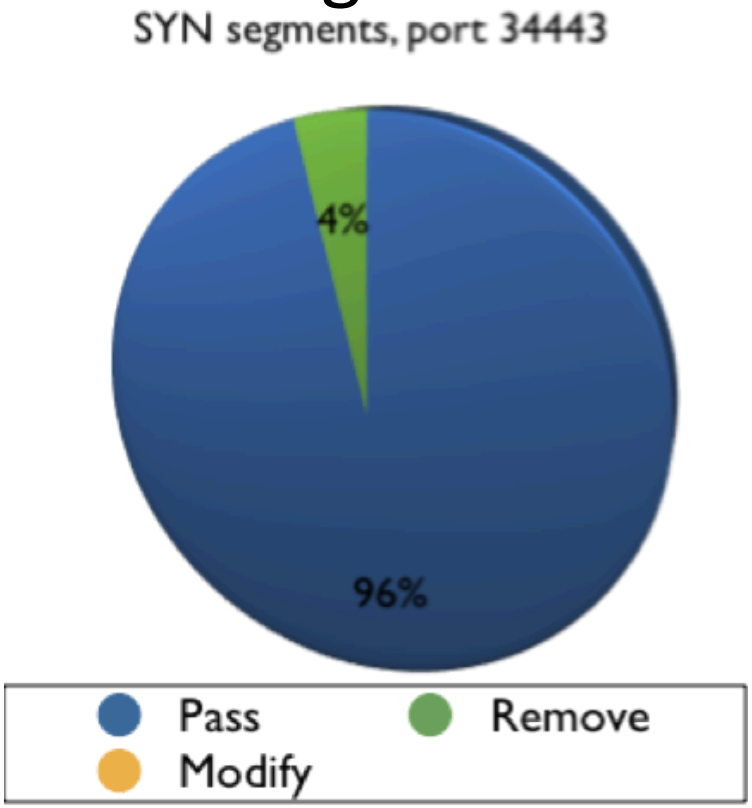


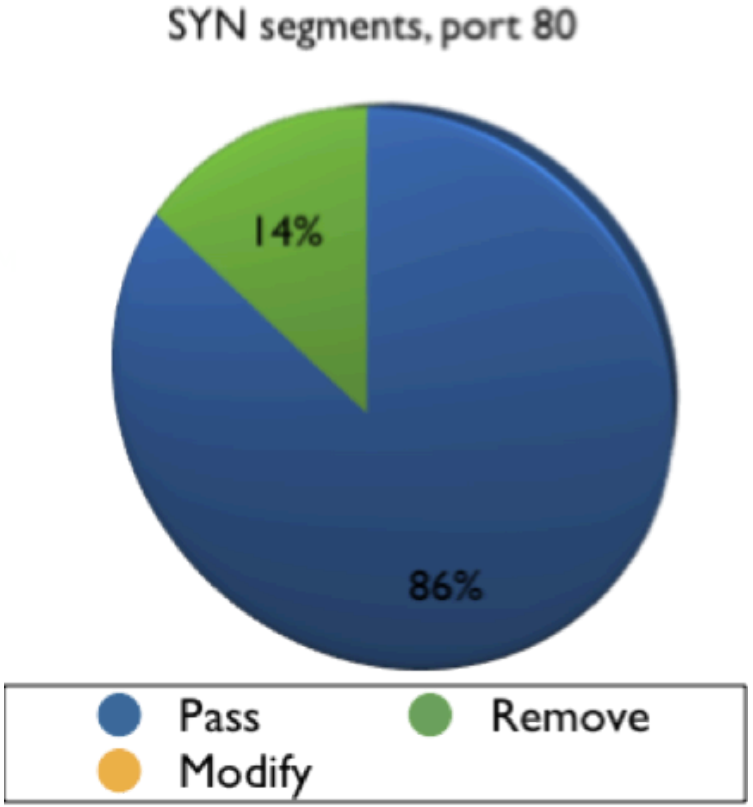
Multipath TCP handshake



TCP options

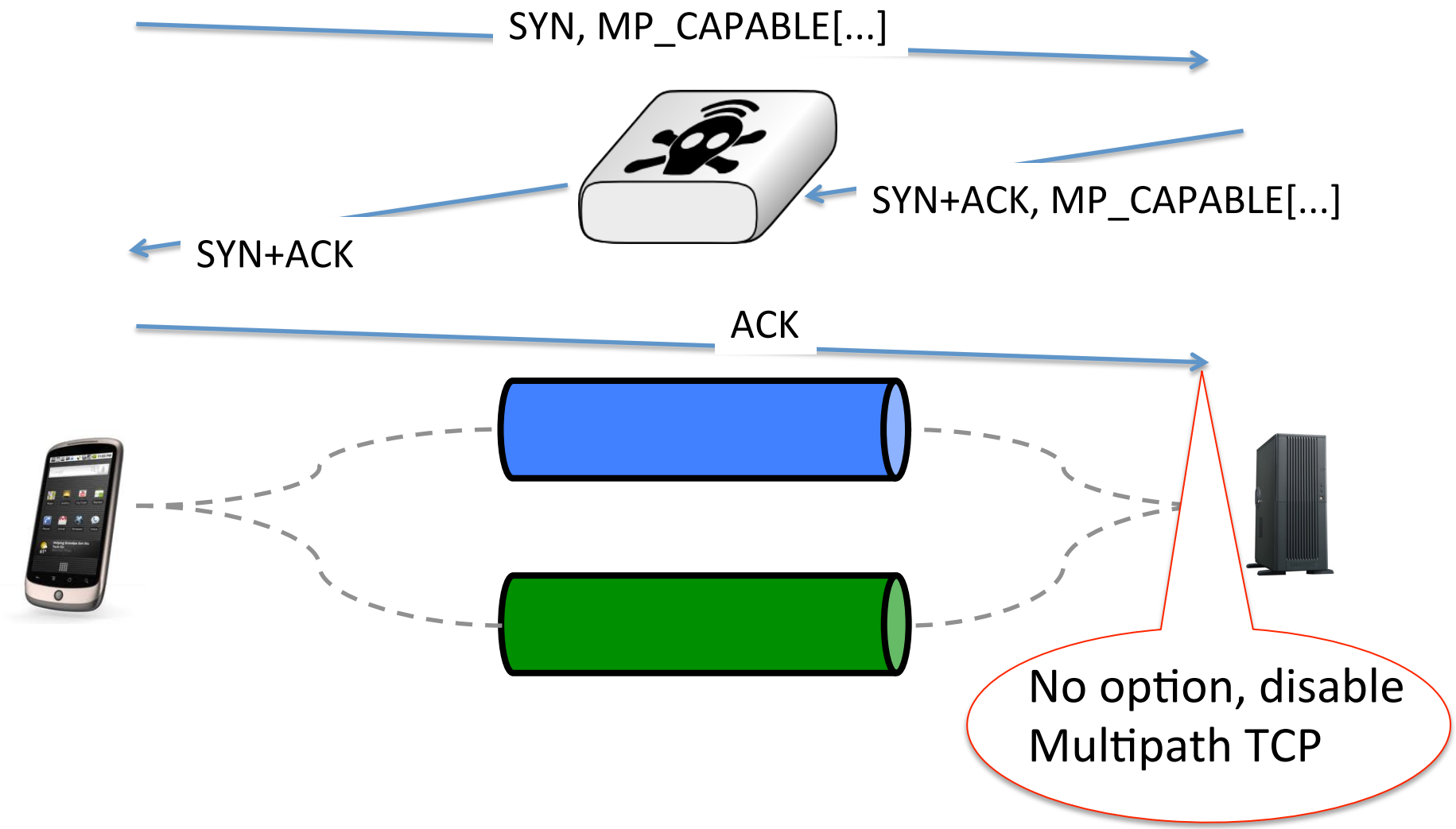
- In SYN segments





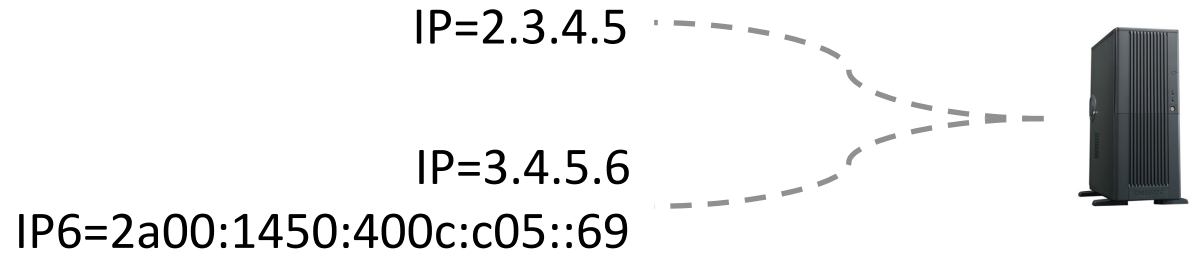
Honda, Michio, et al. "Is it still possible to extend TCP?." Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011.

Multipath TCP option in third ACK

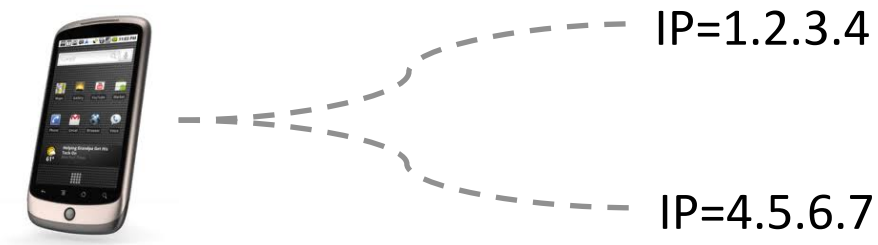


Multipath TCP
Address dynamics

- How to learn the addresses of a host ?

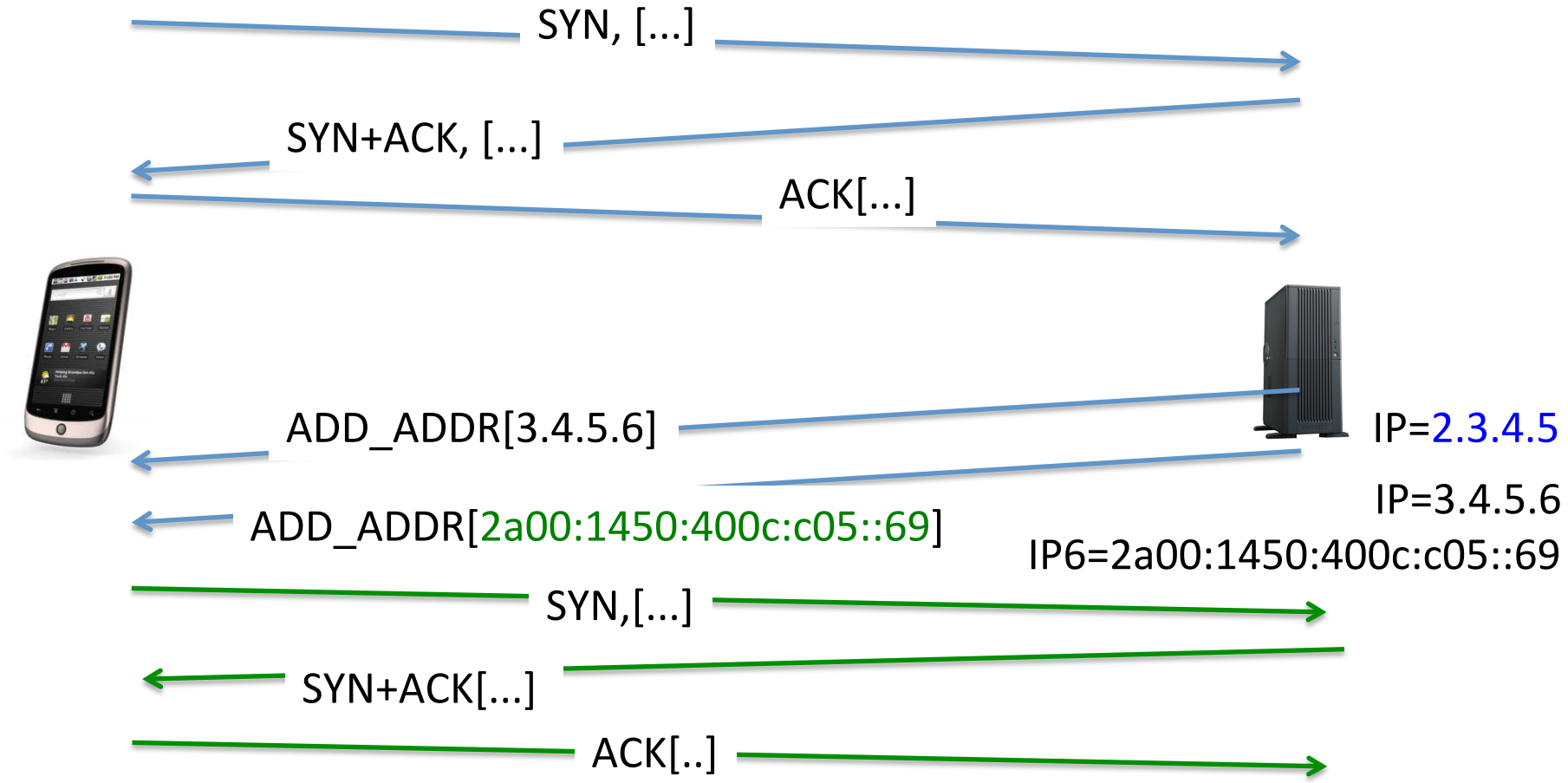


- How to deal with address changes ?



