IAB Workshop on

Stack Evolution in a Middlebox Internet (SEMI)

26-27 January 2015, Zurich

—and—

Substrate Protocol for User Datagrams (SPUD) BoF

25 March 2015, Dallas

Transport Area Open Meeting, 23 March 2015, Dallas

Brian Trammell <ietf@trammell.ch>

Background

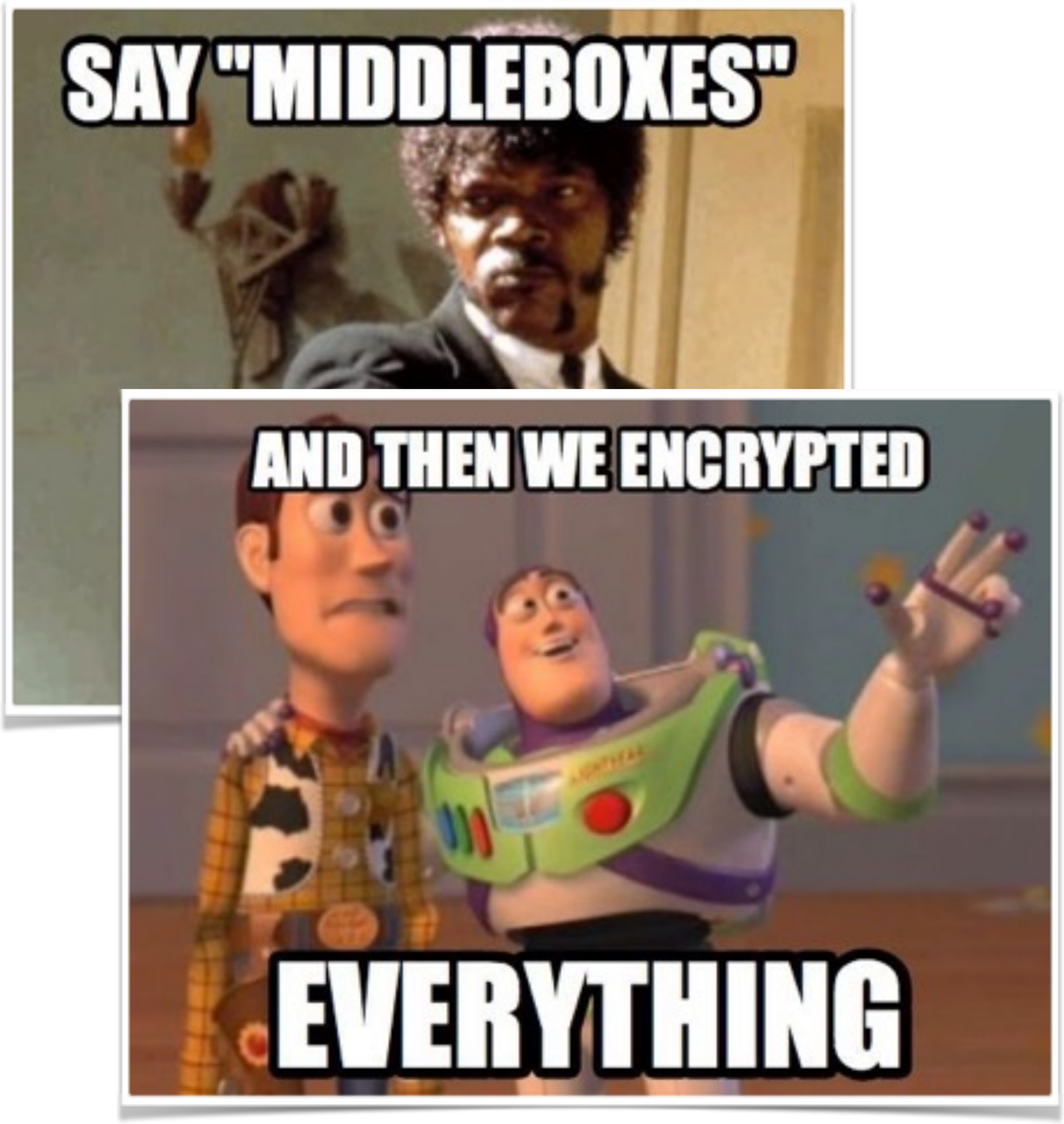
- IAB IP Stack Evolution Program currently focuses on two broad areas:
  - evolution of interfaces to transport and network-layer services beyond `SOCK_STREAM` and `SOCK_DGRAM`
  - Improving path transparency in the presence of firewalls and middleboxes.
- Follows the IAB's interest in general issues of protocol evolution (RFC 5218, ITAT workshop)
- Within the program, the IAB convened a workshop in January to discuss ossification of the transport layer…
  - …and how to fix it for emerging applications (e.g. rtcweb)

