Enabling Renewed Innovation in TCP by Establishing an Isolation Boundary

Umar Kalim[*+], Eric Brown[+], Mark K. Gardner[+], Wu-chun Feng[*]

Department of Computer Science[*], Office of IT[+]

Virginia Tech

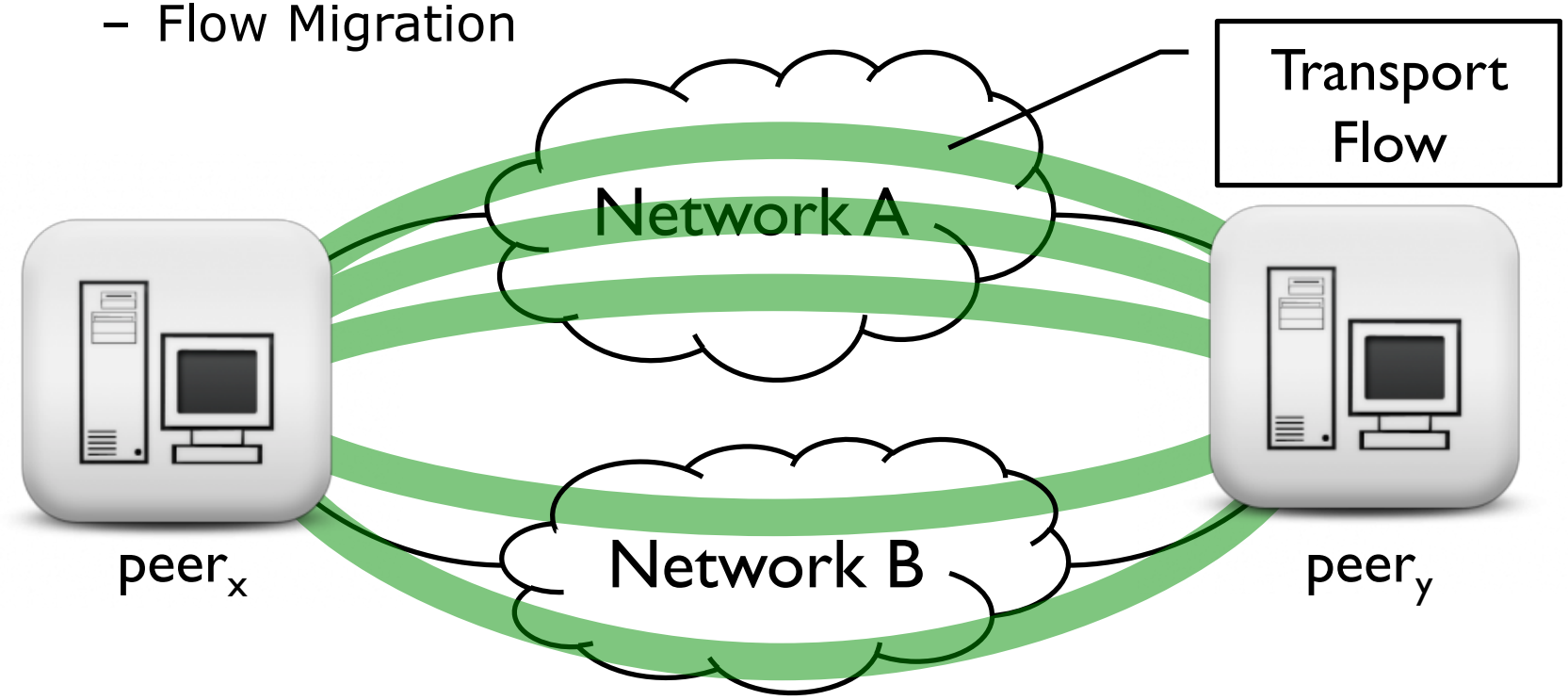