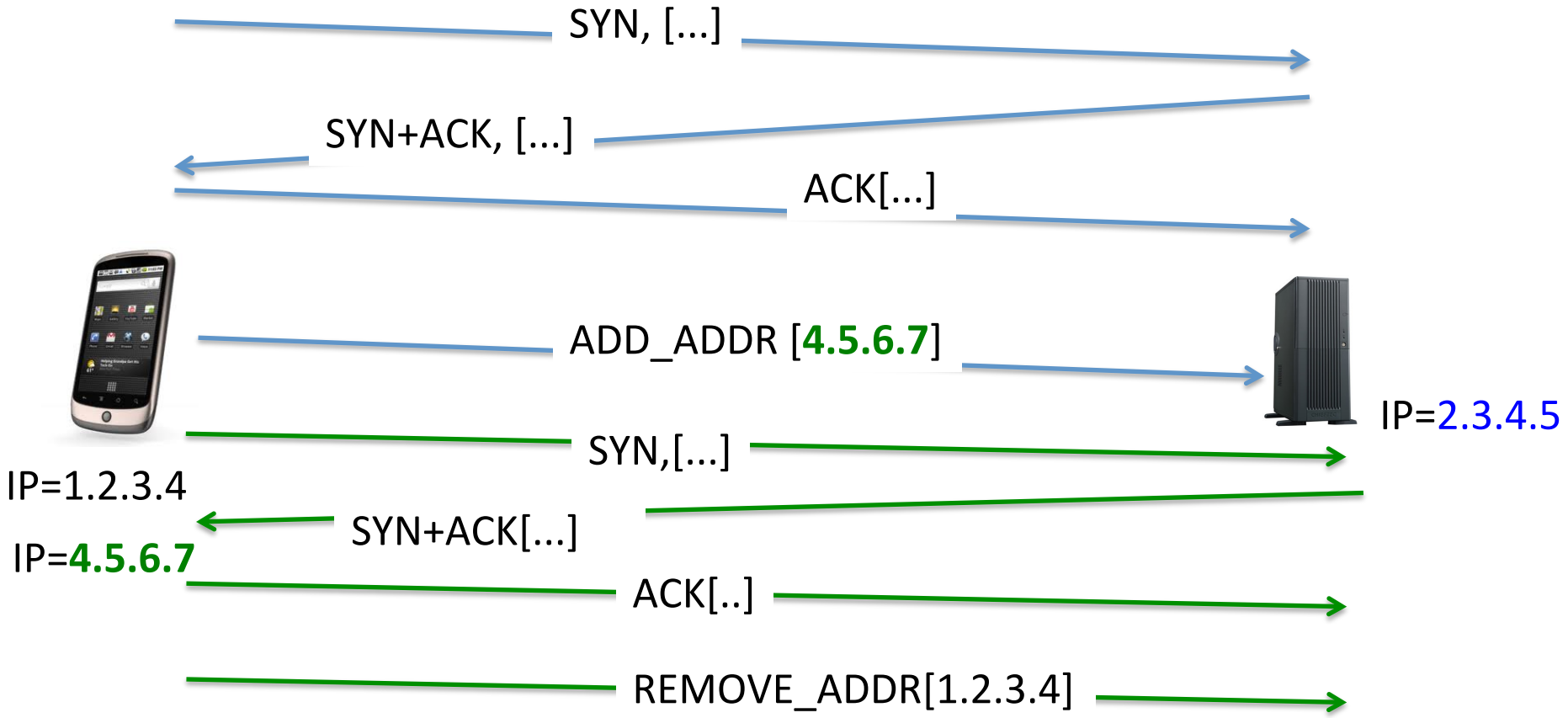
Address dynamics

- Basic solution : multihomed server

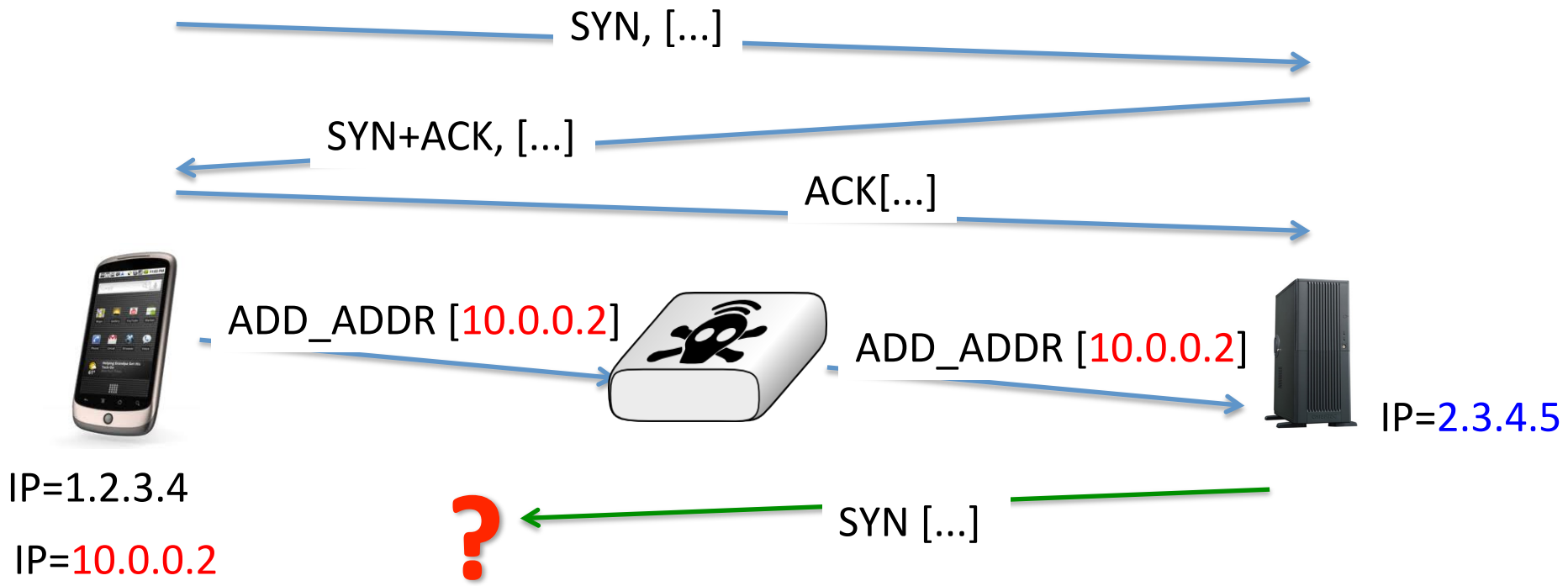


Address dynamics

- Basic solution : mobile client



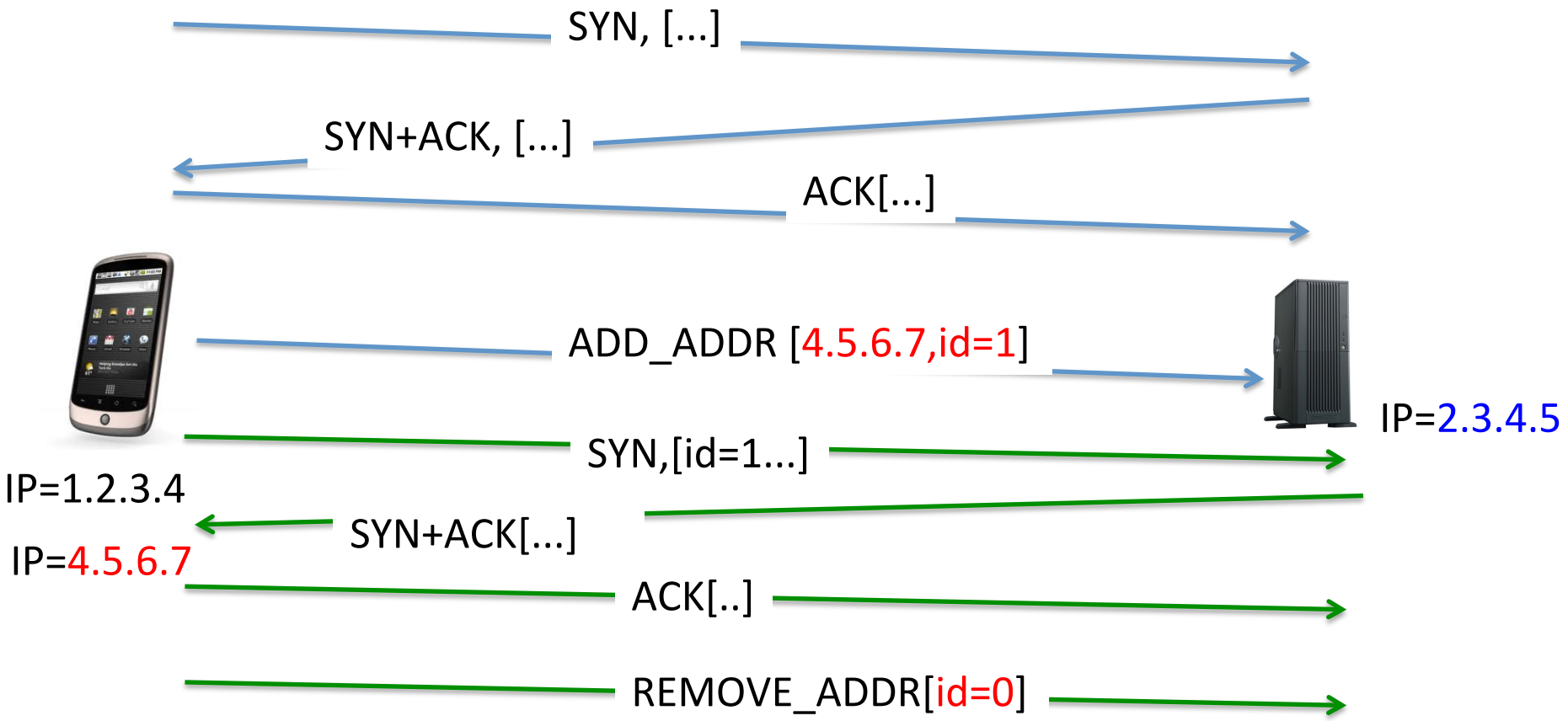
Address dynamics in today's Internet



Address dynamics with NATs

- Solution
  - Each address has one identifier
    - Subflow is established between id=0 addresses
  - Each host maintains a list of <address,id> pairs of the addresses associated to an MPTCP endpoint
  - MPTCP options refer to the address identifier
    - ADD_ADDR contains <address,id>
    - REMOVE_ADDR contains <id>

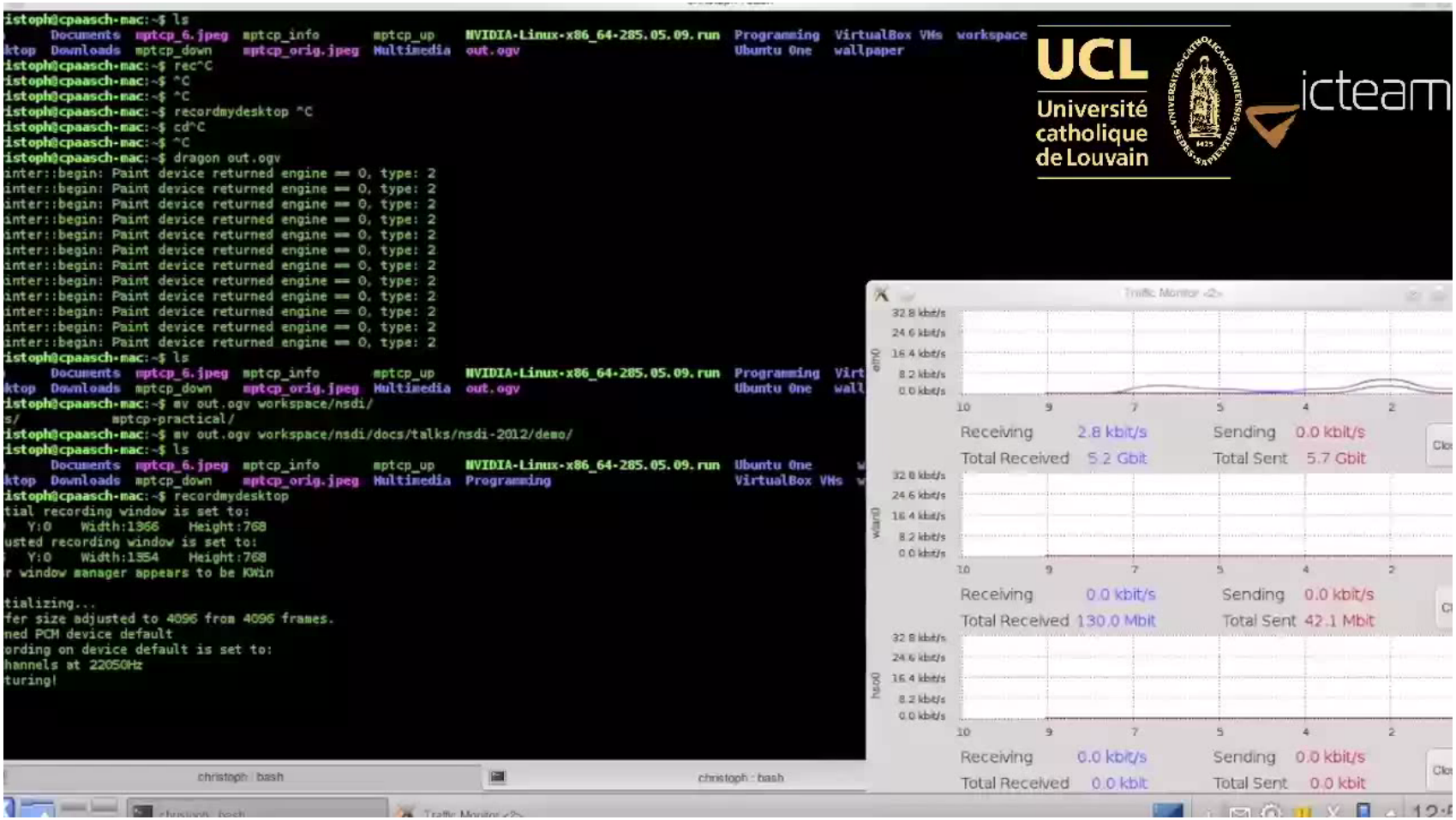
Address dynamics



Agenda

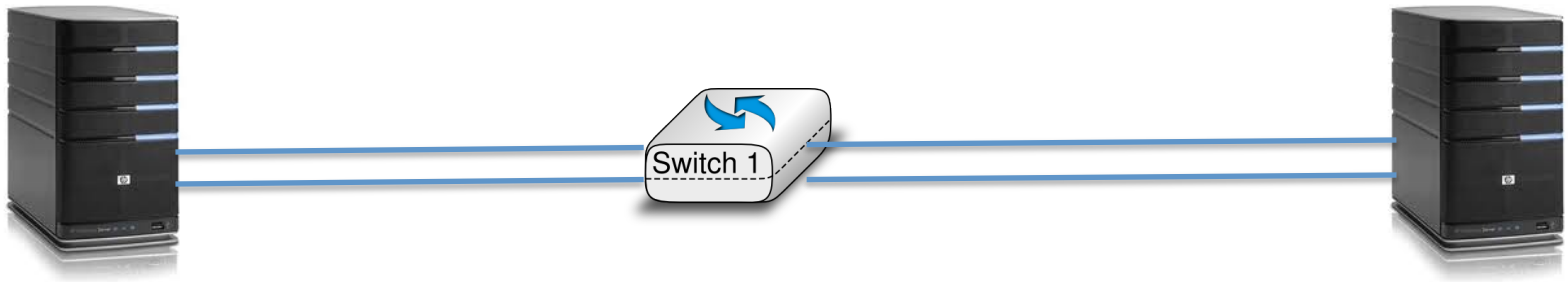
- The motivations for Multipath TCP
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases
  - Datacenters
  - Smartphones
  - IPv4/IPv6 coexistence

ssh with Multipath TCP



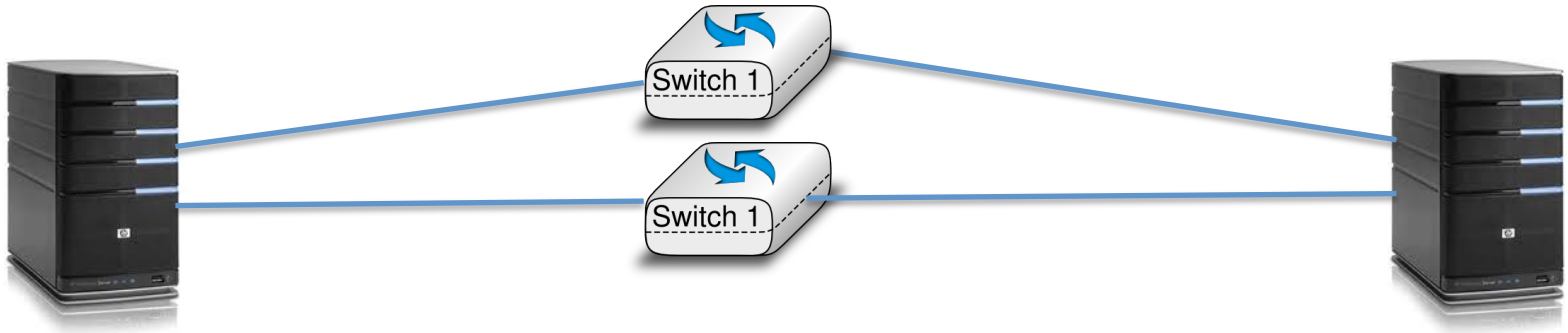
TCP on servers

- How to increase server bandwidth ?



