

Exploiting Causality and Communication Patterns in Network Data Analysis

Pekka Pietikäinen (pp@ee.oulu.fi)



Oulu University
Secure Programming
Group

Applications Actions it-pc61-eth1-leap-6-vlan-1111.pcap - Ethereal Thu Oct 28, 10:23:23

janek's Home Trash

it-ap1-leap-6.pcap - Ethereal

File Edit View Go Capture Analyze Statistics Help

Filter: Expression... Clear Apply

No. -	Time	Source	Destination	Protocol	Info
110	11.33022	D-Link_84:aa	Cisco_51:33	EAPOL	Start
111	11.33235		D-Link_84:a	IEEE 802.11	Acknowledgement
112	11.33416	Cisco_51:33	D-Link_84:a	EAP	Request, Identity [RFC3748]
113	11.33628		Cisco_51:33	IEEE 802.11	Acknowledgement

Frame 112 (78 bytes on wire, 78 bytes captured)
 IEEE 802.11
 Logical-Link Control
 802.1x Authentication

it-pc61-eth1-leap-6-vlan-1111.pcap - Ethereal

File Edit View Go Capture Analyze Statistics Help

Filter: Expression... Clear Apply

No. -	Time	Source	Destination	Protocol	Info
13	7.855474	Cisco_7a:cc:18	Broadcast	ARP	Who has 193.209.91.251? T
14	8.001727	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8
15	8.854188	Cisco_7a:cc:18	Broadcast	ARP	Who has 193.209.91.251? T
16	9.998743	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8
17	11.14917	193.209.91.252	193.209.91.241	RADIUS Access Request(1)	(id=214,
18	11.18254	193.209.91.241	193.209.91.252	RADIUS Access challenge(11)	(id=2
19	11.20673	193.209.91.252	193.209.91.241	RADIUS Access Request(1)	(id=215,
20	11.24006	193.209.91.241	193.209.91.252	RADIUS Access challenge(11)	(id=2
21	11.30683	193.209.91.252	193.209.91.241	RADIUS Access Request(1)	(id=216,
22	11.33063	193.209.91.241	193.209.91.252	RADIUS Access Accept(2)	(id=216,
23	11.33702	D-Link_84:aa:8b	01:40:96:ff:ff:0	0x872 PRI: 0 CFI: 0 ID: 1111	
24	12.00134	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8

