25 Years of ntop

How we started, and where we're heading

Luca Deri

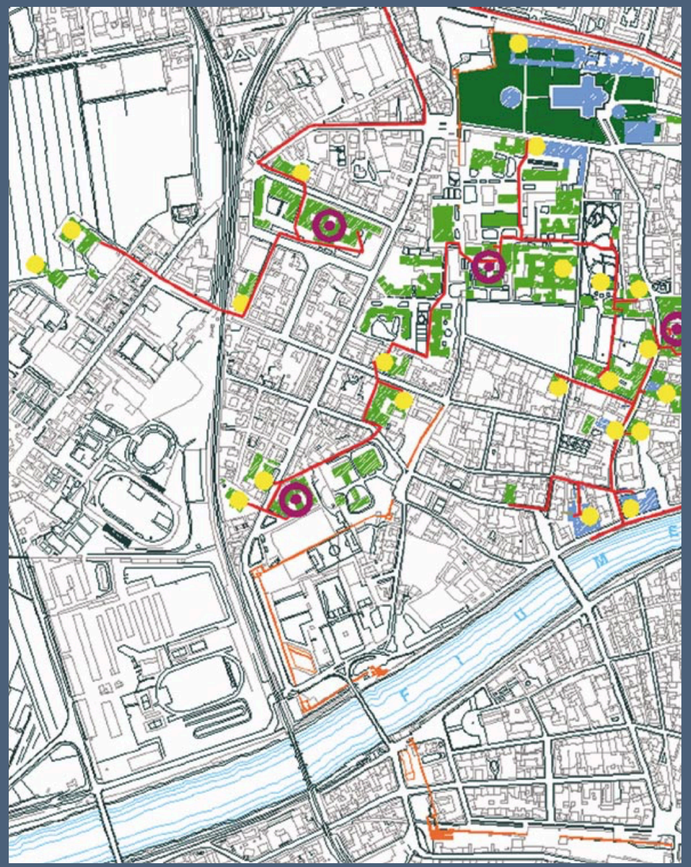
deri@ntop.org
@lucaderi

ntop

PacketFest'25

How we started

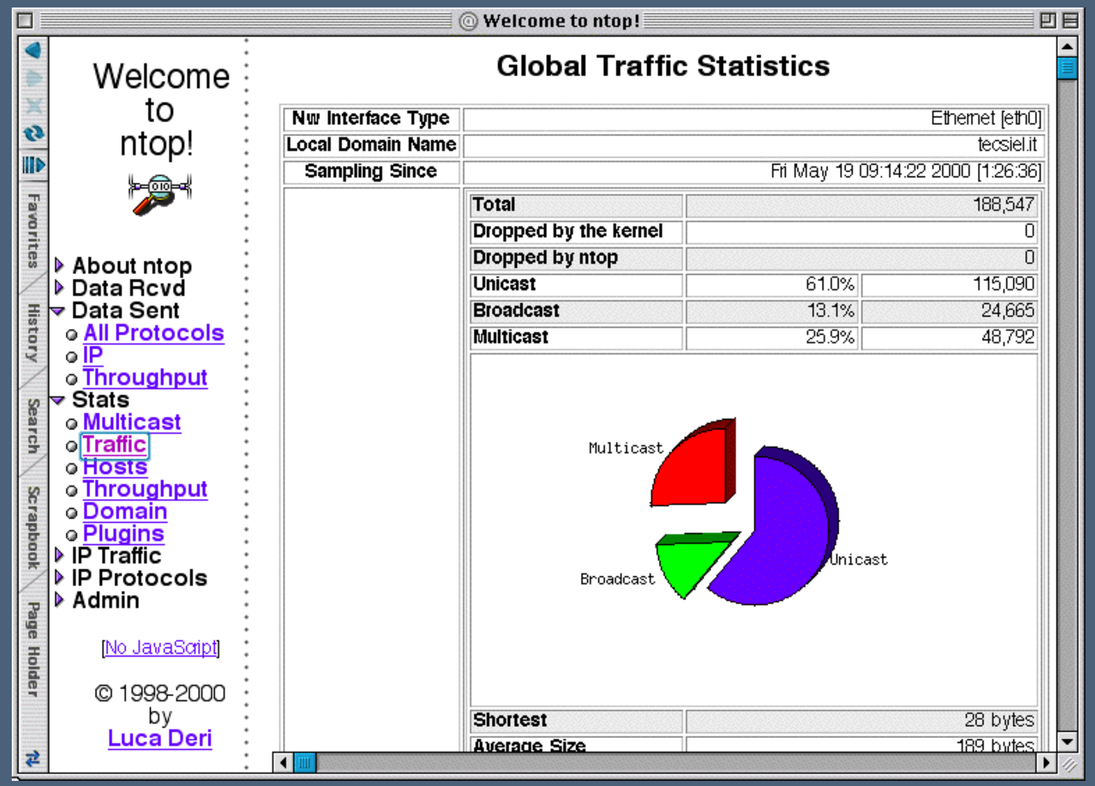
- Everything has started in 1998, when back from Zürich.
- Initial project goal was to monitor the traffic of Unipi.it: gopher, ftp, and www.
- No available tool for traffic analysis, costly commercial licenses, not designed for university needs.
- ntop was born as a short term project and eventually turned into a life-long tool, project, life....

