Network Monitoring for Cyber Security

Paul Krystosek, PhD

CERT Network Situational Awareness

What’s Coming Up

The scope of network monitoring

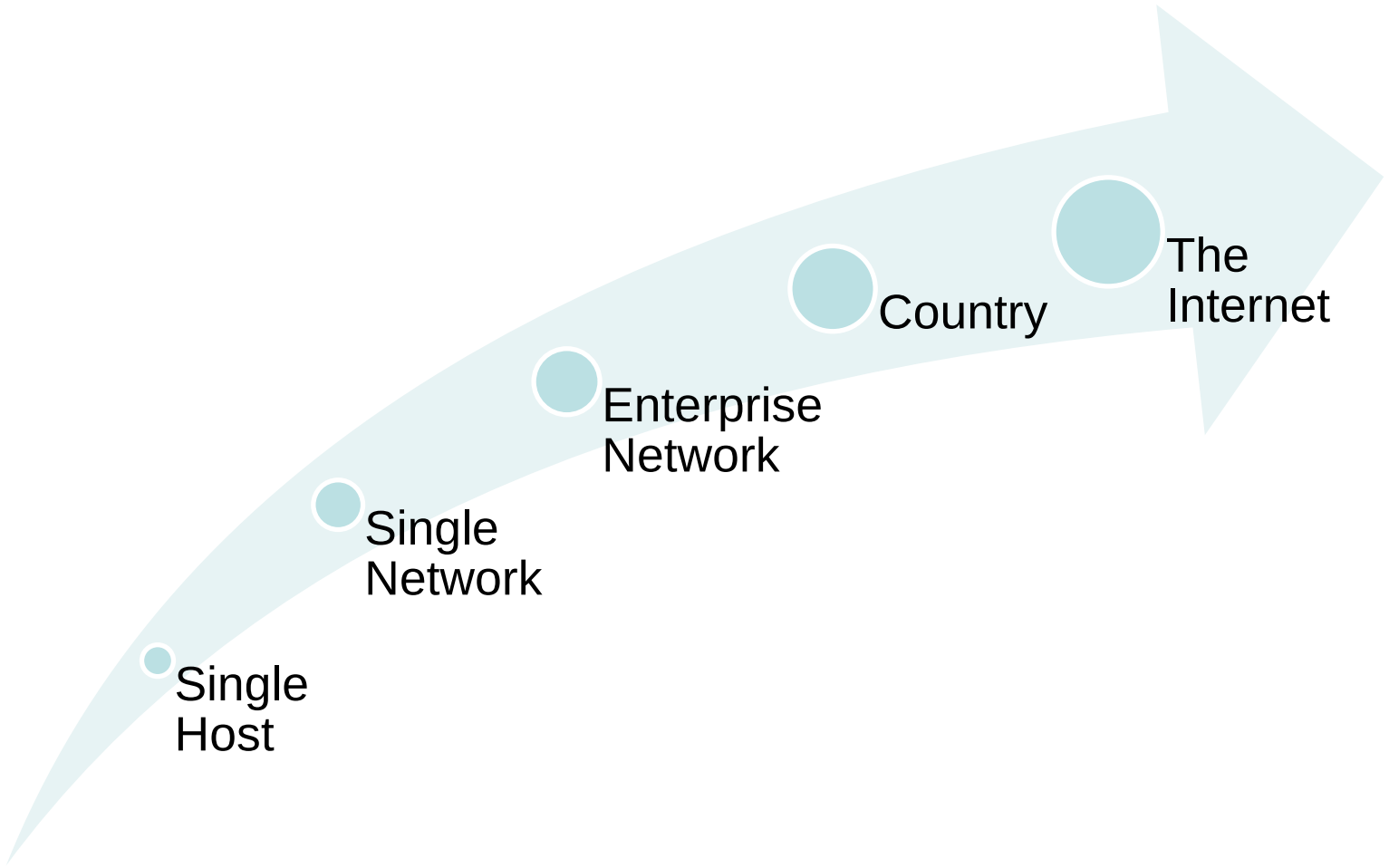
Cast of characters

Descriptions

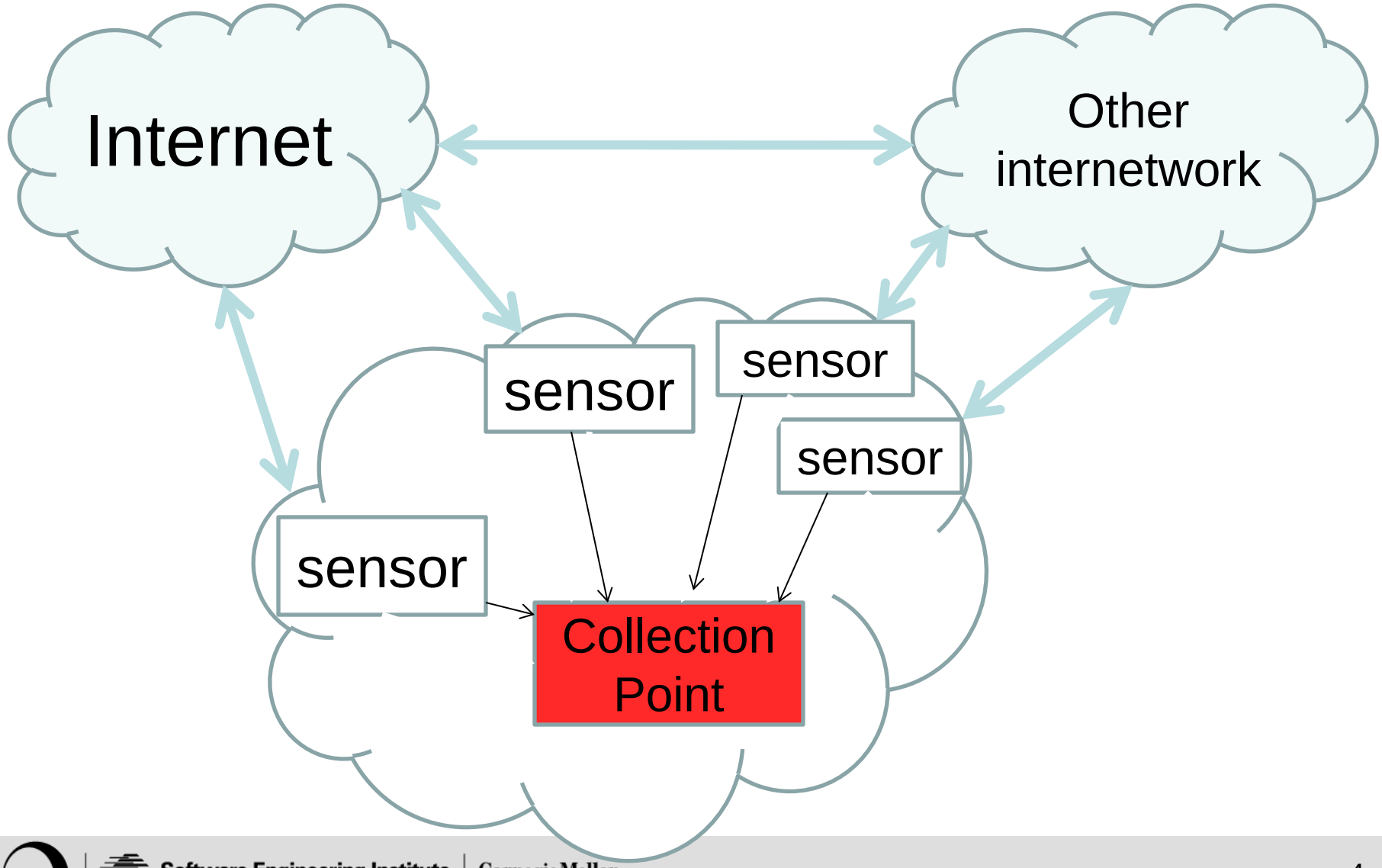
Comparisons

Usages

The Range of Network Monitoring



Network Monitoring



Cast of Characters

Full Packet Capture

Meta Data Capture

Intrusion Detection System

NetFlow

