

Can transport protocols be different?

Gorry Fairhurst

University of Aberdeen

Talk to READY Workshop, Kista, Sweden 30th Nov 2016



measurement and architecture for a middleboxed internet

measurement

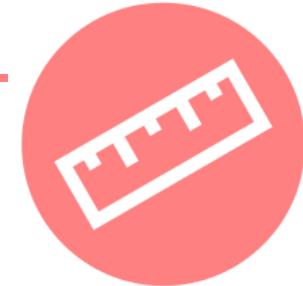
architecture

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..



Overview



The reality of the Internet

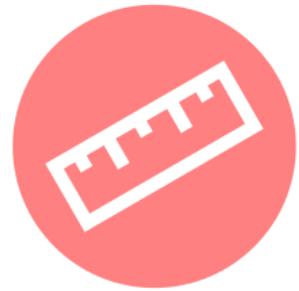
Finding out what is going on

Let's make a Bestiary

What is to be done?

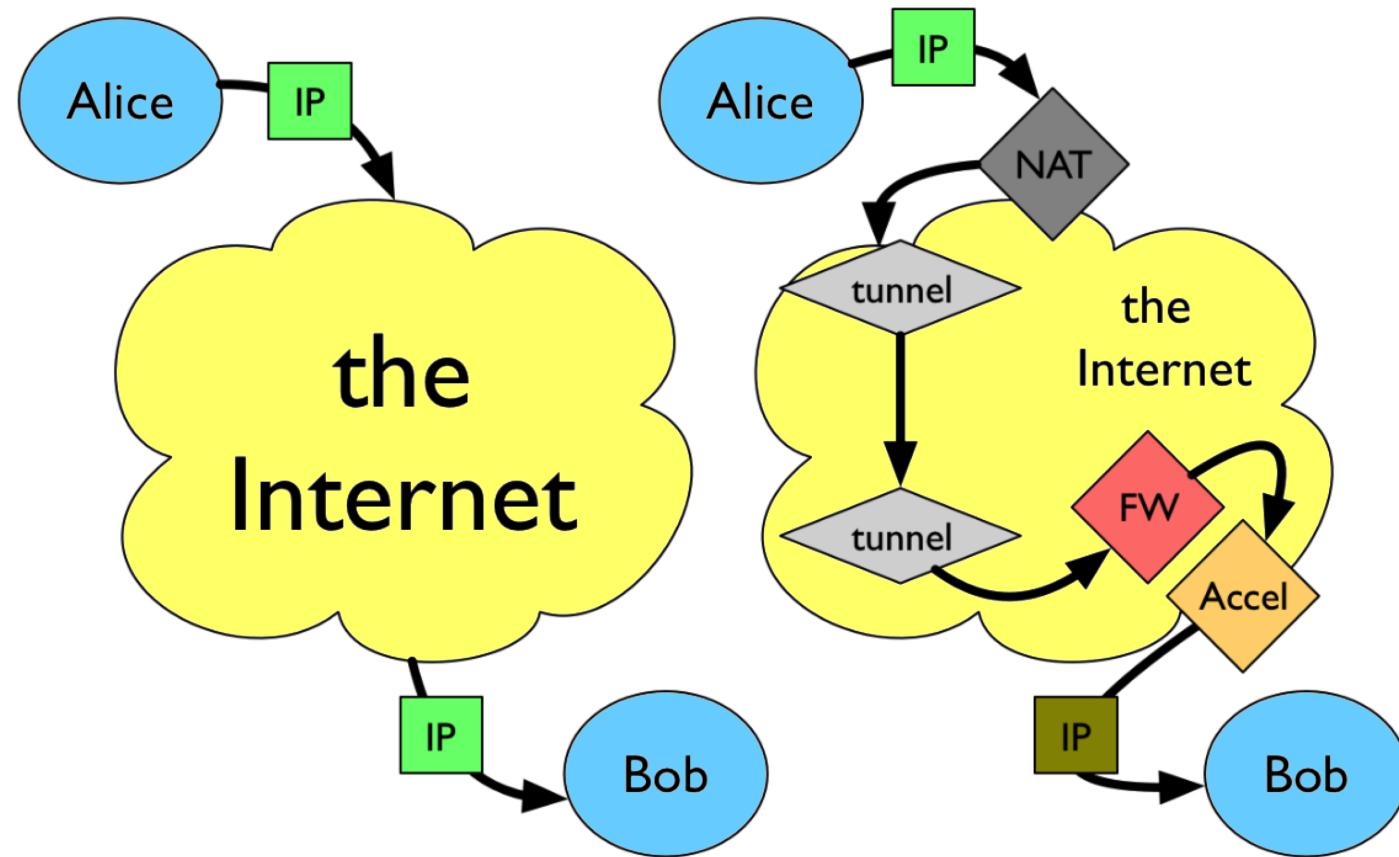
I want to work differently!