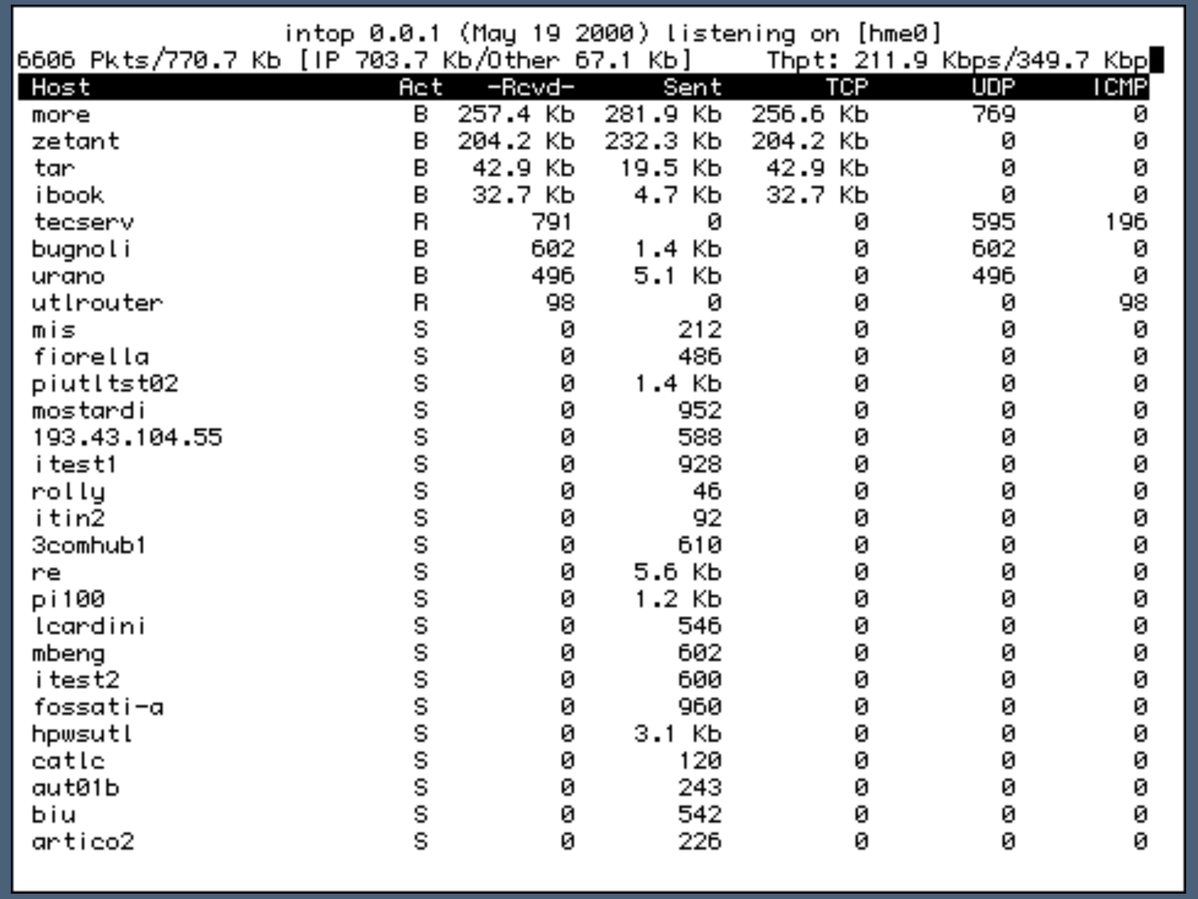


Initially It Was a Side Project...

- Initially, it was a side project, carried out during evenings and weekends.
- Open source was (and it's still) the driving force behind all this:
  - Be open and rewarding for the tools I was using.
  - Good chance to test the code on uncharted locations.
  - Your contribution to make the world a better place.
  - Learn a lot from contributors and feedback: books are great, but bugs, lessons from seasoned users, and positive criticism are way better.

Early Days


Welcome to ntop!
Welcome to ntop!
About ntop
Data Rcvd
Data Sent
All Protocols
IP
Throughput
Stats
Multicast
Traffic
Hosts
Throughput
Domain
Plugins
IP Traffic
IP Protocols
Admin
[No JavaScript]
© 1998-2000 by Luca Deri
Global Traffic Statistics
Nw Interface Type Ethernet [eth0]
Local Domain Name tecsiel.it
Sampling Since Fri May 19 09:14:22 2000 [1:26:36]
Total 188,547
Dropped by the kernel 0
Dropped by ntop 0
Unicast 61.0% 115,090
Broadcast 13.1% 24,665
Multicast 25.9% 48,792
Multicast
Broadcast
Unicast
Shortest 28 bytes
Average Size 189 bytes