Squeezing the Transport Layer

- The problem is one of narrow interfaces as well as narrow paths:
  - `SOCK_STREAM` and `SOCK_DGRAM` not enough for every situation
  - Rolling your own transport can be dangerous
  - Middleboxes won’t pass traffic they can’t understand (new protocols, protocol extensions, flows with tricky state)
- Increasing deployment of encryption adds another dimension.

Why now?

1. new energy in the IETF:
   - work which requires flexibility we don't appear to have (RTCWEB, TCPINC)
   - work to provide that flexibility at the interface (TAPS)
2. pressure created by increasing deployment of encryption:
   - "Everything over TLS" will brick lots of deployed middleboxes
   - Opportunity to strike a balance between endpoint and midpoint requirements.



Workshop Positions

- 20 position papers accepted, 38 invitations sent.
- Stated goals of participants included:
  - deeper understanding of architecture and incentives,
  - broadening of transport interfaces
  - further research and community education on the issue
  - definition of middlebox cooperation approaches.
- On transport evolution, there were two camps:
  - “TCP is broken, burn it to the ground and start over”
  - “Long live TCP!”

Identified Goals

- Future work (WG/RG) on middlebox cooperation (protocol/functionality/etc.), including:
  - mechanisms for detection of path characteristics
  - measurement for path impairment detection and troubleshooting
- Better understanding of how transport should/must evolve, including applicability of present transports to specific use cases.
- Interface improvement: expose more to applications about transport (in the right way)
- Identify trust issues and deployment incentives in cooperation and evolution approaches (this is hard)

Measurement

- We need to make data-driven engineering decisions about transport protocol extension
    - If a protocol works in 99.5% of the Internet, why not use when you can?
    - If a feature breaks in 0.5% of the Internet, how much complexity to work around that is too much?
- Service providers and platform developers have access to a great deal of data which, in aggregate, could better inform these decisions.
- **HOPS BarBoF, 21:30 Sunday** (see Aaron's talk)