Frame 16 (68 bytes on wire, 68 bytes captured)
 Ethernet II, Src: 00:d0:58:21:a7:0e, Dst: 01:00:0c:cc:cc:cd
 802.1q Virtual LAN
 Logical-Link Control
 Spanning Tree Protocol

```

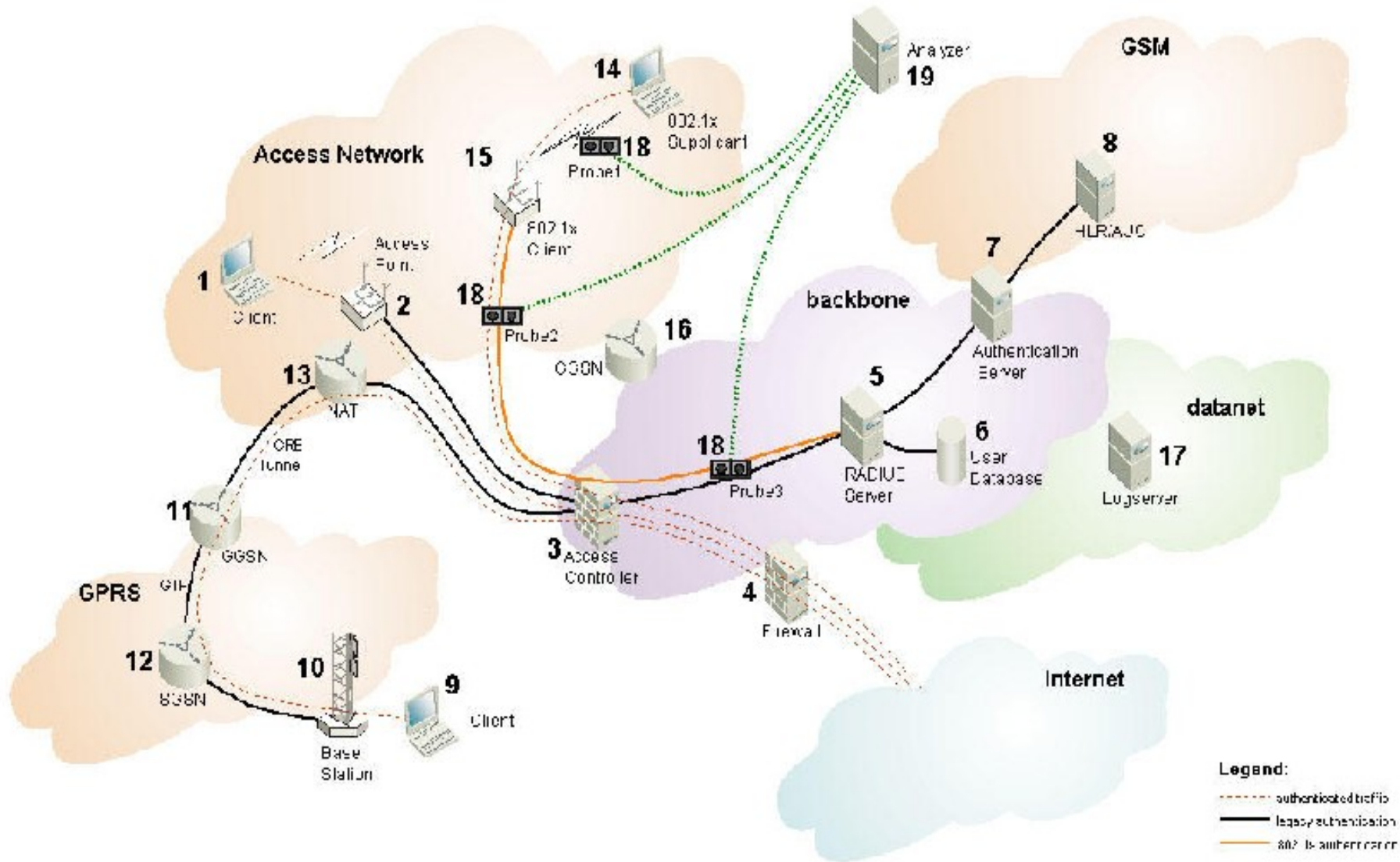
0000  01 00 0c cc cc cd 00 d0 58 21 a7 0e 81 00 04 57  ....X!....W
0010  00 32 aa aa 03 00 00 0c 01 0b 00 00 00 00 00 80  .2.....
0020  00 00 30 85 47 0f 89 00 00 00 17 80 00 00 d0 58  ..0.G.....X
0030  21 a7 02 80 1b 02 00 14 00 02 00 0f 00 58 00 00  !.....X..
0040  00 02 00 03  ....
  
```

File: it-pc61-eth1-leap-6-vlan-1111.pcap P: 26 D: 26 M: 0

By Jilin Yaron / May 2003

Network analysis is fun...

Frontier Network Overview



Let's add some complexity.