intop 0.0.1 (May 19 2000) listening on [hme0]
6606 Pkts/770.7 Kb [IP 703.7 Kb/Other 67.1 Kb] Thpt: 211.9 Kbps/349.7 Kbp
Host Act -Rcvd- Sent TCP UDP ICMP
more B 257.4 Kb 281.9 Kb 256.6 Kb 769 0
zetant B 204.2 Kb 232.3 Kb 204.2 Kb 0 0
tar B 42.9 Kb 19.5 Kb 42.9 Kb 0 0
ibook B 32.7 Kb 4.7 Kb 32.7 Kb 0 0
tecserv R 791 0 0 595 196
bugnoli B 602 1.4 Kb 0 602 0
urano B 496 5.1 Kb 0 496 0
utlrouter R 98 0 0 0 98
mis S 0 212 0 0 0
fiorella S 0 486 0 0 0
piutltst02 S 0 1.4 Kb 0 0 0
mostardi S 0 952 0 0 0
193.43.104.55 S 0 588 0 0 0
itest1 S 0 928 0 0 0
rolly S 0 46 0 0 0
itin2 S 0 92 0 0 0
3comhub1 S 0 610 0 0 0
re S 0 5.6 Kb 0 0 0
pi100 S 0 1.2 Kb 0 0 0
lcardini S 0 546 0 0 0
mbeng S 0 602 0 0 0
itest2 S 0 600 0 0 0
fossati-a S 0 960 0 0 0
hpwsutl S 0 3.1 Kb 0 0 0
catlc S 0 120 0 0 0
aut01b S 0 243 0 0 0
biu S 0 542 0 0 0
artico2 S 0 226 0 0 0

Business or Part-time Project ?

- As project grows, more time is needed to spend on the tool but the employee does not like "unfocused" workers.
- Alone on a project is fine, but development cycles are long.
- At some point I had to make a decision:
  - Look for funding, eventually making a foundation as Wireshark and Suricata do (but ntop doesn't have a Janice or Kelly equivalent).
  - Turn this project into a business, make revenues, pay bills and hire people. This is what I did.

What Kind of Company is ntop?

- In short, not a standard company for various reasons.
  - Innovate, and don't settle.
  - Opensource from the source: don't pack open source, make open source !
  - European mindset: designed to last.
  - Open source with premium editions.
  - Completely free (not just discounted) for education, and no-profit.
  - No marketing team (nothing against marketing folks) but we're hackers (*"What's a network? A hacker with a modem"*, Linux Network Administrators Guide).

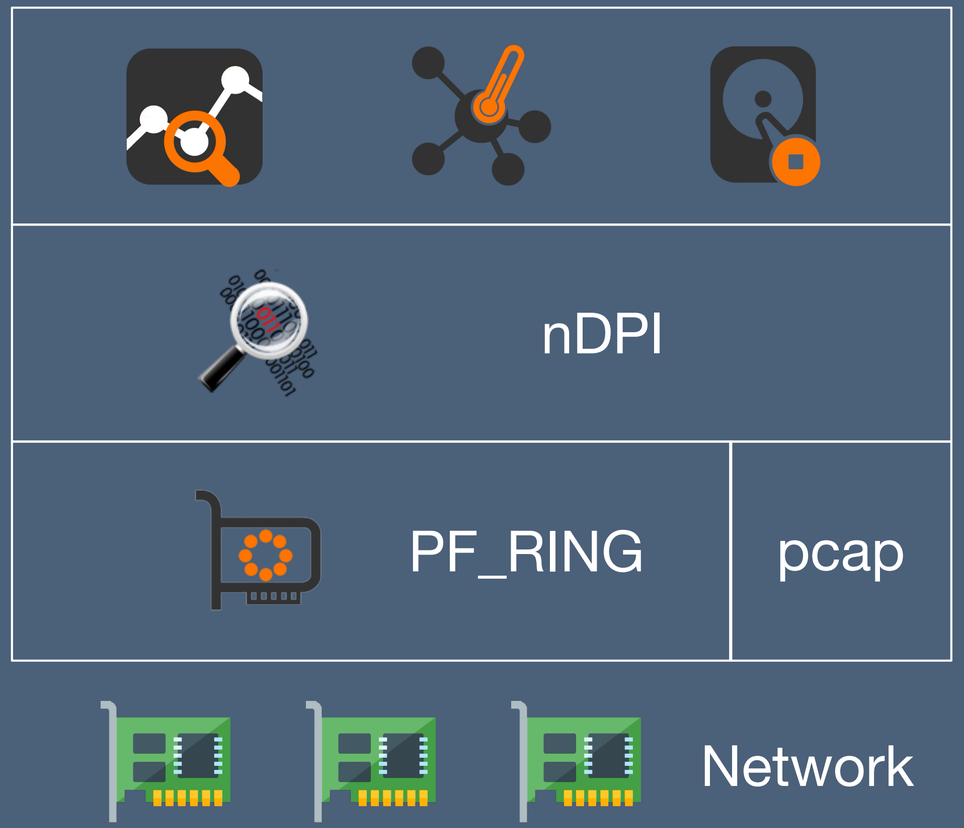
ntop Today

- Two locations: Italy (Pisa, Tuscany) and Switzerland (Ticino).
- Small development team (< 10 people): everything must be under control, quick development cycle, you can speak with developers.
- 30+ partners and resellers in all the continents.
- ~40% revenues coming from North America, ~30% from Europe, ~30% from the rest of the world.
- Most customers are mid-large size companies, governments, OEMs.
- Community channels in addition to GitHub and premium support.