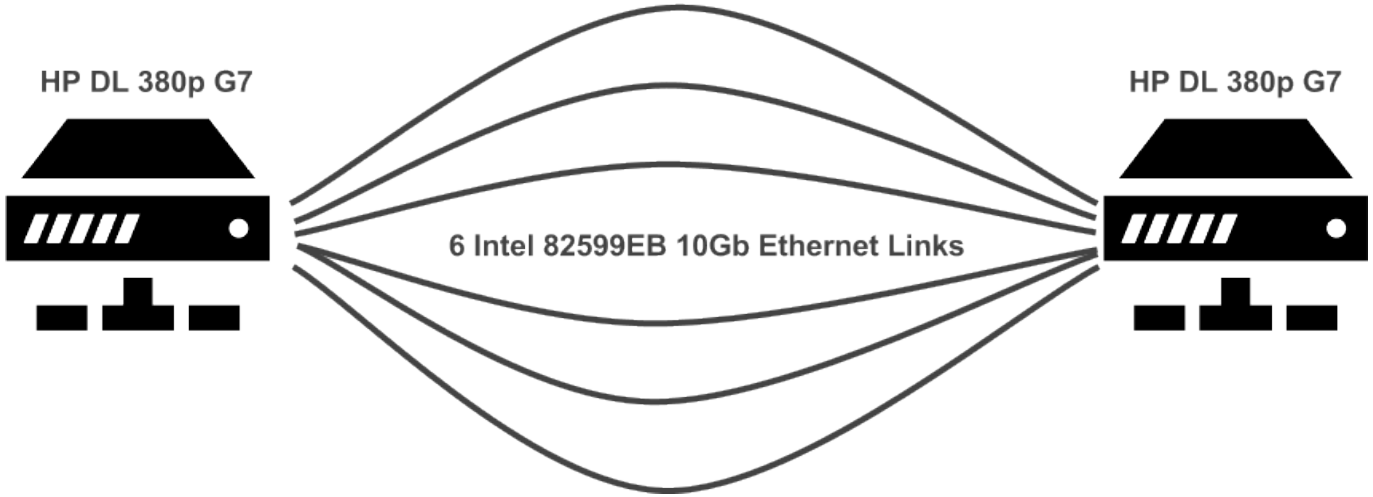
- Load balancing techniques
  - packet per packet
  - per flow load balancing
    - each TCP connection is mapped onto one interface

Increasing server bandwidth with Multipath TCP



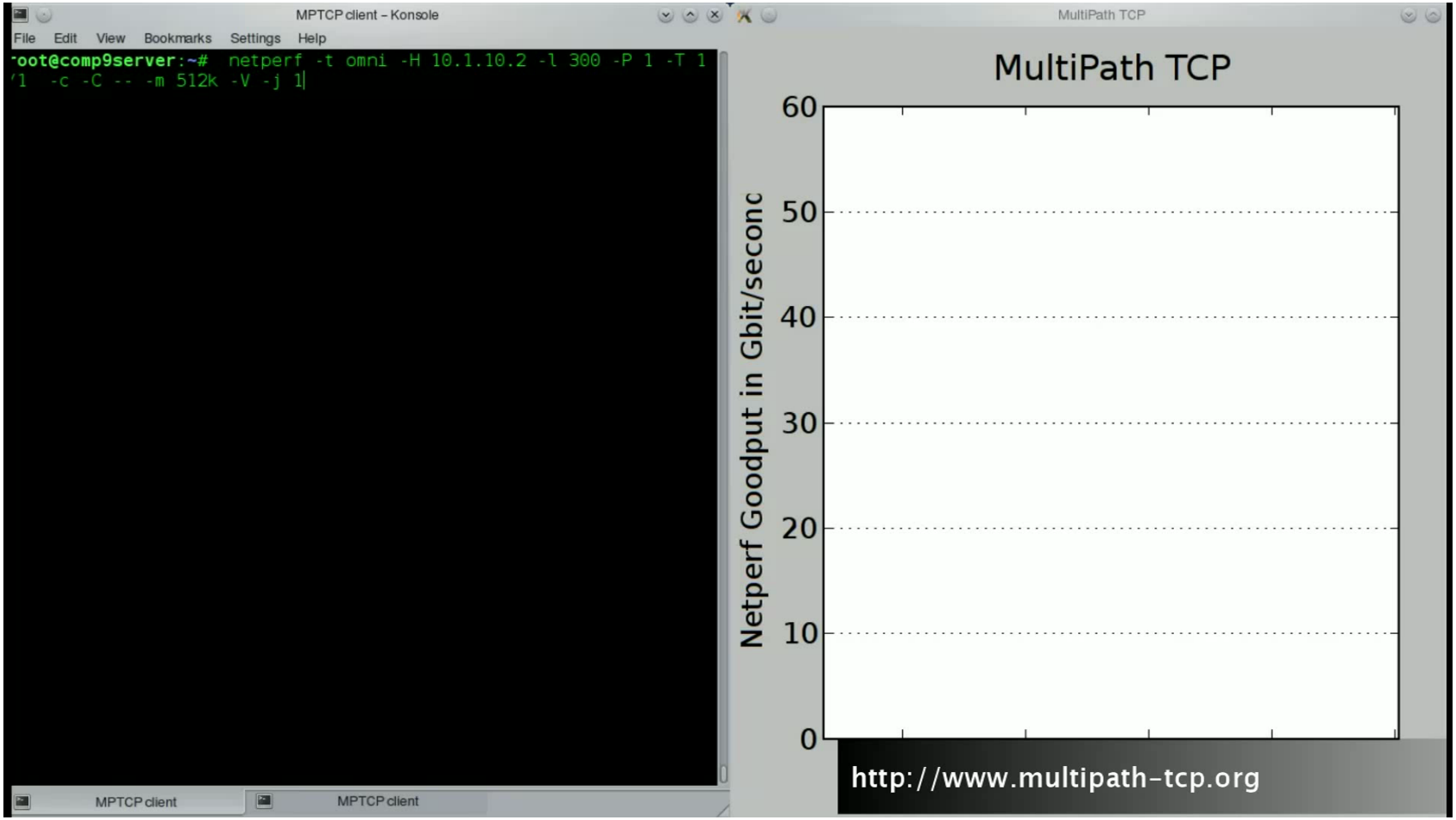
- Load balancing with Multipath TCP
  - Congestion control efficiently uses the two links for each MPTCP connection
  - Automatic failover in case of failures

How fast can Multipath TCP go ?



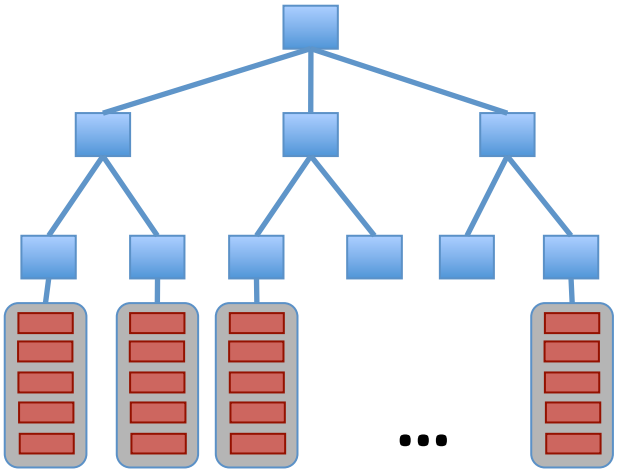
http://linux.slashdot.org/story/13/03/23/0054252/a-50-gbps-connection-with-multipath-tcp

How fast can Multipath TCP go ?



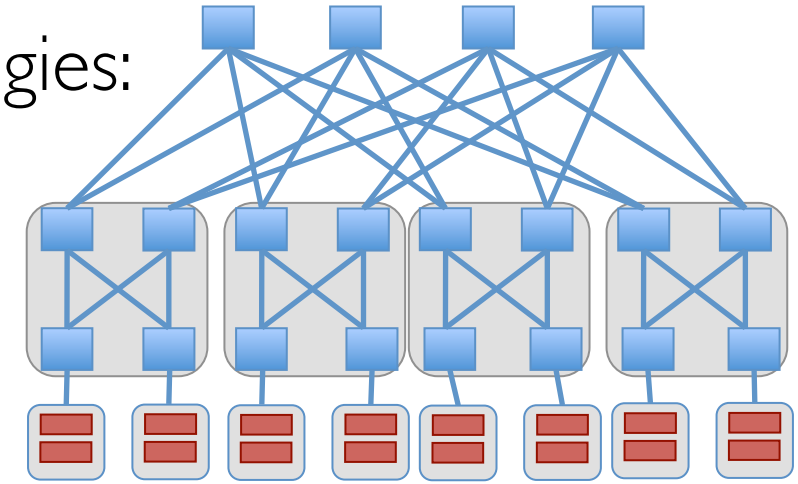
Datacenters evolve

- Traditional Topologies are tree-based
  - Poor performance
  - Not fault tolerant



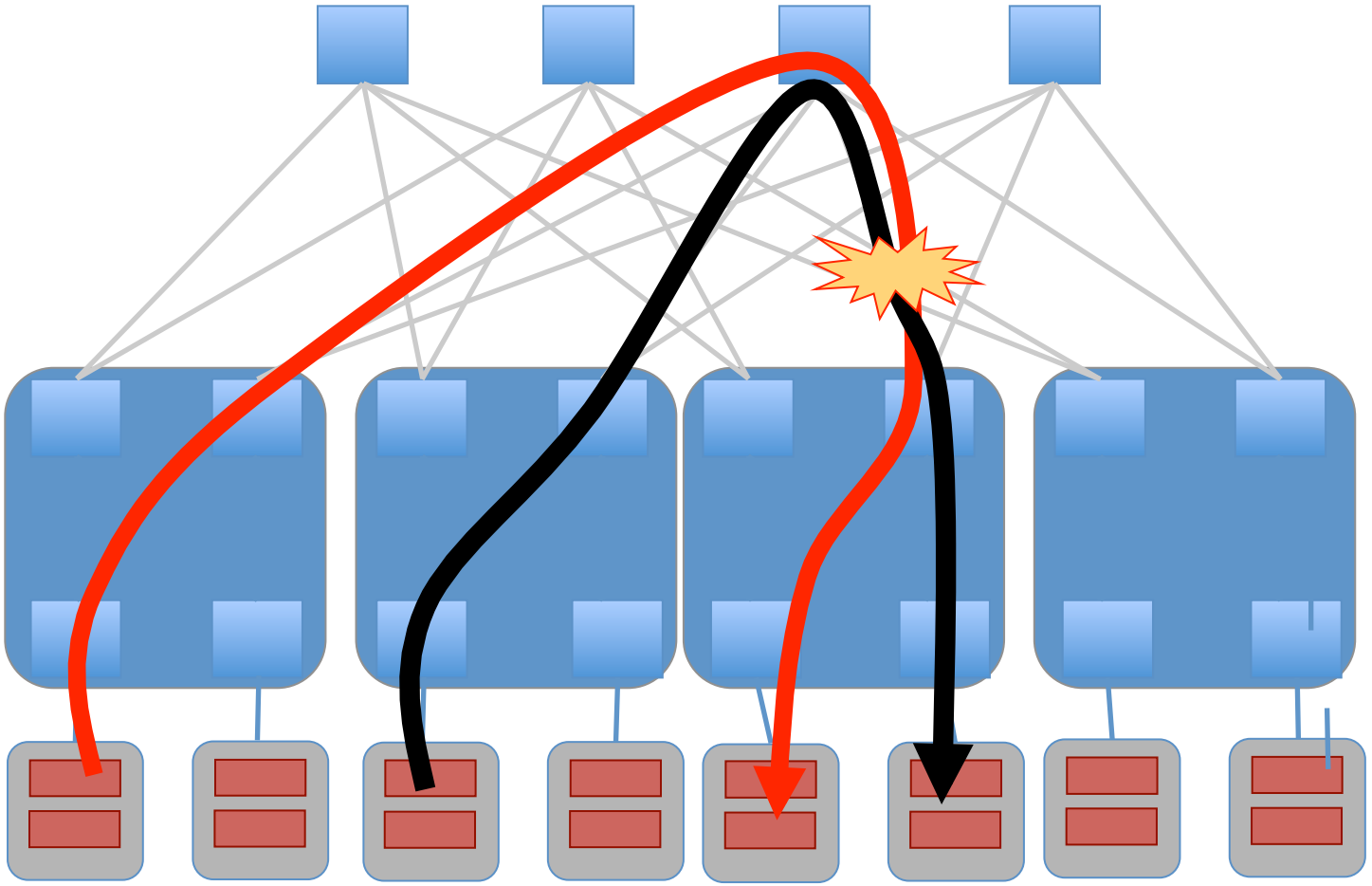
- Shift towards multipath topologies: FatTree, BCube, VL2,