11/28/2010

VirginiaTech

Invent the Future

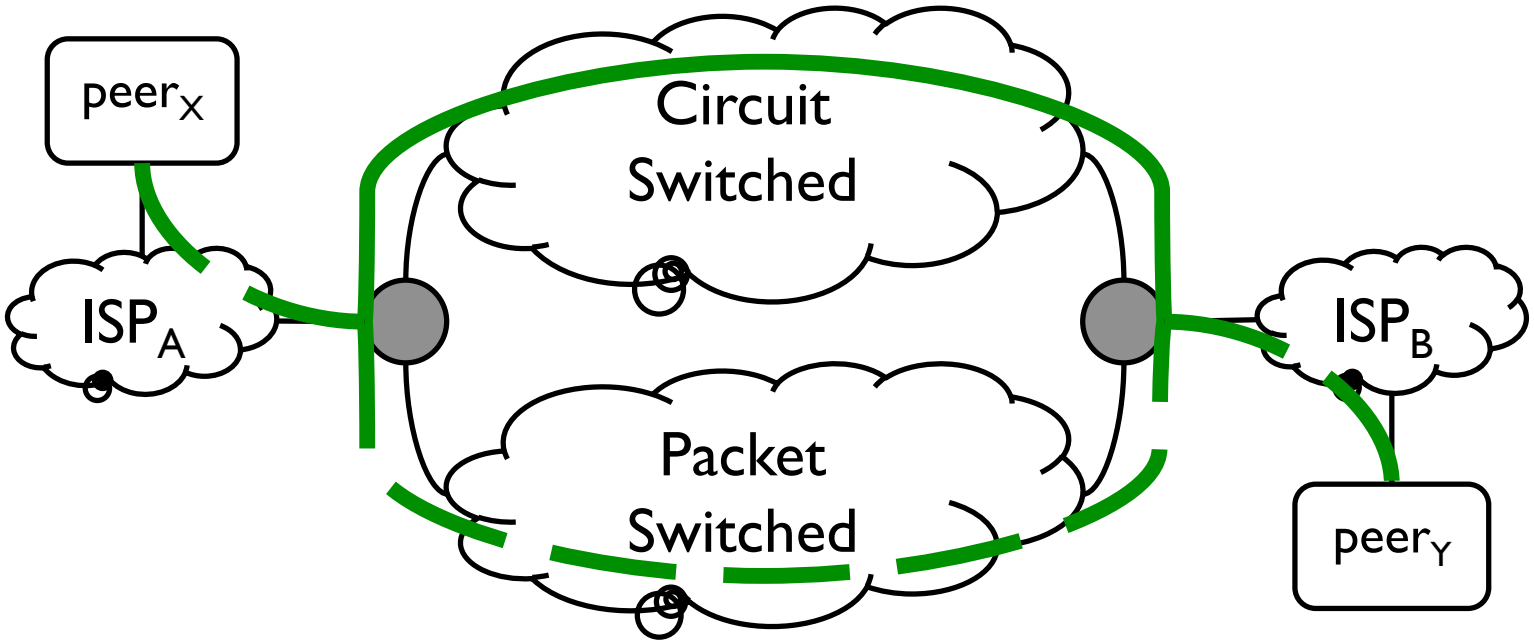
Why Renewed Innovation?

- Desirable functionality
  - Multipath TCP / Multihoming
  - Flow Migration



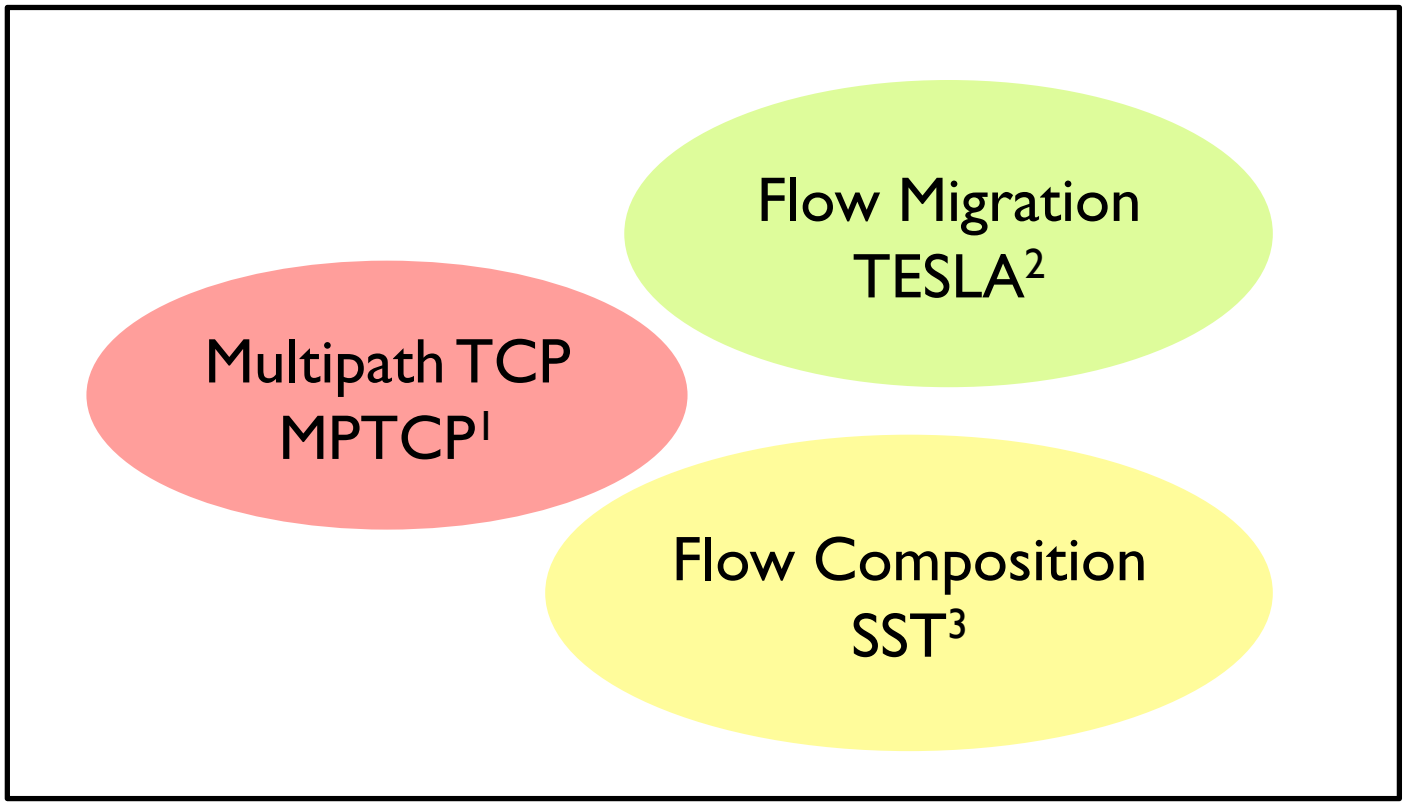
Why Renewed Innovation?