Training and Support

- We believe that as a small company we need to be different, providing premium service to an international user base (i.e. think in terms of timezones, national holidays, week workdays).
- We spend a large part of our time with support both direct and community.
- We offer remote (twice a year) and on-site training (on demand).

ntop Application Architecture

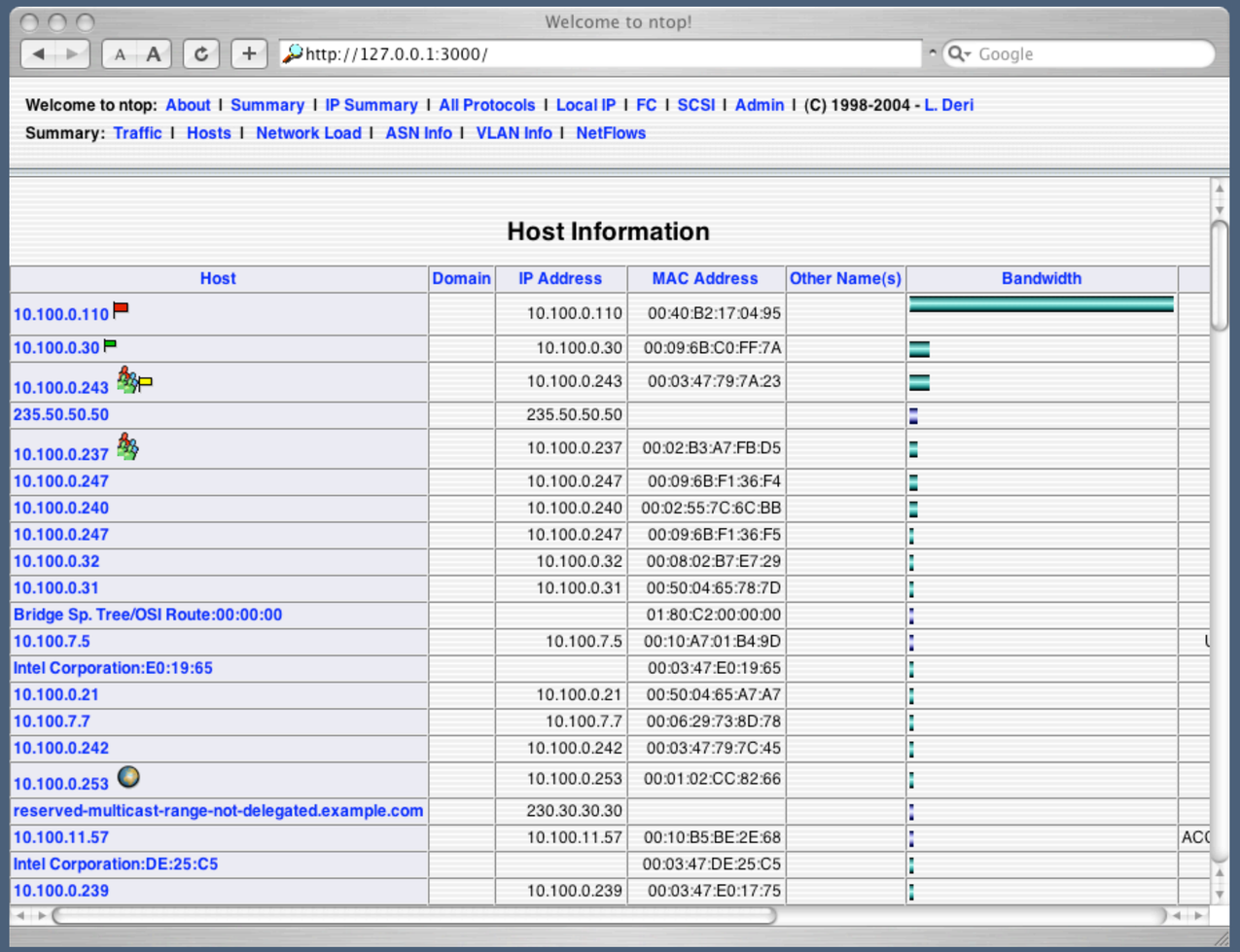


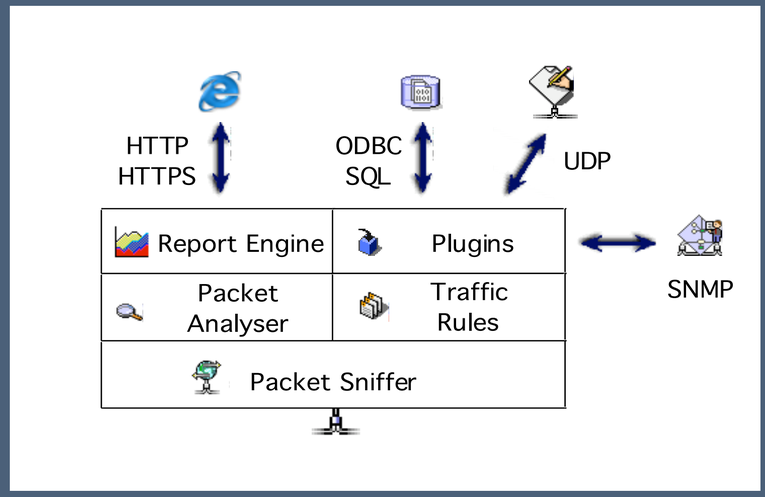
Where We "Were" Heading: 1998-2005 [1/2]

"If you obey the rules you'll miss all the fun" (ntop motto, 1998)

- When we started our focus was on visibility: develop tool able capable of keeping up with network speed on commodity hardware.
- Wireshark is also a good example of an open source tool that changed the rules of packet sniffing, both in terms of features and price.
- What was the meaning of "visibility" 25 years ago (~100 Mbit links) ?
  - Top talkers, top protocols, LAN connection matrix, multicast groups statistics.
- We developed the original ntop (now ntop-ng) and nProbe (packet-to-flow software probe) for complete visibility. nProbe has been the first probe on a market saturated by NetFlow/sFlow collectors.

