Continuous Packet-to-Disk Recording for Security and Troubleshooting

Luca Deri <deri@ntop.org>
Introduction

• Dumping network traffic to disk is a well established technique for troubleshooting.
• Often problems occur in remote locations and it is not unusual to remote connect in order to create a traffic dump of the interesting traffic.
• Pcap is the de-facto standard for dumping traffic packets to disk supported by all monitoring tools.
• Tools such as tcpdump (tcpdump -i eth1 -w dump.pcap) are traditionally used to produce packet dumps to be analysed later on with tools such as wireshark.
Packet Dump with Wireshark

- Capture options:
  - Wi-Fi: en0: Ethernet, promiscuous mode enabled, snapshot length 262144, buffer size 2 MiB, monitor mode disabled
  - Thunderbolt Bridge: Ethernet, promiscuous mode enabled, snapshot length 262144, buffer size 2 MiB, monitor mode n/a
- Capture filter: 
- Capture files:
  - Use multiple files
  - Next file every 1 megabyte(s)
  - Next file every 1 minute(s)
  - Ring buffer with 2 files
- Stop capture automatically after:
  - 1 packet(s)
  - 1 megabyte(s)
  - 1 file(s)
  - 1 minute(s)
Packet Analysis with Wireshark
Why “Continuous” Packet Recording?

- Troubleshooting isn’t always an easy activity as problems can occur at any time.
- Sometimes it’s not possible to predict when start recording and thus it’s better to start recording until the problem occurs.
- Continuous packet recording is the activity of:
  - Recording network traffic permanently (24x7).
  - Capturing without any packet drop: the problem can occur during traffic bursts so dropping isn’t an option.
  - Oldest packet dumps are overwritten as disk space fills up.
Limiting packet dump to troubleshooting can be misleading as there is more than that.

Large companies are often protected by a firewall + IDS (Intrusion Detection System): these tools do not keep traffic history but just log security events.

Such as in real life when housebreaking occurs, VCRs can help understanding the genesis of the attack (if from the outside) or information leak (if from the inside).

Therefore we need a “network VCR tool” for network monitoring able to provide evidence of the problem for both troubleshooting and tracing those who made the crime.
Packet Shuffling Isn’t an Option

- Modern computing architectures and network adapters support multi-core/multi-queue:
  - Incoming traffic is spread across multiple RX queues where a cores can operate in parallel.
- The above practice improves the overall performance but it has the side effect of shuffling ingress traffic and thus changing the order of network packets.
- This is something that packet records must avoid as:
  - Problems can occur due to packet shuffling.
  - Traffic traces that have been manipulated by any mean are not useful for providing real evidence.
What about Network Speed?

- All modern computers (from laptops to HPCs) are connected at least with a 1 Gbit connection.
- 1 Gbit switches are interconnected by 10 Gbit links.
- Core network switches are interconnected by 40 Gbit (interim) or 100 Gbit (long term) links.
- For practical reasons it isn’t practical to capture on many places just to capture at lower speeds, thus we will probably need to place our traffic recorders onto the backbone.
- Max Traffic volume (Hint: Multiply x2 for Full Duplex):
  - 10 Gbit: 1.25 GB/sec (~4.4 TB/hour)
  - 40 Gbit: 5 GB/sec (~17.6 TB/hour)
What about Disk Space?

- The more you capture, the more space is necessary to accommodate pcap dumps.
- Packet compression can help depending on traffic type:
  - Internet traffic is already compressed (JPEG, MP3)
  - LAN traffic is often uncompressed (SQL, FTP…)
  - You can save ~5% on Internet, and > 50% on LAN.
- RAID is a good option for increasing disk bandwidth:
  - SATA/SAS 10k/15k RPM drives are a good compromise in terms of price/number, SSDs can be fewer/faster but more expensive.
  - At least 10 drives for 10 Gbit, 32 drives for 40 Gbit.
Saving Disk Space: Filtering Traffic

- Filtering can occur during or after capture:
  - During capture it allows traffic dumps to be reduced as unwanted traffic is discarded and thus disk space is saved. Caveat: interesting packets can be in the traffic portion you have dropped (e.g. dropped traffic can cause spikes)
  - Filtering after capture might require GB of data to be scan in order to find the traffic we’re interested in, thus some kind of traffic indexing is necessary in order to avoid long waiting times during search (pcap dumps can be scan only sequentially).
Saving Disk Space: Traffic Slicing

- Packet filtering can be rude (it’s an on/off activity) so slicing can be a better option sometimes.
- Packet slicing is the ability to reduce packet size by cutting them dynamically at specific sizes (e.g. up to the IP, or up to the TCP/UDP header).
- As network packets are in average ~512 bytes, this practice can help saving space.
- The best is to combine traffic slicing + filtering
  - Slice un-interesting traffic.
  - Full capture of packets we care about.
Introducing n2disk
Every pcap file comes with a companion index file created during packet capture (no post-processing).

On-the-Fly Packet Indexing
On-the Fly Packet Compression

• In addition to packet indexing it is possible to instruct n2disk to also compress pcap’s.
• Traffic compression depends on its nature:
  ◦ Almost no compression (-6%) with YouTube.
  ◦ -25% … -40% average compression rate.
  ◦ 1:5 High-frequency trading.
• Compression in implemented as yet-another-stage in the pipeline after packet indexing.
• PF_RING-aware pcap library supports transparent decompression.
Time-Indexing with Packet Timeline
Combining Realtime Monitoring with n2disk

Ingress Packet Aggregation

Balancing
0-Copy Packet Fanout
Additional n2disk Features

- Ability to operate on multi 10Gbit.
- Selective packet slicing.
- BPF-based packet filtering during capture and in post-capture.
- Support for hardware packet timestamps (if network adapter supports it)
- On-the-fly (i.e. during capture) packet indexing and compression.
Traffic replay is key to reproduce on a remote place the same network conditions observed during capture.

pcap files written by n2disk can be reproduced using popular tools such as tcpreplay or pfsend.

n2disk comes with a companion tool named disk2n that allows to

- Reproduce pcap files at the same rate as they were received (1/10 Gbit).
- Use the same sw timestamping technology used by n2disk to send packets at a high precision rate.
- Reproduce multiple pcap files (multi-TB) for long-run traffic replay.
Packet Recording Using the nBox

- The nBox GUI is a web-user interface used to manage all the ntop monitoring and packet-to-disk applications.
- Through the web GUI it is possible to start packet dump, extract packets and download packet dumps.
WÜRTHPHOENIX Recorder

- Turn-key solution for packet-to-disk needs
- Ability to operate at 1, 10 and 40 Gbit.
- Integrated with NetEye traffic monitoring console
- Ability to dump, extract, and visualise traffic being written on disk.

WÜRTHPHOENIX Recorder

based on n2disk™
Final Remarks

- Thanks to PF_RING and n2disk/disk2n it is possible to create a powerful yet price effective solution to packet recording and replay.
- Continuous packet recording is becoming a mandatory activity in various contexts not limited to just troubleshooting and security.
- The ability to combine on the same box packet recording and traffic monitoring is key to provide a compact solution to modern monitoring needs.