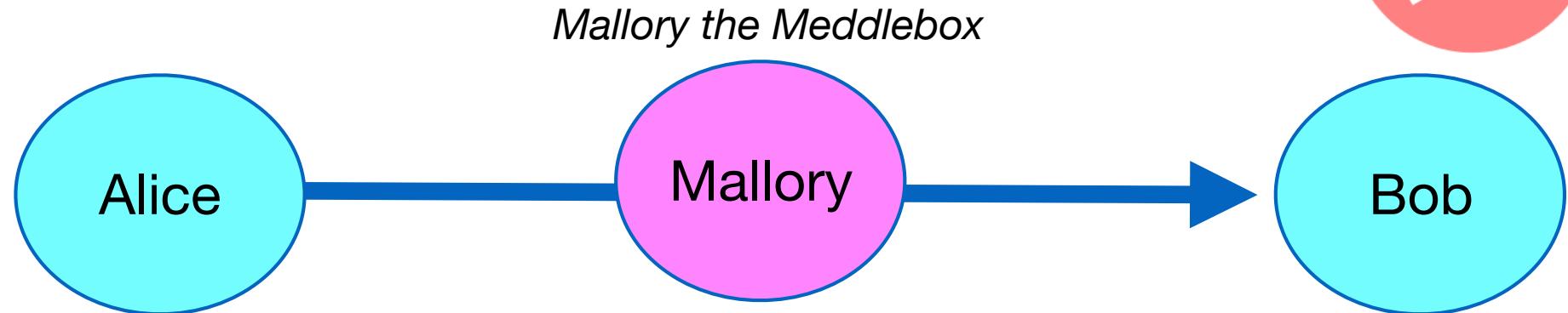
The reality of the Internet

What is going on?



Routing, ACLs, Firewalls, Tunnels, ECMP, etc

What is really going on?



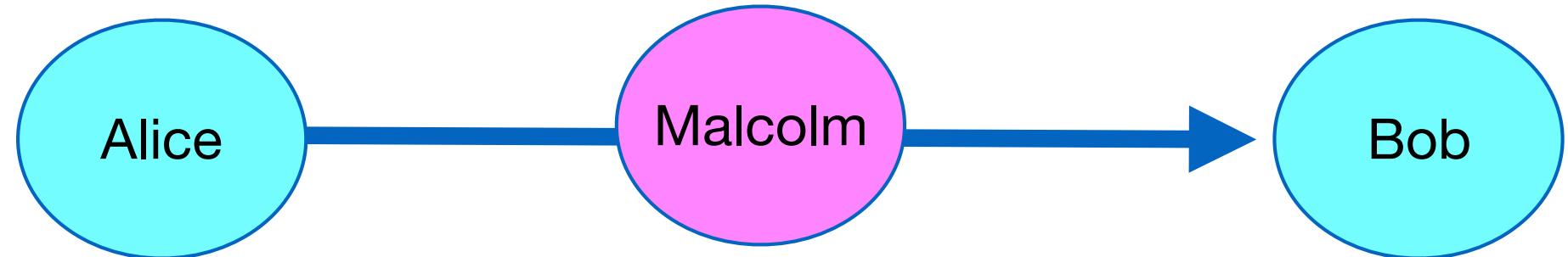
Filtering and tracking

- UK government pressured ISPs to filter traffic to prevent children and young people from content deemed unsuitable
- Results for Top 100,000 UK websites:
 - TalkTalk Strict: 12982 blocked (13%)
 - Sky: 7006 blocked (7%)
 - TalkTalk Kidsafe: 5587 blocked (6%)
 - Virgin Media: 4320 blocked (4%)



