PLUS and QUIC: Deploying the MCP

Brian Trammell / Mirja Kühlewind, ETH Zürich

3. MAMI Plenary, Seville, 6 February 2017

mami

measurement and architecture for a middleboxed internet

measurement | architecture | experimentation

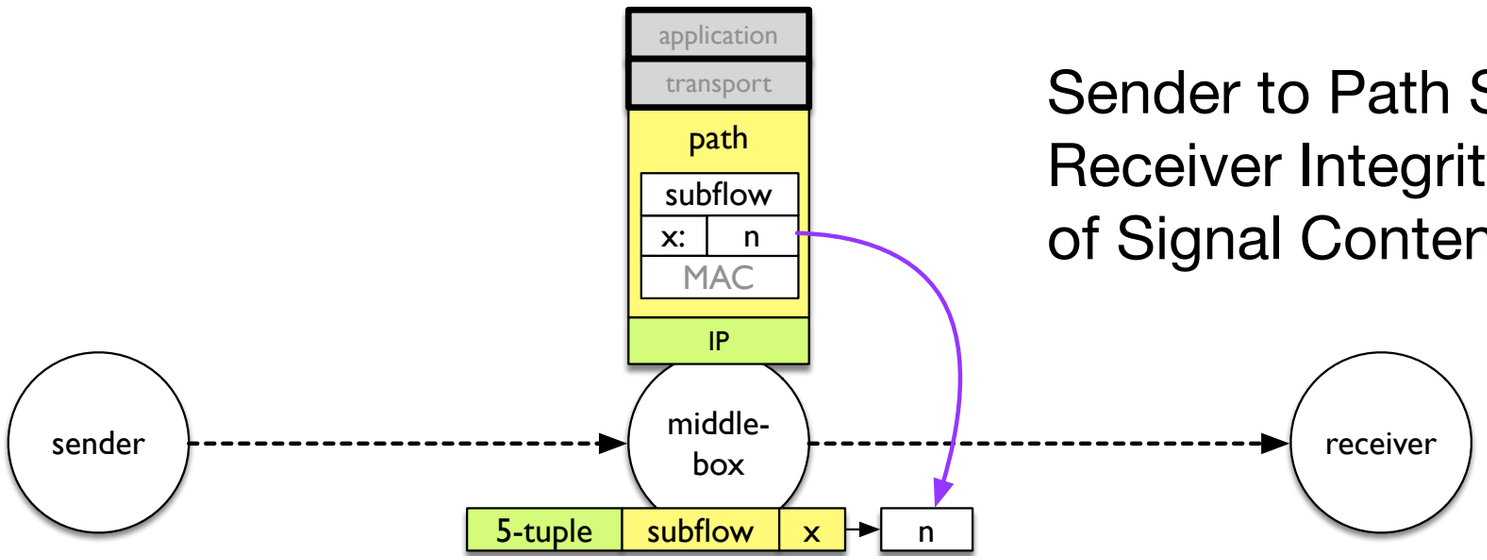
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688421.The opinions expressed and arguments employed reflect only the authors' view. The European Commission is not responsible for any use that may be made of that information.

Supported by the Swiss State Secretariat for Education, Research and Innovation under contract number 15.0268. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government.