Sampled NetFlow

Full Packet Capture

Capture “everything” that goes across the network

Typically used on a single network

Example

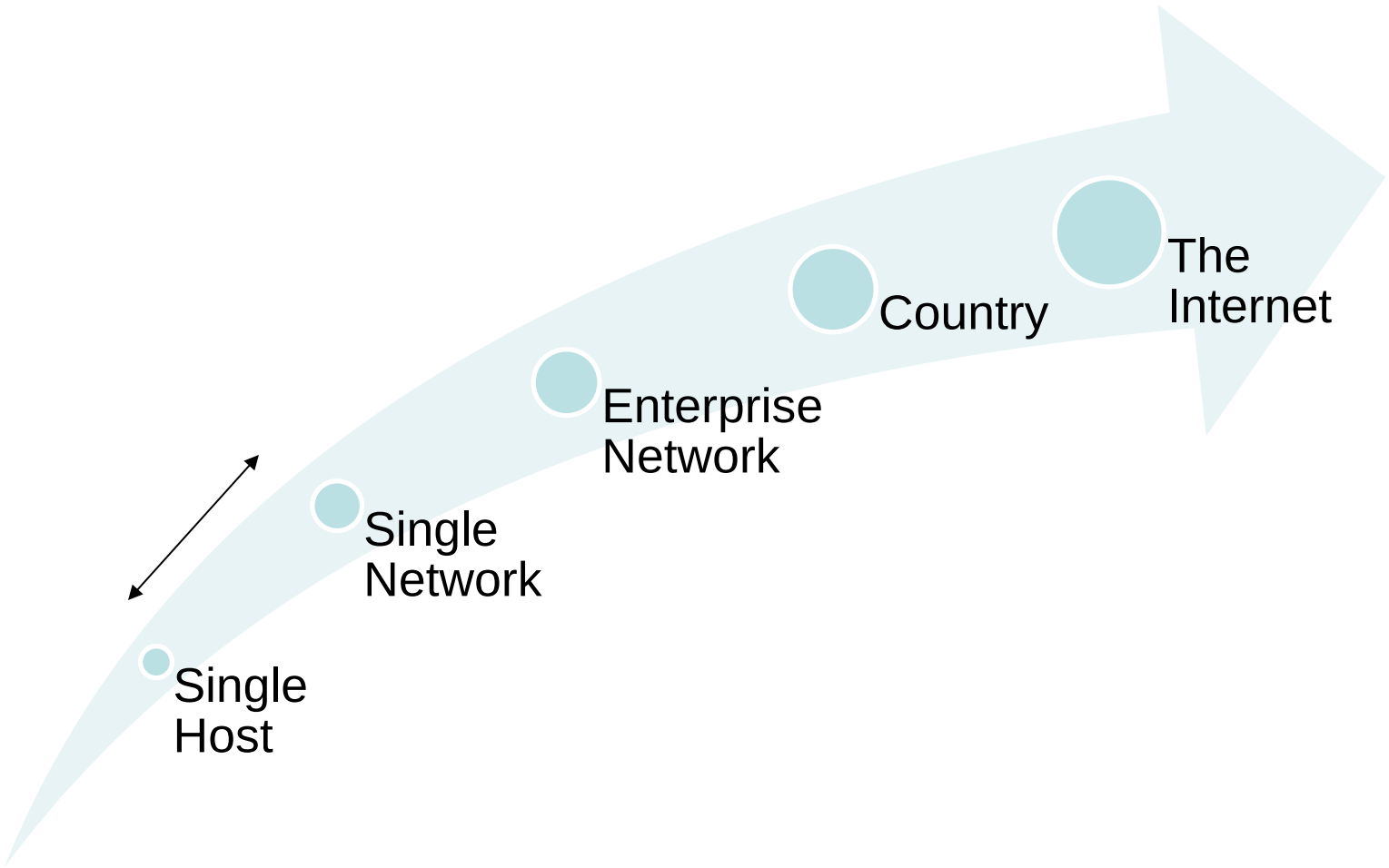
PCAP

Full Packet Capture

Not unlike recording all telephone conversations in a building

But, you don’t have each call in one place in it’s entirety, instead, you have a collection of one sided fragments each labeled with both phone numbers, the time and duration