  Cisco, EC2

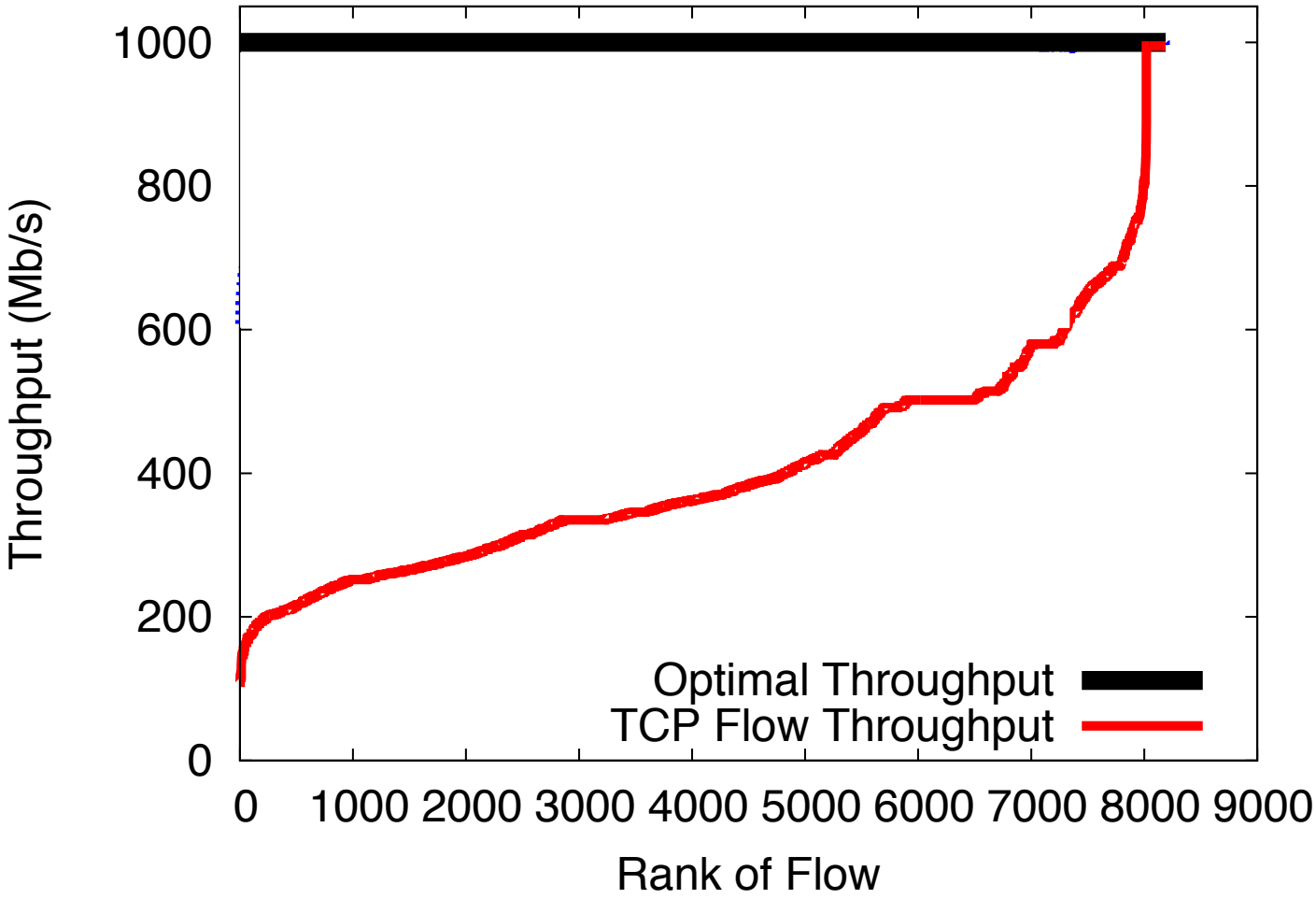


C. Raiciu, et al. "Improving datacenter performance and robustness with multipath TCP," *ACM SIGCOMM* 2011.

TCP in data centers



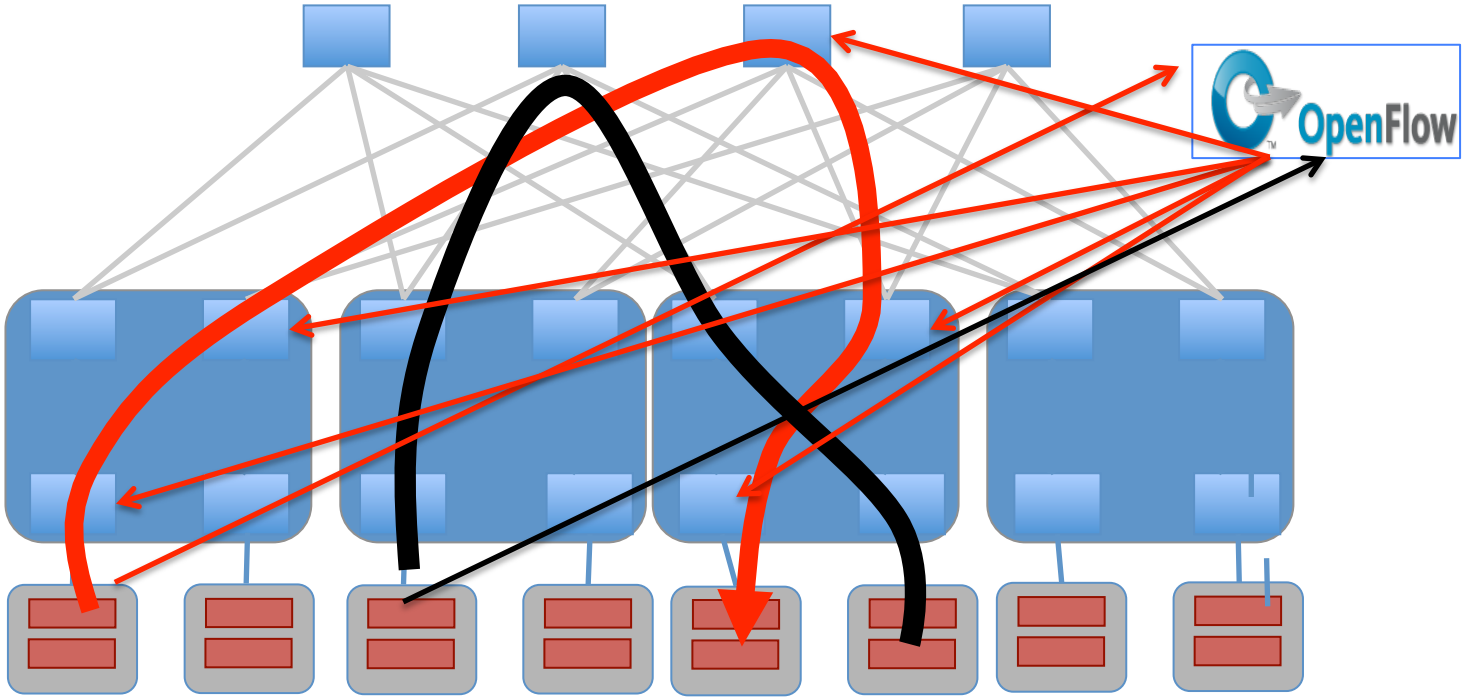
TCP in FAT tree networks
Cost of collissions



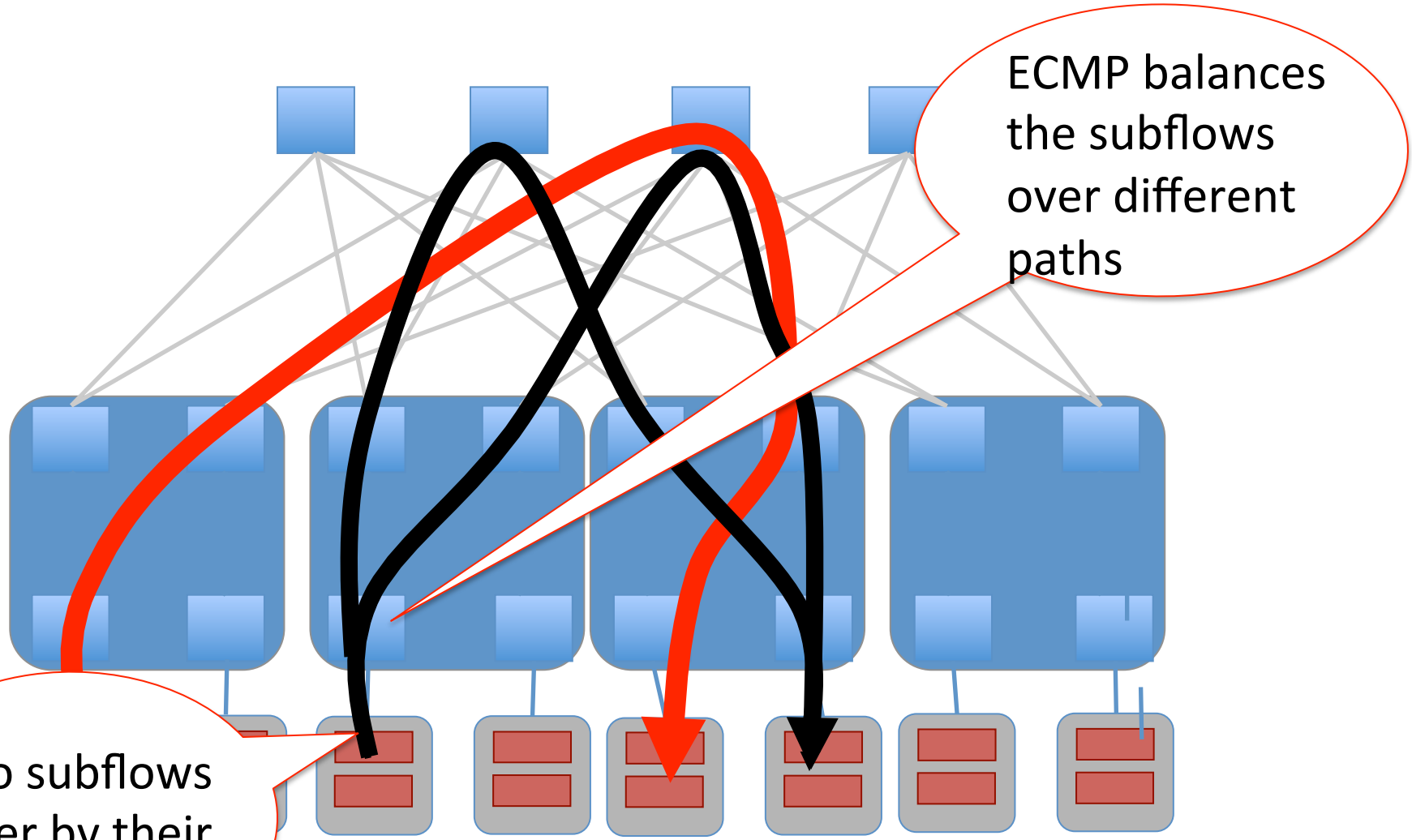
C. Raiciu, et al. "Improving datacenter performance and robustness with multipath TCP," *ACM SIGCOMM* 2011.

How to get rid of these collisions ?

- Consider TCP performance as an optimisation problem

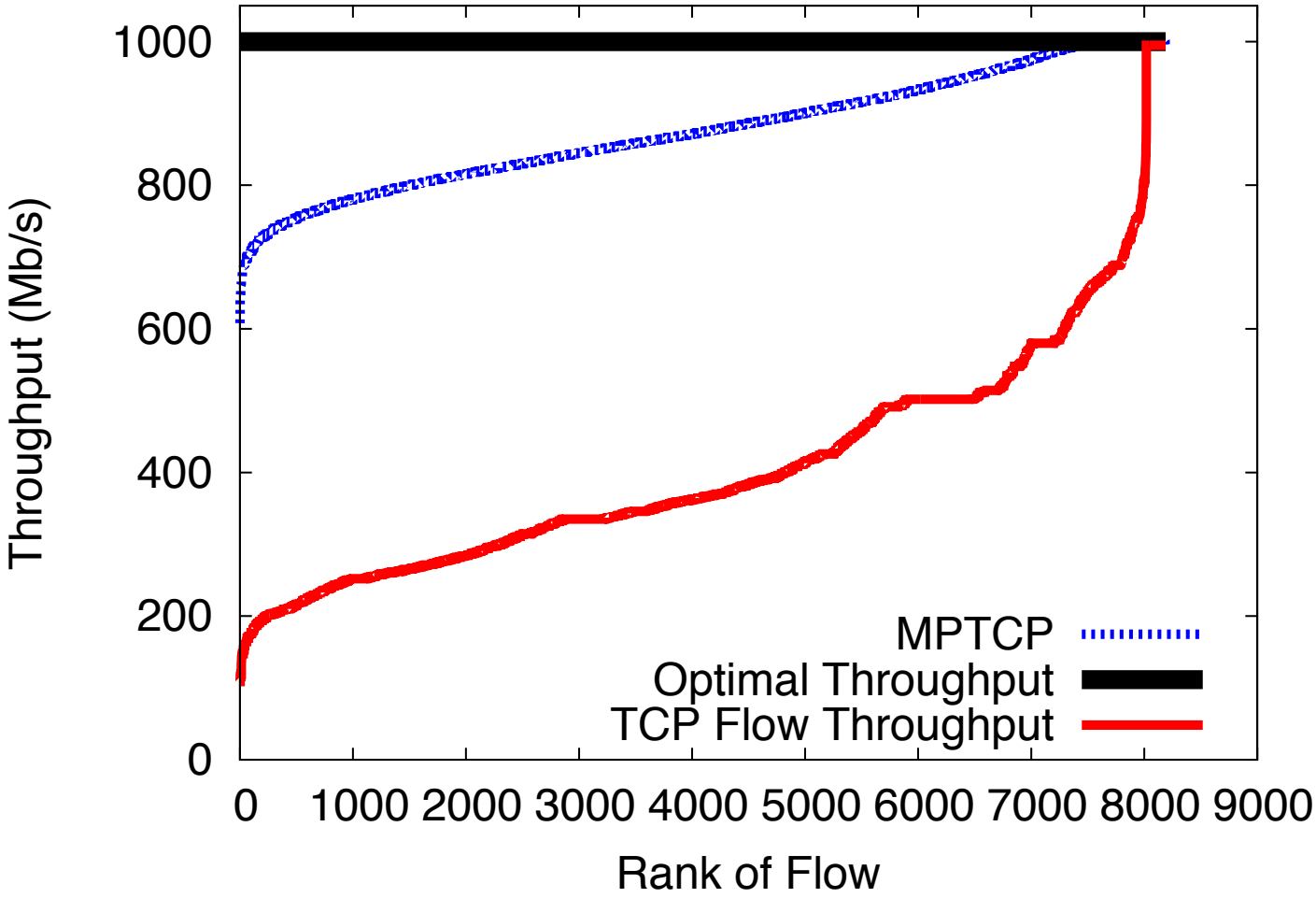


The Multipath TCP way



Improving datacenter performance and robustness with multipath TCP," *ACM* 011.