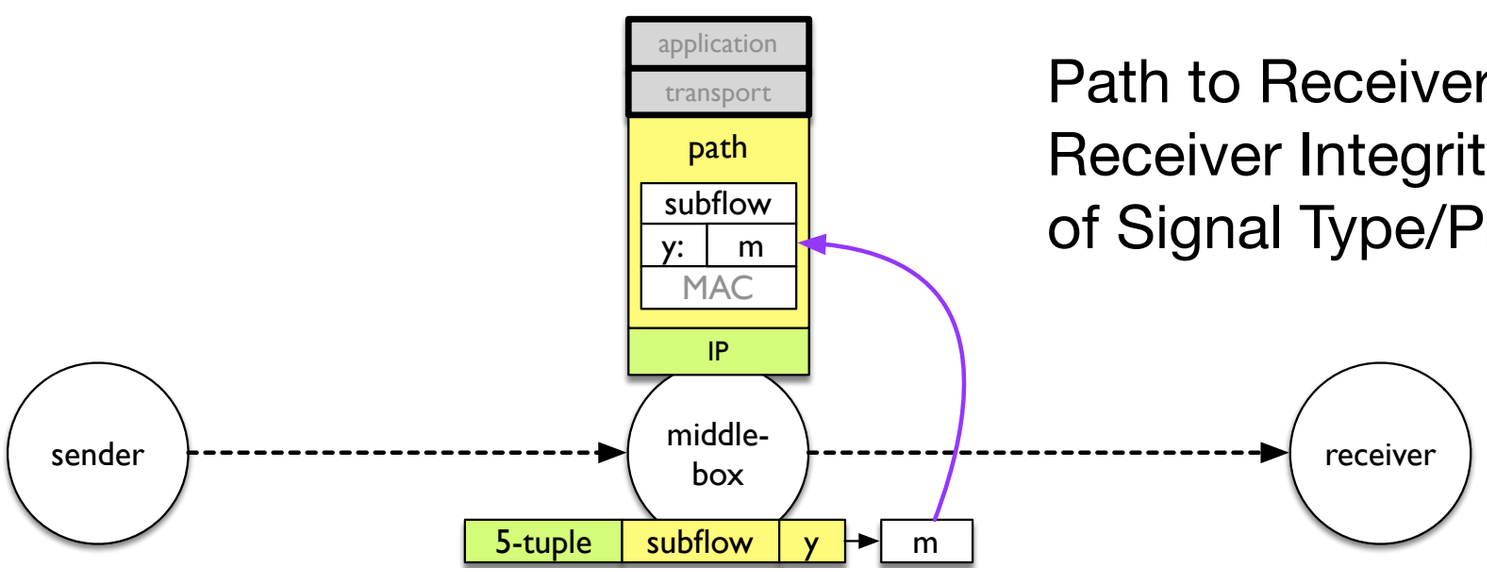
Overview Internet-drafts and preprints

- SPUD (expired)
  - draft-kuehlewind-spud-use-cases-00: Use Cases for a Substrate Protocol for User Datagrams (SPUD)
  - draft-trammell-spud-req-04: Requirements for the design of a Substrate Protocol for User Datagrams (SPUD)
- **PLUS**
  - **draft-trammell-plus-statefulness-02:** Transport-Independent Path Layer State Management
  - **draft-trammell-plus-abstract-mech-00:** Abstract Mechanisms for a Cooperative Path Layer under Endpoint Control
  - **draft-trammell-plus-spec-00:** Path Layer UDP Substrate Specification
- **QUIC**
  - **draft-kuehlewind-quic-appman-00:** Applicability and Management of the QUIC Transport Protocol
- **IPIM**
  - **Principles for Measurability in Protocol Design:** describes an abstract In-Protocol Internet Measurement (IPIM) facility to be implemented in PLUS/MCP, https://arxiv.org/pdf/1612.02902v1.pdf

Abstract Mechanisms for the Path Layer (PLUS BoF, Berlin, “-abstract-mech”)