The Range of Packet Capture



Meta Data Capture

Capture data associated with a particular network activity

Typically in the form of logs

Examples:

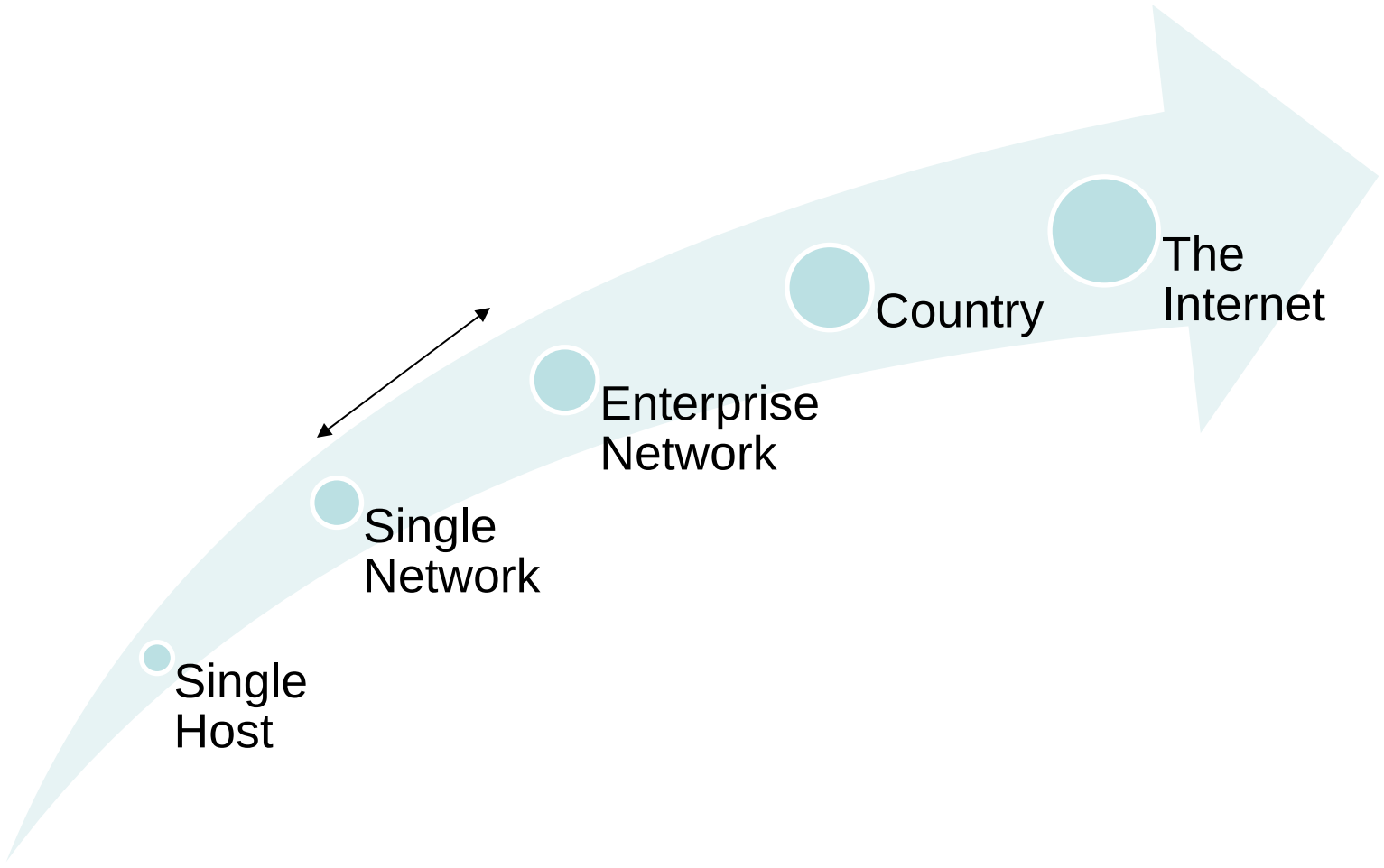
- For email traffic capture:
  - *from, to, subject, date, attachments*
- For web traffic capture:
  - *Source IP, destination IP, URL, User Agent String*

Meta Data Capture

Still working on a good physical analogy…

Street corner survey

The Range of Metadata Logging



Intrusion Detection System (IDS)