MPTCP better utilizes the FatTree network



C. Raiciu, et al. "Improving datacenter performance and robustness with multipath TCP," *ACM SIGCOMM* 2011.

See also G. Detal, et al. *, Revisiting Flow-Based Load Balancing: Stateless Path Selection in Data Center Networks*, Computer Networks, April 2013 for extensions to ECMP for MPTCP

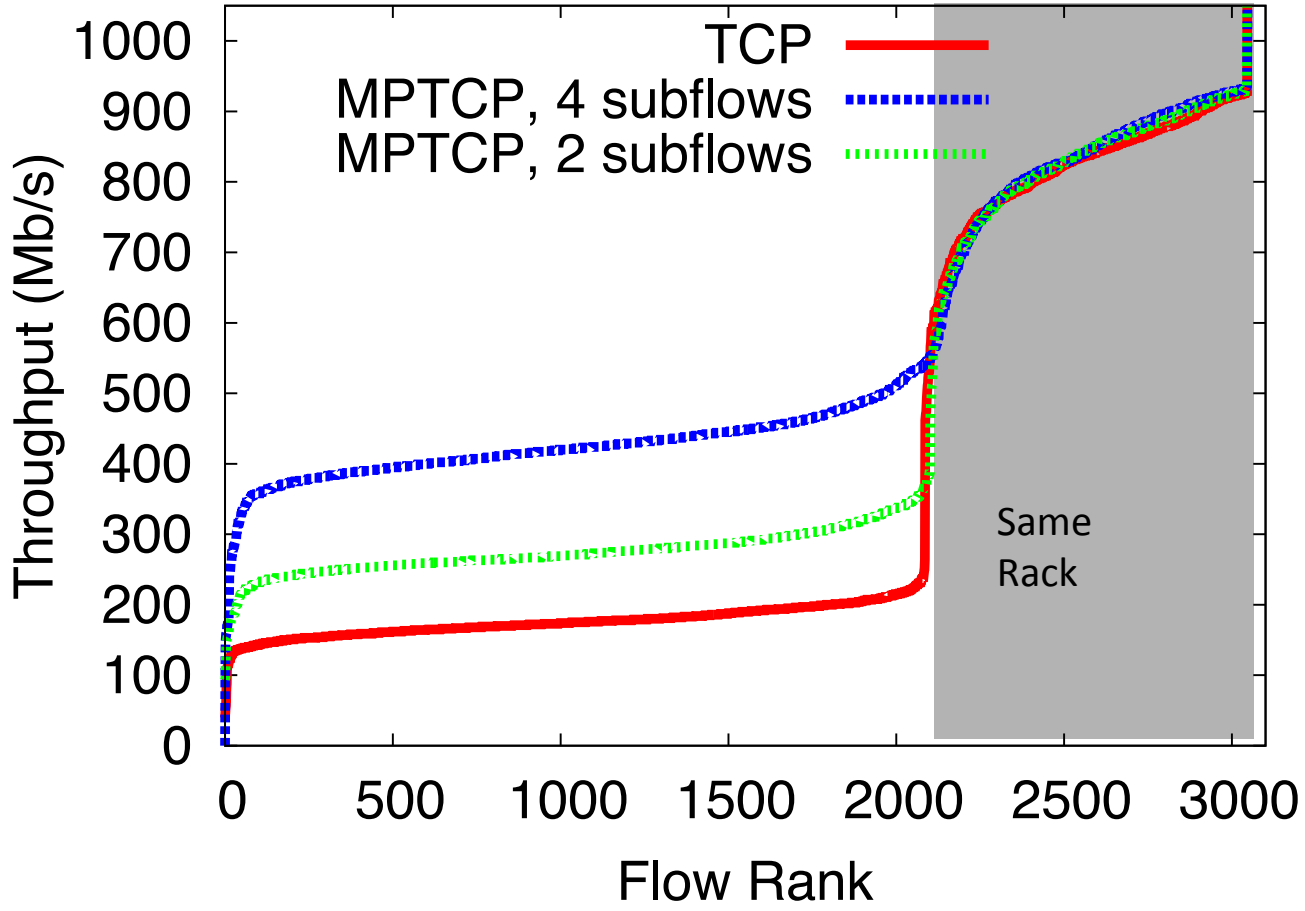
Multipath TCP on EC2

- Amazon EC2: infrastructure as a service
  - We can borrow virtual machines by the hour
  - These run in Amazon data centers worldwide
  - We can boot our own kernel
- A few availability zones have multipath topologies
  - 2-8 paths available between hosts not on the same machine or in the same rack
  - Available via ECMP

Amazon EC2 Experiment

- 40 medium CPU instances running MPTCP
- During 12 hours, we sequentially ran all-to-all `iperf` cycling through:
  - TCP
  - MPTCP (2 and 4 subflows)

MPTCP improves performance on EC2



C. Raiciu, et al. "Improving datacenter performance and robustness with multipath TCP," *ACM SIGCOMM* 2011.

Agenda

- The motivations for Multipath TCP
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases
  - Datacenters
  - Smartphones
  - IPv4/IPv6 coexistence

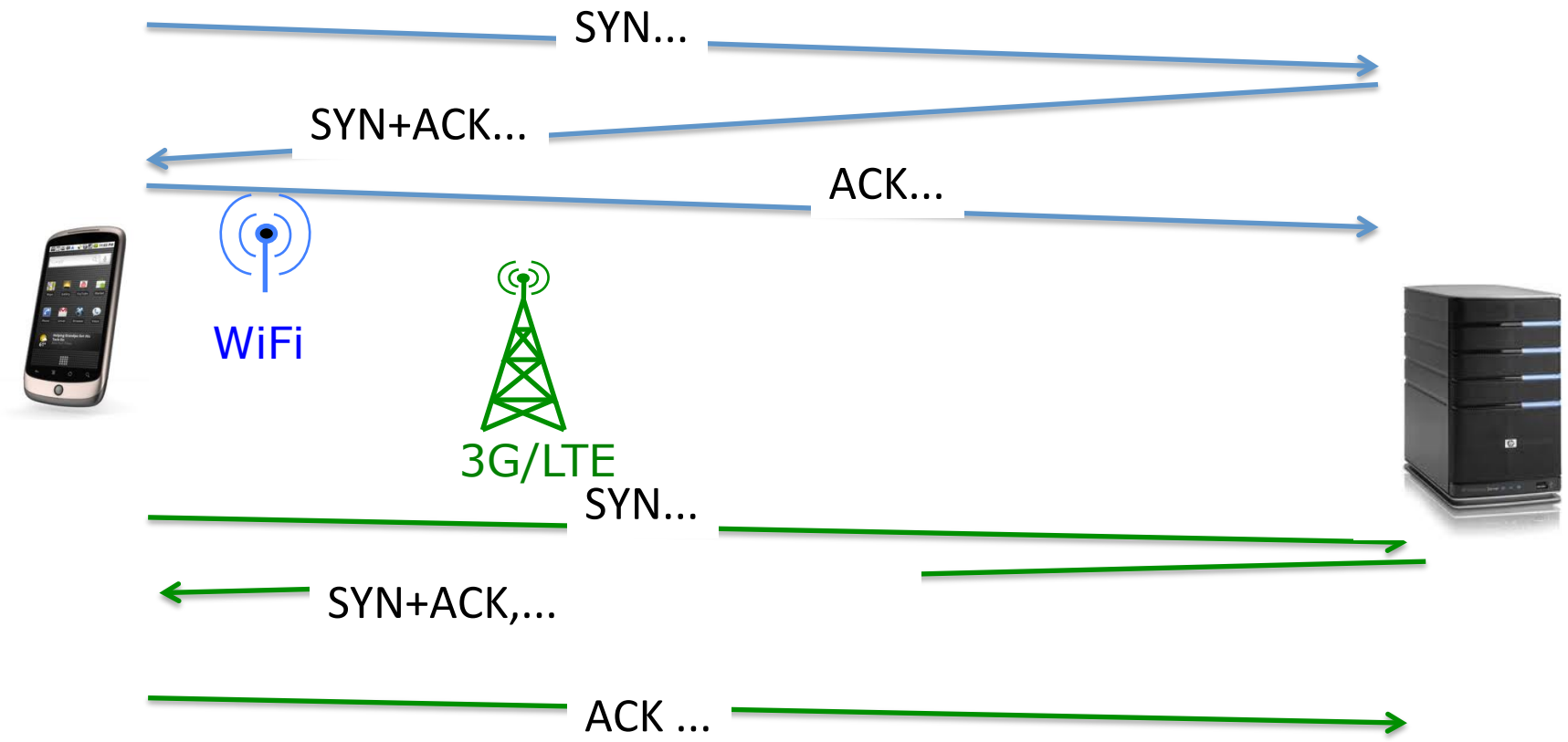
Motivation

- One device, many IP-enabled interfaces

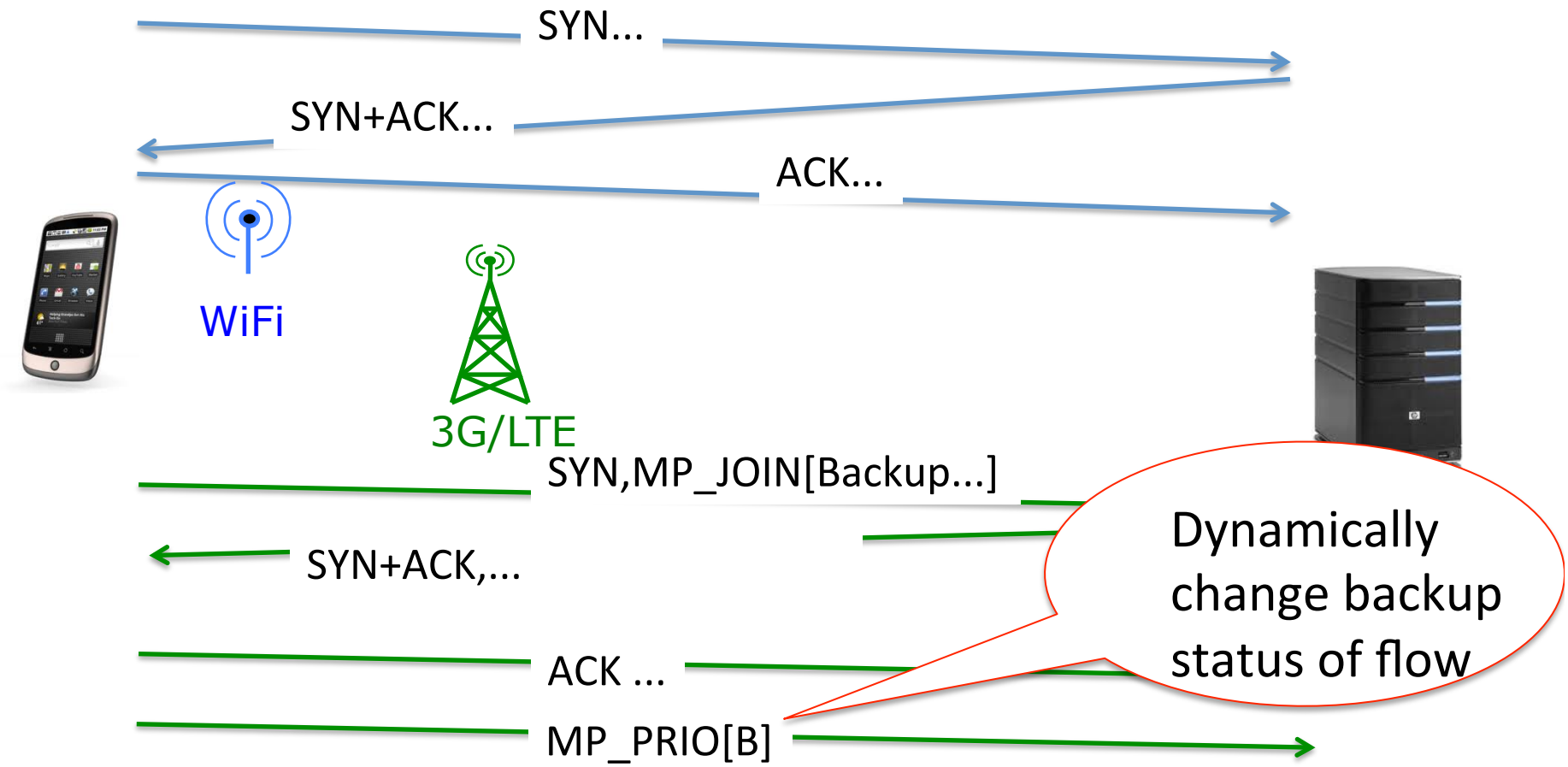
Usage of 3G and WiFI

- How should Multipath TCP use 3G and WiFi ?
  - Full mode
    - Both wireless networks are used at the same time
  - Backup mode
    - Prefer WiFi when available, open subflows on 3G and use them as backup
  - Single path mode
    - Only one path is used at a time, WiFi preferred over 3G

