Advanced Computer Networking (ACN)

IN2097 – WiSe 2018-2019

Prof. Dr.-Ing. Georg Carle

Sebastian Gallenmüller, Max Helm, Benedikt Jaeger, Patrick Sattler, Johannes Zirngibl

Chair of Network Architectures and Services
Department of Informatics
Technical University of Munich
Network Measurements
Network Measurements

Introduction

Active Measurements

Passive Measurements

Hybrid Measurements

Summary
Network Measurements

Introduction

Active Measurements

Passive Measurements

Hybrid Measurements

Summary
Why do we measure the network?

Network provider view
• Manage traffic
  • Model reality
  • Predict future
  • Plan network
  • Avoid bottlenecks in advance
• Reduce cost
• Accounting

Service provider view
• Get information about clients
• Adjust service to demands
• Reduce load on servers
• Accounting
Why do we measure the network?

Client view
• Get the best possible service
• Check the service
  • „Do I get what I paid for?”

Researcher view
• Understand Internet better
• Performance evaluation
  • “Could our new routing algorithm handle all this real-world traffic?”
• …

Security view
• Detect malicious traffic
• Detect malicious hosts
• Detect malicious networks
• …
But why should we do it at all?

Do we really have to?

• The network is well engineered
• Well documented protocols, mechanisms, …
• Everything built by humans → no unknowns (compare this to physics)
• In theory, we can know everything that is going on

⇒ No need for measurements?!

But:

• Distributed multi-domain network → information only partially available
• Moving target
  • Requirements change
  • Growth, usage, structure changes
• Highly interactive system
• Heterogeneity in all directions
• The total is more than the sum of its pieces
• Built, driven, and used by humans → errors, misconfigurations, flaws, failures, misuse, …

⇒ Need for network measurements
Network Measurements

Introduction

Active Measurements

Passive Measurements

Hybrid Measurements

Summary
Active Measurement Overview

Performance measurements
• Latency, Throughput, Loss

Topology measurements
• Internet mapping (L2, IP layer, AS level,…)
• Alias resolution
• IPv4-IPv6 interdependency

Security measurements
• TLS certificates & Certificate Transparency
• SSH server keys
• Internet-of-Things & Out-of-Band Management protocols
Internet-wide measurements

Challenges:

• Scanning tools
  • Performance (should finish in hours to days)
  • Resource constraints (RAM, storage)

• Target selection
  • Whole IPv4 address space (feasible with ZMap)
  • IPv6 needs a hitlist ($3.4 \times 10^{38}$ addresses)

• Ethical considerations
  • Reduce impact on targets and networks
  • Take possible misbehaving targets into account
Performance Measurements

Methodology

• Probe packets exchanged between specific nodes
• Measurement of packet loss, one-way delay, round-trip times, packet inter-arrival time

Analysis

• Complete packet loss
  → link down, invalid route, router defect
• Partial packet loss
  → available bandwidth, level of congestion
• Delay = propagation time + buffer time
  → distance estimation, filling level of buffers
• Interarrival times of packet pairs/trains
  → path capacity
Topology Measurement: Traceroute

Traceroute: possible anomalies due to load balancing

- **Approach:** „Paris traceroute“
  
- **Scamper**
  - All-in-one tool
  - IPv4 & IPv6
  - Built-in alias resolution

- **Alternative:** yarrp
Recap: Solution for multipath routing

Hash “consistently” and use packet headers as “random” values

```
From: 10.0.0.1   To:   10.9.8.7
Src port: 31377   Dst port: 80
```

(\text{payload})

\text{hash( )}

\begin{align*}
\begin{cases}
\text{h} &= 0 \Rightarrow \text{use Route A} \\
\text{h} &= 1 \Rightarrow \text{use Route B} \\
\text{h} &= 2 \Rightarrow \text{use Route C}
\end{cases}
\end{align*}