[illegible][illegible][illegible]

PPP RFC1661, RFC1662; Frame relay FRF.1, RFC1490
 HDLC Cisco; IP RFC791; Ipv6 over Ipv4 supported,
 RFC2529; Ipv6 over Ipv4 tunneling supported,
 RFC2185; Network time protocol RFC1305;
 Network address translation (NAT); DHCP RFC2131;
 CIDR RFC1519; ICMP Router Discovery (server portion) RFC1256;
 ICMP RFC792; ARP RFC826 Route aggregation;
 Requirements for IPv4 routers RFC1812;
 Route redistribution; DVMRP RFC1075;
 IGMPv2 RFC2236; PIM-SM; Multicast tunnels;
 PIM-DM (multicast); RIPv1 RFC1058; VRRP RFC2338;
 OSPFv2 RFC2328; RIPv2 (with authentication);
 RFC1723; IGRP (optional) Cisco; Static routing
 BGP4 (optional, available; only for IP330) RFC1771;
 Supports IEEE802.1x authentication framework
 GRE tunneling; SSL versions 2 and 3, TLS;
 version 1 supported; Native IPsec (IKE, AH, ESP);
 SSH server, versions 1 and 2; supported;
 MD5 Routing Authentication; (RIPv2) RFC1723;
 SNMPv3 with User-Based; Security Model; Radius client RFC2865
 Radius accounting client; RFC2866; Proxy Radius RFC2865;
 Virtual Router Redundancy; Protocol RFC2338;
 Traffic management;
 SSL/TLS RFC2246; SSL/TLS RFC2216; SSH server, versions 1
 and 2 supported; SNMP, SNMPv2 and SNMPv3
 CLI via Telnet RFC854; RFC959; SMTP mail (send) RFC821;
 RFC1760; SNMP and SNMP MIB II RFC1213; RADIUS auth.client MIB
 RFC2618; RADIUS acc.client MIB RFC2620; P022 MIB;
 DiffServ, EF) RFC2598; 1350 The TFTP Protocol

```

WTF [0x5] BADPTR
Inchworm BOZ in EMP ERM
EMP: macthink:load:in: macthink:
Reluctant:the:de:commission:from:EMP:WTF
EMP: macthink:load:in: EMP: ERM:and:in:
EMP:WTF:and:Emp: [0x5] macthink:in: macthink:
Supports:EMP:algorithms:and:credentials:from:EMP:
Uses:credentials:for:auth:manip: dynamic:for:permissions:and:
referring:to:encrypt:the:KAR:as:an:interface
31R: macthink:
Use:time:password:via: [0x5]
EMP:WTF:macthink:used:open:space
COW:and:and:EMP:WTF: macthink: according:to:to:EMP:WTF: macthink: COW
COW: for:EMP:WTF: macthink: charging: via: [0x5]
Should: macthink: to: macthink: COW: as: a: killing: system
WTF: (don't: macthink: to:EMP:WTF: charging: macthink:
WTF: macthink: to: macthink: down: macthink: system
COW: macthink:
Secure: transfer: of: COW: via: WTF.

```

[illegible][illegible]

... and related protocols

Applications Actions File Edit View Go Capture Analyze Statistics Help

Filter: Expression... Clear Apply

No. -	Time	Source	Destination	Protocol	Info
16	9.998743	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8
17	11.14917	193.209.91.252	193.209.91.241	RADIUS	Access Request(1) (id=214,
18	11.18254	193.209.91.241	193.209.91.252	RADIUS	Access challenge(11) (id=2

Frame 7 (64 bytes on wire, 64 bytes captured)

0000 ff ff ff ff ff ff 00 0f 24 7a cc 18 81 00 04 57 \$z.....w

0010 08 06 00 01 08 00 06 04 00 01 00 0f 24 7a cc 18 \$z...

0020 c1 d1 5b f8 00 00 00 00 00 00 c1 d1 5b fb 00 00 ..[...].....[...

0030 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 [.....]