- Desirable functionality
  - Hybrid Transports



1. Kissel, E., Brown, A., Swany, M. Improving GridFTP Performance Using the Phoebus Session Layer. *SC 2009.*
2. Veeraraghavan, M., Zheng, X., Lee H., Gardner, M., Feng, W. CHEETAH: Circuit-switched High-speed End-to-End Transport ArcHitecture. *OPTICOMM 2003.*
3. Lehman, T., Sobieski, J., Jabbari, B. DRAGON: A Framework for Service Provisioning in Heterogenous Grid Networks. *IEEE Communications Magazine March 2006.*

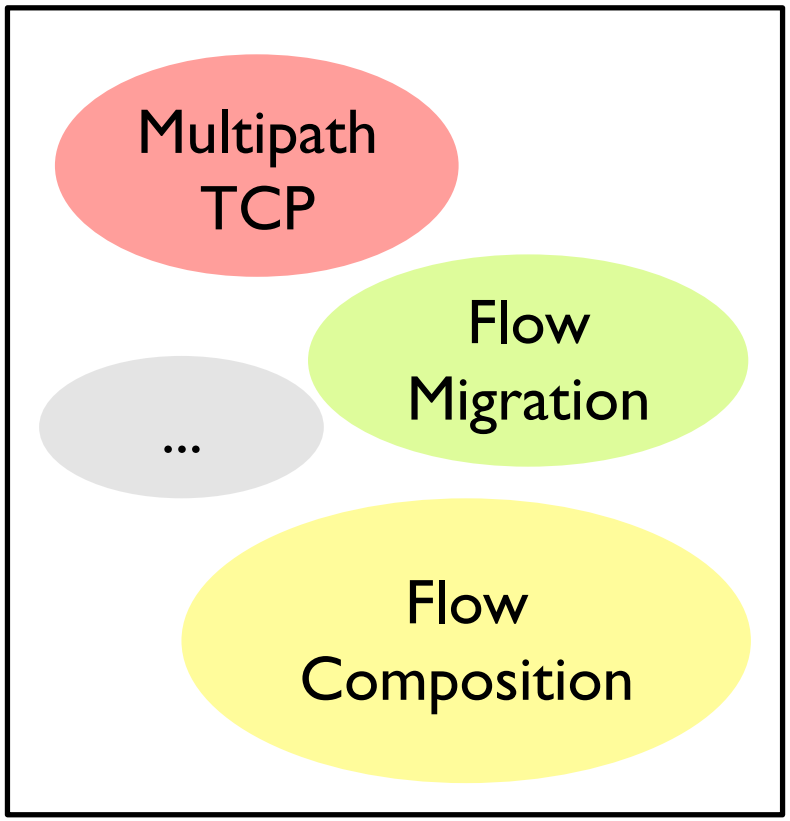
Motivation



1. A. Ford, C. Raiciu, M. Handley, and S. Barre, "TCP Extensions for Multipath Operation with Multiple Addresses," RFC (Experimental), Work in Progress, Internet Engineering Task Force, Jul. 2010.
2. J. Salz, A. C. Snoeren, and H. Balakrishnan, "TESLA:A Transparent, Extensible Session-Layer Architecture for End-to-end Network Services," in 4th USENIX Symposium on Internet Technologies and Systems (USITS), 2003, pp. 211–224.
3. B. Ford, "Structured Streams: A New Transport Abstraction," in ACM SIGCOMM CCR, vol. 37, no. 4.ACM, 2007, pp. 361–372.

