# PLUS and QUIC: Deploying the MCP

Brian Trammell / Mirja Kühlewind, ETH Zürich

3. MAMI Plenary, Seville, 6 February 2017



measurement and architecture for a middleboxed internet

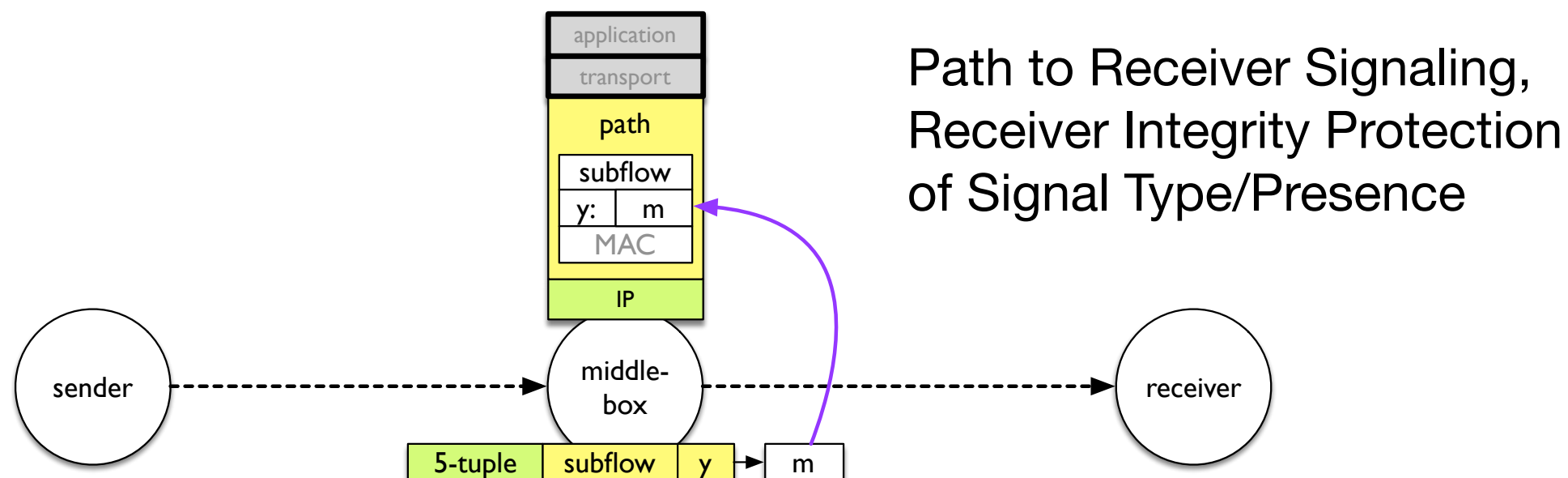**measurement**    **architecture**    **experimentation**

# Overview Internet-drafts

- 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

WP3: PLUS and QUIC

2

# 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
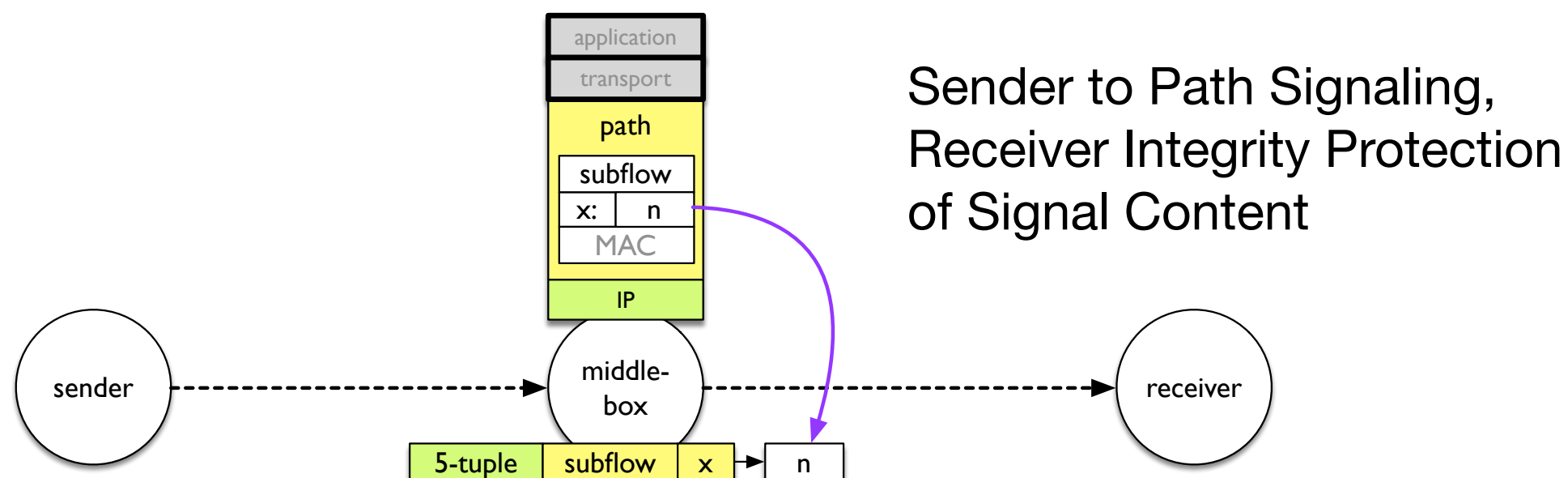
# Anatomy of the Path Layer (PLUS BoF, Berlin)