blocked.org.uk

Actually, far from the total story

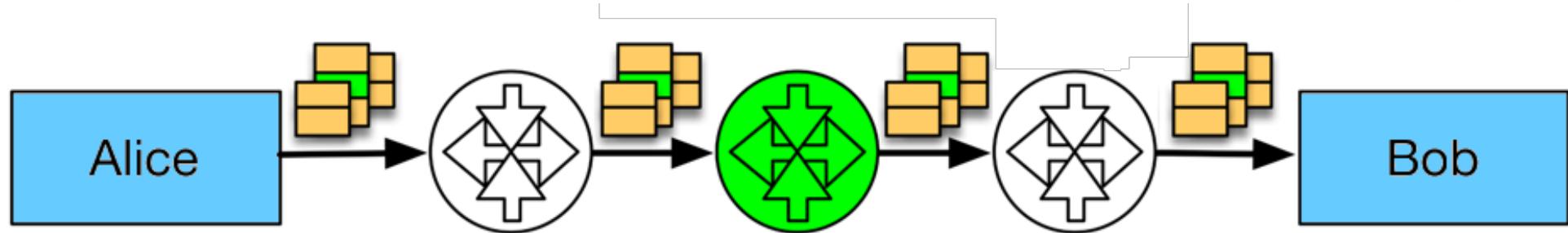
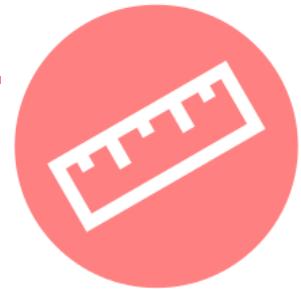


- **Support for billing and operational reasons**
 - Classify to service platforms/bearers/profiles
 - Policers/Shapers/Volume limits

- ***Provide acceleration:***
 - Transparently cache content
 - Transcode content
 - Redirect connections
 - Load balance

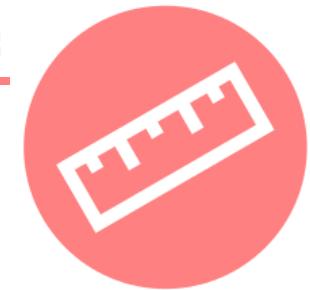


Finding out what is going on



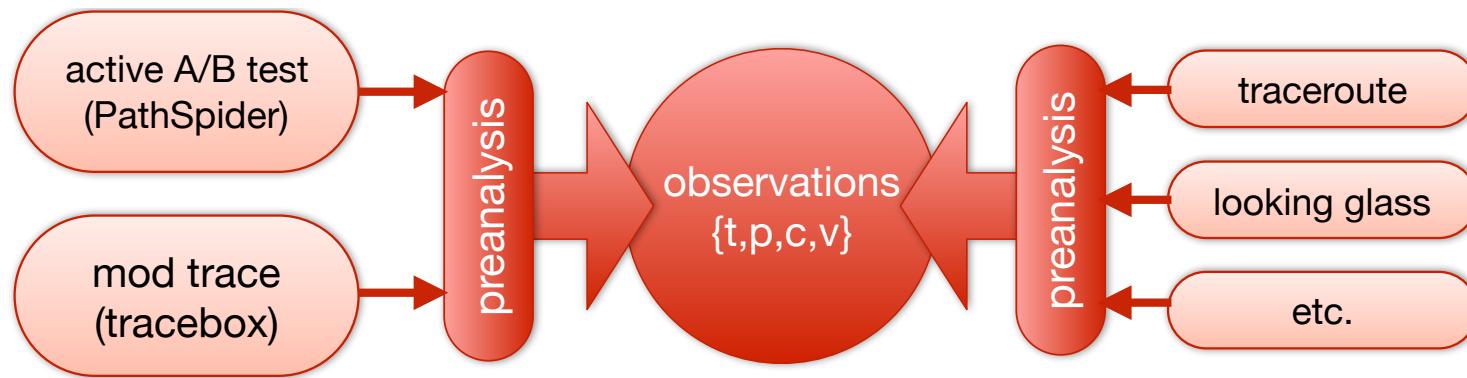
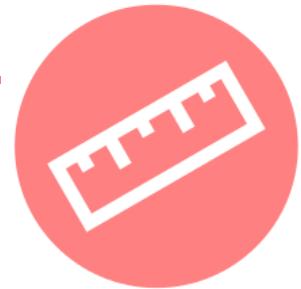
- Simple conclusion is the network is now broken
- Network measurement is hard...
 - “Big five” metrics (loss, latency, jitter, rate, reordering)
 - Getting measurement right is even harder!
 - Measurements are essential to effectively operating networks
 - Measurements crucial to protocol development and peer review