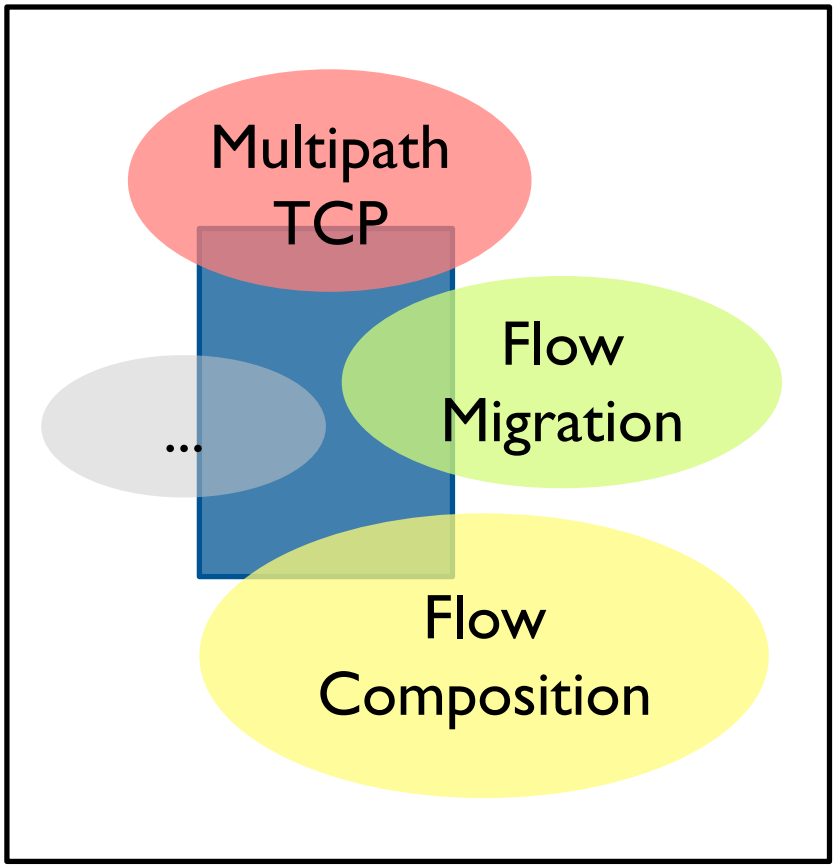
Motivation

- Existing solutions
  - TCP option space
    - MPTCP addresses the issue
  - Custom libraries
    - Duplication of effort
  - Clean-slate approach
- Legacy behind TCP
  - Experience and tools built to support TCP
  - Best to adopt an incremental approach



Proposed Solution

- Goals
  - Extract commonality in current work
  - Lay the foundation for higher layer services
  - Admit incremental adoption
- Common mechanism
  - Decouple application stream from transport flows
  - Construct a control channel
- Leverage TCP options



Outline

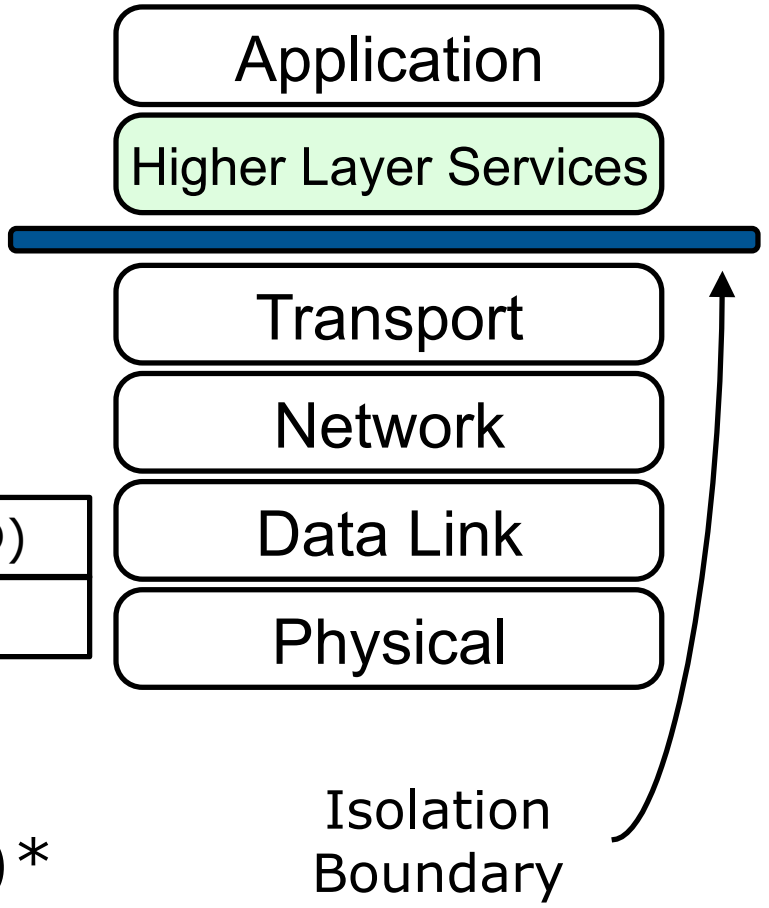
- Motivation
- Proposed Solution
  - Isolation Boundary: An Overview
  - Connection Setup
  - Control and Data Option
- Community Engagement
- Contribution and Implications