File: it-pc61-eth1-leap-6-vlan-1111.pcap P: 26 D: 26 M: 0

it-pc61-eth1-leap-6-vlan-1112.pcap - Ethereal

Filter: Expression... Clear Apply

No. -	Time	Source	Destination	Protocol	Info
11	8.011201	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8
12	8.426753	Cisco_7a:cc:18	Broadcast	ARP	Who has 193.209.91.251? T
13	9.425487	Cisco_7a:cc:18	Broadcast	ARP	Who has 193.209.91.251? T
14	10.01394	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8
15	12.01698	Cisco_21:a7:0e	01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:30:8

Statistics Help

Destination	Protocol	Info
01:00:0c:cc:cc:c	STP	Conf. Root = 32768/00:d0:5
Broadcast	ARP	Who has 193.209.91.251? T
Broadcast	ARP	Who has 193.209.91.251? T

it-pc30-eth1-leap-6.pcap - Ethereal

Filter: Expression... Clear Apply

No. -	Time	Source	Destination	Protocol	Info
1	0.000000	Cisco_21:a7:06	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
2	0.000653	Cisco_21:a7:07	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
3	0.006709	Cisco_21:a7:16	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
4	2.005160	Cisco_21:a7:06	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
5	2.005806	Cisco_21:a7:07	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
6	2.011886	Cisco_21:a7:16	Spanning-tree-(f	STP	Conf. Root = 32768/00:d0:5
7	3.643706	212.213.170.5	193.209.91.230	ICMP	Echo (ping) request

Frame 1 (64 bytes on wire, 64 bytes captured)

Ethernet II, Src: 00:40:96:58:e2:9b, Dst: 00:40:96:58

802.1q Virtual LAN

0000 00 40 96 58 e2 9b 00 00 96 58 e2 9b 81 00 00 01 X!...&BB

0010 80 00 00 00 01 00 00 00 00 00 00 00 00 00 00 X!....

0020 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 X!.....

0030 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 X.....

802.1q Virtual LAN (vlan) P: 2 D: 2 M: 0

P: 47 D: 47 M: 0

End result

- Analyzer shows what it can decipher from the data
- Filters are simple, efficient rules that reduce the amount of data that is shown
 - BPF etc.
- You have to know what you're looking for
 - Sometimes it's data that's **MISSING**
 - Firewall blocking packets, MTU issues etc.
- Multiple networks, multiple addresses for different components