Define signatures which indicate a specific activity

An IDS will look for packets on the network that match those signatures

The IDS will issue an alert when a match is found

Examples

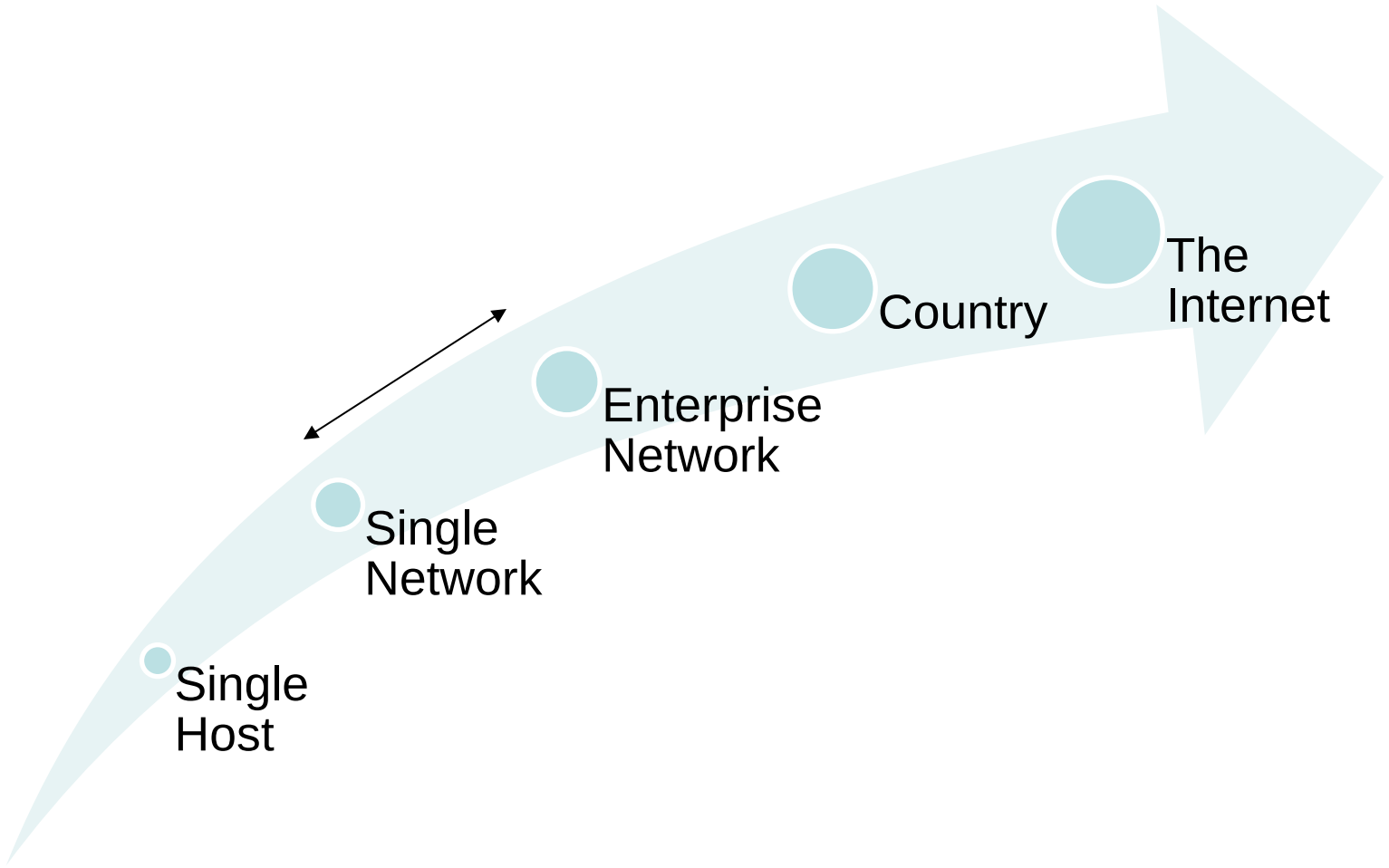
- Snort
- Suricata
- Bro

Intrusion Detection System (IDS)

A little like a collection of police records; traffic tickets, arrests, complaints…

Also similar to “profiling” (driving too fast for conditions leaving a location where a crime recently occurred)

The Range of Intrusion Detection



NetFlow

NetFlow aggregates related packets into unidirectional flows

Some systems aggregate into “biflows”

The flow records are collected and stored for later analysis

Examples

- SiLK
- Argus

Sampled NetFlow

NetFlow in which not all flows are stored

Sample rates could be 1 in 10 or 1 in 100 stored

NetFlow and Sampled NetFlow

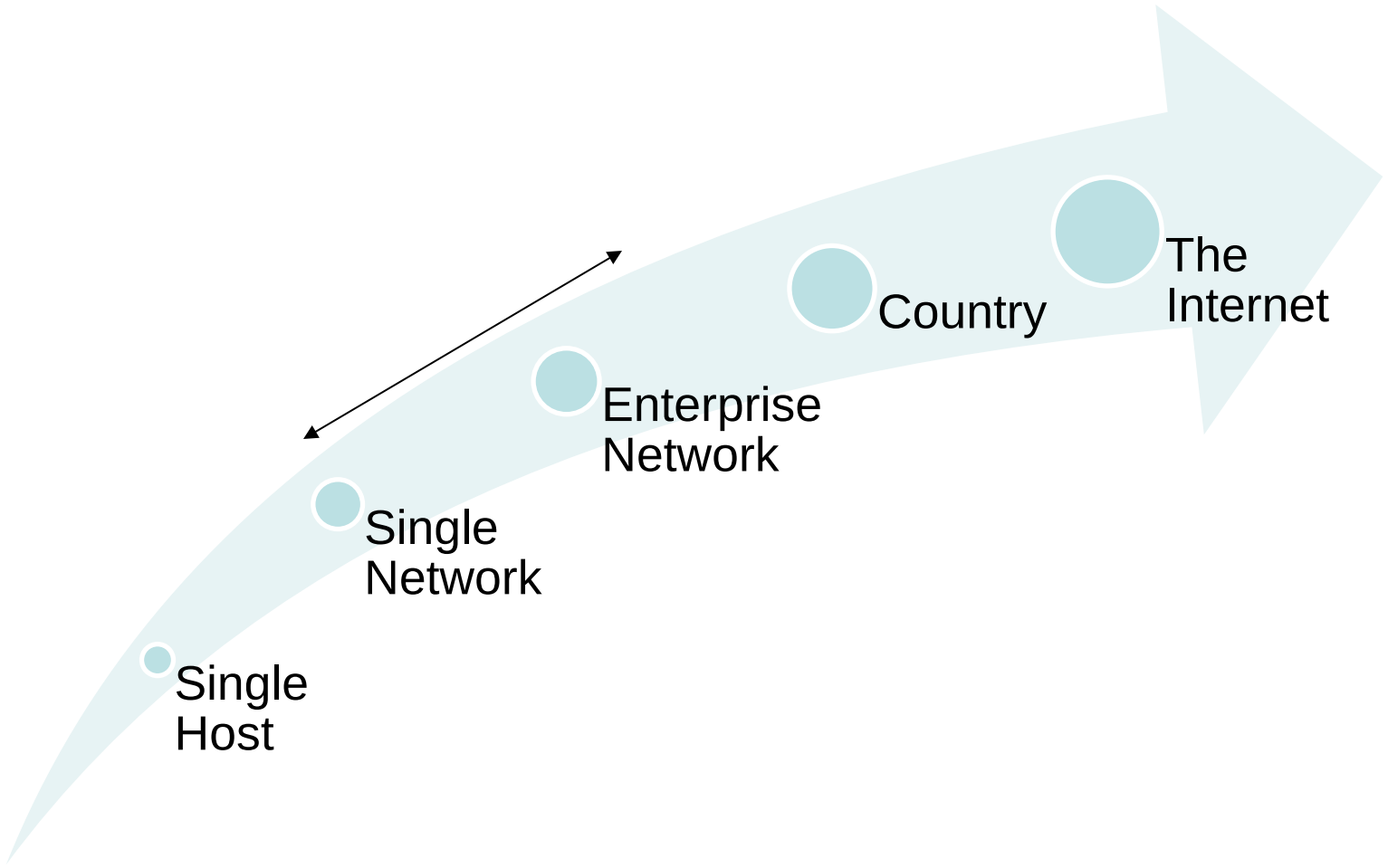
Similar to

"Pen Register" and

"Trap and Trace"

for calls entering and leaving a whole building or even an entire telephone exchange.