Result

- Packets from same TCP connection yield same hash value
- No reordering within one TCP connection
Paris Traceroute

- Idea: Vary header fields that are within the first 28 octets
  - TCP: sequence number
  - UDP: checksum field
  - Requires manipulation of payload to ensure correctness of checksum
  - ICMP: combination of ICMP identifier and sequence number

- Experiment results
  - Certain routers use first four octets after IP header combined with IP fields for load balancing

- Still fails on per packet load balancing
  - MDA [1] and yarrp [2] try to cover this problem

Chapter 8: Network Measurements

Comparing IPv4 and IPv6 Paths in the Internet

• Traceroute a set of known sibling pairs

• Compare paths and form sibling candidates from these paths

• Apply sibling detection techniques [1] to evaluate sibling status of IPv4/IPv6 routers

• Develop a set of path comparison metrics and evaluate against these

Internet-Wide Measurements

- Goal: Evaluate properties of a service deployed on the public Internet

- Challenges:
  - Which scanning tools to use?
  - How to reach (all) targets?
  - Performance
  - Ethical considerations and intrusiveness
  - Evaluation metrics
Zipfs Law

- Internet traffic is assumed to follow Zipfs law [1]

\[ w_k = \frac{1}{k^s} \div \frac{1}{\sum_{n=1}^{N} \frac{1}{n^s}} \]

\( k \) rank of object
\( s \) slope of distribution

- \( s \) is set to 1 based on related work [2]

Zipf's Law

- Log-Log graph with different values for $s$
Prefix Top Lists [1]

- Aggregate top lists over a week [2]
- Collect A and AAAA records for domain based top lists
- Assign Zipf weight of domain to IP addresses
- Aggregate on prefixes and ASes

Useful for:
- Prefix prioritization
- Security issue valuation


https://prefixtoplists.net.in.tum.de/
Security Measurements at i8

Internet-wide security measurements:

- TLS
- Certificate Transparency
- SSH
- IPMI
- BACnet
Security Measurement: TLS Certificate Scanning

• TLS: Transport Layer Security

• Multiple versions
  • SSL 3.0
  • TLS 1.0
  • TLS 1.1
  • TLS 1.2
  • TLS 1.3

• Security foundation for HTTPS, IMAPS, SMPTPS, DoT, DoH, …

→ Evaluate TLS deployment
Security Measurement: TLS Certificate Scanning

Methodology
1. Identify hosts offering TLS service (HTTPS, IMAPS, …)
2. Download certificate chains
3. Analyze and validate chains

Challenges
• Targets (0/0?)
• Performance
• Evaluation metrics
Security Measurement: TLS Certificate Scanning

Analysis of the TLS landscape [1]

• Active and passive measurements

• Analyses of
  • Certificate chains
  • Expiry
  • Algorithms

• Conclusion: TLS landscape in sorry state (expired, no root cert,...)
  • But: situation improves over time [2]

---


Security Measurement: TLS Certificate Scanning

Evolution of TLS scanning

<table>
<thead>
<tr>
<th></th>
<th>Holz et al. (2011)</th>
<th>Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targets</td>
<td>• Alexa Top 1M</td>
<td>• Full IPv4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Server Name Indication (SNI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Alexa Top 1M</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Zone files</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reverse DNS</td>
</tr>
<tr>
<td>Software stack</td>
<td>• Nmap</td>
<td>• ZMap</td>
</tr>
<tr>
<td></td>
<td>• openssl</td>
<td>• Custom-built protocol scanner for TLS and HTTPS</td>
</tr>
<tr>
<td>Performance</td>
<td>Weeks for 1M hosts</td>
<td>Day(s) for complete Internet</td>
</tr>
<tr>
<td>Frequency</td>
<td>Single measurement</td>
<td>Continuously running measurement service</td>
</tr>
</tbody>
</table>
New features in TLS 1.3