Finding out what is going on



```
2016-10-18T16:16:29.127972263Z tracebox to 139.133.217.73 (exile.erg.abdn.ac.uk): 64 hops max
2016-10-18T16:16:29.158536934Z 1: 192.168.0.1 (192.168.0.1) 0ms
2016-10-18T16:16:32.235704787Z 2: *
2016-10-18T16:16:32.360431055Z 3: 10.9.14.121 (10.9.14.121) 31ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:32.524192494Z 4: 81-228-103-0.link.se.telia.net (81.228.103.0) 17ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:32.641387877Z 5: hy-peer4-link.se.telia.net (81.228.86.35) 21ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:32.754444139Z 6: s-b6-link.telia.net (62.115.114.16) 21ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:32.905966875Z 7: s-bb4-link.telia.net (62.115.114.170) 18ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.058814051Z 8: hbg-bb4-link.telia.net (62.115.136.5) 30ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.210298313Z 9: ldn-bb2-link.telia.net (62.115.115.75) 53ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.368238117Z 10: ldn-b5-link.telia.net (213.155.132.197) 48ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.506478489Z 11: jisc-ic-318436-ldn-b5.c.telia.net (62.115.148.161) 49ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.687750079Z 12: ae29.londtw-sbr2.ja.net (146.97.33.9) 46ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:33.887723786Z 13: ae31.lowdss-sbr1.ja.net (146.97.33.29) 49ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:34.031737247Z 14: ae29.leedaq-sbr1.ja.net (146.97.33.49) 47ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:34.171431026Z 15: ae0.dund-ban3.ja.net (146.97.37.182) 50ms IP::DiffServicesCP IP::TTL IP::CheckSum
2016-10-18T16:16:34.335674991Z 16: ael.dund-ban1.ia.net (146.97.64.98) 55ms IP::DiffServicesCP IP::TTL IP::CheckSum
```

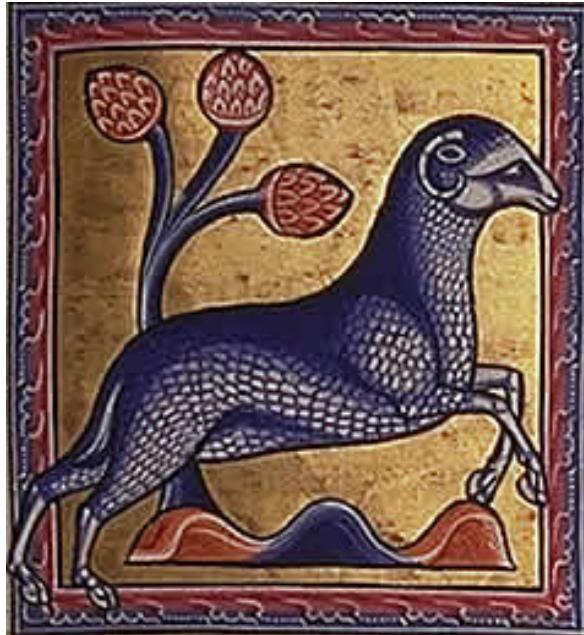
Collecting Measurements



Empirical measurements concerning what kinds of middleboxes perform what kinds of manipulations on what paths...

- Link **observations** about **conditions** on a **path** at a given **time**
- Publish data in an open **observatory** (Public end 2016)