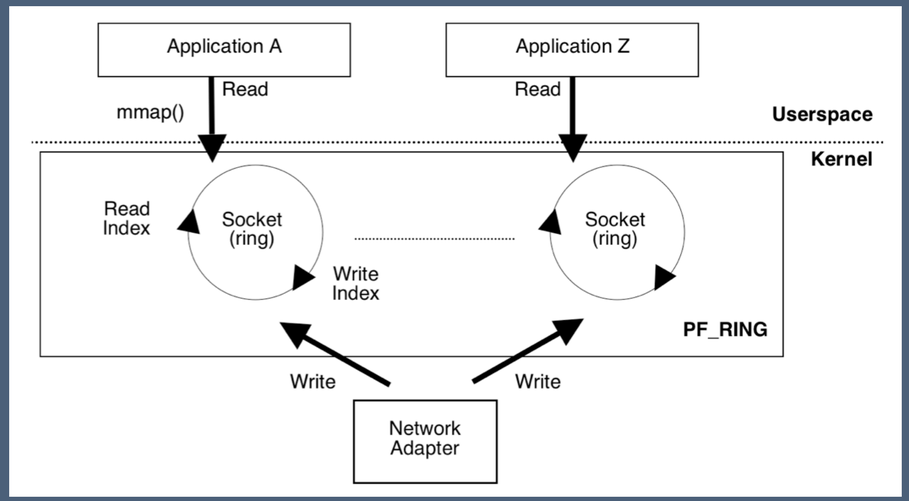
Where We "Were" Heading: 1998-2005 [2/2]





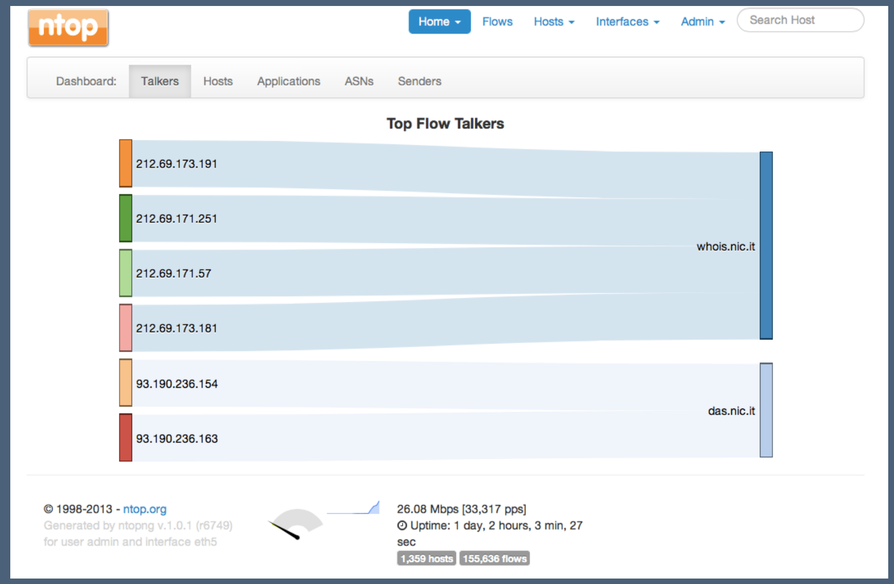
Where We "Were" Heading: 2005-10

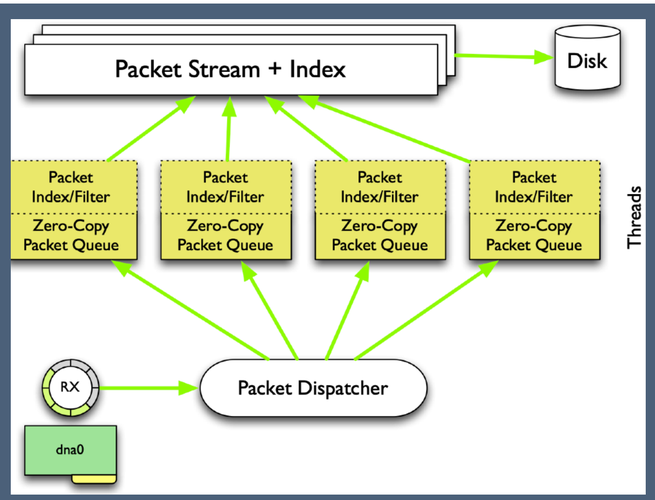
- Network started to become faster, CPUs/OSs not much faster, packet capture became a problem.
- Options: use (costly) FPGA accelerated cards able to handle 1 Gbit, or....
- We took option 2 and PF_RING was created:
  - Ability to capture traffic at 1 Gbit on Linux
  - No changes in kernel drivers.



