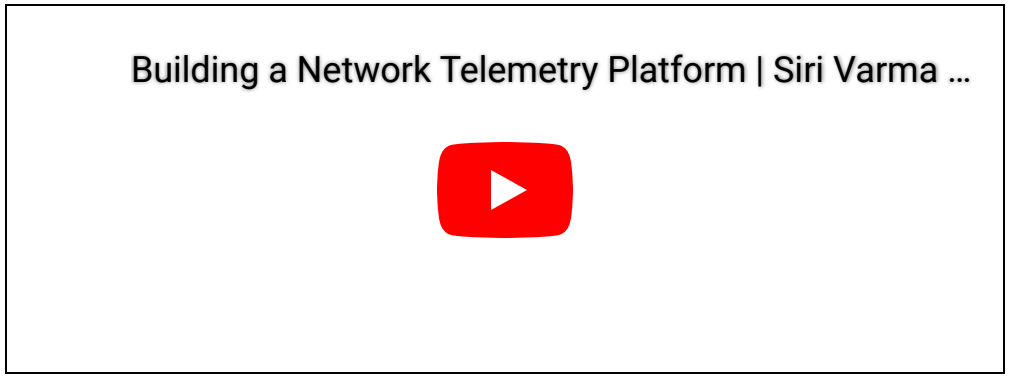
Conf42 Platform Engineering 2024 - Online

September 05 2024 - premiere 5PM GMT

Subscribe to watch

Watch this talk



Building a Network Telemetry Platform to Minimize Security Threats

Video size:

Abstract

In today's digital landscape, network security is paramount. Imagine a platform that gathers telemetry from virtual machines, scrutinizes traffic for anomalies, and alerts teams to take swift action. This approach minimizes the attack surface, ensuring security across the entire infrastructure.

Summary

Transcript

This transcript was autogenerated. To make changes, submit a PR.
Hello, everyone. Welcome to my talk on building a network telemetry platform to minimize security threats. As part of this presentation, we'll dive into network telemetry, the need for a platform and what are the things from the platform needed to minimize security threats. With that, let's get started. First, we'll take a look at why we need network telemetry and the motivation for building a platform. Second, we'll look into the fundamental building blocks required for creating such a platform. Third, we'll use the blocks and create an architecture that will be able to ingest billions of network events every day. And Fourth, once we have the data from the platform, how we can use it to minimize security threats. That is what we look into. Last but not the least, we'll take a look at how we can extend our network telemetry platform to other feature sets. Why we

need network telemetry. Gartner predicts by end of, by 2028, 50 percent of the enterprise companies will be on cloud. With the, with such rapid adoption, companies are also facing increased cyber security risks. Zero Trust, a famous cloud security model, also emphasizes the need for having, also emphasizes the need for having network segmentation so that during an attack, or, yeah, during an attack, the attack surface is minimized. Zero Trust. So how can we achieve this segmentation? One way is you have to first understand the network topology of your organization. And in order to understand network topology, that is where network telemetry will help. Second, threat detection and anomaly detection. Imagine a threat actor has entered your infrastructure. How will you detect that? Either using either by identifying a malicious or an unknown IP. Next, let's say the threat actor is trying to access a bunch of different resources in a short duration, which is definitely an anomaly or an abnormality. You can run machine learning algorithms on your network telemetry data to even identify these. So you have telemetry to identify unknown or malicious IPs as well as anomalies. Lastly, compliance. Imagine you are a network service provider or a service provider in the cloud, and you are offering services to a ton of government agencies or you run in the national security regions. Network telemetry is an important part of your audit. Why we need a platform. Think of platform as a framework That you're setting up so that others can build on it. So it is not just one solution, but it is an enabler for endless solutions Now that you have built a platform if there are a bunch of scenarios that you want to enable on this It is easy for the service teams to onboard what they have to do is Come to your platform take the data you have and run with it. So you're Saving a lot of cost and time for the company Third, we, you have also achieved consistency and standardization because there is only one source of truth. You now have consistency with the way how resources are accessed because you have one single store where customers can come in and access the data and standardization because your schema and contracts are fixed and whatever customers have and any partner teams that want to onboard have to run with the schema. So what