Let's make a *Bestiary*



Delay beasts
Rate beasts

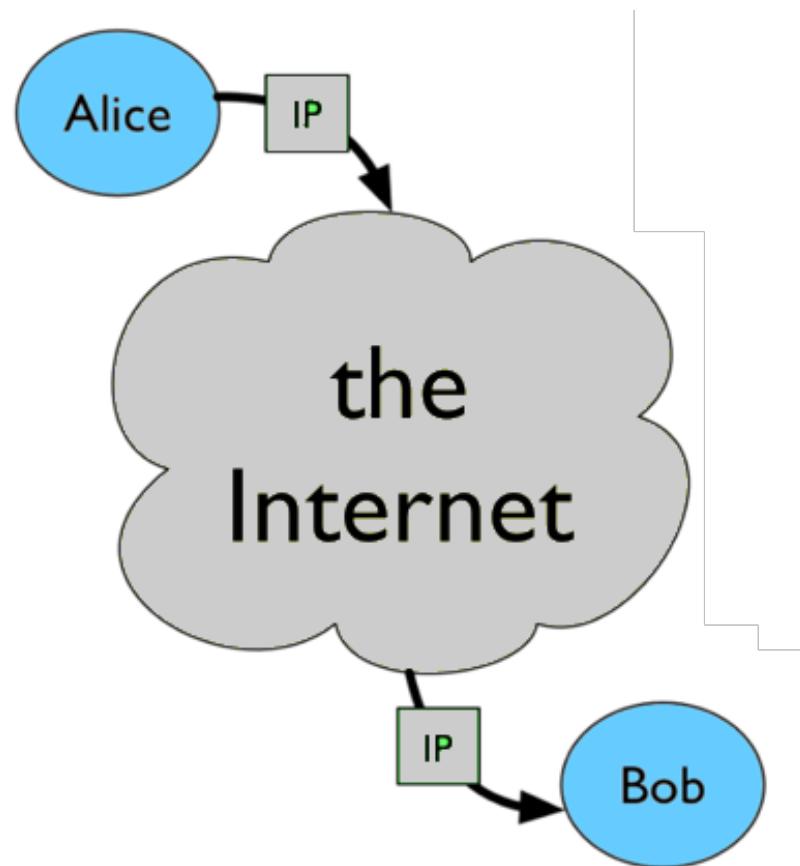
Loss beasts
Reordering beasts

Fragile beasts

I would love to create a *Middlebox Bestiary*

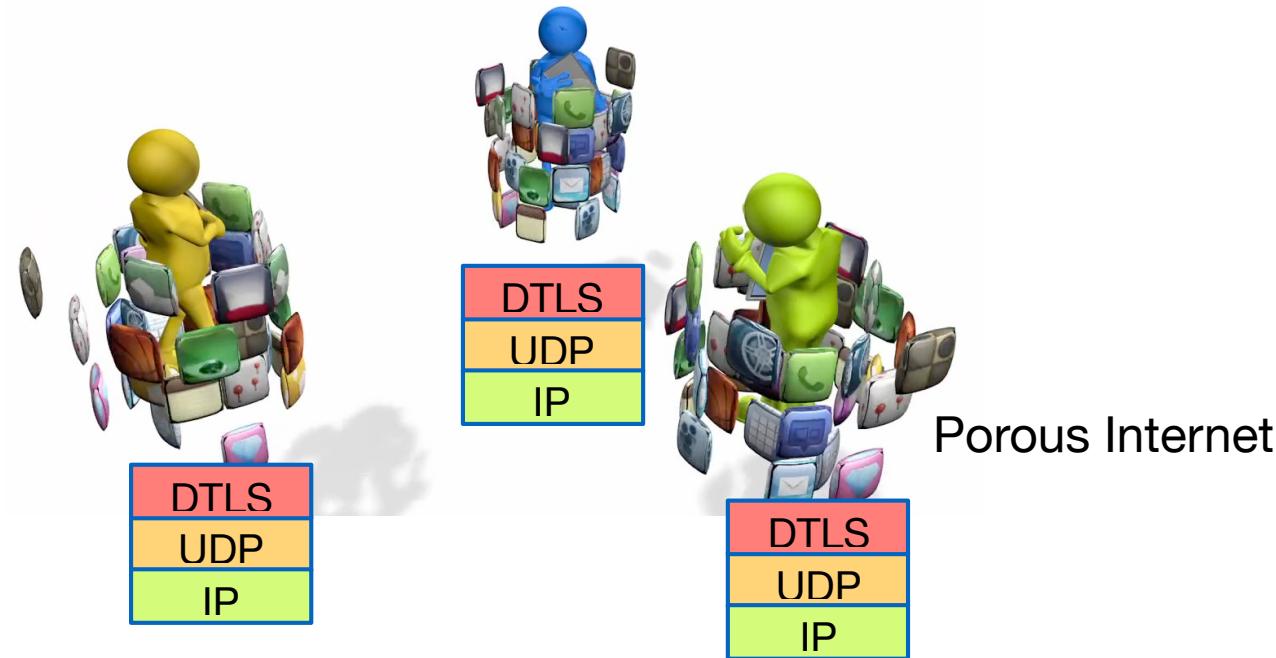
What is to be done?

