A Ten Minute Introduction to Middleboxes

Justine Sherry, UC Berkeley
This Talk: Three Questions!

What is a middlebox?

What are some recent trends in middlebox engineering?

What research challenges do middleboxes present?
What is a middlebox?

Also called a “network appliance” or a “network function.”

“A middlebox is defined as any intermediary device performing functions other than the normal, standard functions of an IP router on the datagram path between a source host and destination host.”


Fun fact: the term “middlebox” was coined by Lixia Zhang.
What is a middlebox?

Primarily deployed for security and performance benefits.

- Firewalls
- Application Firewalls
- Intrusion Detection Systems (IDS)
- Intrusion Prevention Systems (IPS)
- Proxy/Caches
- WAN Optimizers
- Protocol Accelerators

Many other uses too!

Billing and usage monitoring, asset tracking, Network Address Translation, protocol converters (6to4/4to6)…
Example: Intrusion Prevention System

Security Appliance.

Monitors all open connections to detect and block suspicious activity.

What defines suspicious activity? Traditionally: “signatures”.

alert tcp $HOME_NET 20034 -> $EXTERNAL_NET any (flow:to client,established; content:"BN|10 00 02 00|"; depth:6; content:"|05 00|"; depth:2; offset:8; classtype:trojan-activity; sid:115; rev:15;)

-> This signature represents that a host is infected with a botnet.
Example: Web Proxy

Performance-Improving Appliance.

Caches web content to improve bandwidth consumption and page load times.
What is a middlebox?

Key differences between middleboxes and routers:

Middleboxes are often **stateful**. They remember fine-grained data that is updated as frequently as every packet or every connection.

Middleboxes perform **complex and varied operations** on packets. There are new categories of middleboxes on the market every year.
Who uses middleboxes? (Primarily)

Enterprises: “1/3 : 2/3 Rule”...
Sherry et al. “Making Middleboxes Someone Else’s Problem” SIGCOMM 2012
Potharaju and Jain. “Demystifying the Dark Side of the Middle” IMC 2013

...and ISPs...
Kreibich et al. “Netalyzer: Illuminating the edge network.” IMC 2010
Xu et al. “Investigating Transparent Web Proxies in Cellular Networks” PAM 2015

...and even your home router likely has some middlebox capabilities!
What are some recent trends in middlebox engineering?

“Network Functions Virtualization”

“Network functions virtualization (NFV) is an initiative to virtualize the network services that are now being carried out by proprietary, dedicated hardware.”

—SearchSDN.com

This is definitely a “buzzword” you will hear at SIGCOMM!
What are some recent trends in middlebox engineering?

Enables innovation and experimentation!

- Dedicated customized hardware
- x86 middleboxes implemented in software
- virtualized software middleboxes running in a datacenter (NFV)
An aside: Why I Think Middleboxes are *Fun*

Breadth and Generality = Your Imagination is the Limit

Rise of Software Implementations = Easy to Build and Experiment With

“1/3 : 2/3 Rule” and Rise of NFV = Opportunities for industrial impact
Compatibility: Do middleboxes harm our ability to deploy new protocols?

What if I want to use HTTP 2.0, but my web proxy only knows how to use 1.5?

Justine’s reading list:
Raiciu et al. “How Hard Can It Be? Designing and Implementing a Deployable Multipath TCP” NSDI 2012
RFC 6886: “NAT Port Mapping Protocol (NAT-PMP)”
What research challenges do middleboxes present?

(2) State. How does network management change — e.g., in terms of scalability and fault tolerance — when state is involved?

If a NAT crashes during my connection, does my connection get reset — and all of my neighbors too?

Justine’s reading list:
Rajagopalan et al. “Split/Merge: System support for elastic execution in virtual middleboxes.” NSDI 2013
What research challenges do middleboxes present?

If a NAT crashes during my connection, does my connection get reset — and all of my neighbors too?

Check out “Rollback Recovery for Middleboxes” on Wednesday!
What research challenges do middleboxes present?

(3) Privacy. Should users have to give network operators the ability to read all of their network traffic in order to receive network services?

Middleboxes today either do not operate on TLS/SSL traffic or perform a “man in the middle” (attack) on the connection!

Justine’s reading list:
What research challenges do middleboxes present?

Middleboxes today either do not operate on TLS/SSL traffic or perform a “man in the middle” (attack) on the connection!

Check out two papers on Wednesday!

“multi-context TLS (mcTLS): Enabling Secure In-Network Functionality in TLS”

and

“BlindBox: Deep Packet Inspection over Encrypted Traffic”
What research challenges do middleboxes present?

(3) Censorship. How can users detect that middleboxes are used to censor them — and how can they avoid it?

Justine’s reading list:
Marzak, Weaver, et al. “China’s Great Cannon” University of Toronto, April 2015
(4) NFV: Management. How do we build frameworks for NFV like cloud computing has for compute? (e.g. OpenStack, EC2)

Justine’s reading list:
What research challenges do middleboxes present?

(4) NFV: Management

Check out “Scaling Up Clustered Network Appliances with ScaleBricks” on Wednesday!
What research challenges do middleboxes present?

(5) How do middleboxes “fit in” with Software Defined Networking?

Justine’s Reading List:
Qazi et al. “SIMPLE-fying Middlebox Policy Enforcement with SDN” SIGCOMM 2013
What research challenges do middleboxes present?

So. Much. More.

What policies are different ISPs enforcing using middleboxes?

Do middleboxes break the end to end principle? Should we care?

Can we get rid of middleboxes and do all the same work at the edge?
Where to learn about Middleboxes at SIGCOMM!

(1) The Middlebox Session at the Main Conference. Wednesday, 8:50 AM

(2) The HotMiddlebox Workshop
Friday, All Day — it’s not too late to register!
Fin.

slides: http://cs.berkeley.edu/~justine/mbpreview.pdf
me: justine@eecs.berkeley.edu
@justinesherry in the Twitterverse.