Research Faculty Summit 2018
Systems | Fueling future disruptions
Towards Self-managing Networks

Behnaz Arzani
Post doctoral researcher at Microsoft
Today, managing networks is expensive

- 1096 reports
- The engineer’s time is largely spent on debugging/configuring the network
- This time could be spent on improving the network itself
  - Designing new protocols
  - Adding new functionality
  - Upgrading to new technology: e.g. P4.

Albert Greenberg
Head of Azure Networking
There has been a lot of networking research.

So why do we still need people to manage/configure our networks?

Vint Cerf and Bob Kahn
Publish first TCP/IP paper in 1973
In this talk:
I will describe my research and what I’ve learned about why we still don’t have a self-managing network.
Talk Outline

• Diagnosis

• Security

• What are the common themes?

• How can AI help?
Availability is important!

- Performance
- Stranding
- Velocity
- Manageability
- Availability

Borrowed from Amin Vahdat from Google PhD summit talks 2017
Failures are disruptive
Why is diagnosis hard?
What happens if no one takes responsibility for the failure?

Someone accepts responsibility

Each blames the other
A real example of this happening: Event17

• Azure uses virtual hard drives for storage
  • VMs connect to remote storage for read/write to disk
• Failures often can result in a VM to panic and reboot
• What happened?
  • Storage blames network
  • Network blames storage
  • But who was it?
NetPoirot
(SIGCOMM 2016)
A solution to this problem should

- Allow for monitoring the client, service, and the network

NetPoirot:

Use TCP statistics to identify whether the problem was because of the client, server, or the network
NetPoirot, an overview

Fault injector

Learning Agent
Not all faults are the same

lower drop rate

Associate failed links with problems they cause

High drop rate
Mapping complaints to faulty links

But operators don’t always know where the failures are either
Clouds operate at massive scales

Problems can and will happen*


Each Data center has millions of devices
Low congestion drop rates add noise

One-off, transient, drops do occur on many links and add *noise* to diagnosis*

Solution Requirements

- Detect short-lived failures
- Detect concurrent failures

Failure: any systemic cause of packet drop whether transient or not
Want to avoid infrastructure changes

• Costly to implement and maintain
• Sometimes not even an option
  • Example: changes to flow destinations (not in the DC)
A “strawman” solution

• Suppose
  • we knew the path of all flows
  • we knew of every packet drop
• Tomography can find where failures are

*If we assume there are enough flows*
Example of doing tomography

\[ x_{14} + x_{43} = 0 \]
\[ x_{14} + x_{42} = 1 \]
\[ x_{34} + x_{42} = 1 \]

\[ x_{ij} \in \{0, 1\} \]

\[ x_{ij} = 0 \] if link not dropping packets
\[ x_{ij} = 1 \] if link dropping packets

Only solvable if we have \(N\) independent equations

\(N = \text{number of links in the network}\)
How 007 works

Monitoring agent: Deployed on all hosts
Notified of each TCP retransmission (ETW)
Path discovery agent finds the path of the failed flows
Flows vote on the status of links
How 007 works
Talk Outline

• Diagnosis

• Security

• What are the common themes?

• How can AI help?
Detecting compromised VMs

Ideally...
In practice…

There is a need for agent-less compromise detection systems

The operator needs customer permission to install antimalware
Talk Outline

• Diagnosis

• Security

• What are the common themes?

• How can AI help?
Why do we need operators managing our networks?

- The ideal data needed to solve the problem can be missing.
- Gathering the right data can be expensive.
- Sometimes it's not even clear what the right data is?
- Sometimes, there are datasets that indirectly point to the solution but it's hard to derive that such correlation exists.
Talk Outline

• Diagnosis

• Security

• What are the common themes?

• How can AI help?
Where can AI help?

• Prediction
  • Monitoring can be expensive
  • Prediction can help turn on expensive monitoring when it is needed

• Identifying (complex) correlations
  • Sometimes there is data that can help solve the problem
  • The relationship between the data and the problem may be unintuitive
  • E.g. NetPoirot – using TCP statistics to find the cause of client/server problems

• Identifying when operator help is really needed
  • Despite our best efforts, sometimes an operator should intervene
  • AI can help reduce the noise
Thank you!