Multipath TCP : Full mode

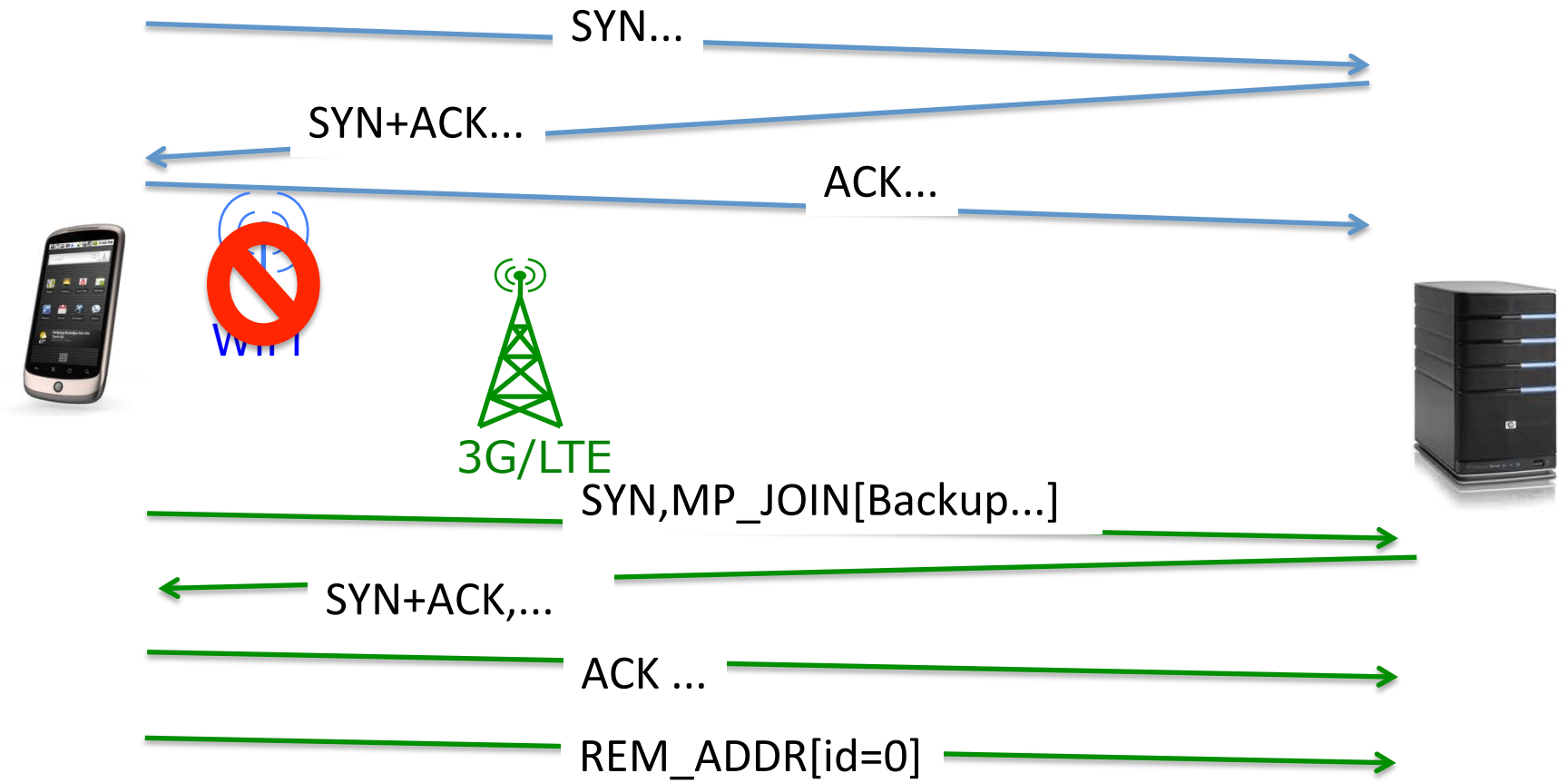


Multipath TCP : Backup mode

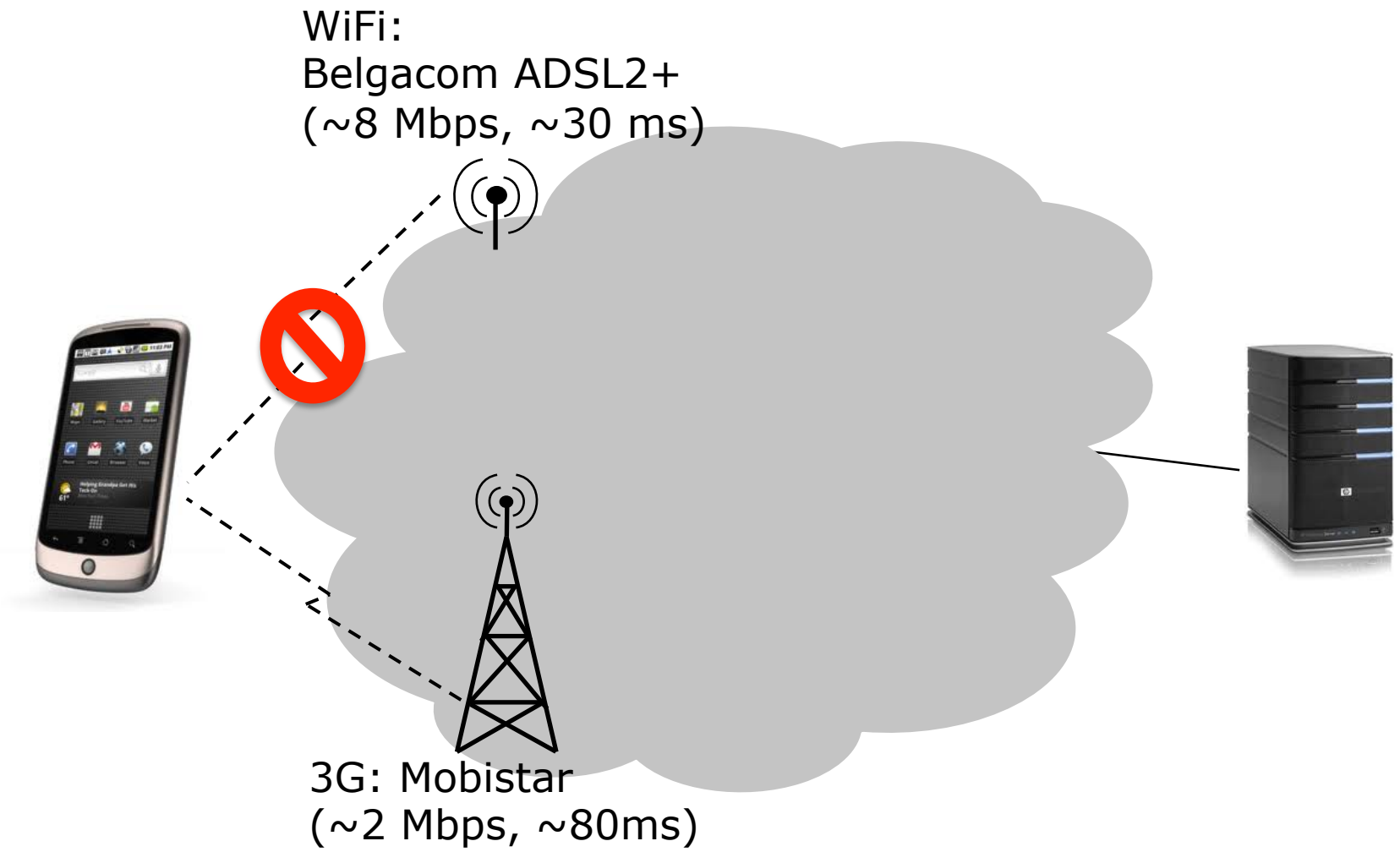


Multipath TCP : Backup mode

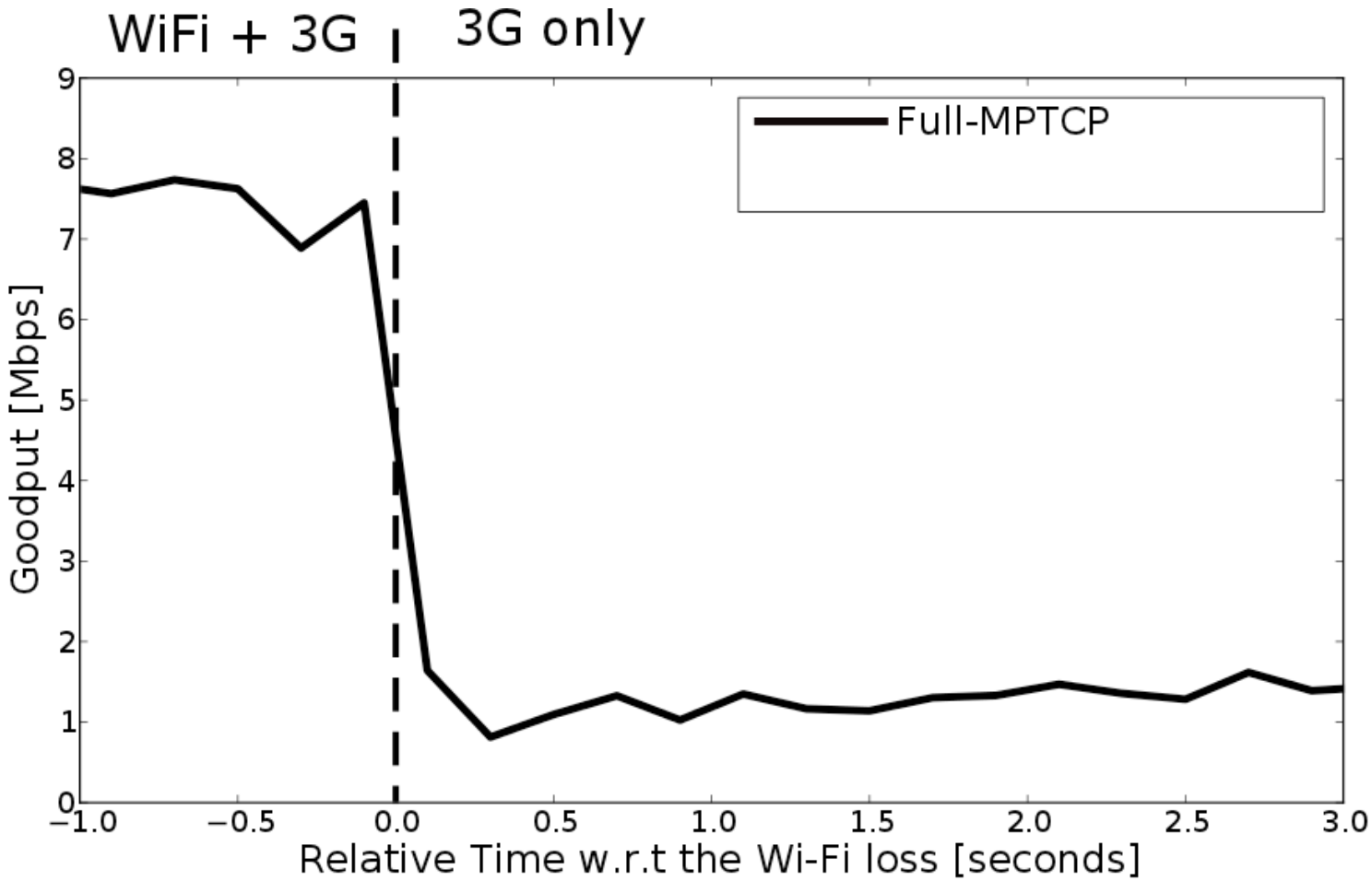
- What happens when link fails ?



Evaluation scenario

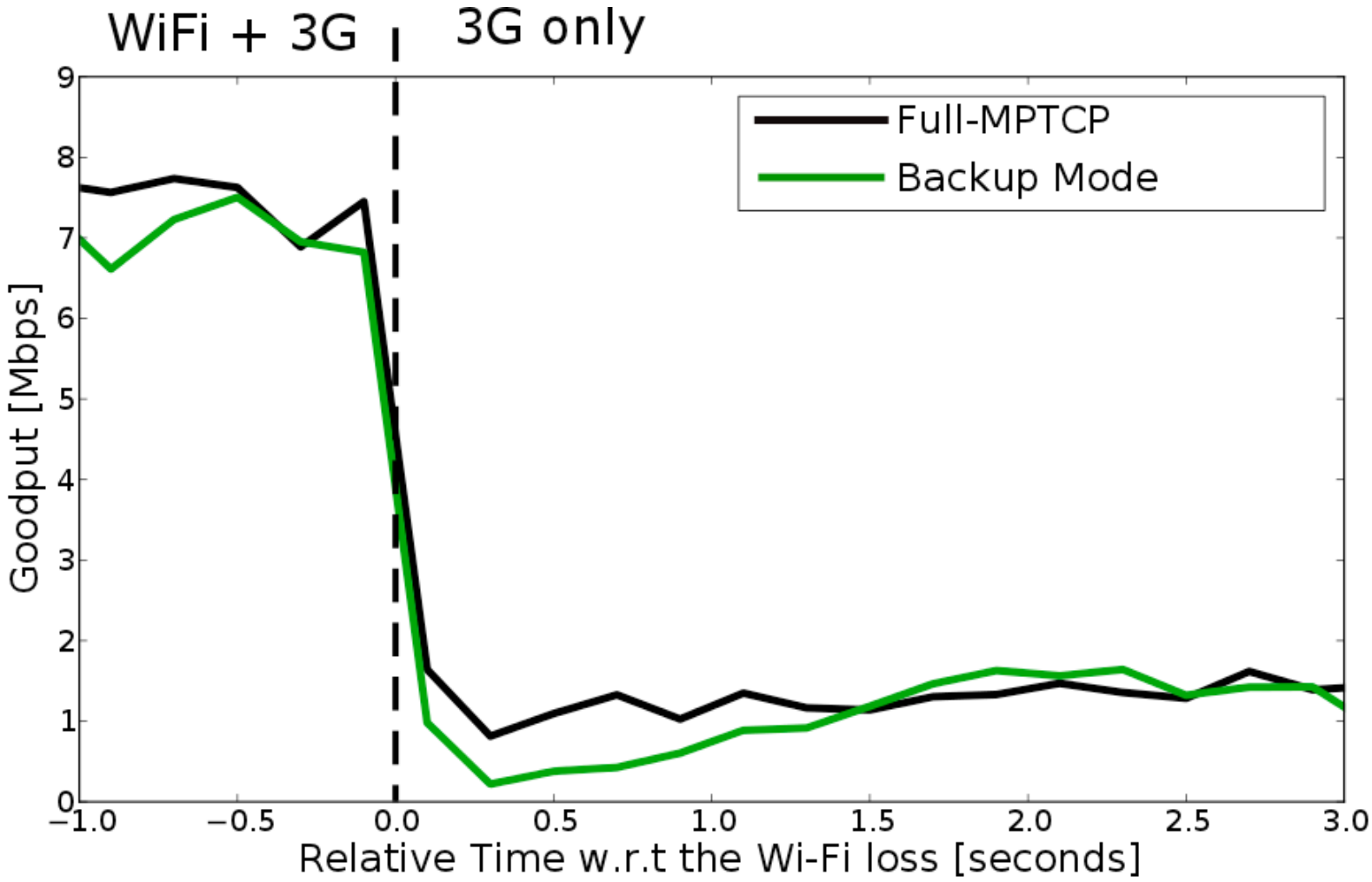


Recovery after failure



C. Paasch, et al. , “Exploring mobile/WiFi handover with multipath TCP,” presented at the CellNet '12: Proceedings of the 2012 ACM SIGCOMM workshop on Cellular networks: operations, challenges, and future design, 2012.

Recovery after failure

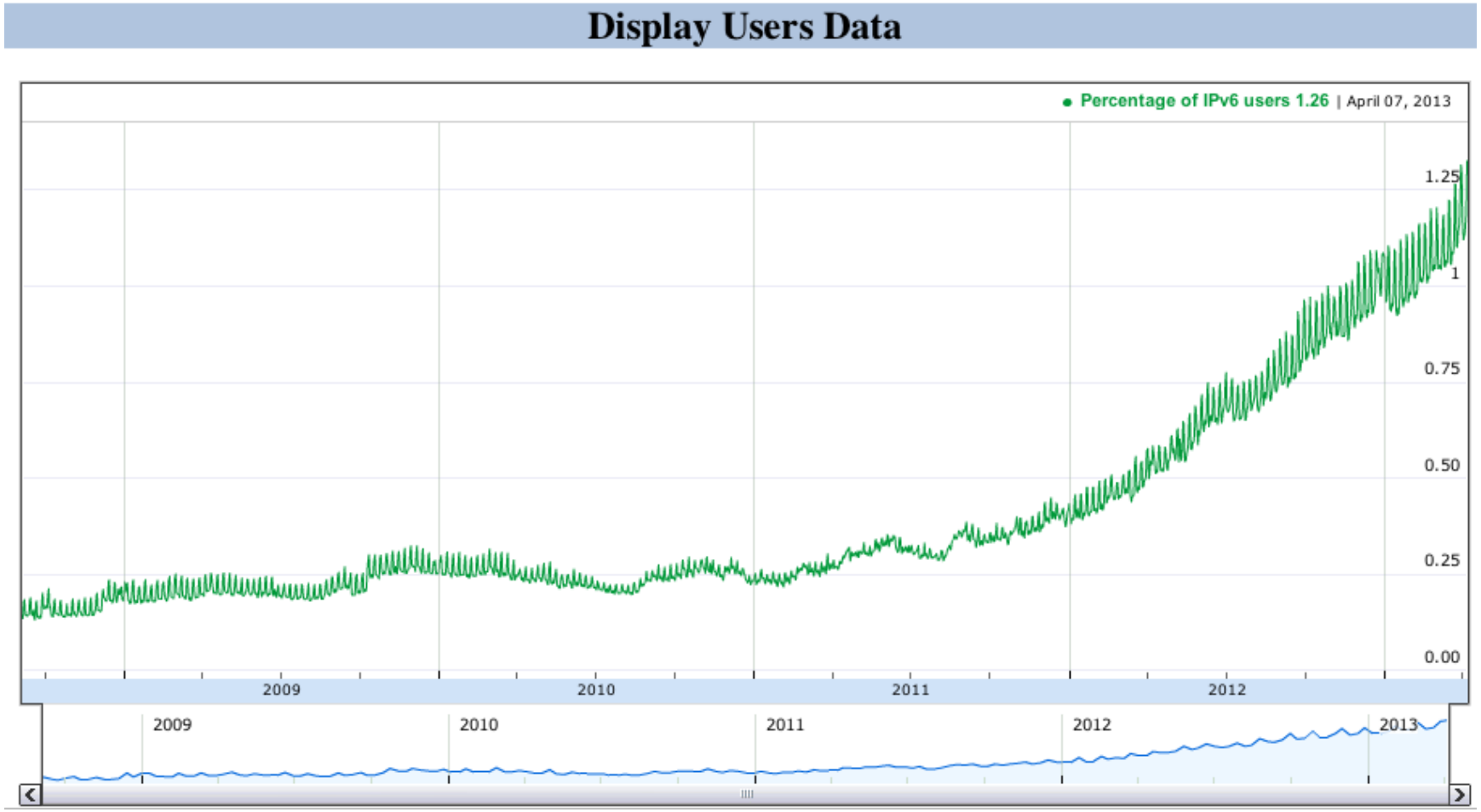


C. Paasch, et al. , “Exploring mobile/WiFi handover with multipath TCP,” presented at the CellNet '12: Proceedings of the 2012 ACM SIGCOMM workshop on Cellular networks: operations, challenges, and future design, 2012.