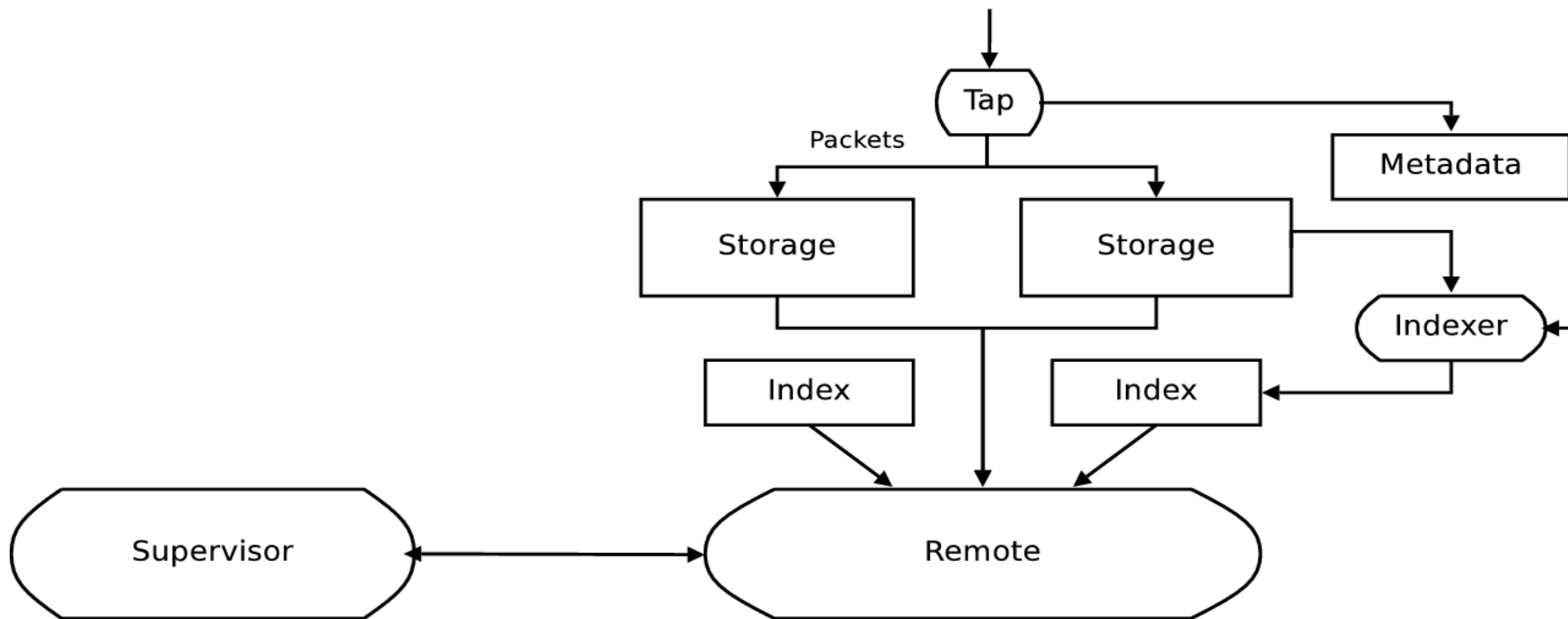
In other words

- Treat network traffic as flows consisting of packets
- When a new flow is discovered on the network, find candidates that are potentially related
- Candidates are chronologically preceding flows, for which one of the endpoints of the new flow is involved.
- Have to have understanding on what addresses different components have

Causality in network traffic

- In first iteration all packets were transmitted to central location for analysis
- Maintaining a graph for all traffic is resource intensive.
- Usually you can work on the flow level, sometimes you need the packets.
 - Architecture should allow quick retrieval of flows and corresponding packets
- Causality is an useful abstraction

Drawbacks and lessons learned



Architecture of probe

- “Regular expressions” for network traffic
- Simple Python library for fetching flows/packets and matching attributes.
- Packets are fetched only if necessary

```
@matcher
def matchexample((flow, packet)):
    def flowMatcher():
        return flow.src=='10.0.0.1' and
            flow.probe=='probe-eth2'
    def packetMatcher():
        # Accept any packet belonging to
        # a matching flow
        return True
    return flowMatcher, packetMatcher
```