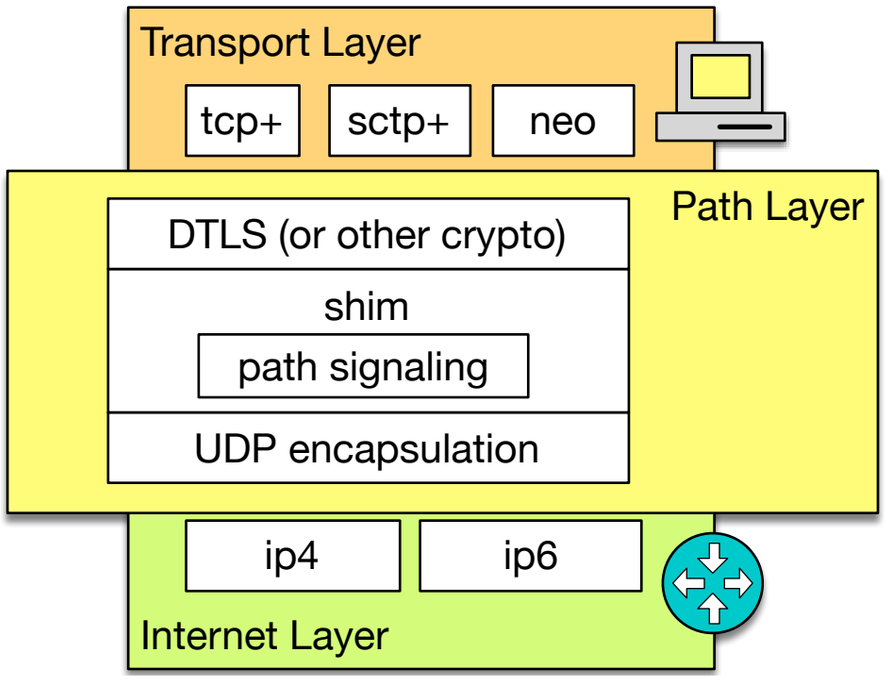
Sender to Path Signaling, Receiver Integrity Protection of Signal Content



Path to Receiver Signaling, Receiver Integrity Protection of Signal Type/Presence

Anatomy of the Path Layer (PLUS BoF, Berlin)

- UDP encapsulation
  - userspace implementation
  - ports for NAT
  - ~95% deployable today
- encoding for abstract signaling mechanisms
- crypto (unspecified) to protect transport headers and above