Agenda

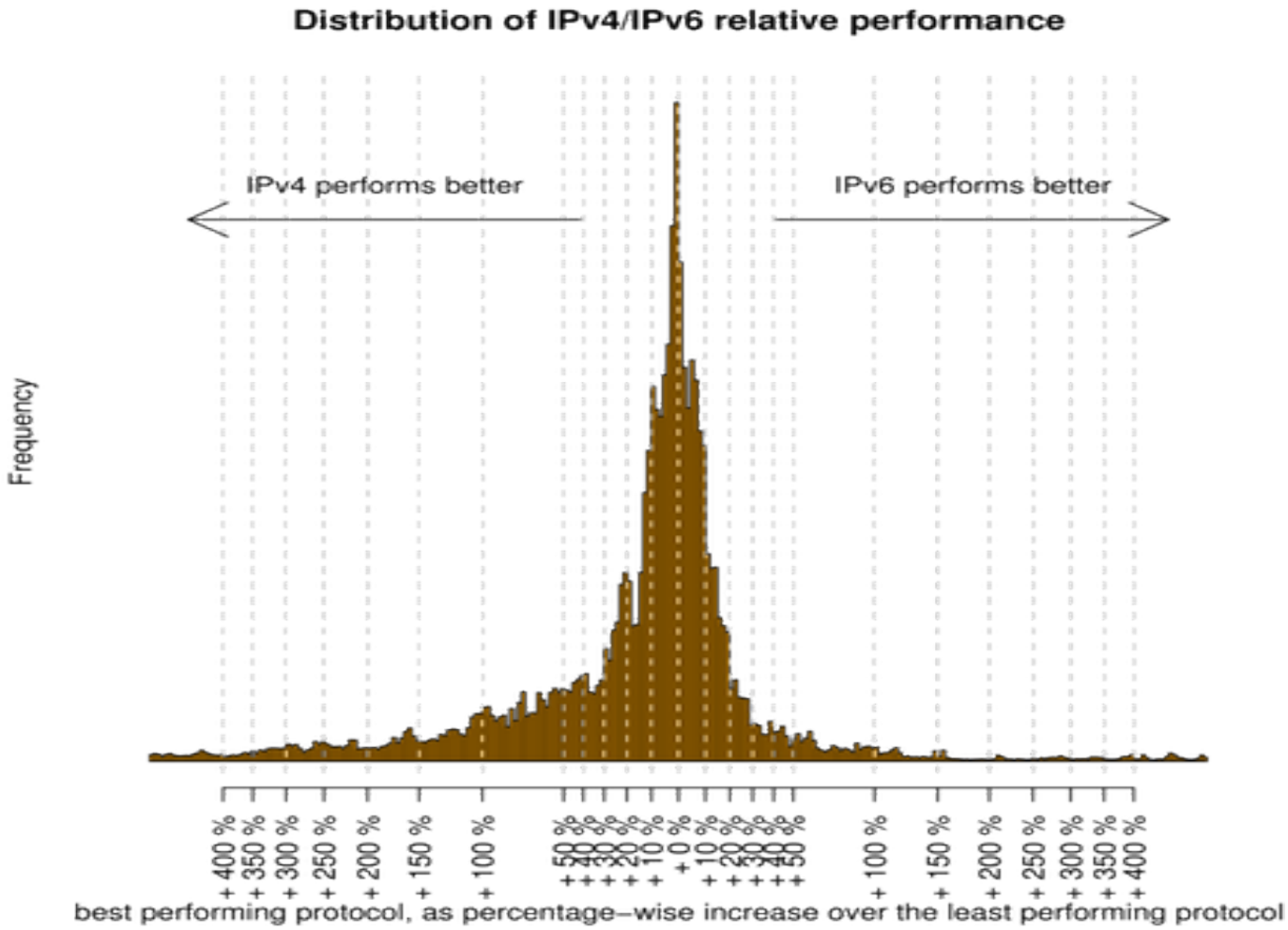
- The motivations for Multipath TCP
- The changing Internet
- The Multipath TCP Protocol
- Multipath TCP use cases
  - Datacenters
  - Smartphones
  - IPv4/IPv6 coexistence

IPv6 is coming ...



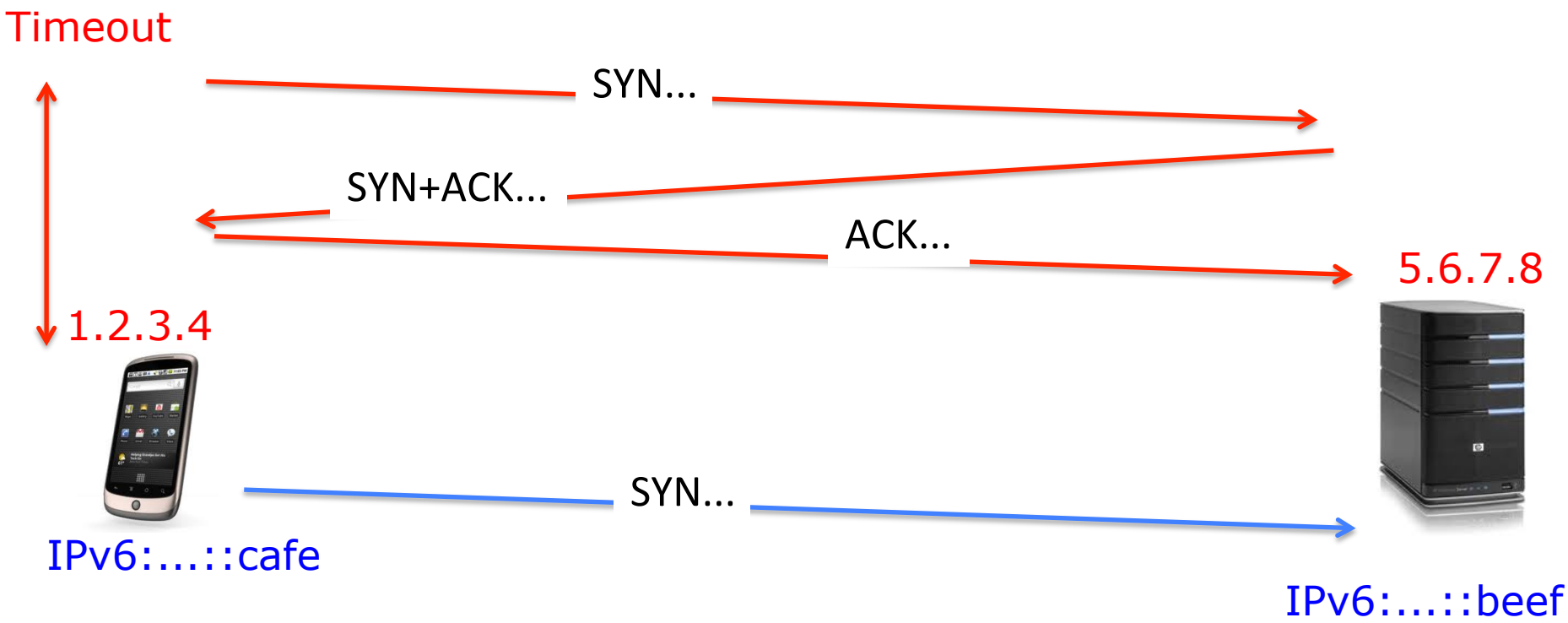
Source http://6lab.cisco.com/stats/cible.php?country=world

But IPv4 and IPv6 perf. may differ

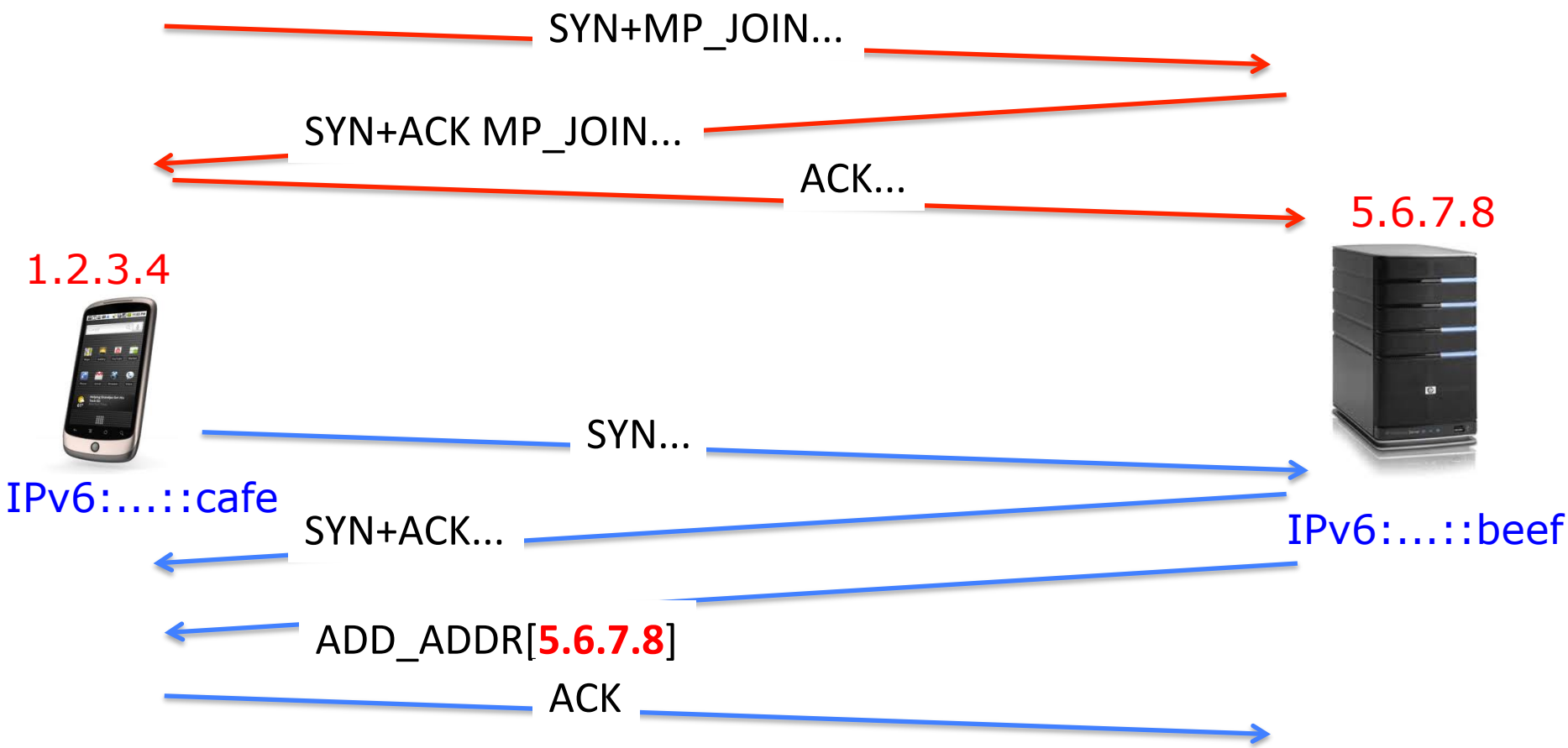


E. Aben, *Measuring World IPv6 Day - Comparing IPv4 and IPv6 Performance*, https://labs.ripe.net/Members/emileaben/measuring-world-ipv6-day-comparing-ipv4-and-ipv6-performance

Happy eyeballs

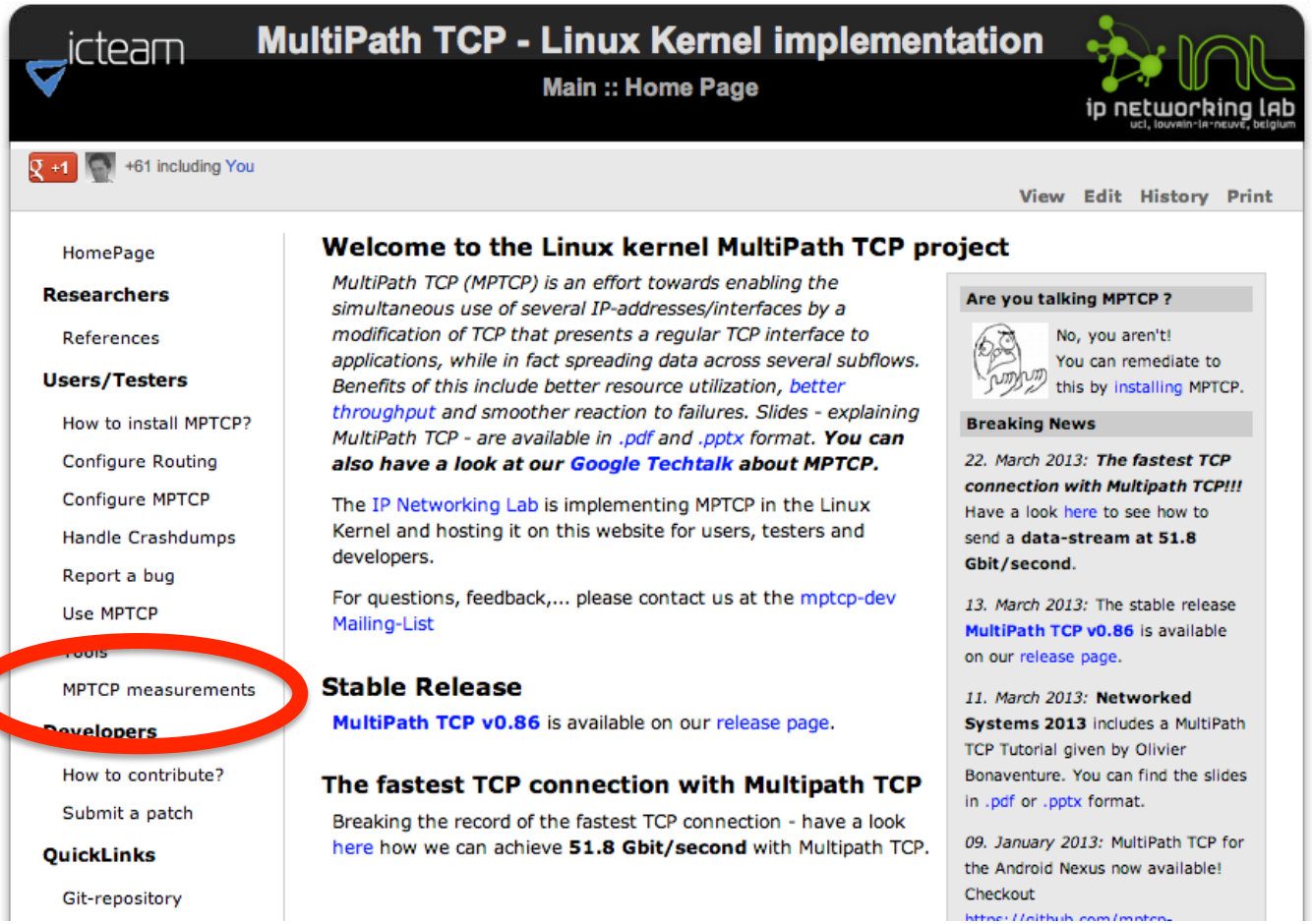


How to get best of IPv4 and IPv6 ?