Isolation Boundary: An Overview

- Decouple application-data stream from transport-endpoint identification
  - Avoid use of 4-tuple for transport endpoint ID
- TCP Isolation Boundary Option

Transport Independent Flow ID (TIFID)	
TI Seq. #	TI Ack. #

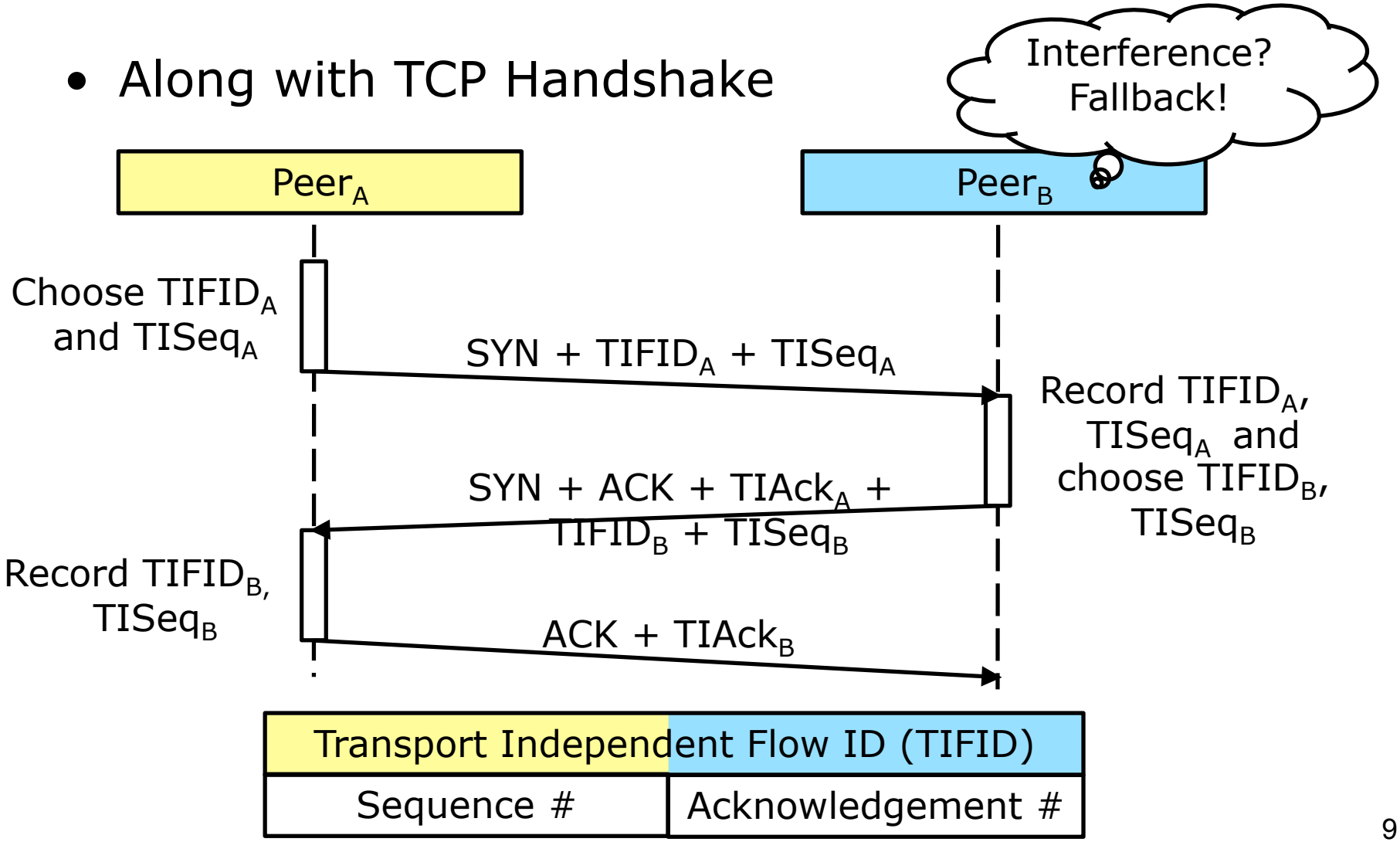
- Partially adapt Isolation Layer proposed by Next Generation Transport (T*ng*)*