• 1-RTT handshakes by default
  • Use presumed cipher suite selection

• 0-RTT handshake with resumption possible
  • PSK for early data
  • Forward secrecy after early data

• Privacy
  • Client certificates are encrypted
  • SNI not encrypted (RFC Draft for encrypted SNI in TLS 1.3)

• Grease mechanism
  • Send random version data to increase robustness

[1] Valsorda and Sullivan: Deploying TLS 1.3: the great, the good and the bad. 33C3.
Security Measurement: Certificate Transparency

Certificate Transparency

- Name owner correctly gets a certificate for their site
Certificate Transparency

- Name owner correctly gets a certificate for their site
- Evil actor incorrectly gets a certificate for third party site
Certificate Transparency

- Name owner correctly gets a certificate for their site
- Evil actor incorrectly gets a certificate for third party site
- Both Certificates are publicly recorded in CT logs → incorrectly issued certificate can be detected
- Approach: Use public, append-only Merkle trees at CT logs
Certificate Transparency

- Allows us to get a more complete picture of the TLS certificate ecosystem

- Use Certificate Transparency logs as source for certificates (which themselves are a source for domain names → cf. IPv6 Hitlist Sources section)

- Goals
  - Understand CT ecosystem
  - Analyze certificates obtained through CT
Security Measurement: Certificate Transparency Scanning

Certificate Transparency ecosystem evolution [1]

- Large increase of log entries before April 2018 deadline (dashed orange line)
- Let’s Encrypt CA dominates
- Strong rise overall

Certificate Transparency ecosystem log distribution [1]

- System overly reliant on few log servers
- Almost all CAs use few logs for their certificates
- CAs are not distributing certificates over logs evenly → limits reliability

Chapter 8: Network Measurements

Security Measurement: Certificate Transparency Scanning

Analyzing certificates in CT logs [1]

- How do we define “quality”?
  - E.g. Baseline Requirements

- How “high-quality” are certificates?
  - Many low quality

- Does the “quality” change over time?
  - Enforcing stricter rules helps improve the situation

Certification Authority Authorization [1]

- DNS Record to specify CAs allowed to issue certificate
- Example:

<table>
<thead>
<tr>
<th>Domain</th>
<th>Type</th>
<th>Flags</th>
<th>Tag</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>tum.de</td>
<td>CAA</td>
<td>0</td>
<td>issue</td>
<td>“letsencrypt.org”</td>
</tr>
<tr>
<td>tum.de</td>
<td>CAA</td>
<td>0</td>
<td>issue</td>
<td>“pki.dfn.de”</td>
</tr>
<tr>
<td>tum.de</td>
<td>CAA</td>
<td>0</td>
<td>issuewild</td>
<td>“;”</td>
</tr>
<tr>
<td>tum.de</td>
<td>CAA</td>
<td>0</td>
<td>iodef</td>
<td>“mailto:a@b”</td>
</tr>
</tbody>
</table>

Tags:
- **issue**: allowed CAs to issue
- **issuewild**: optionally to define CAs for wildcard issueance
- **iodef**: contact information for incident reports

[1] Scheitle et al. *A First Look at Certification Authority Authorization (CAA)*, CCR 2018
Analysis of malicious web servers use of TLS

Assumption: malicious web servers are not configured with maximum security in mind [1]

Approach:
- Actively scan a set of benign and malicious web servers
- Collect all sorts of identifiers (selected cypher suit, extensions, …)
- Compare the results and the differences

Related work:
- Passive fingerprinting of TLS

Contact sattler@net.in.tum.de

[1] Blake et al. Deciphering Malware’s use of TLS (without decryption)
Available Open Theses

Performant Certificate Transparency Monitoring

Tasks:
• Familiarize with concepts of CT
• Retrieve data from CT logs continuously
• Analyze the results

You should:
• be experienced with Linux, Bash, and Go
• have knowledge in SQL and preferably also in PostgreSQL