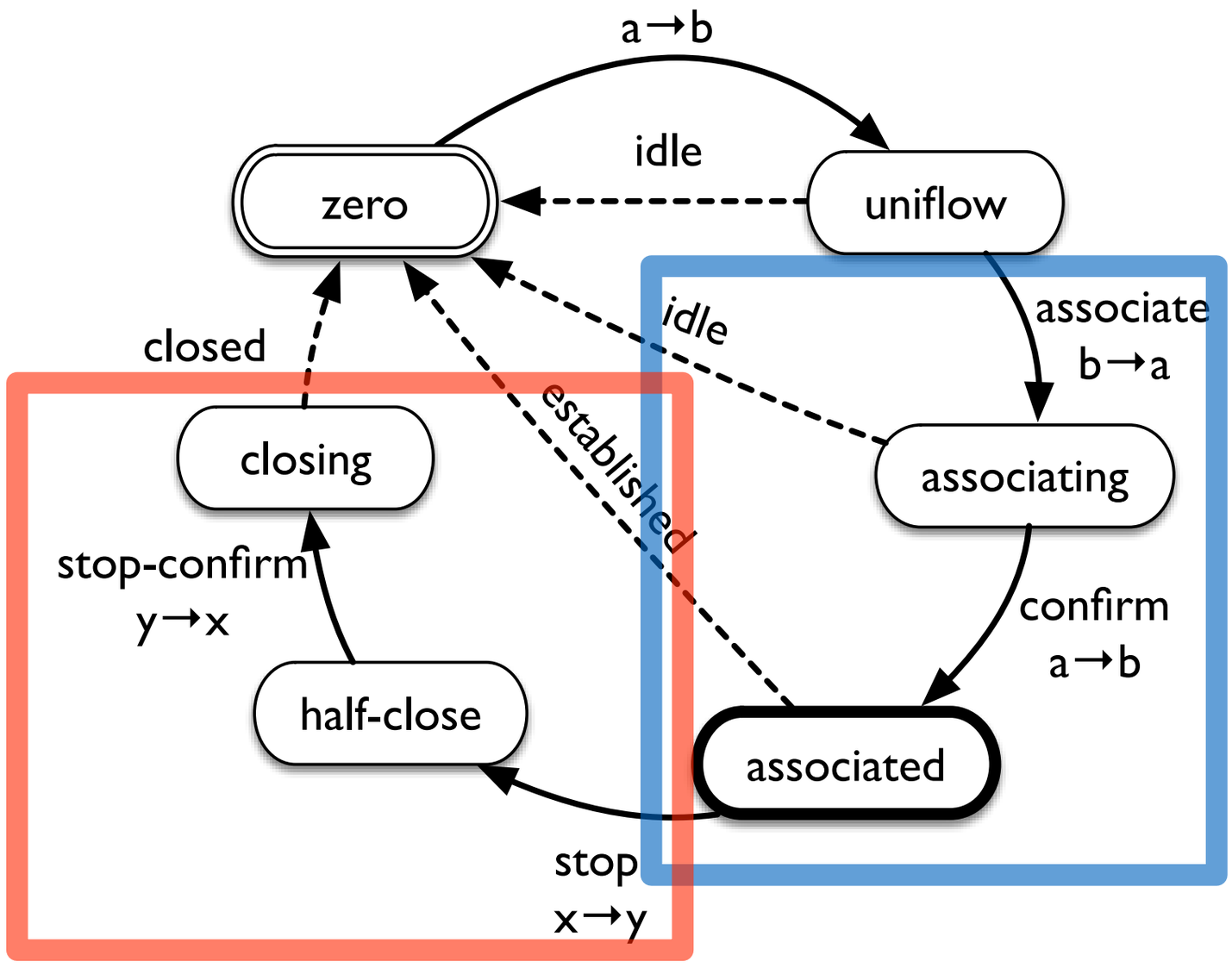


- Unable to achieve IETF consensus due to concerns about privacy and operator abuse of path signaling.

Reframing the MCP problem

- We really want three different sets of features:
  - TCP wire image replacement for encrypted transport protocols — state exposure and basic measurement
  - Sender to path signaling for one-bit signals (e.g. LoLa)
  - Additional sender-to-path and path-to-receiver signals for future troubleshooting, management, provisioning network functions.
- So let's define a wire image with these features in mind.

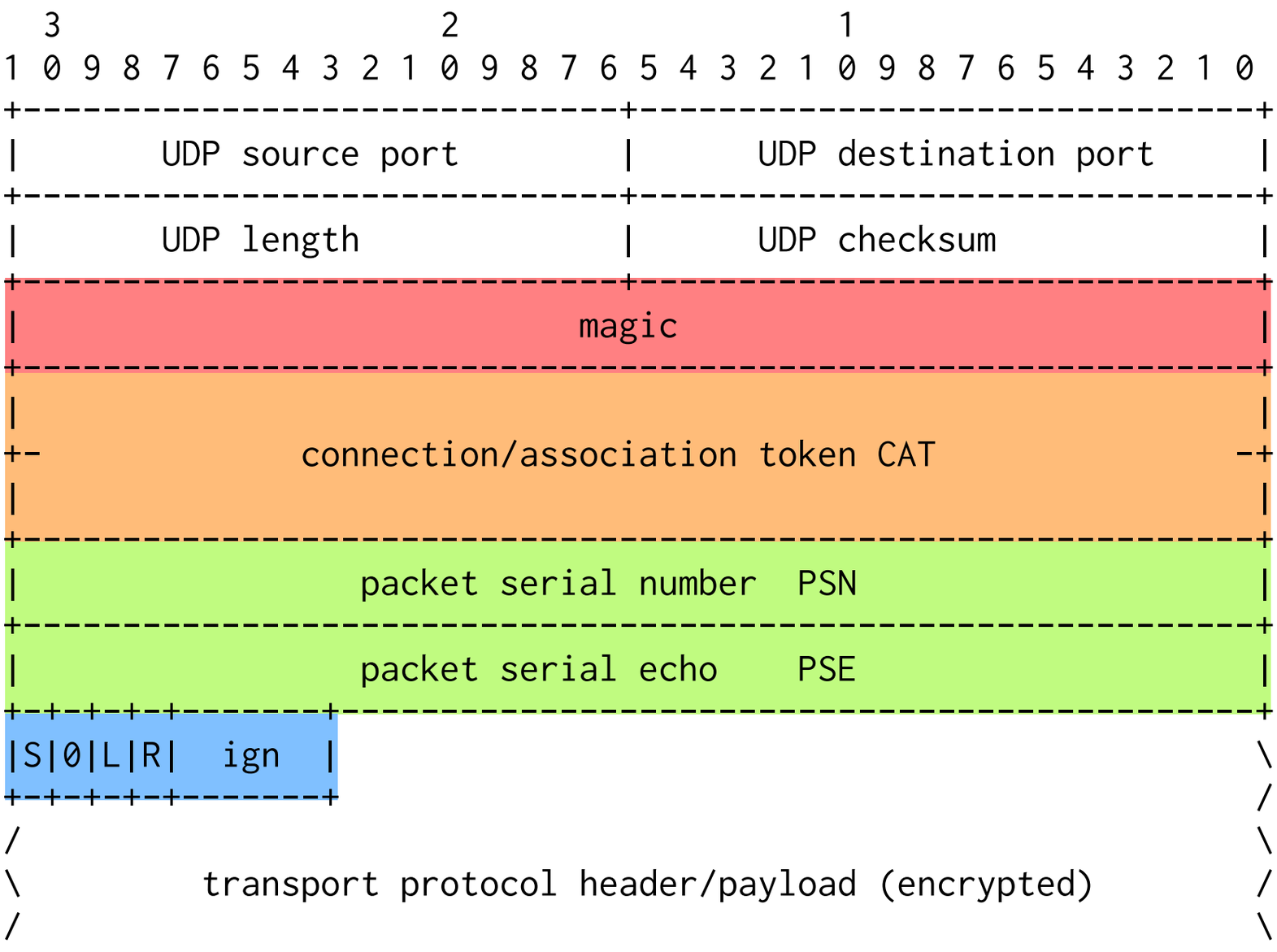
Transport-Independent Exposure of Transport State ("-statefulness")