The Range of NetFlow Monitoring



Uses of Network Monitoring

Network Operations
Network Defense
Forensics
Intelligence
Research

Network Operations

Bandwidth utilization

- Current status
- Short term cyclical
- Longer term trends

Network service usage

- Descriptive statistics
  - How is bandwidth used

Network outages

- May only tell you something is wrong
- Not what or where

Network Operations

Metadata and NetFlow work well here

Full packet capture is too much detail

IDS only looks for certain events

Network Defense

Retrospective analysis

- Is an event an isolated occurrence or more wide spread?
- How long has this activity been going on?

Exploratory analysis (aka Data Mining)

- What can I learn about my network?
- Activity X is possible…
  - Can it be found, reliably, on the network with the current monitoring infrastructure?

Network Defense

A well tuned IDS is a logical choice for defending a network.

NetFlow provides a broader view of the network

Full packet capture can be used in conjunction with IDS and NetFlow

Network based Forensics

Who, what, where, when, why, how

Reconstructing events that occurred on the network in an effort to establish actions in time

Network Forensics

Similar to network defense, IDS and NetFlow can be used together to provide tremendous detail on network events

Intelligence

Find probe attempts

Find data exfiltration

Examples

Intelligence

Intel → IDS signatures

Data mining of Metadata and NetFlow

Once you are close use Full Packet Capture

Research

Develop new analysis techniques

Establish future network monitoring systems

A little more detail on our cast of characters

Now that we know what they can be used for

Full packet capture

With

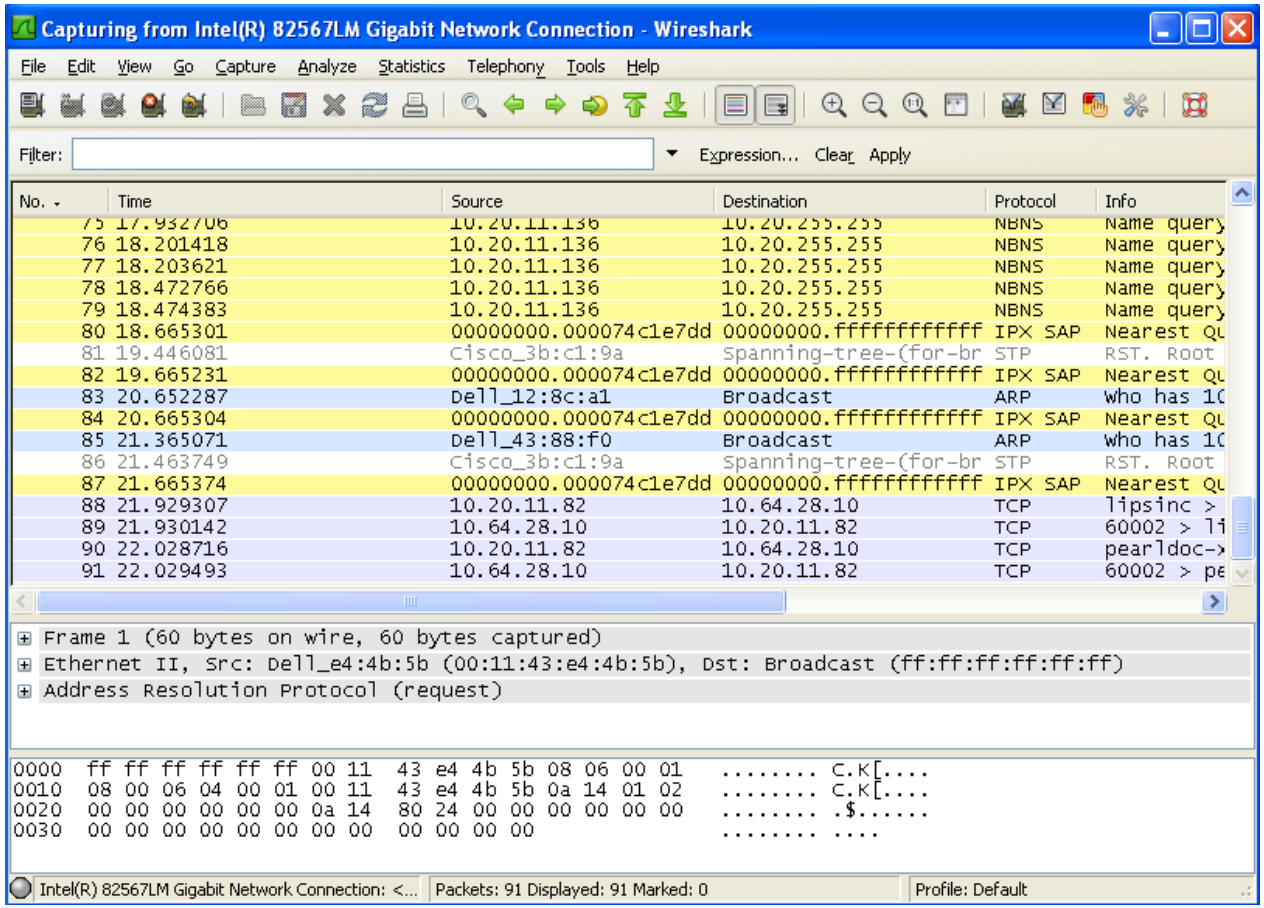
- captured packets,
- the proper tools,
- expertise
- and time