- Use encryption of transport to enforce end-to-end-ness (prevent meddling) - actually not new (IPSEC/IPv6!)





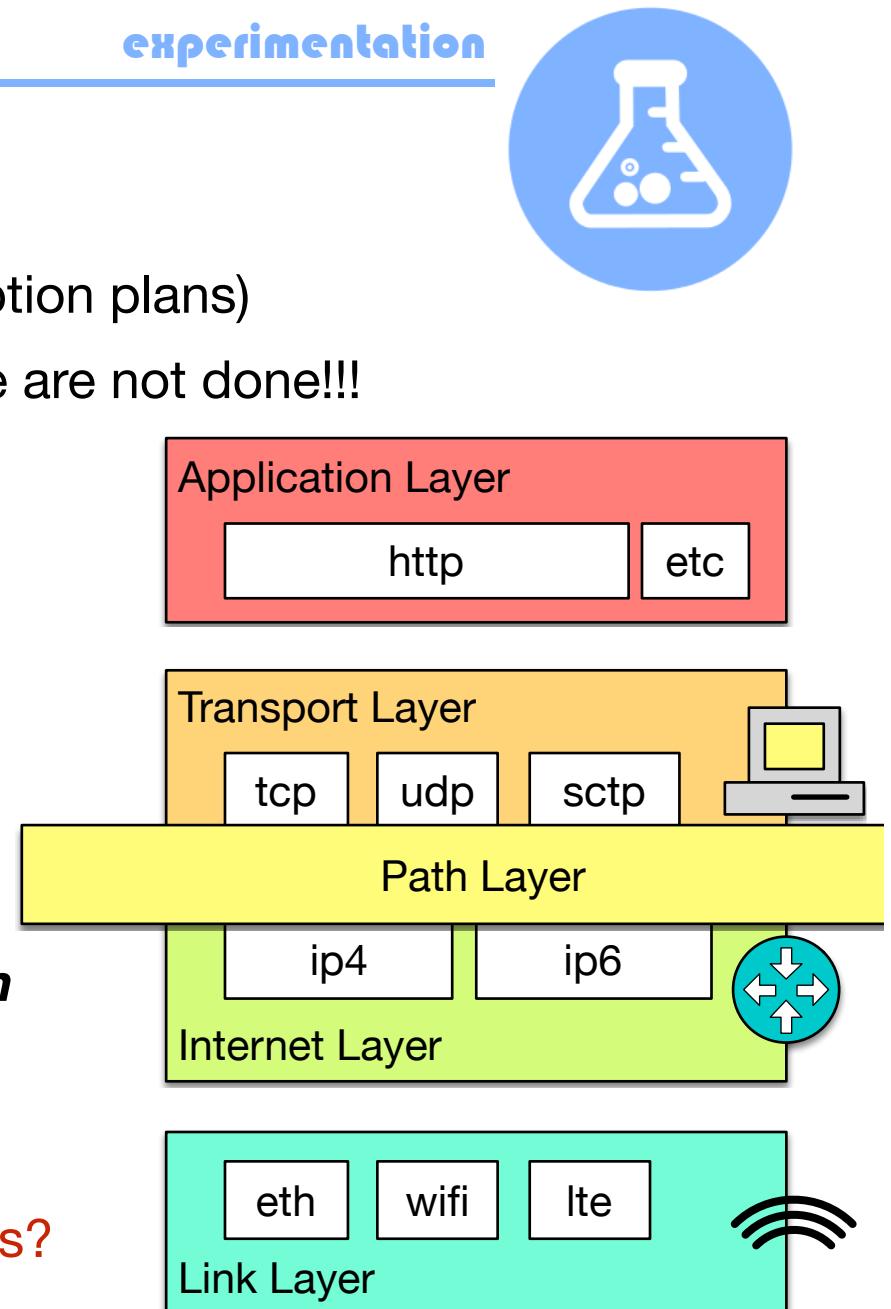
Encryption with Application-layer updates