Where We "Were" Heading: 2010-15 [1/2]

- In 2013 ntopng 1.0 was released.
- After a short TNAPI -> DNA (Direct Network Access) transition, PF_RING ZC (Zero Copy) was released that guaranteed 1 and 10 Gbit line rate packet capture (DPDK was "inspired" by our work on PF_RING).
- We have created n2disk 1.0 for dumping traffic to disk at line rate, our first network traffic recorder.
- In 2014 we have also create nScrub for DDoS traffic mitigation.





Where We "Were" Heading: 2010-15 [2/2]

- In 2011 it became clear that TCP/80=HTTP or TCP/443=HTTPS was no longer holding.
- The Internet changed and we had to change with it as visibility was seriously compromised. In essence it was time for DPI (Deep Packet Inspection).
- From unmaintained OpenDPI code we have created nDPI. The first few years were spent fixing many code design limitations, leaks, insane memory usage, but we finally reimplemented visibility on our tools.

Where We "Were" Heading: 2015-24 [1/2]

- These years have been the 100 Gbit decade:
  - We have been busy creating new ZC drivers for modern/faster network adapters.
  - Rethinking monitoring by complementing existing standard formats (SNMP, NetFlow/IPFIX/sFlow) with less efficient but more modern ones such as those JSON-based.
  - Data was delivered using message brokers (Kafka).
  - Timeseries databases became popular so we had to support them (e.g. Influx, Prometheus) as the new market of data "observability" fostered by the cloud became popular.
- nDPI moved from ~100 protocols to ~500 and it features various algorithms and tricks to make it 100 Gbit friendly. It has been integrated with popular tools such as Wireshark and Suricata.

Where We "Were" Heading: 2015-24 [2/2]

- After some attempts to design our own data indexing system (nIndex), with the advent of columnar databases we decided to scale up and become "data-lake oriented": today nProbe/nProbe Cento/ntopng are used by many users world-wide to fill up their lakes.
- We've added support of popular tools such as Elastic and Grafana.
- Support for mobile networks (GTP) and industrial networks (OT) is now a stable feature of all our tools. Cybersecurity is a first class citizen on all our tools.
- A new tool nTap (virtual network tap) has been created and ntopng is radically different from what it used to be 10 years ago and it now features an inline version (ntopng Edge).
- We have started to create (beta) ntop Cloud for interconnection ntop applications.
- We did many other things in the last decade, but the question is: what are our future steps?

ntop Roadmap

- We need to squeeze performance even further as 400 Gbit is already used in some datacenter.
- As PF_RING has been the cornerstone of ntop apps for 20+ years, nDPI is pervasive across all the applications, and this is the tool where we should focus for both visibility and cybersecurity.
- AI is here to stay: let's use it (not for chatbots please) !

ntop Roadmap: Towards 400 Gbit [1/2]

- Host CPUs need to offload tasks to network adapters in order to scale up.
- We are trying to port code to smart/superNICs since 5+ years (P4, ARM cores) in order to simplify software application design. This technology does not seem to be mature yet, so for the time being we're using them from x64 host.
- The main limitation for scalability is the flow cache, but fortunately modern NICs are implementing hw mechanisms for assisting sw applications.
- Currently the most promising solution is Napatech Flow Manager that we are using successfully on nProbe Cento.
- Possible alternatives (not yet mature) nVidia BlueField and Broadcom TruFlow.

ntop Roadmap: Towards 400 Gbit [2/2]

- We have on the pipeline a 200 Gbit packet recorder using little CPU cycles based on Napatech FPGA network adapters. Currently the main limitation is the storage cost that requires mixed-use NVMEs
- We're planning to leverage flow caches in some other products such as nScrub, and eventually create a component for building a hw-accelerated extension to nDPI where it's necessary to have an efficient flow cache.

ntop Roadmap: nDPI [1/3]

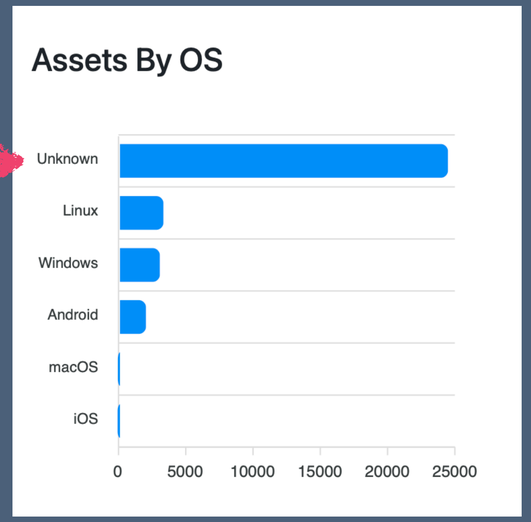
- 10+ years ago we have started to develop nDPI for characterizing network protocols.
- Today we need to continue along that line and better characterise "generic" protocols such as TLS and QUIC. For instance classify domain names in categories (business, sport, software...) to detect what is a allowed/forbidden on certain locations (e.g. betting on horses or watching adult content at work is not expected).
- Creating protocol dissectors requires skilled developers to code this into nDPI: as the number of "application protocols" is very high compared to nDPI developers we need to automate this task.

ntop Roadmap: nDPI [2/3]