One can reconstruct an entire network conversation

Tool for Examining Packets

Wireshark

- Headers
- Content
- Footers



Metadata

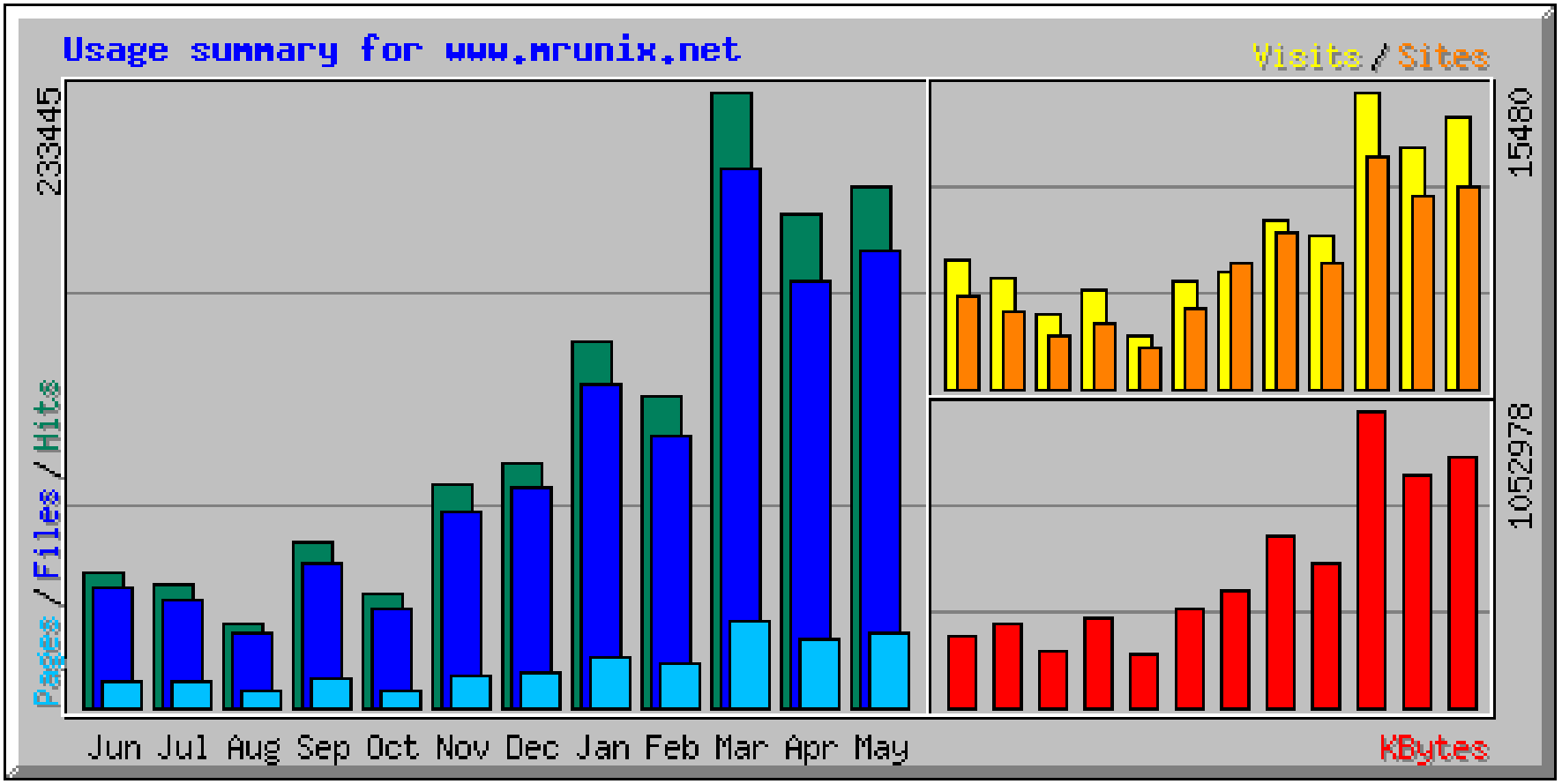
Web logs permit you to learn about

- Browsing habits
- Client capabilities
- Referring pages

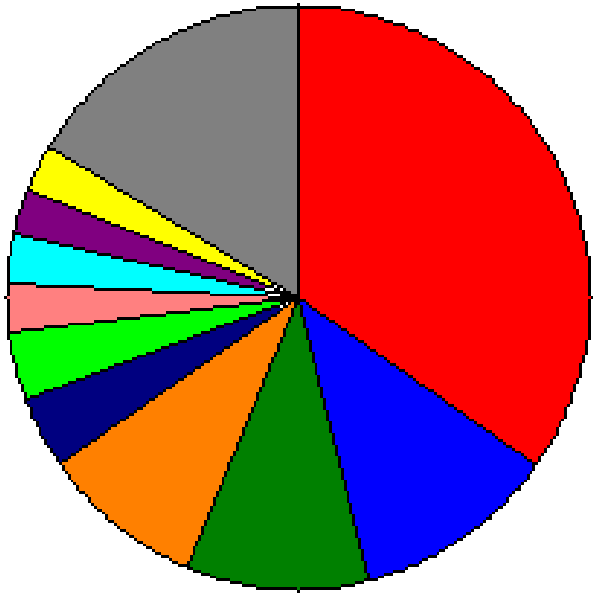
Email logs will tell you

- What security services (if any) are in use outside
- Email paths, number of hops
- Fingerprint external email servers

Sample analysis from Webalyzer



Sample Analysis from Analog



/~sret1/stats/
/~sret1/stats/reportmagic/
/~sret1/stats/reportmagic/GENERAL.html
/~sret1/stats/reportmagic/navfile.html
/~sret1/analog/olddocs/
/~sret1/stats/reportmagic/DAILYREP.html
/~sret1/stats/reportmagic/BROWSERSUM.html
/~sret1/stats/reportmagic/DAILYSUM.html
/~sret1/stats/reportmagic/REQUEST.html
/~sret1/stats/reportmagic/HOURLYSUM.html
Other

The wedges are plotted by the number of requests.