* J. Iyengar and B. Ford, "A Next Generation Transport Services Architecture", RFC (Informational), Work in Progress. http://tools.ietf.org/html/draft-iyengar-ford-tng-00

Transport Independent Flow Setup

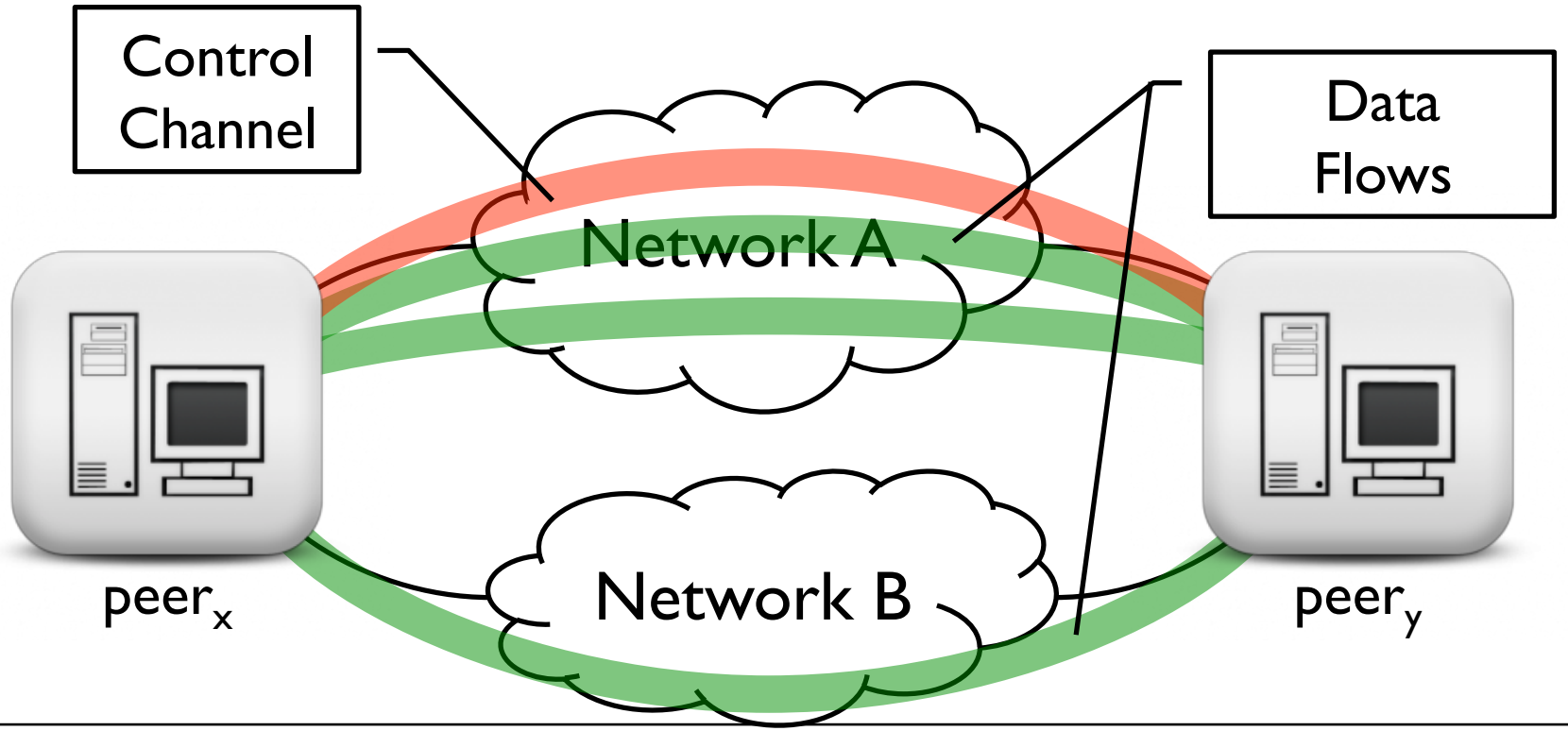
- Along with TCP Handshake



TISeq – Transport Independent Sequence #
TIAck – Transport Independent Acknowledgement #