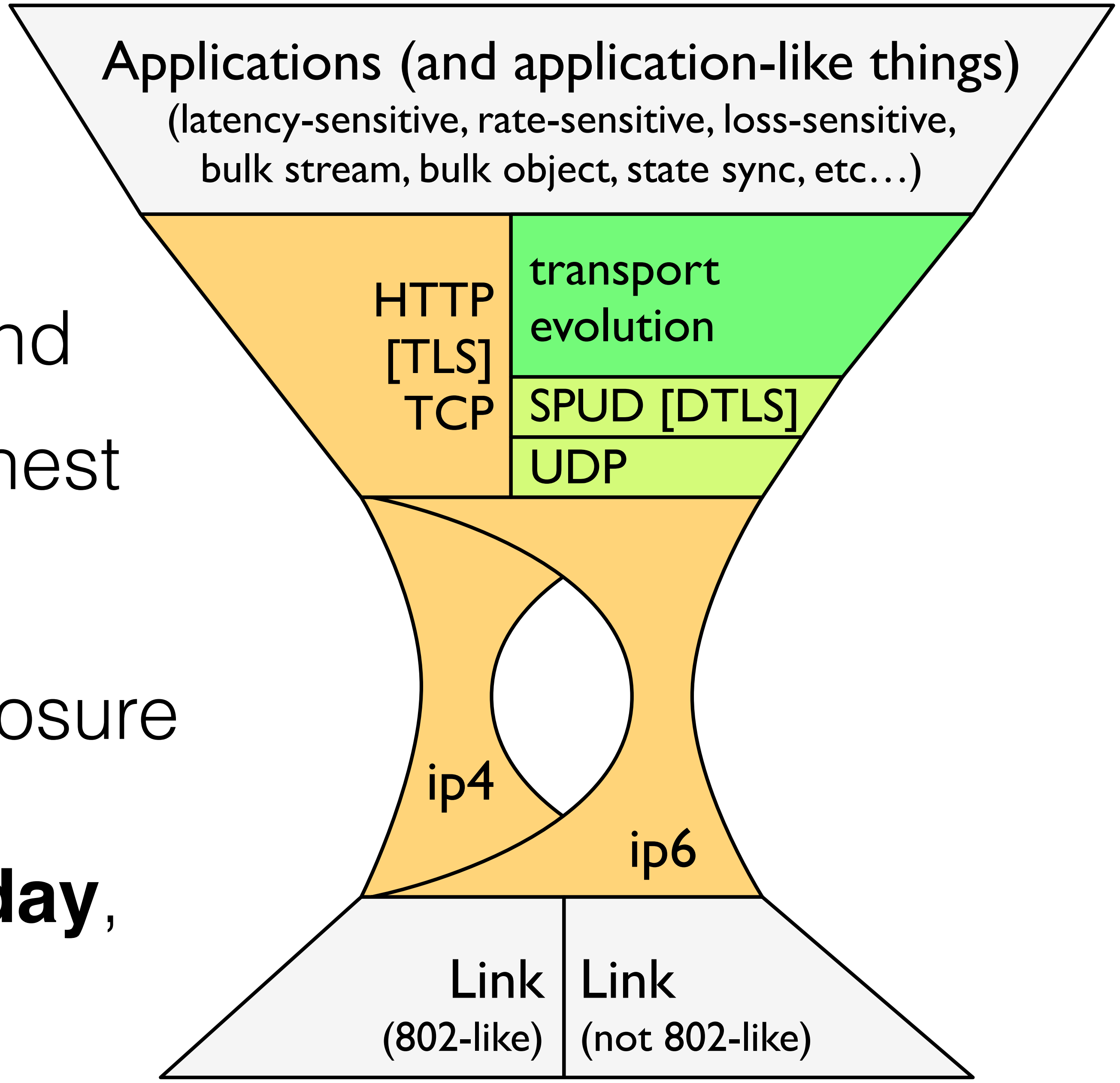
Discussion points on cooperation

- We spent a great deal of time talking about endpoint/middlebox cooperation:
  - Contracts about how packets are handled are currently implicit: should these be made explicit?
  - Encryption provides a tool to enforce a balance of power between endpoints and the middle.
  - Incentives to deployment of any new transport protocol or encapsulation are key.

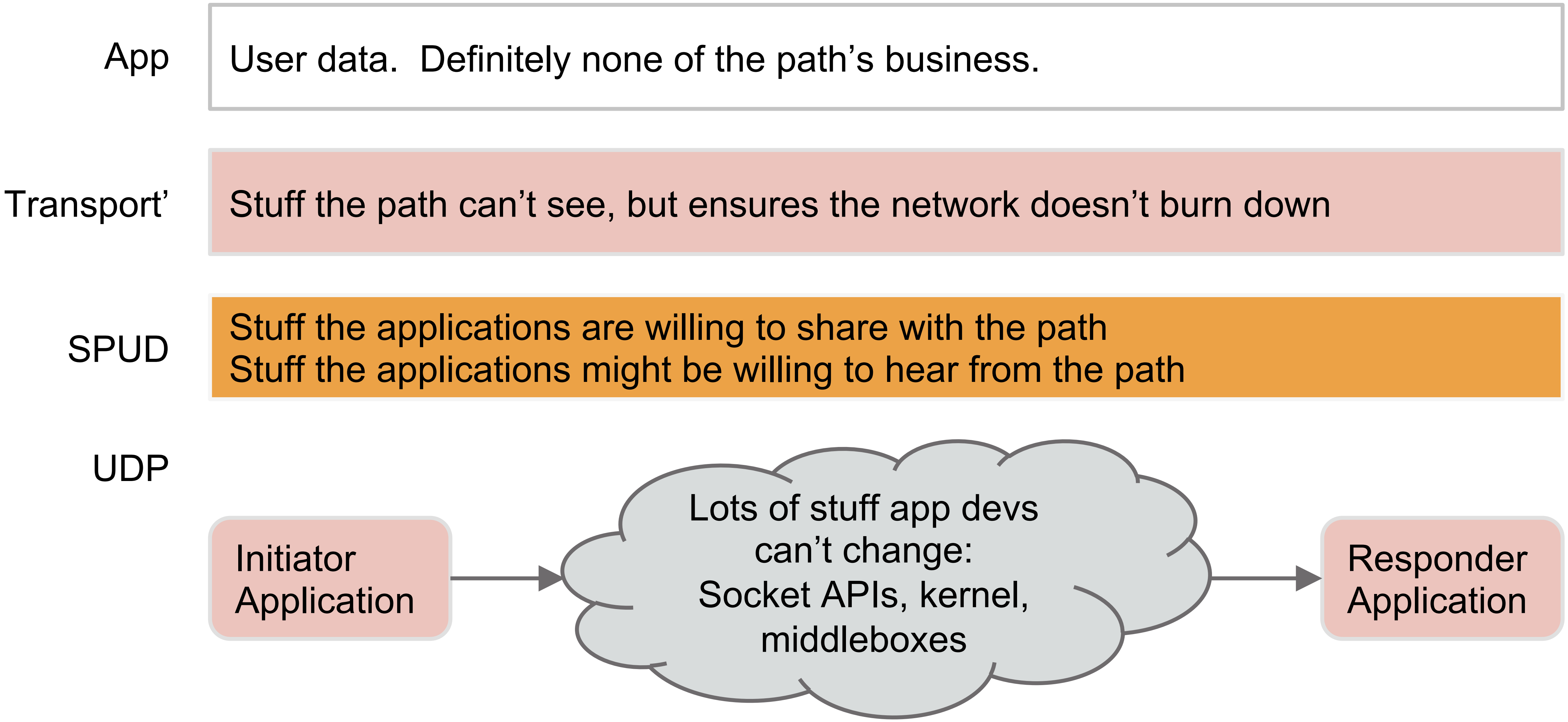
Cooperation: A new view of the two-stemmed Internet martini glass