are the critical fundamental blocks in a network telemetry platform? Like any other platform you have Services that are collecting data Then you have enrichment going on which is responsible for adding meaningful data to whatever you have right now And then standardization as we were discussing previously One single way or one single pane of glass experience for our customer teams Let's look at a typical architecture on the left. You see there are a bunch of vms And the VMs have something called an agent running in the behind. This is a background agent which runs at regular intervals, maybe every two minutes or every one minute, and identify all the ingress and decrease traffic and log that onto your system. If it is a Linux based system, you can use tcpdump to take the snapshot, and if it is a Windows system, you can use windowsdump. Now, with all that information, logged onto the disk, you have another agent that is responsible for shipping the telemetry data from there to a data store or a SIC. One critical thing to remember here is, it is better to think about the schema design from the beginning itself. So what you see below is a schema that emit, that is emitted from the TCP dump, right? Source IP, source port, destination IP, destination port. And then the VM ID is something you attach. Now with all this information available in the data store, also one more thing to emphasize is these are billions of events. Now you need a big data processing pipeline to capture all this data. So either you can use streaming or a bad job to Start processing the data and add additional enrichments. These enrichments can be anywhere from adding more VM metadata like VM names, the account the VM belongs to, or the VM SKU details to mapping this mapping the IP address, either the source or destination to a particular IP range. generally have all the data stores hosted between a particular IP range. And once you have a particular IP, determining which IP range it belongs to will definitely smoothen your operations later. So your big data or your streaming jobs can also do that. And once you have enriched all this information, you can put this in a final data store, which will become the source of truth for your consumers. To reiterate, our platform architecture consists of collectors, which is your agent and fluency agent gathering all the data, sending it to a data

store. And sending it to a data store. Then you have enrichment, which is going on in your big data processing pipeline, where you add bulk of data or additional metadata, and then you have your destination where all the events are present so that the consumers can use that. Now with that information available, what we have here is an alerting service that is able to go through the telemetry data and alert the security teams based on any vulnerability threats. Let's look at how we can implement the network security itself. First and foremost, for implementing network security, you need baselines. At its core, baselines are the minimal security policies that every team on the organizations should follow. These policies can range anywhere from mandating Static analysis tools and secret analyzers into your build process to blocking high risk ports like 22 on your virtual machines where SSH brute force can happen. Now let's dive into our network security architecture. So on the left, this is the data store that we have created previously using a platform. And then you have your traffic analyzer, which is reading data from the store And then it evaluates this data against a baseline the baseline here can have rules something like Here is an example where we are saying This is conformant traffic. That means the destination here, let's say It belongs to a SQL server. Anyone communicating to the SQL server must belong, must have this IP only, or must be from this VM only. If there is any other communication going on to this destination, that is not confirm it. And then we immediately alert those security teams. To reiterate, our rule is basic, simply stating any communication to the destination, which is an ingress, should be only from source 1. 2, which can be a VM or maybe an app service. Anything, any other communications are non conformant and immediately be taken care of by the security team. Now, of course, this is a very simple scenario. Let's took look, take a look at a little complicated one. So here we have bunch of data stores that are hosted in a particular IP test range, right? 2.3 to two point 30. And these data stores might be, assigned in an IP address from one of these ranges. Now I go and tag this IP address ranges data store. Next, what I have is a VM that is trying to communicate to this IP address range, which also has another tag, which is

called VM for simplicity. So your baseline rule can simply say, if it is an ingress data, the source should be VM and the destination should be a data store. And if the destination is the data stored, and if the source is not in. In the ones you're looking for, you can immediately mark the data as non compliant. Actually, you can take a step further saying, if the source or destination is not a service tag, I immediately mark the data as non compliant. In this way, by step, I can improve the overall security posture of my organizations. Now, one more thing we've, I forgot to discuss this. This telemetry platform we see here is, will take time to build. Now, if you're on cloud, you can leverage existing solutions. For example, AWS, Azure, or Oracle have flow logs that give you information about ingress and egress data in your cloud. Now, with that information, what you have to store is, what you have to do is just store the data in a, maybe an S3 bucket or an Azure storage, and then pull the data from there for your processing. So in that way, you have completely get rid of the agent and the Fluentd that is responsible for sending your data to the data sink. You can just rely on the cloud providers to send them to your sink. And from there, you can do your additional processing. With that, let's get to extensions. now you have a very good platform, right? Now you can add more scenarios to it. So one scenario is security recommendation. Let's say with the telemetry data you have, you are able to achieve network segmentations. And what you observe is some of the service teams are communicating across these segments. Which you want to drastically reduce. Now you can provide recommendations based on this traffic to the team saying how they can reduce the segmentation so that how they can reduce the cross traffic communication across the segment so that the overall security posture improves. Similarly, because you have the data right now, You can identify if your network is reaching to a capacity and immediately look for capacity provisioning. For example, let's say there is a bunch of communication between the virtual machines and your data store, and you see the bandwidth is at 60 to 70 percent capacity. That is when you will go and decide that, okay, we need more capacity. So we'll add it there. These are the further, these are further extensions that you