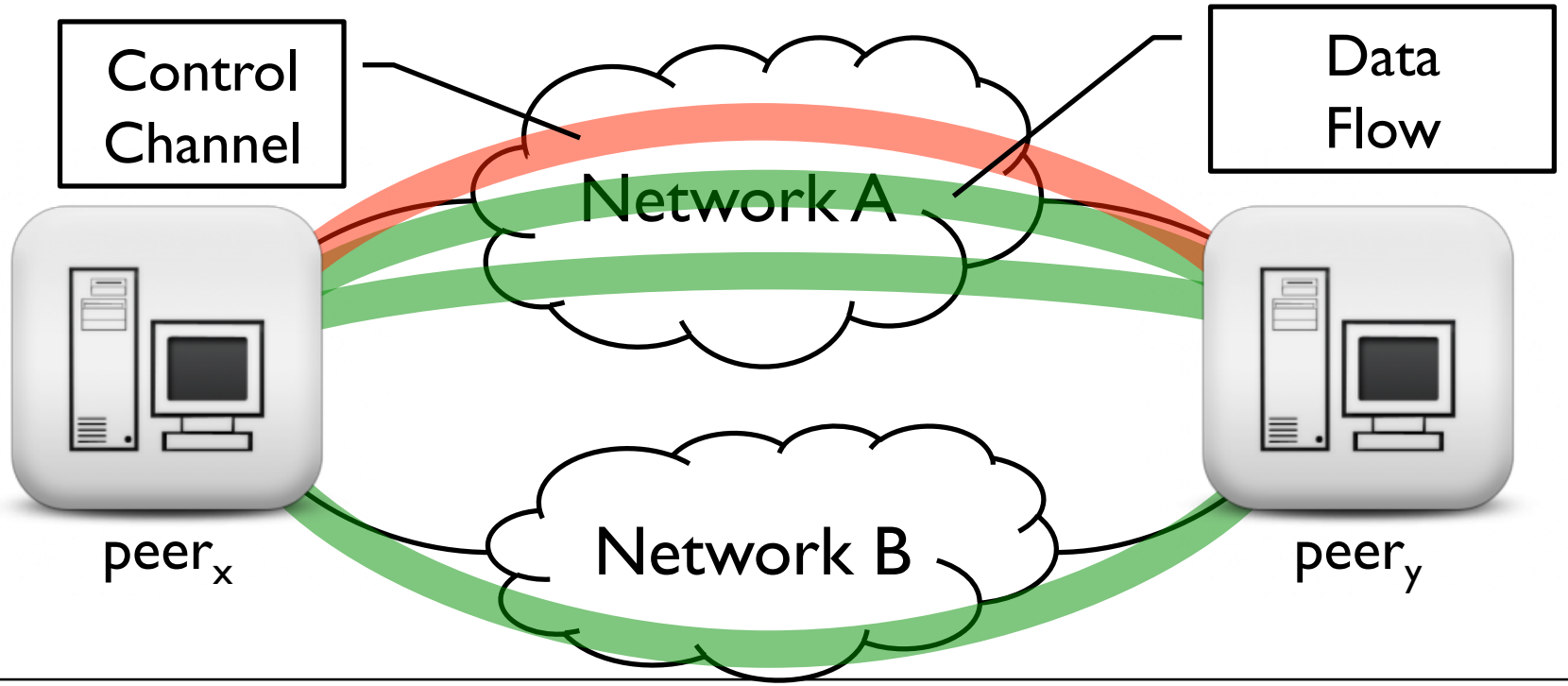
Isolation Boundary Options

- Isolation Boundary Option - Control
  - Admits out-of-band control channel
- Control protocol to be defined by community