Example #1

```
@matcher
def icmpMatcher(flow, prev):
    def flowMatcher():
        return flow.dst == prev.src and
            flow.proto == 'ICMP'

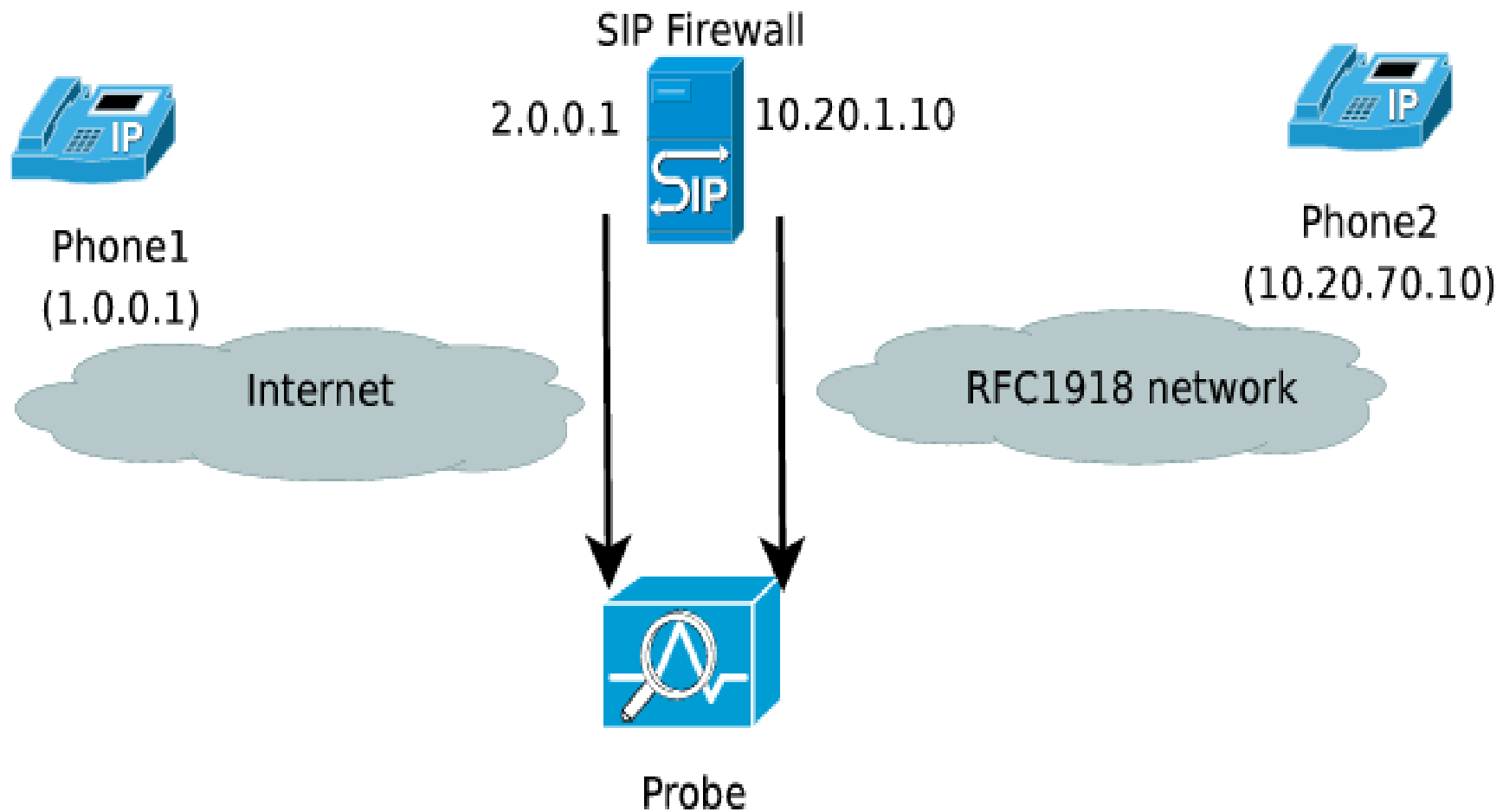
matcher2 = sequence(matchexample,
                    icmpMatcher)
```

Chained matchers

- NoPair
 - Triggers when a given sequence does not occur within a given time window
 - eg. NoPair(matchexample, icmpResponse)

```
gi = GeoIP.new(GeoIP.GEOIP_MEMORY_CACHE)
def foreignDstMatcher(flow):
  def flowMatcher():
    return gi.country_code_by_addr \
      (flow.dst) != 'US'
```