can add to the platform. To conclude, as cloud adoption grows, network telemetry will start becoming important to monitor and protect organizations. So it is important that we start investing in such platforms so that when the time comes, we are resilient from infrastructure failures and hacks. With that, I conclude my talk. If you have any further questions or if you have any feedback, Good luck. Please feel free to reach out to me. This is my LinkedIn email and Twitter. Thank you everyone. Thank you for listening to the talk

Slides

Siri Varma Vegiraju

Technical Leader @ Microsoft

Post

Share

Join the community!

Learn for free, join the best tech learning community for a price of a pumpkin latte.

Annual Monthly

NEWSLETTER

$0/mo

- ✓ Event notifications, weekly newsletter
- ✓ **Delayed access** to all content
- ✓ Immediate access to Keynotes & Panels

COMMUNITY

$8.34/mo

- ✓ Access to Circle community platform
- ✓ **Immediate access** to all content
- ✓ **Live events!**

Email address

First Name

Last Name

Company

Job Title

Phone Number

Country

United States

I consent to the following terms: Terms and Conditions & Code of Conduct

Subscribe to free newsletter →

- ✓ Regular office hours, Q&As, CV reviews
- ✓ Courses, quizes & certificates
- ✓ Community chats

Join the community (7 day free trial) →

Online tech events

EVENTS 2025

DevOps 2025

Python 2025

Chaos Engineering 2025

Cloud Native 2025

Large Language Models (LLMs) 2025

Golang 2025

Site Reliability Engineering (SRE) 2025

Machine Learning 2025

Observability 2025

Quantum Computing 2025

Rustlang 2025

Platform Engineering 2025

MLOps 2025

Incident Management 2025

Kube Native 2025

JavaScript 2025

Prompt Engineering 2025

Robotics 2025

DevSecOps 2025

Internet of Things (IoT) 2025

EVENTS 2024

DevOps 2024

Chaos Engineering 2024

Python 2024

Cloud Native 2024

Large Language Models (LLMs) 2024

Golang 2024

Site Reliability Engineering (SRE) 2024

Machine Learning 2024

Observability 2024

Quantum Computing 2024

Rustlang 2024

Platform Engineering 2024

Kube Native 2024

Incident Management 2024

JavaScript 2024

Prompt Engineering 2024

DevSecOps 2024

Internet of Things (IoT) 2024

EVENTS 2023

DevOps 2023

Chaos Engineering 2023

Python 2023

Cloud Native 2023

Golang 2023

Site Reliability Engineering 2023

Machine Learning 2023

Observability 2023

Quantum Computing 2023

Rustlang 2023

Platform Engineering 2023

Kube Native 2023

Incident Management 2023

JavaScript 2023

DevSecOps 2023

Internet of Things (IoT) 2023

EVENTS 2022

Python 2022

Mobile 2022

Chaos Engineering 2022

Golang 2022

Cloud Native 2022

Machine Learning 2022

Site Reliability Engineering 2022

Quantum Computing 2022

Rustlang 2022

Incident Management 2022

Kube Native 2022

JavaScript 2022

DevSecOps 2022

Web 3.0 2022

EVENTS 2021

Chaos Engineering 2021

Enterprise Software 2021

Cloud Native 2021

Python 2021

Golang 2021

Machine Learning 2021

Site Reliability Engineering 2021

JavaScript 2021

DevSecOps 2021

EVENTS 2020

Chaos Engineering 2020

Open Source Showcase 2020

Site Reliability Engineering 2020

JavaScript 2020

COMMUNITY

Support us

Speakers

Hall of fame

Discord

LEGAL

Code of Conduct

Terms and Conditions

Privacy policy

About the team

SPONSORS

Sponsorship

Request the Prospectus

Media kit