- 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.

**Transport Layer**
tcp+ | sctp+ | neo

**Path Layer**
DTLS (or other crypto)
shim
path signaling
UDP encapsulation

**Internet Layer**
ip4 | ip6
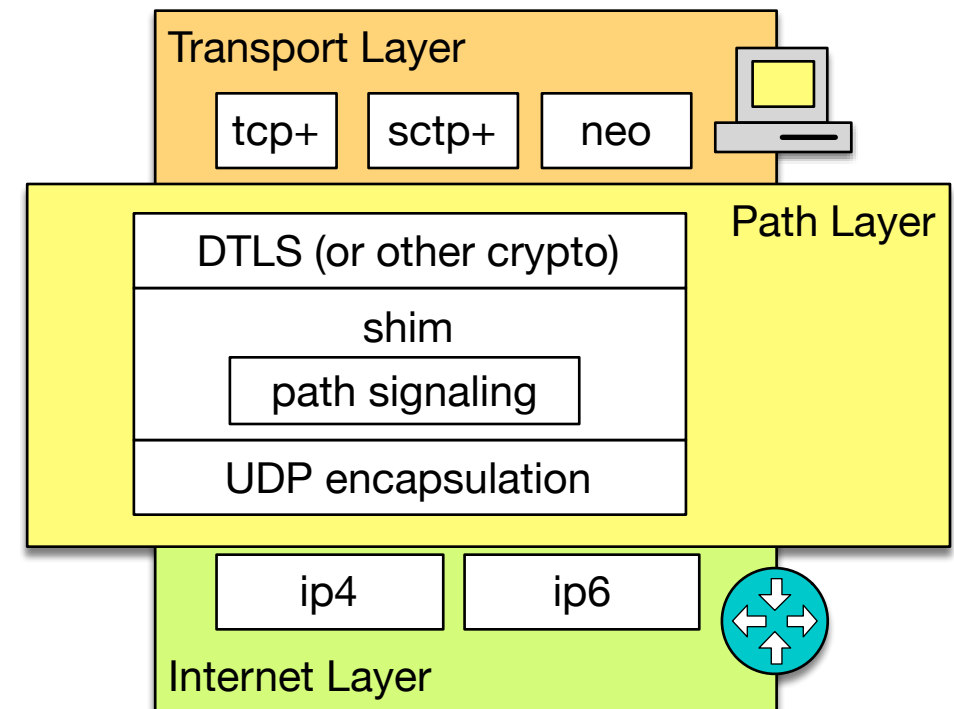
# Reframing the MCP problem

- 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.