- Since a few months, we have been developing (currently and internal project, but soon public) an nDPI+ component able to:
  - (Semi)automatically analyze a packet trace of a host (including a tablet or a mobile phone), cluster traffic traces generated by applications (e.g. divide the traffic generated by app A from the one generated by app B), and generate an nDPI configuration file to be loaded in nDPI for detecting new applications (no coding required).
  - Implement accurate multimedia call tracing (merge individual flows into a single "Luca called Miriam" call) and in general transform flows into activities meaningful to humans (e.g. host X downloaded a file).

ntop Roadmap: nDPI [3/3]

- nDPI+ (cont)
  - Network Intelligence: identify OSs, host "inconsistencies" (e.g. a hotspot), operating systems (via fingerprints).
  - High-speed Linux traffic labelled and enforcer: user-space component used to nDPI-analyze the first few flow packets and report monitored data to ntopng.
  - Detailed traffic metadata extraction, including cryptocurrencies information.
  - Extend nDPI to implement a Zeek-like mechanism for generating rich metadata stored on protocol files.
  - Offer professional services for those who want to deploy nDPI on their tools.



ntop Roadmap: Artificial Intelligence [1/4]

- AI has been often used for chatbot-like tasks that often replace GUIs with text. We don't believe that this is the future.
- AI has to be used to consolidate and analyze monitored data. Examples:
  - Correlate alerts, silencing non-relevant ones (e.g. leveraging on assets information), and highlighting the most important.
  - Generate code for creating nDPI dissectors and ntopng flow/host checks to reduce the development cycle and lower the bar for extending tools.
  - Enrich monitored data by providing context by means of automatic web search, prior knowledge (e.g. during LLM training), and acting as an expert in data analysis and cybersecurity.

ntop Roadmap: Artificial Intelligence [2/4]

DeepWiki

powered by Devin | Share

← return to ntop/nDPI

write the skeleton of a new nDPI protocol dissector

Fast

Searched across ntop/nDPI | Go deeper

I'll help you create a skeleton for a new nDPI protocol dissector. Based on the code context provided, I can see that nDPI is a deep packet inspection library that identifies network protocols through protocol dissectors.

Answer

To create a new protocol dissector for nDPI, you need to follow a specific structure that integrates with the nDPI framework. Here's a skeleton for implementing a new protocol dissector:

Step 1: Create the protocol header file

First, create a new file in `src/lib/protocols/your_protocol.c`:

```
/*
```

C ntop/nDPI src/lib/protocols/stun.c

```
1140
1141 /* ************************************************************ */
1142
1143 static void ndpi_int_stun_add_connection(struct ndpi_detection_module_struct *ndpi_struct,
1144                                          struct ndpi_flow_struct *flow,
1145                                          u_int16_t app_proto,
1146                                          u_int16_t master_proto) {
1147   ndpi_confidence_t confidence = NDPI_CONFIDENCE_DPI;
1148   u_int16_t new_app_proto;
1149
1150   /* In monitoring the classification can't change again */
```

C ntop/nDPI src/include/ndpi_private.h

```
419   struct cache *tinc_cache;
420
421   /* NDPI_PROTOCOL_OOKLA */
422   struct ndpi_lru_cache *ookla_cache;
423
424   /* NDPI_PROTOCOL_SIGNAL */
425   struct ndpi_lru_cache *signal_cache;
426
427   /* NDPI_PROTOCOL_BITTORRENT */
428   struct ndpi_lru_cache *bittorrent_cache;
```

```
766 void init_diameter_dissector(struct ndpi_detection_module_struct *ndpi_struct, u_int32_t *id);
```

ntop Roadmap: Artificial Intelligence [3/4]

- In our view AI can be used in two areas:
  - [Déjà Vu] Automate human operations, make decisions based on prior knowledge (e.g. I have seen alert X: is it important or I can ignore it?) and in essence "turn most users into experts".
  - [New] Use AI to provide monitored data (e.g. a pcap file of a device) and generate some metadata that can be used in nDPI/nProbe/ntopng to "teach" the tool without requiring code development.

ntop Roadmap: Artificial Intelligence [4/4]

Some use cases:

- Detection of new application protocols (or apps installed on mobile devices) or things like "this is a PLC: look at its traffic, find out what methods can be used to detect its traffic and generate a config file for spotting these devices".

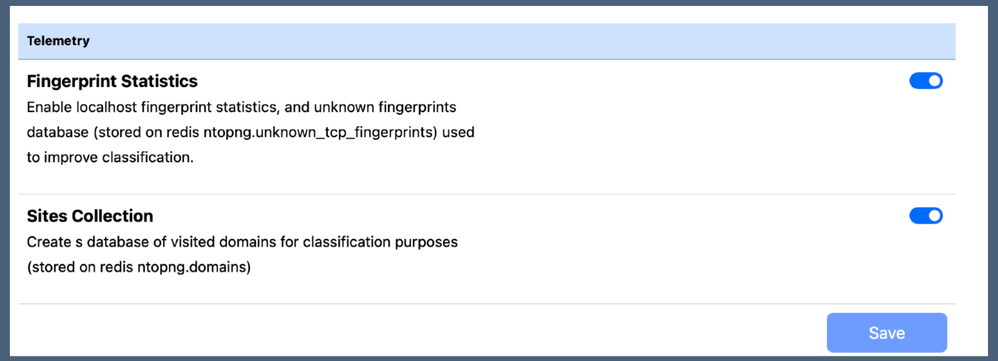
- Misbehavior detection: "ntopng has observed the traffic of my network for a month. Identify corner cases or weak network design. Provide suggestions on how to fix them, and in the meantime generate an ACL configuration file for my device model X for blocking the most dangerous issues that have been observed".