Require closing signals in both directions to prevent injection

Force three-way handshake to prove return routability

PLUS Basic Header (“-spec”)



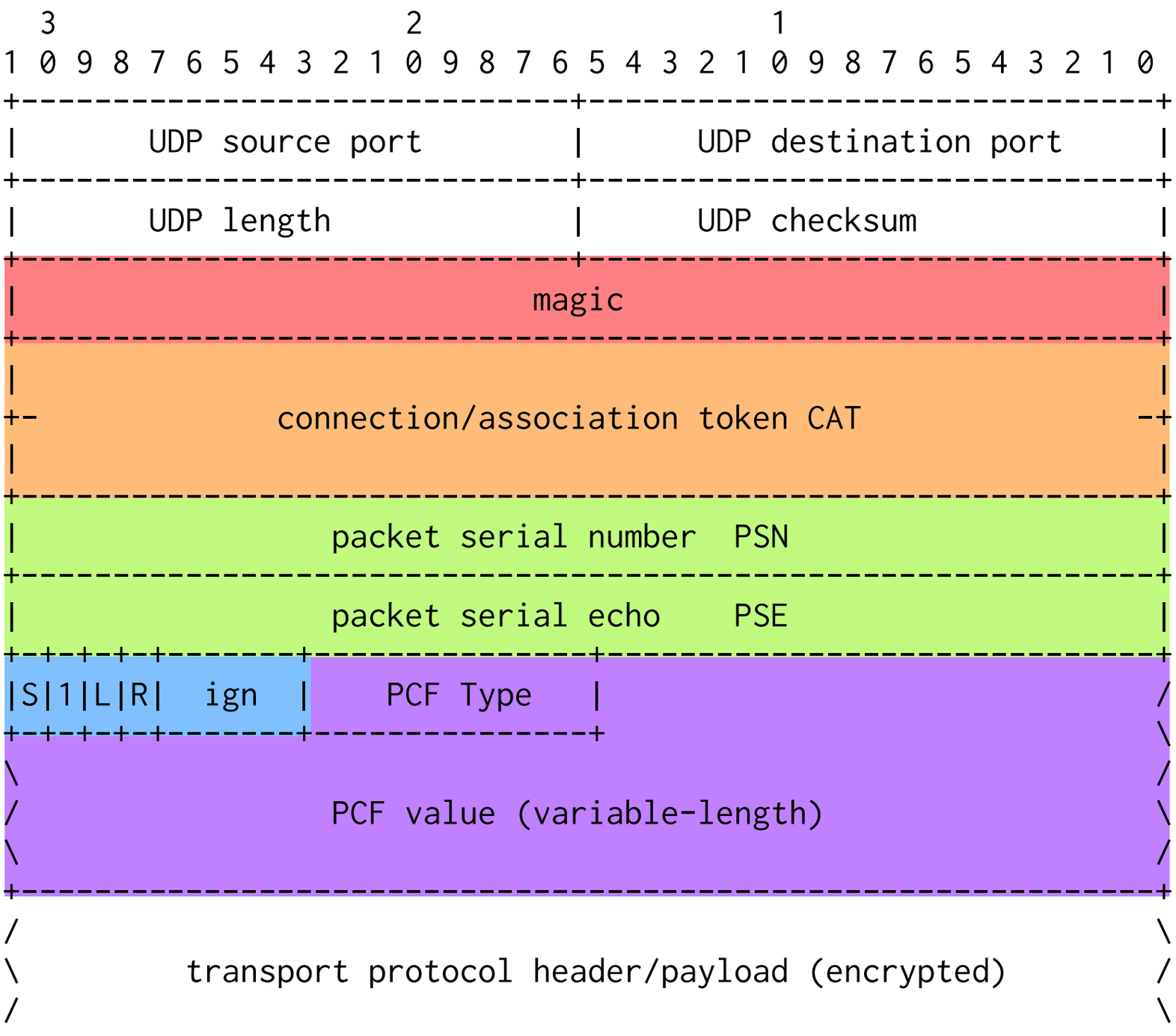
identifies packet as PLUS, collides with other UDP application protocols.