Contact helm@net.in.tum.de and sattler@net.in.tum.de
Security Measurement: SSH Scanning

- **SSH**: Secure Shell protocol

- Provides encrypted & authenticated remote shell access to devices
  - Other use-cases:
    - X-Forwarding
    - Tunneling, port forwarding
    - git
    - sshfs

- Mostly used on servers and routers to provide administrative access

- Security critical protocol → evaluate SSH’s security
Security Measurement: SSH Scanning

Cautionary measures for SSH scans

• SSH is security and infrastructure critical protocol
  • Mostly used by administrators and not regular users (different from TLS)
  • Can provide access to a system

• Take additional cautionary measures for SSH scans
  • Notify CERTs, watchlist services, blocklist operators
  • Use separate subnet with own WHOIS abuse email contact
  • Avoid accidental login
    • Abort connection before login OR
    • Send inexistent username and authentication method
Security Measurement: SSH Scanning

Internet-wide SSH scans [1]

• Three complete IPv4 scans for SSH servers over 7 months

• Found ~15 M servers

• 42 k servers offer SSH 1 only

• Downloaded > 25 M SSH host keys
  • Host keys identify a server similar to a certificate in TLS

Security Measurement: SSH Scanning

Internet-wide SSH scans [1]

- Weak SSH host keys
  - Coprime weakness [2]
    - Keys can be factorized due to low entropy during key-generation
    - \( \sim 0.015 \% \) (2.4 \% for SSH1)
  - Debian-weak keys
    - Keys are not randomly generated due to OpenSSH bug in Debian from 2006 to 2008
    - \( \sim 0.05 \% \)
- Man-in-the-Middle (MitM) attack possible by imitating a server with a weak host key

Security Measurement: SSH Scanning

Internet-wide SSH scans [1]

• Duplicate keys pose similar MitM threat as weak keys
  • Anyone could imitate identity of host with the same key

• Analysis of duplicate keys
  • Strong differences based on used SSH server version

• Possible reason:
  • Deployment of systems/devices with default or pre-generated keys

![Graph showing share of server versions in % for different SSH versions: OpenSSH 4.3, Dropbear 0.46, OpenSSH 5.9, Dropbear 0.51, OpenSSH 1.1, Cisco 1.25. The graph compares unique and duplicate keys.]](image)
Security Measurement: SSH Scanning

Internet-wide SSH scans [1]

• Duplicate keys pose similar MitM threat as weak keys
  • Anyone could imitate identity of host with the same key

• Analysis of duplicate keys
  • Clustering of duplicate keys based on Autonomous System

• Possible reason:
  • Web-hosting providers deploy systems with pre-generated keys or SSH gateway
Security Measurement: SSH Scanning

Internet-wide SSH scans [1]

- Short keys pose similar MitM threat as weak keys
  - Anyone who cracked a key could imitate identity of this host

- Analysis of key length
  - SSH 2
    - <1024 bits: ~5%
    - 1024 bits: ~50%
    - >1024 bits: ~44%
  - NIST & BSI recommend key length of at least 2048 bits

Security Measurement: IPMI Scanning

IPMI scanning:

- IPMI: Intelligent Platform Management Interface

- Used in (rack-mounted) servers for out-of-band management

- Separate minimal operating system with complete access to main OS

- Known vulnerabilities → should not be reachable from public Internet

- Security critical protocol → evaluate IPMI deployment
Security Measurement: IPMI Scanning

Internet-wide IPMI scans [1]

- Internet-wide scan for IPMI-over-IP devices

- New scanning method to detect *dark* IPMI devices
  - IPMI disabled in configuration but still detectable as IPMI device
  - → RMCP Ping requests

Security Measurement: IPMI Scanning

Internet-wide IPMI scans [1]

- Number of publicly reachable IPMI devices seems to be declining

- Combine IPMI responses with HTTPS reachability to detect IPMI web-interfaces

- Web-interfaces can be vulnerable, e.g. buffer overflow bug [2]