Intrusion Detection Systems

There are two primary types of IDS

- Signature based
- Anomaly detection

Signature based is more common

Anomaly detection is a generalized and looser form of signature

Intrusion Detection Systems

A good signature is a thing of beauty

A signature that is too simple will have many false positives

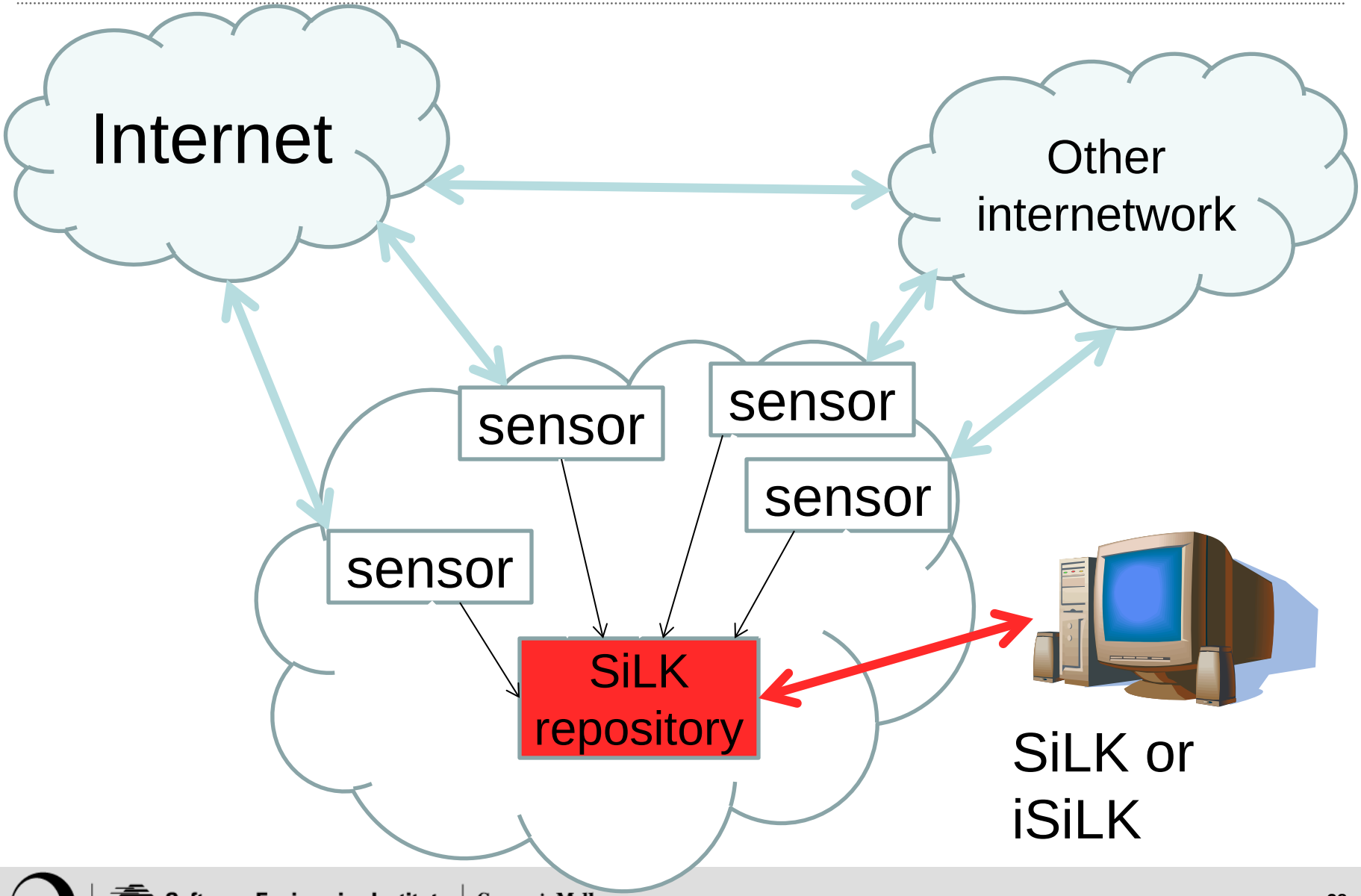
A signature that is too complex may have many false negatives

Overly complex signatures may tax the computer on which the IDS runs to the point of missing traffic

NetFlow

There are now many software packages that capture, store and analyze NetFlow records

We'll discuss SiLK as a representative of command line software packages


Internet
Other internetwork
sensor
sensor
sensor
sensor
SiLK repository
SiLK or iSiLK

Where did SiLK come from

SiLK

the System for Internet-Level Knowledge

A collection of traffic analysis tools developed by the Carnegie Mellon University (CMU)

Software Engineering Institute (SEI)

Network Situational Awareness Team (CERT NetSA)

to facilitate security analysis of large networks.

What Is a Flow?

A flow is an aggregated record of packets.

SiLK flows are defined by five unique attributes:

- internet protocol (any of about 130 in use)
- source address
- destination address
- source port
- destination port

SiLK flows are unidirectional.

These five keys form a “tuple”

- Similar to a “primary key” in a database record

What is SiLK

SiLK is a collection of Unix command line tools that manipulate SiLK records

The SiLK tool suite supports

- **the efficient collection**
- **storage**
- **analysis**

of network flow data, enabling network security analysts to rapidly query large historical traffic data sets.

Why is it written funny?