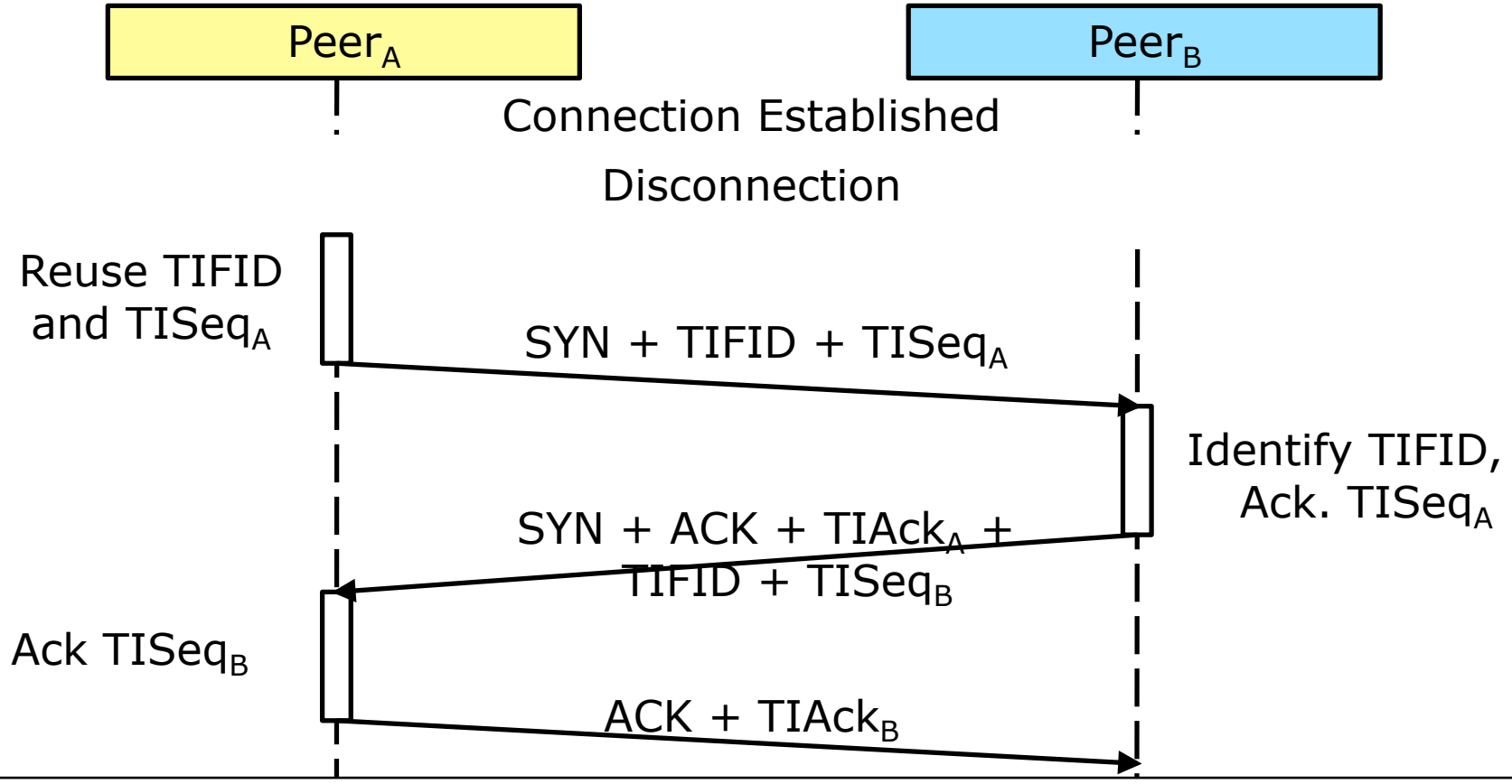
Isolation Boundary Options

- Not all applications need a control channel
- Isolation Boundary Options
  - Control
  - Data



Isolation Boundary Options

- Isolation boundary enables reconnection
  - From network fault



Flow Termination

- Control Channel to clear state
- Exchange of TCP FINs implies cleanup
- Rely on timeouts for network faults

Community Engagement

- Standardization
  - Specification of the extensible control channel
  - Reference implementation
- Wire Protocol Specification
- Isolation Boundary for other transports

Contribution and Implications

- Develop an Isolation Boundary
  - Decouple the entity naming from transport-endpoint identification
  - Construct a control channel to facilitate higher layer services
- Infrastructure support for
  - Hybrid transport
  - Multihoming
  - Flow migration over networks
  - Flow migration between processes
  - Reconnection after network fault
  - … and other innovations as the community thinks of them

Thank you

- Contact
  - Umar Kalim - umar@cs.vt.edu
  - Eric Brown - brownej@vt.edu
  - Mark Gardner - mkg@vt.edu
  - Wu-chun Feng – feng@cs.vt.edu
- SyNeRGy
  - http://synergy.cs.vt.edu

This research is supported in part by
Juniper Networks and Virginia Tech

Analysis

- TCP Options Space
  - 3-way handshake ~ 20 octets available
    - MSS, Window Scaling, SACK, Timestamp
- Incompatible Options
  - Alternate checksum, Partial Ordering, Transactional TCP, TCP MD5, TCP Authentication, Quick Start Response
- Performance
  - Exchange is off the critical data path (3-way handshake)
- SYN Cookies
  - IBO not preserved when under attack

Analysis

- Middleboxes
  - Fall back to legacy TCP if options are stripped
- Security
  - No worse than TCP
  - To hijack a session the attacker must know:
    - Transport Independent Flow ID
      - Exchanged only during 3-way handshake
    - Sequence numbers for unacknowledged data
      - Each TCP segment must be accounted for, to derive current state from Initial Seq. Nos.
- Application Compatibility
  - Backward compatible