# PLUS Basic Header ("-spec")

```
 3                   2                   1
 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0
+---------------------------------+---------------------------------+
|          UDP source port        |       UDP destination port      |
+---------------------------------+---------------------------------+
|            UDP length           |          UDP checksum           |
+---------------------------------+---------------------------------+
|                             magic                                 |
+-------------------------------------------------------------------+
|                                                                   |
+-              connection/association token  CAT                  -+
|                                                                   |
+-------------------------------------------------------------------+
|                    packet serial number   PSN                     |
+-------------------------------------------------------------------+
|                    packet serial echo      PSE                    |
+-+-+-+-+-------+----------------------------------------------------+
|S|0|L|R|  ign  |                                                    \
+-+-+-+-+-------+                                                    /
/                                                                   \
\          transport protocol header/payload (encrypted)           /
/                                                                   \
```
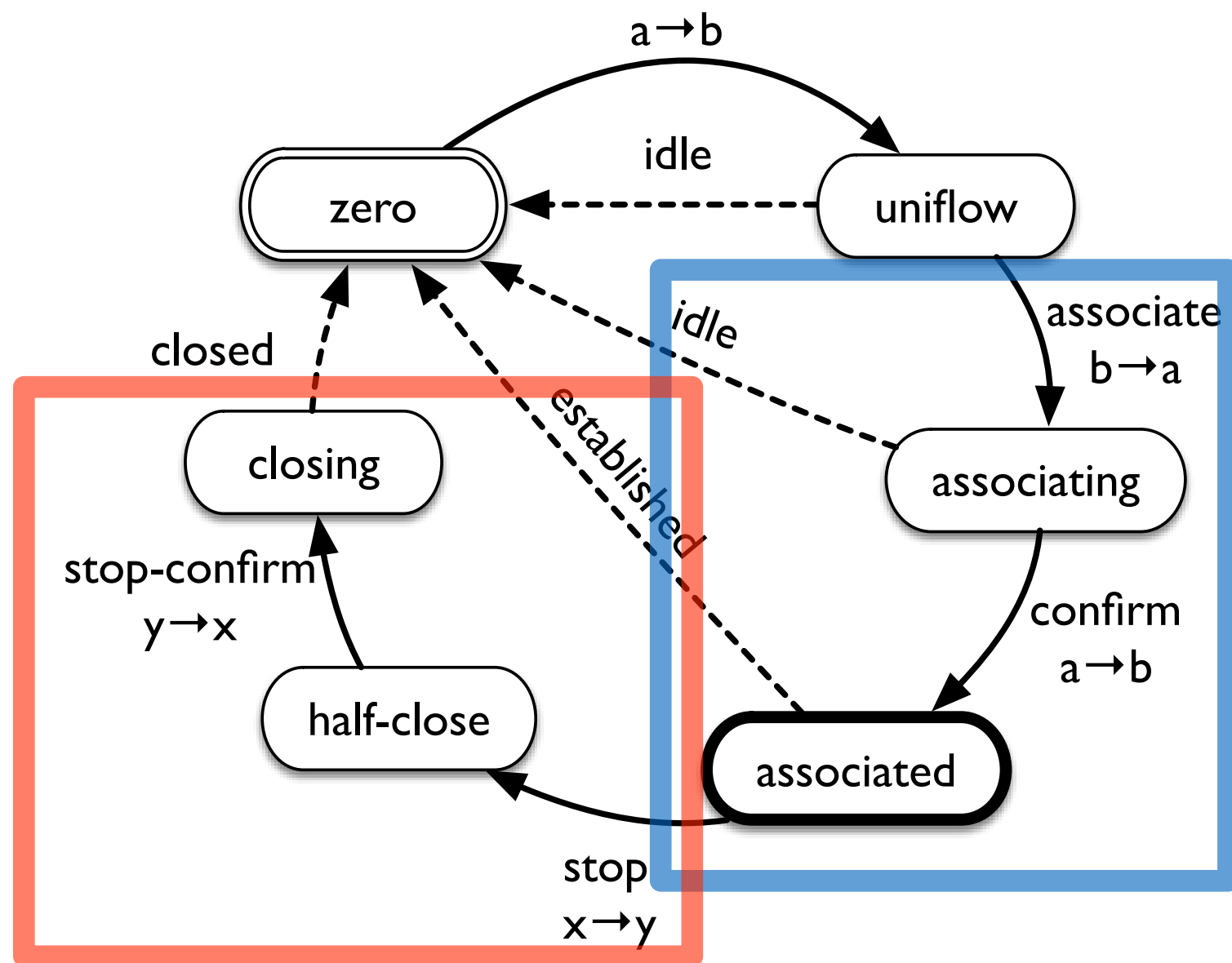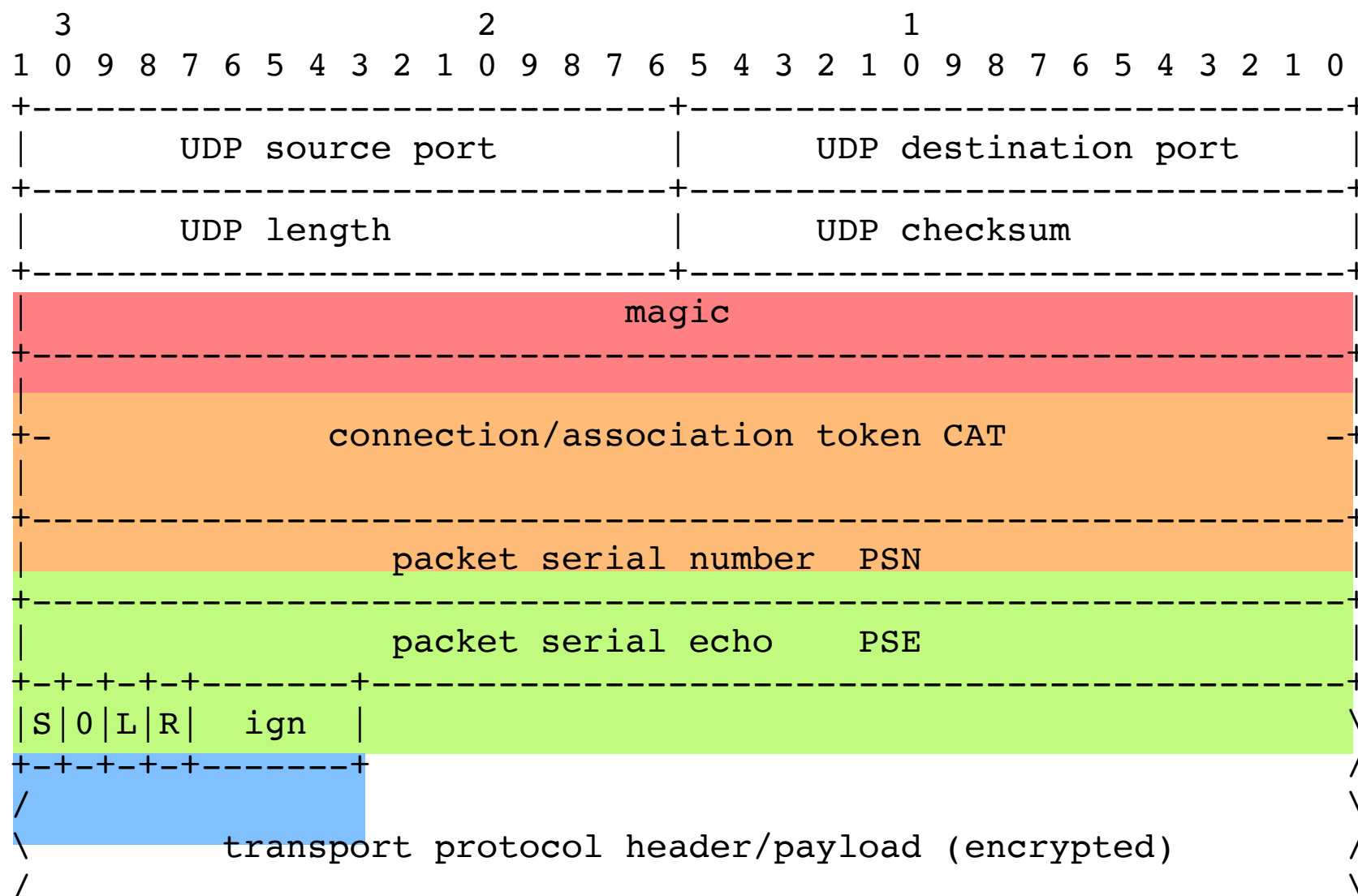
**identifies packet as PLUS**