ntop Roadmap: ntopng [1/3]

- ntopng is the most popular tool we develop, often used by non-experts.
- How can we exploit ntopng popularity? Some ideas:
  - (Upon user consent) send IP addresses of heavy attackers via the ntop Cloud to ntop that will then redistribute this information in realtime to all active ntopng instances.
  - Characterize unknown traffic fingerprints leveraging on other learned information and send them to ntop for improving nDPI. Example if there is an un known TCP fingerprint originated by an Android device (learnt by other means), such fingerprint is automatically assigned to Android devices too.

ntop Roadmap: ntopng [2/3]

- We have studied large networks for several months and the result is:
  - During the first few days (<< 1 week) we can "auto train" the system in terms of observed domains and fingerprints.
  - In the following weeks, very few new domains/fingerprints are observed, so we can avoid classifying "the world" as the problem to tackle is much simpler.

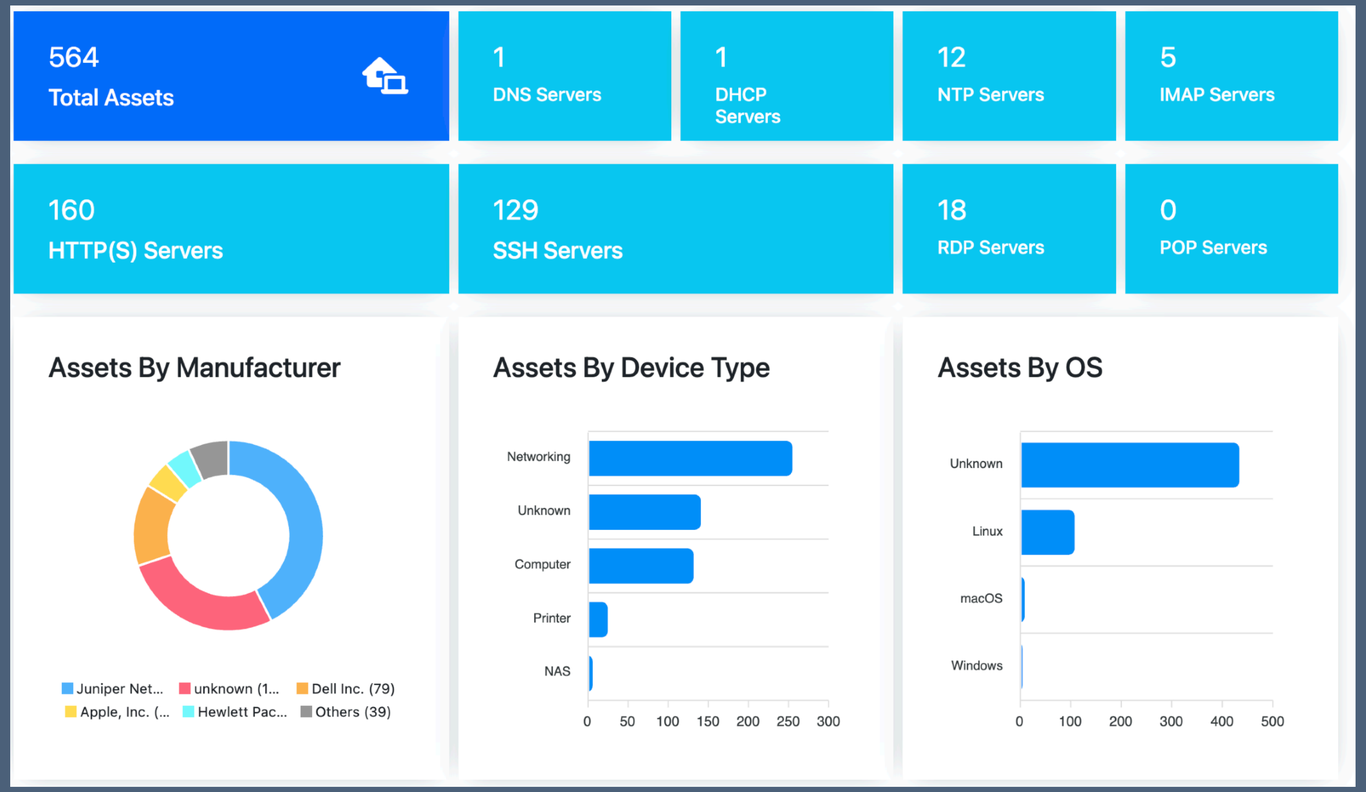


ntop Roadmap: ntopng [3/3]

- Improve asset inventory by leveraging on active monitoring (e.g. OT protocol probing). Likewise, use asset inventory to detect unexpected traffic flows (e.g. an iPhone that connects to a Windows server via RDP).
- Use AI to detect inconsistencies , and automatically label devices with multiple tags (e.g. printer, windows device)
- Scale-up for monitoring large-distributed networks and for users whose network do not belong to them (e.g. IXPs and service providers)

Our Vision in a Nutshell

- Who is using my network and for what purpose ?
- Are network communications efficient and of good quality ?
- Is my network secure ?
- Use AI to self-adapt monitoring tools to my network.



Thank you!