token for load balancing, NAT rebinding, association and confirmation signals

for loss and RTT measurement, and stop signal confirmation

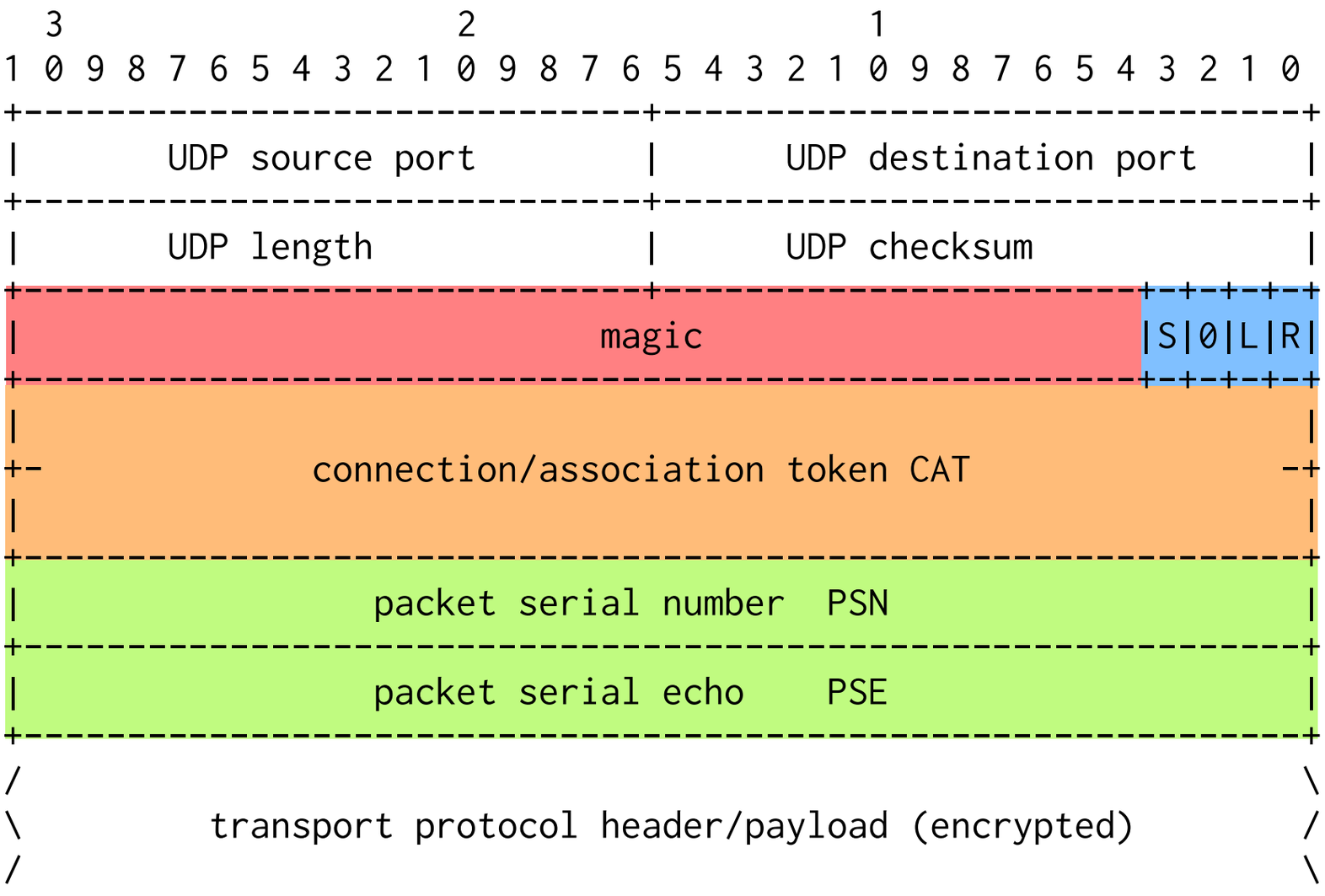
stop, extended header, latency sensitivity, and reordering OK flags