I don’t know either

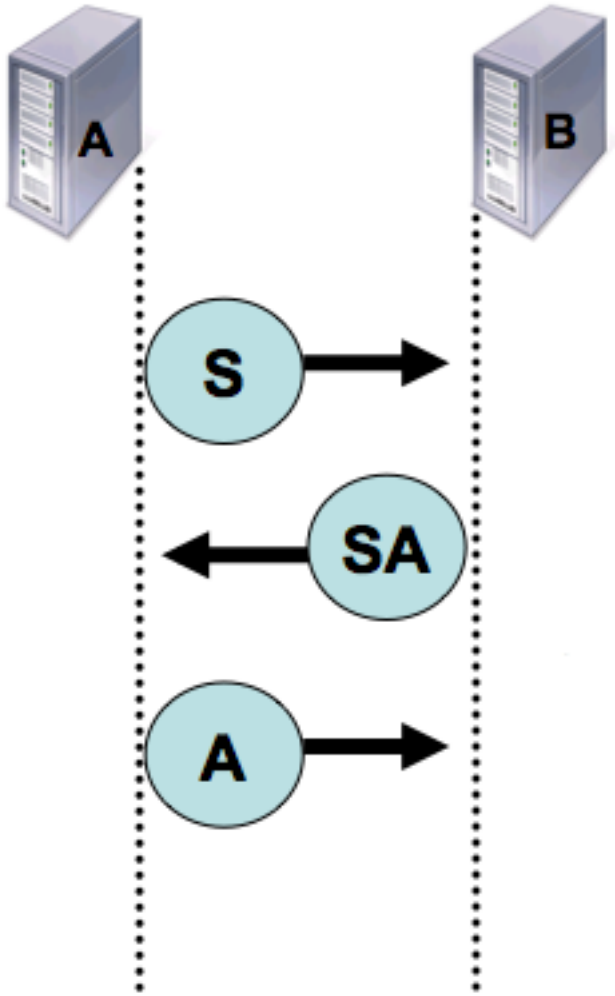
it was lost to the mists of history

in this case, way back around 2002

Flows Are Half Duplex

For the TCP three-way handshake, consider how flows are counted:

- Flow 1 is created when the sensor observes the first packet between hosts A and B.
- Flow 2 is created with the second packet. Swapped IPs means a new flow.
- The third acknowledgement (ACK) packet updates flow 1, since the source and destination addresses and ports match.



Questions SiLK can help answer

What's on my network?

What happened before the event?

Where are policy violations occurring?

What are the most popular websites?

How much volume would be reduced with a blacklist?

Do my users browse to known infected web servers?

Do I have a spammer on my network?

When did my web server stop responding to queries?

Who uses my public servers?

Specific Concerns

Privacy

Full Packet Capture is as intrusive as it can get. Whatever was sent over the network is captured and stored

IDS looks at some content, but records that a signature was matched

Privacy

Metadata will provide “traffic analysis” and a some facts about users. Email subjects would be considered content.

Netflow stores no content. Several organizations concerned with privacy have done legal research and approved its use.

Comparison: Analysis

	Analysis						
Monitor type	Traffic	Retrospective	Exploratory	Descriptive	Automation	Malware	Correlation
Full Packet Capture	*	**	*	**	*	****	-
Meta Data Capture	**	***	*	**	***	*	**
IDS	*	***	-	*	**	***	***
NetFlow	****	****	****	***	****	**	***
Sampled NetFlow	**	**	**	**	****	**	*

Comparison: Efficiency

	Efficiency		
Monitor type	**Capture**	**Storage**	**Manipulation**
Full Packet Capture	*	*	*
Meta Data Capture	***	***	***
IDS	***	***	**
NetFlow	****	****	****
Sampled NetFlow	*****	*****	****

Comparison: Ease of …

	Ease of			
Monitor type	**Design**	**Deployment**	**Analysis**	**Monitoring**
Full Packet Capture	****	****	**	-
Meta Data Capture	***	***	****	**
IDS	*	**	**	****
NetFlow	*	*	*	-
Sampled NetFlow	*	*	*	-

Comparison: Privacy

Monitor type	Privacy
Full Packet	*
Meta Data	**
IDS	***
NetFlow	*****
Sampled NetFlow	*****

Other Forms of Network Monitoring

DNS response packet capture

- what domains are being looked up?
- what address does <some name> resolve to?

WiFi monitoring

- Network Operation
- Over subscription
- Rogue access points
- Triangulation of resources

Other Forms of Network Monitoring

Internet route monitoring

- BGP Updates

Darkspace monitoring

- External users trying to access unused addresses are probably up to no good
- Internal hosts on unallocated addresses are probably misconfigured

AS 112

Conclusions

Conclusions: Full Packet Capture

The good

- Excellent for finding out exactly what was inside packets on the network

The not so good

- Requires massive storage and a way to index it
- In other words it doesn't scale well
- Requires great skill and patience to find the nugget you need
- Correlation is difficult

Conclusions: Metadata

The good

- Excellent for known traffic types
- The applications that produce that traffic probably already provide the metadata
- Scales reasonably well

The not so good

- Each application will have it’s own specific type of metadata
- Can be challenging to integrate disparate ones

Conclusions: Intrusion Detection Systems

The good

- Excellent if you know what you are looking for
- Typically works well with Security Information and Event Managers (SIEMs aka SIMs)
- Signatures are **easy** to write
- Scales reasonably well

The not so good

- If you don't tell it to look for something, it won't
- Can be difficult to correlate with some other data types
- **Good** signatures are **difficult** to write
- False sense of security

Conclusions: NetFlow

The Good

- With no content there are few, if any, privacy issues
- When analyzed, it provides useful information
- Scales well
- Many analysis tools

The Not So Good

- NetFlow is “low information content” data and so requires a lot of analysis
- You either have to know what you are looking for or aggregate a lot of data to make it useful

For More Information

SiLK:	http://tools.netsa.cert.org/silk/
Argus:	http://www.qosient.com/argus/
PCAP:	http://www.tcpdump.org/
Snort:	http://www.snort.org
Wireshark:	http://www.wireshark.org/
Web Log:	http://www.webalizer.org/

NO WARRANTY

THIS MATERIAL OF CARNEGIE MELLON UNIVERSITY AND ITS SOFTWARE ENGINEERING INSTITUTE IS FURNISHED ON AN “AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

Use of any trademarks in this presentation is not intended in any way to infringe on the rights of the trademark holder.

This Presentation may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.

Contact Information

Paul Krystosek pnk@cert.org

Software Engineering Institute

Carnegie Mellon University

Pittsburgh, PA