Try it by yourself !

http://multipath-tcp.org

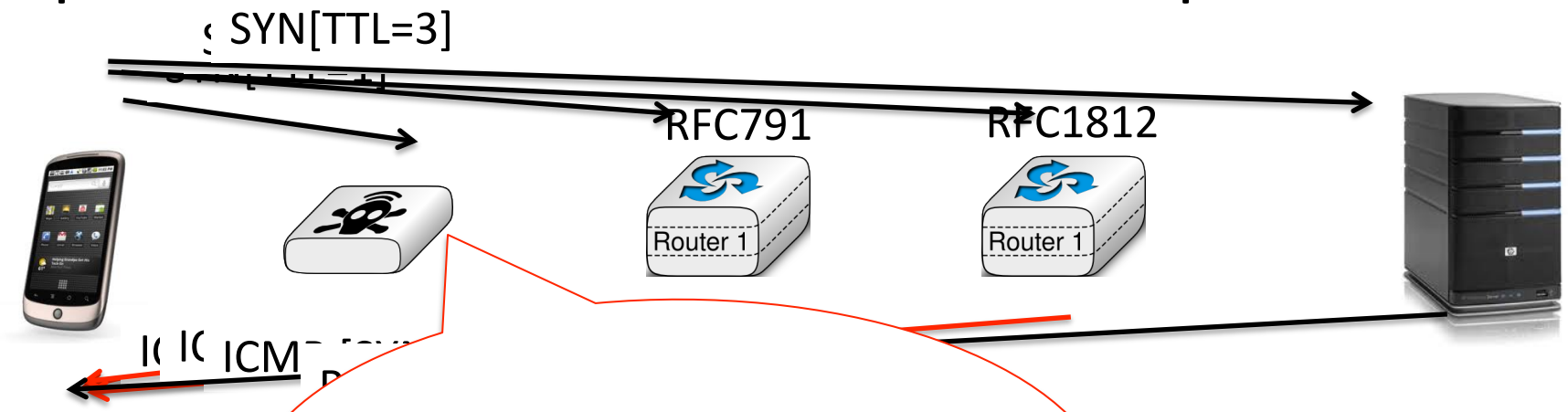


Testing Multipath TCP through **your** network

- Tests included
  - tracebox
  - HTTP, HTTPS, SCP, FTP over MPTCP with
    - a single TCP subflow
    - 4 TCP subflows
      - segments sent over best subflow
      - segments sent in round-robin
      - segments duplicated over all subflows
- Theses tests include corner cases that might trigger reactions from firewalls/ DPI/IDS/...

tracebox

Work in progress, but will probably have wide operational use to debug middlebox problems

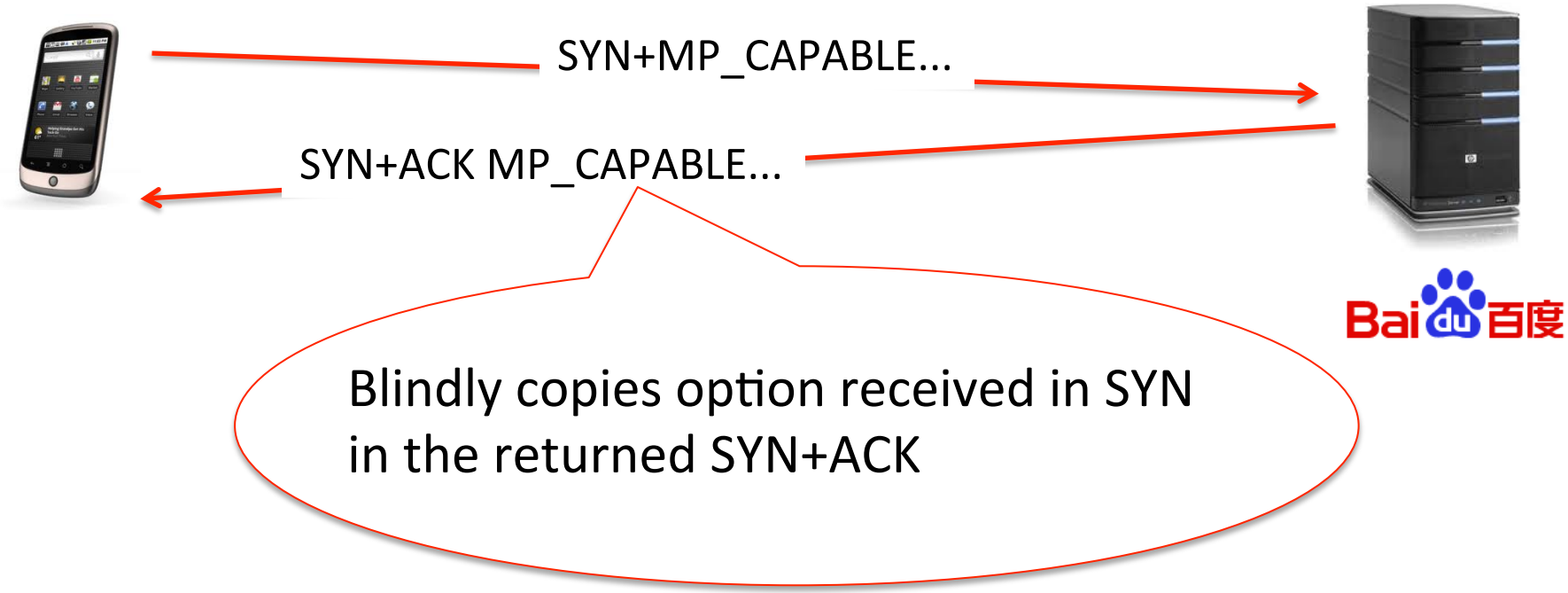



```
tracebox -
tracebox -
1: 2.2.2.2
2: 3.3.3.3 [
3: 4.4.4.4 [TCP op                    ged] [Reply ICMP full pkt]
4: 5.5.5.5
```

First results

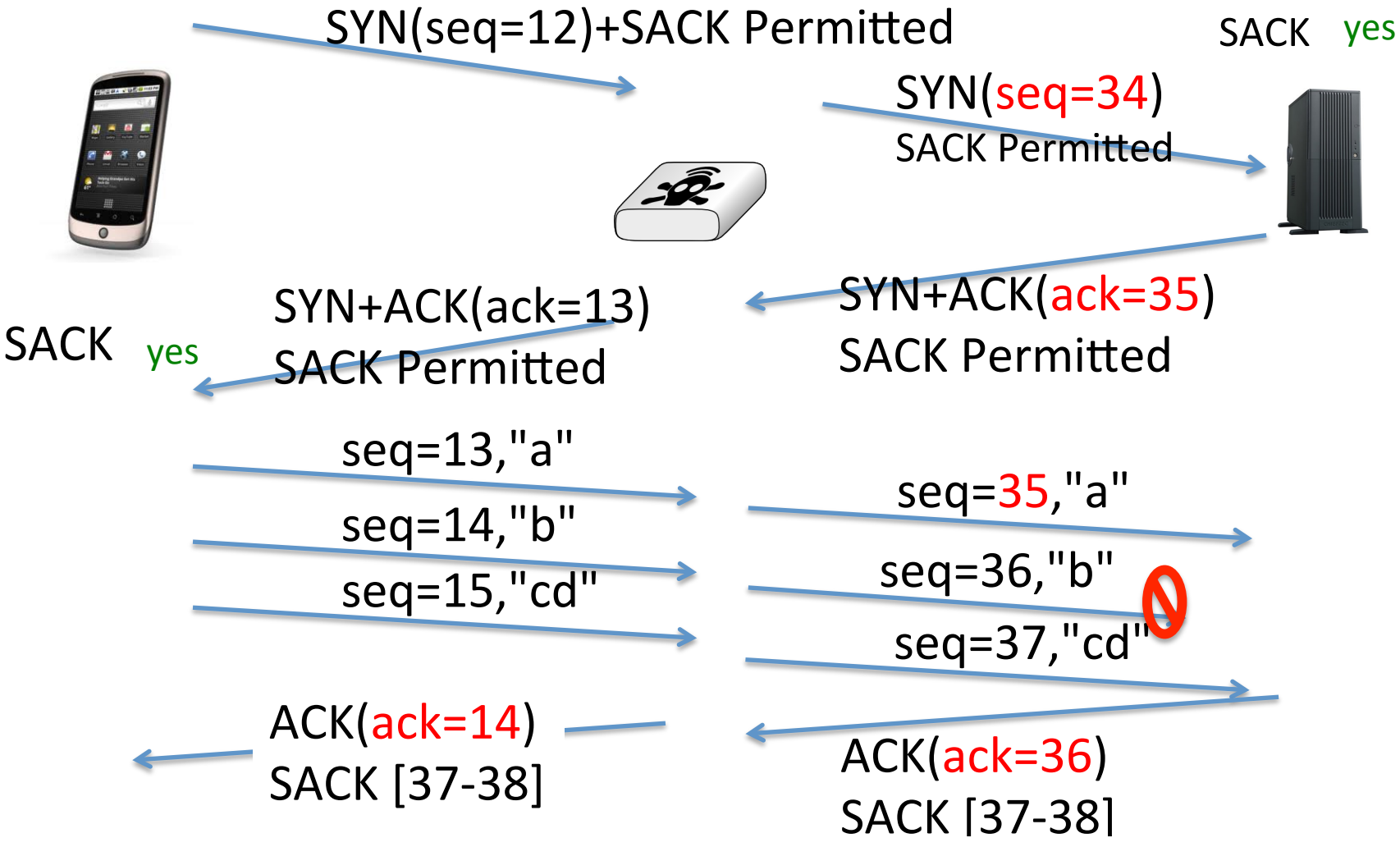
- Multipath TCP has been tested in a few dozen of networks
  - Works in most tested networks, no major problem identified
    - We are looking for tests from networks with known middleboxes
  - In some cases, Multipath TCP fallbacks to regular TCP
    - Middlebox dropping options in SYN
    - NAT translating PORT command for FTP
    - Fallback works as expected

Hard to debug problems



- Be liberal in what you accept conservative in what you send

TCP sequence number randomization and SACK



Conclusion

- Multipath TCP is becoming a reality
  - Due to the middleboxes, the protocol is more complex than initially expected
  - RFC6824 has been published
  - there is running code !
  - Multipath TCP works over today's Internet !
- What's next ?
  - More use cases
    - BGP over MPTCP, anycast, VM migration, ...

References

- The Multipath TCP protocol
  - http://www.multipath-tcp.org
  - http://tools.ietf.org/wg/mptcp/

A. Ford, C. Raiciu, M. Handley, S. Barre, and J. Iyengar, “Architectural guidelines for multipath TCP development", RFC6182 2011.

A. Ford, C. Raiciu, M. J. Handley, and O. Bonaventure, “TCP Extensions for Multipath Operation with Multiple Addresses,” RFC6824, 2013

C. Raiciu, C. Paasch, S. Barre, A. Ford, M. Honda, F. Duchene, O. Bonaventure, and M. Handley, “How hard can it be? designing and implementing a deployable multipath TCP,” NSDI'12: Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation, 2012.

Implementations

- Linux
  - http://www.multipath-tcp.org

    S. Barre, C. Paasch, and O. Bonaventure, "Multipath tcp: From theory to practice," *NETWORKING 2011*, 2011.

    Sébastien Barré. Implementation and assessment of Modern Host-based Multipath Solutions. PhD thesis. UCL, 2011
- FreeBSD
  - http://caia.swin.edu.au/urp/newtcp/mptcp/
- Simulators
  - http://nrg.cs.ucl.ac.uk/mptcp/implementation.html
  - http://code.google.com/p/mptcp-ns3/

Middleboxes

M. Honda, Y. Nishida, C. Raiciu, A. Greenhalgh, M. Handley, and H. Tokuda, “Is it still possible to extend TCP?,” IMC '11: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, 2011.

V. Sekar, N. Egi, S. Ratnasamy, M. K. Reiter, and G. Shi, “Design and implementation of a consolidated middlebox architecture,” *USENIX NSDI*, 2012.

J. Sherry, S. Hasan, C. Scott, A. Krishnamurthy, S. Ratnasamy, and V. Sekar, “Making middleboxes someone else's problem: network processing as a cloud service,” SIGCOMM '12: Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication, 2012.

Multipath congestion control

- Background

  D. Wischik, M. Handley, and M. B. Braun, "The resource pooling principle," *ACM SIGCOMM Computer …*, vol. 38, no. 5, 2008.

  F. Kelly and T. Voice. Stability of end-to-end algorithms for joint routing and rate control. ACM SIGCOMM CCR, 35, 2005.

  P. Key, L. Massoulie, and P. D. Towsley, "Path Selection and Multipath Congestion Control," INFOCOM 2007. 2007, pp. 143–151.

- Coupled congestion control

  C. Raiciu, M. J. Handley, and D. Wischik, "Coupled Congestion Control for Multipath Transport Protocols," *RFC*, vol. 6356, Oct. 2011.

  D. Wischik, C. Raiciu, A. Greenhalgh, and M. Handley, "Design, implementation and evaluation of congestion control for multipath TCP," NSDI'11: Proceedings of the 8th USENIX conference on Networked systems design and implementation, 2011.

Multipath congestion control

– More

R. Khalili, N. Gast, M. Popovic, U. Upadhyay, J.-Y. Le Boudec , MPTCP is not Pareto-optimal: Performance issues and a possible solution, Proc. ACM Conext 2012

Y. Cao, X. Mingwei, and X. Fu, “Delay-based Congestion Control for Multipath TCP,” ICNP2012, 2012.

T. A. Le, C. S. Hong, and E.-N. Huh, “Coordinated TCP Westwood congestion control for multiple paths over wireless networks,” ICOIN '12: Proceedings of the The International Conference on Information Network 2012, 2012, pp. 92–96.

T. A. Le, H. Rim, and C. S. Hong, “A Multipath Cubic TCP Congestion Control with Multipath Fast Recovery over High Bandwidth-Delay Product Networks,” *IEICE Transactions*, 2012.

T. Dreibholz, M. Becke, J. Pulinthanath, and E. P. Rathgeb, “Applying TCP-Friendly Congestion Control to Concurrent Multipath Transfer,” Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on, 2010, pp. 312–319.

Use cases

- Datacenter

  C. Raiciu, S. Barre, C. Pluntke, A. Greenhalgh, D. Wischik, and M. J. Handley, "Improving datacenter performance and robustness with multipath TCP," *ACM SIGCOMM* 2011.

  G. Detal, Ch. Paasch, S. van der Linden, P. Mérindol, G. Avoine, O. Bonaventure, *Revisiting Flow-Based Load Balancing: Stateless Path Selection in Data Center Networks*, Computer Networks, April 2013

- Mobile

  C. Pluntke, L. Eggert, and N. Kiukkonen, "Saving mobile device energy with multipath TCP," *MobiArch '11: Proceedings of the sixth international workshop on MobiArch*, 2011.

  C. Paasch, G. Detal, F. Duchene, C. Raiciu, and O. Bonaventure, "Exploring mobile/WiFi handover with multipath TCP," CellNet '12: Proceedings of the 2012 ACM SIGCOMM workshop on Cellular networks: operations, challenges, and future design, 2012.