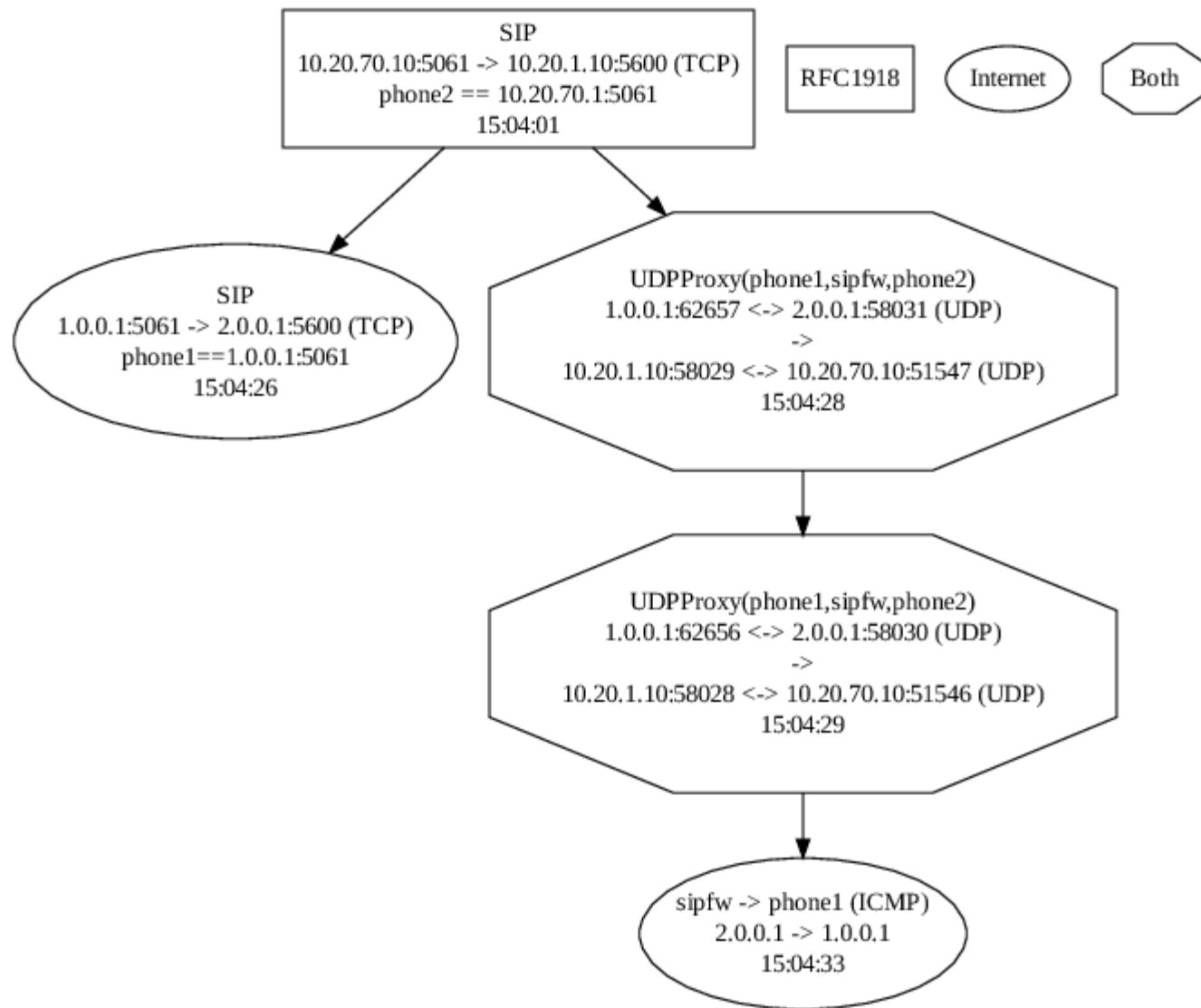
More possibilities



Simple example

- Isolating traffic that involves a call between phone1 and phone2
 - Addresses are dynamically assigned
- SIP
 - matches TCP port 5060 traffic on the firewall, parses session setup with a simple string regexp to find identities of phones
- UDPProxy
 - Sequence matcher that matches flow pairs that are both UDP, have the firewall as an intermediary AND matches phones
- Causal

Pieces needed



Result

- Capture: speed of disk (Multiple GigE)
- Indexing: 40kpps
- Fetching flows: 2500 flows/s
- Simple matches: \sim = same
- Sequence matches: \sim = 60s to process 10s of traffic at 250Mbps

Performance etc.

- Add more understanding of network topology to analysis
- Optimize
- Move more processing to edges
 - Do packet level matching there?
- Don't limit to just network traffic, there are other sources that provide useful information (e.g. WLAN association logs can be used to pinpoint IP/MAC address)

Conclusions and current/future work

THE END

<http://www.ee.oulu.fi/research/ouspg/>