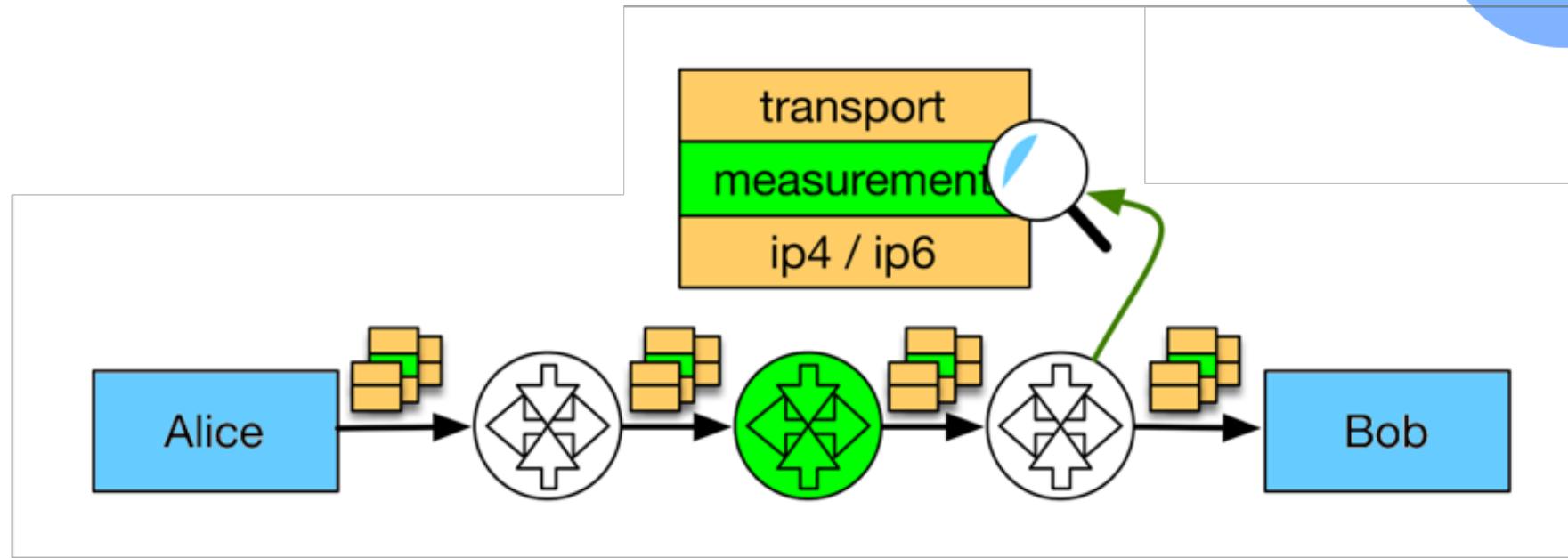
- Eliminates meddling by middleboxes (***don't trust the network***)
- User land stacks reduce time to deploy mechanisms
- Fragmentation of the stack ***and research experience***
- Could encourage ***performance engineering*** for a product...
- Could encourage ***engineering for a popular path***...
- Will make it hard for operators to diagnose network problems

I want to work differently

- Middleboxes are not going away (despite encryption plans)
 - Operators are planning CDN, SDN, NFV ... we are not done!!!
 - Do operators ***trust the applications/users?***
- Is it possible then to combine....
 - Encryption to enforce end-to-end-ness
 - Explicit ***exposure of appropriate information***
 - Let's re-enable measurement in the net
 - Can a middlebox trust more...
 - What may encourage a box trust flows labels?



Can transport protocols be different?



- **Revealing path info helps...**
 - Assist middlebox flow-state management
 - Let operators/managers know what is working
 - Understand if methods perform as designed (fall-back)
 - Gather measurements and help understand problems

Project details



- Research and Innovation Action **688421**
- 30 months from 1 January 2016 (End June 2018)
- <https://mami-project.eu/>
- <https://github.com/mami-project>