- Compromise of web-servers → compromise of host OS


Security Measurement: IPMI Scanning

Internet-wide IPMI scans [1]

- Analyze deployment practices of publicly reachable IPMI-over-IP devices
- Hilbert space-filling curve
  - Maps density in network address space (blue to red)
  - Neighboring prefixes are neighbors in curve
  - Not restricted to one prefix length
  - One pixel is /18 IPv4 prefix
  - \(\rightarrow\) Sparsely populated, but clustered

Security Measurement: IPMI Scanning

Internet-wide IPMI scans [1]

- Analyze deployment practices of publicly reachable IPMI-over-IP devices
- Use CAIDA’s Prefix-to-AS mapping to map IP addresses to Autonomous Systems
- ~30% of IPMI devices in only 10 ASes
  - → IPMI deployment is heavily clustered

Security Measurement: IPMI Scanning

Internet-wide IPMI scans [1]

- Analysis of TLS certificates on web server of IPMI devices
- Analyze Common Name and certificate hash
  - >80% of devices use default certificates
  - Problem?
    - Key can be extracted → MitM
- Short TLS keys on HTTPS web-interfaces
  - ~90% have 1024 bits or shorter

Security Measurement: BACnet Scanning

BACnet scanning:

- **BACnet**: Building Automation and Control Networks

- Used to control heating, solar panels, ventilation and other building automation aspects

- Unsolicited access can have real-world consequences
  - Presence detection → Break into home
  - Manipulate heating, water flow,…

- Security & safety critical protocol → evaluate BACnet deployment
Security Measurement: BACnet Scanning

BACnet scanning:

- BACnet protocol
  - Simple UDP-based request-response protocol

- Default port: UDP/47808

- BACnet devices have properties (e.g. device name, temperature, heating level) which can be set and retrieved
  - SingleProperty message
  - MultiProperty message

- No security built in
Security Measurement: BACnet Scanning

Internet-wide BACnet scans [1]

- Conducted two Internet-wide scans (SingleProperty, MultiProperty)
  - Found 13 k devices

- Evaluated deployment
  - Vendors: Top 5 → ~65 %
  - ASes: Top 5 → ~30 %
  - Countries → see figure

Security Measurement: BACnet Scanning

Internet-wide BACnet scans [1]

- Amplification attack vulnerability characteristics
  - Stateless $\rightarrow$ UDP ✓
  - No authentication ✓
  - Larger response $\rightarrow$ client can choose returned property ✓

- Amplification
  - Factor of 10-30x possible
  - Extreme example: Hwy 57; Located in the silver box on the electrical pole in front of Grove Primary Care Clinic. Pole 688

Security Measurements: Conclusion

Active security measurements can help to improve the Internet’s security

- Find insecure device and network configurations and notify affected parties
- Analyze deployment over time to observe remediation
- Find weaknesses in protocols
- Identify protocols vulnerable to amplification attacks before they are being exploited
Internet-Wide Measurements: Performance

Performance is critical in Internet-wide measurements

- Hitlist generation using DNS resolution
  - Was: adnshost
  - Now: massdns + unbound → 170 M domains in 4 hours

- TCP protocol scanning
  - Was: nmap + openssl/openssh → multiple weeks
  - Now: ZMap + protocol-scanner → day(s) for complete Internet

- How to achieve good scanning performance
  - Asynchronous multi-core software
  - Tweak software to match used hardware
Ethical considerations and intrusiveness [1]

- Active measurements can impact devices and networks
- Goal: Reduce negative impact and false alerts
- Measures taken
  - Reduce intrusiveness of scanning technique (e.g. avoid logins, limit scanning rate)
  - Provide information on scanning machine’s website
  - Respond to every inquiry and abuse email
  - Offer possibility for blacklisting IP addresses and subnets