- Expose what you must to the path
- Everything else is end-to-end
- Crypto keeps everyone honest

- Encapsulation for path exposure in user-space transports: **SPUD BoF: 9:00 Wednesday**, International room


Applications (and application-like things)
(latency-sensitive, rate-sensitive, loss-sensitive, bulk stream, bulk object, state sync, etc…)
HTTP
[TLS]
TCP
transport evolution
SPUD [DTLS]
UDP
ip4
ip6
Link
(802-like)
Link
(not 802-like)

SPUD Architecture



App

User data. Definitely none of the path’s business.

Transport’

Stuff the path can’t see, but ensures the network doesn’t burn down

SPUD

Stuff the applications are willing to share with the path
Stuff the applications might be willing to hear from the path

UDP

Initiator Application

Lots of stuff app devs can’t change:
Socket APIs, kernel, middleboxes

Responder Application

Generic mechanism

- Tube identifier + basic semantics on each packet

- In-band channel with extensible syntax to allow endpoints to signal traffic metadata (per-packet + per-tube) to each other, and devices on path

- Mechanism to allow on-path devices to signal back to either endpoint using the same in-band channel

- draft-hildebrand-spud-prototype defines an instance of this generic mechanism for experimentation

- draft-trammell-stackevo-newtea will discuss generic constraints on signaling over UDP-based encapsulations

Cooperation Vocabulary

- Once you have this mechanism, what do you say with it?
  - There need to be incentives to expose information.
  - There need to be incentives not to lie.
- A2P (app to path): problem appears tractable, there is a minimal set of useful information (e.g. session lifetime) which can be exposed, and is anyway useful to the far endpoint.
- P2A (path to app): the way forward is less clear
  - If treated as advisory: problem might be tractable; similar to ICMP, but inband.
  - If treated as authoritative: previously unsolved problem, many trust issues.

SEMI workshop TODO

- Initial workshop report: Real Soon Now (mid-April)
  - Until then: transcripts, slides, position papers at https://www.iab.org/activities/workshops/semi/
- Cooperation with ETSI NFV Forum on middlebox issues (in progress)
- Discussions on transport extensibility in area meetings
- UDP encapsulation guidelines
- Statement on architectural assumptions in transport evolution (referred to program)

Further Discussion

- Middlebox measurement issues
  (“How Ossified is the Protocol Stack”):
  hops@ietf.org
- Substrate Protocol for User Datagrams
  spud@ietf.org (and come to the BoF)
- Transport Services WG
  taps@ietf.org
- Other future work
  stackevo@iab.org