Some links to tools:

www.tracebox.org/ and <https://github.com/tracebox/tracebox>
pathspider.net

Useful references



- End-to-end Arguments in System Design by J. H. Saltzer, D. P. Reed, and D. D. Clark.
 - In ACM Transactions on Computer Systems 2(4), 1984, pages 277–288.
- End-to-end Internet Packet Dynamics by V Paxson. In SIGCOMM Computer Communications Review 27(4), 1997, pages 139–152.
- On Estimating End-to-end Network Path Properties by M Allman and V Paxson. In Workshop on Data Communication in Latin America and the Caribbean, San Jose, Costa Rica, 2001, pages 124–151.
- Middlebox communication architecture and framework by J Kuthan, J Rosenberg, A Molitor, P Srisuresh, and A Rayhan. RFC 3303, 2002.
- Measuring the Evolution of Transport Protocols in the Internet by A Medina. In SIGCOMM Computer Communications Review 35, 2005, pages 37–52.
- Netalyzr: Illuminating The Edge Network by C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. In the Internet Measurement Conference (IMC), 2010.
- Is It Still Possible to Extend TCP? by M Honda, Y Nishida, C Raiciu, A Greenhalgh, M Handley, and H Tokuda. In ACM SIGCOMM Conference on Internet Measurement Conference, Berlin, Germany, 2011, pages 181–194.
- Measuring the State of ECN Readiness in Servers, Clients, and Routers by S Bauer, R Beverly, and A Berger. In ACM SIGCOMM Conference on Internet Measurement Conference, Berlin, Germany, 2011, pages 171–180.2012
- OONI: Open Observatory of Network Interference by A Filasto and J Appelbaum In 2nd USENIX Workshop on FOCI, 2012.
- On the state of ECN and TCP options on the Internet by M Kühlewind, S Neuner, and B Trammell. In the PAM Conference, Hong Kong, China, 2013, pages 135–144.
- Are TCP Extensions Middlebox-proof? by B Hesmans, F Duchene, C Paasch, G Detal, and O Bonaventure. In Workshop on Hot Topics in Middleboxes and Network Function Virtualization, Santa Barbara, California, USA, 2013, pages 37–42.

Useful references



- Global Network Interference Detection over the RIPE Atlas Network by C Anderson, P Winter, and Roya.
 - In the 4th USENIX Workshop on FOCI, 2014.
- Here Be Web Proxies by Nicholas Weaver, C Kreibich, M Dam, and V Paxson. In 15th PAM – LA, CA, USA, 2014, pages 183–192
- Inline Data Integrity Signals for Passive Measurement by B Trammell, D Gugelmann, and N Brownlee. In the Traffic Monitoring and Analysis: 6th TMA 2014, London, UK, 2014, pages 15–25.
- A Middlebox-cooperative TCP for a Non End-to-end Internet by R Craven, R Beverly, and M Allman. In SIGCOMM CCR 44(4), 2014, pages 151–162.
- Evolving Transport in the Internet by B Trammell and J Hildebrand. In IEEE Internet Computing, 2014.
- Explicit Congestion Negotiation (ECN) support based on P2P networks by E Gubser. Masters's thesis, ETH Zurich, 2015.
- Enabling Internet-Wide Deployment of Explicit Congestion Notification by B Trammell, M Kühlewind, Damiano Boppart, Iain Learmonth, Gorry Fairhurst, and Richard Scheffenegger. In PAM Conference, New York, USA, 2015, pages 193–205.
- Experiences Deploying a Transparent Split TCP Middlebox and the Implications for NFV by F Le, Erich Nahum, V Pappas, M Touma, and D Verma. In ACM SIGCOMM Workshop on Hot Topics in Middleboxes and Network Function Virtualization, London, United Kingdom, 2015, pages 31–36.
- Informing Protocol Design Through Crowdsourcing: the Case of Pervasive Encryption by A M Mandalari, M Bagnulo, and A Lutu. In ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big Data (C2B(I)D), August 2015.
- A new transport encapsulation for middlebox cooperation by B Trammell, M Kühlewind, E Gubser, and J Hildebrand. In IEEE Conference on Standards for Communications and Networking, Tokyo, Japan, October 2015.
- Is it possible to run the Internet over UDP? by B Trammell, M Kühlewind, and E Aben. MAP Research Group at IETF 95, 2016.
- Towards an Observatory for Network Transparency Research by S Neuhaus, R Muamp;, K Edeline, B Donnet, and E Gubser. In the ANRW, Berlin, Germany, 2016, pages 71–73.
- PATHspider: A tool for active measurement of path transparency by Iain R. Learmonth, B Trammell, M Kühlewind, and G Fairhurst. In ANRW, 2016, pages 62–64.