**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")

```
 3                   2                   1
 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0
+-------------------------------+-------------------------------+
|          UDP source port      |        UDP destination port   |
+-------------------------------+-------------------------------+
|          UDP length           |          UDP checksum         |
+-------------------------------+-------------------------------+
|                            magic                              |
+---------------------------------------------------------------+
|                                                               |
+-           connection/association token CAT                 -+
|                                                               |
+---------------------------------------------------------------+
|               packet serial number   PSN                      |
+---------------------------------------------------------------+
|               packet serial echo      PSE                     |
+-+-+-+-+-------+---------------+-------------------------------+
|S|1|L|R|  ign  |     PCF Type  |                               /
+-+-+-+-+-------+---------------+                               \
\                                                               /
/                  PCF value (variable-length)                  \
\                                                               /
+-------------------------------------------------------------+
/                                                               \
\          transport protocol header/payload (encrypted)        /
/                                                               \
```
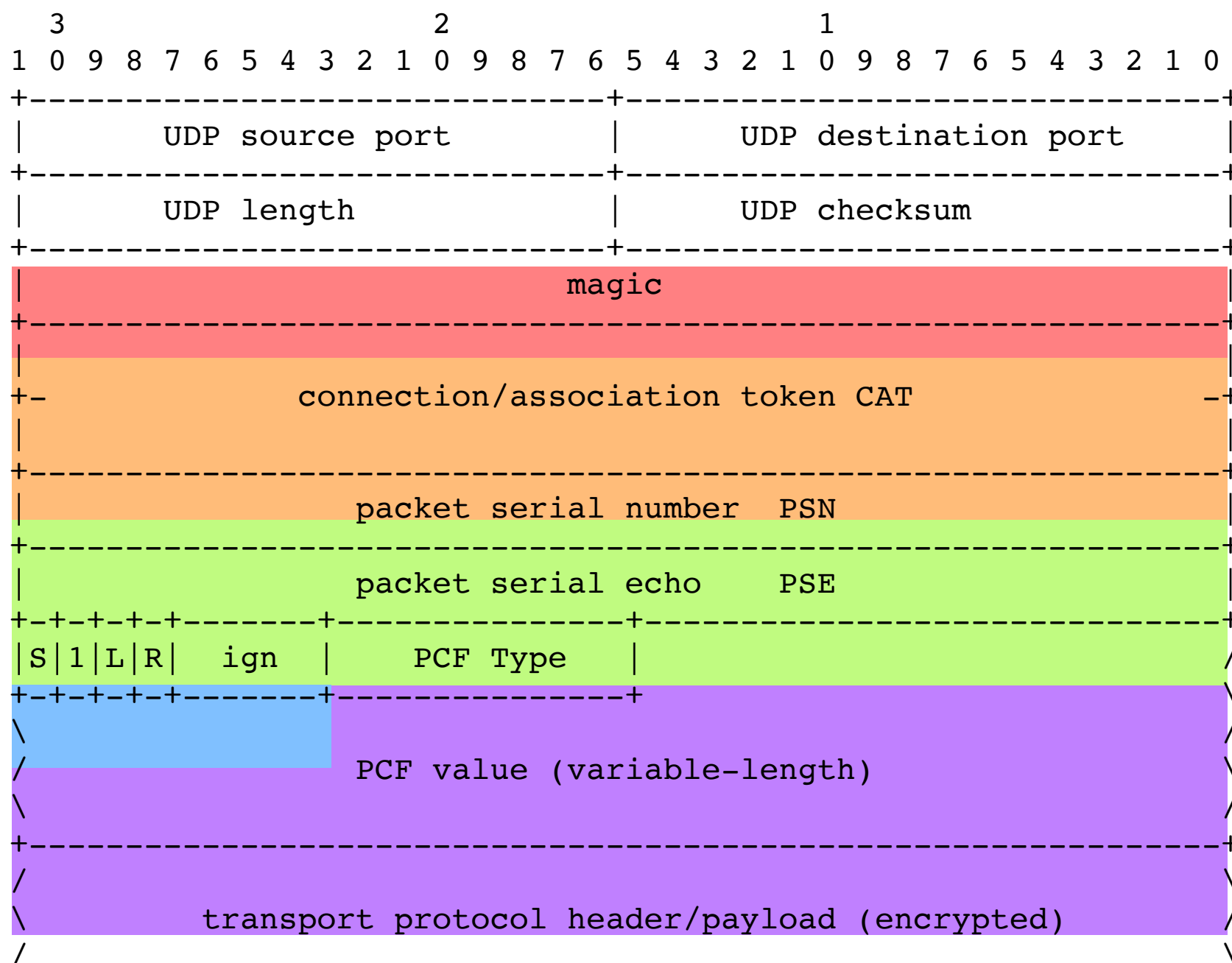
sender-to-path and path-to-receiver signals (length/behavior specified by PCF type)

architecture

# 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)

mami

# 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

  - Basic server, test client

  - Measuring PLUS connectivity towards own traget test server

- Reference PLUS middlebox: https://fd.io?

  - focusing on NAT and 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