PLUS Extended Header (“-spec”)



sender-to-path and path-to-receiver signals (length/behavior specified by PCF type)

PLUS Basic Header, optimized (“-spec”)



Insights:

- 28 bits is enough magic.
- We only need four flags.
- No advantage to overlap with QUIC Special Packet flags byte.

PLUS Path Communication Field vocabulary

- Sender to path:
  - Timestamp/Timestamp Echo (less useful with PSN/PSE)
  - Relative priority *(not in -00)*
  - Congestion exposure *(not in -00)*
- Path to receiver:
  - MTU accumulator
  - State timeout accumulator
  - Rate limit accumulator
  - Path delay accumulator
  - Path element trace (IPIM §4.3)

QUIC (Interim WG Mtg. Tokyo, Jan 2017)

- QUIC will deploy before PLUS — getting some MCP concepts into it is worth the effort.
  - Primary focus: PSN/PSE, two-way stop.
  - Secondary focus: make sure QUIC can layer on PLUS for experimentation.
    - Align PSN/PSE, CAT semantics.
- Applicability and manageability document(s) for QUIC:
  - Description of in-network functions supported by QUIC's design

QUIC+ implementation plan

- Fork https://github.com/lucas-clemente/quic-go
  - first step: update implementation to match current spec (with or without crypto)
  - H2 implementation out of the box, simple to set up basic server, test client
  - Measure PLUS connectivity towards target test servers at DigitalOcean, etc.

- Reference PLUS middlebox: https://fd.io
  - focusing on PLUS-aware NAT, prioritization, traffic diagnostics

- Experimental PLUS passive measurements based on https://github.com/britram/mokumokuren (new PathSpider observer)
  - potential experimentation with extended header to increase measurement accuracy

Path Aware Networking RG?

- PLUS is part of a mass of work in the IETF community on *path-aware* networking
  - MPTCP, TAPS, QUIC, HOMENET, SPUD, PLUS, ACCORD, ARCING, BANANA, ICCRG
- Some questions remain unanswered
  - communication and discovery of information about the properties of a path on local networks and in internetworks, exploration of trust and risk models associated with this information, and algorithms for path selection at endpoints based on this information
  - algorithms for making transport-layer scheduling decisions based on information about path properties
  - exploration of the role of naming in path-aware networking
  - algorithms for reconciling path selection at endpoints with network operations practice and widely deployed routing protocols and best practices
- Interest in contributing?