Automated Malware Analysis The Dissect || PE Project way

Rodrigo Rubira Branco

Founder Dissect || PE – Now the Qualys Vulnerability & Malware Research Lab

rodrigo *noSPAM* kernelhacking.com

http://twitter.com/bsdaemon

Agenda

- Motivation: Vulnerability x Malware Research
- The Feed Server Challenges
- Laboratory Topology
- Scheduler
  - Scalability
- Unpackers
- Dissectors
- Kernel Driver
- Innovations and Community Support
- Interface and Results

QUALYS®

ON DEMAND SECURITY

Before Starting

- I have always been a security researcher in the sense of ‘vulnerability exploitation / mitigation researcher’
- When first had to deal with the Malware Analysis problem I tried to see what is common area and what is the differences

QUALYS®

ON DEMAND SECURITY

Malware x Vulnerability Research

- Many books on automated vulnerability hunting
  - Fuzzers are a common sense requirement for security testing
  - When we say we have some test cases, we usually mean millions
- Everything that you read or that is released on automated malware analysis uses the ‘automated’ word in a way to replace some tasks done by the analyst
  - Automated tools for analysis lack the performance and scalability requirements for real usage
- We all know the AV vendors have complex laboratories, from time to time we see some results, but no one ever documented how it works and what it does, and neither what are the limitations

QUALYS®

ON DEMAND SECURITY

What is missing here?

- Researchers and academics release many new ideas and techniques for malware analysis
- Nobody is really able to apply/test those techniques due to the lack of:
  - Good amount of samples
  - Machine power to process the samples
- THIS TALK IS ABOUT TO GIVE AWAY THE MISSING PARTS
- AND EVEN IF YOU DON'T WANT TO USE IT, YOU CAN USE OUR EXPERIENCE BUILDING THE LAB TO BUILD YOUR OWN!

QUALYS®

ON DEMAND SECURITY

Motivation

- Pretty obvious, hundred thousands of new malwares every week
- Complex systems, professionals developing malcodes
- Submit to the vendor is not always a good option:
  - Targeted attacks?
  - Timely reply?
- Some public options:
  - Missing network dissection
  - Usually just shows binary internals or antivirus results
  - How do I execute my own analysis scripts?

Frontend x Backend

8 machines
Complex Analysis
Network Monitoring

1 machine
Used to exchange samples with partners
Initial analysis (classification) is done here

QUALYS®

ON DEMAND SECURITY

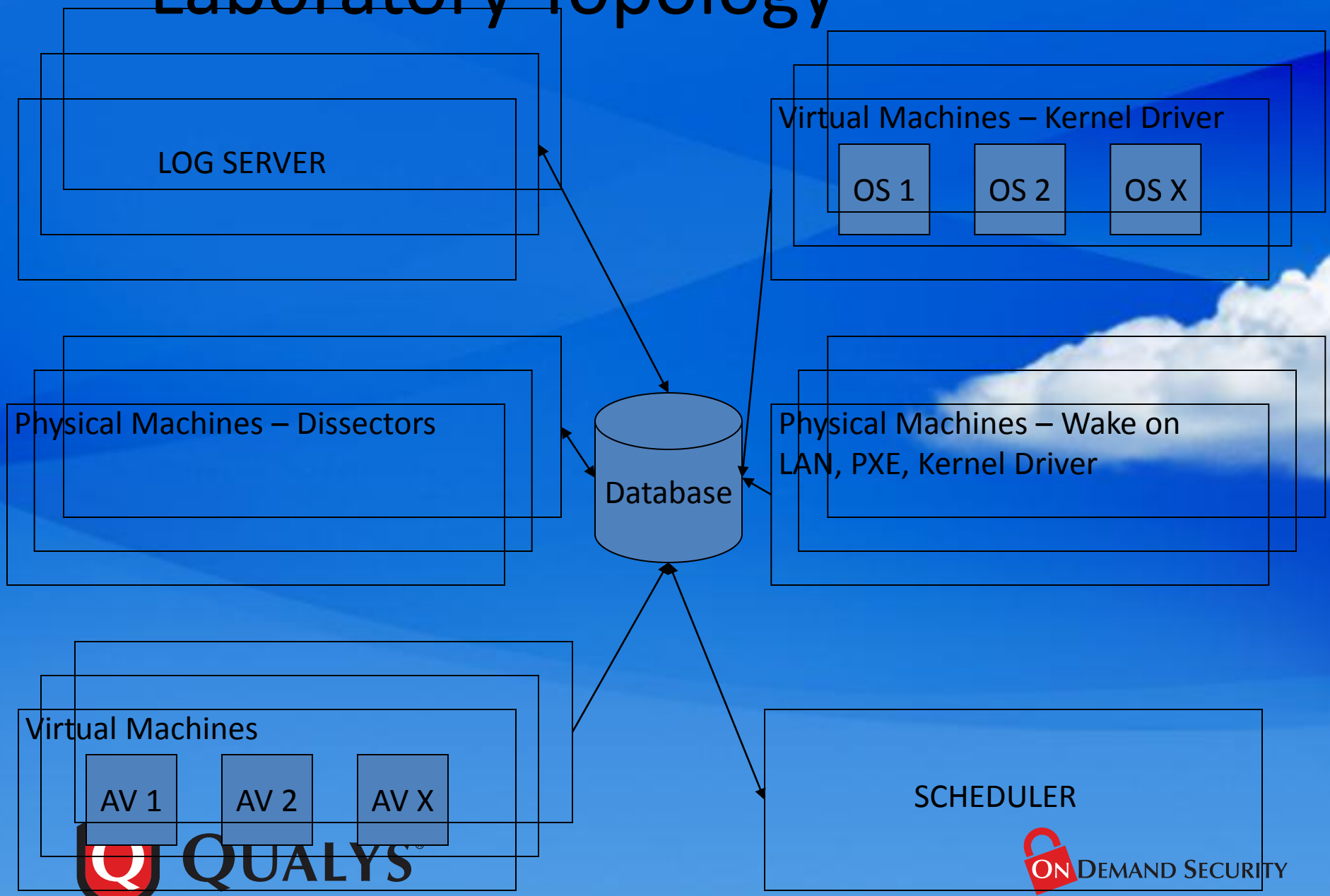
The Feed Server Challenge

- Samples comming from different partners, in different formats (.tar.gz, .tar.bz2, .zip, .rar, .7zip, with and without password, encrypted or not)
- Huge files breaking wget –r (ok ok, we implemented our own downloader with resuming options -> we don't want to try everytime all the files to see if we already downloaded it)
- In different months, we receive the same samples from the same partner
- In the same day we receive multiple copies of the same sample from different partners (and sometimes, from the same partner too!)

QUALYS®

ON DEMAND SECURITY

Laboratory Topology



Scheduler

- Inotify interfaces: Multiple priority queues
- Routing policies:
  - Round Robin(RR) -> Each Unit registers itself into the central Data Base and mark itself as "FREE" state. When the Scheduler needs to dispatch a new malware it simply gets the first "FREE" Unit and dispatch it as a new task. There is no starvation.
  - Least Load(LL) -> Each Unit registers itself into the central Data Base inserting its (INT) "LOAD" number; the insertion is being frequently updated. When the Scheduler needs to dispatch new malware, it simply chooses the lower (INT) LOAD number.
  - Fast Respond(FR) -> Each Unit registers itself into the central Data Base inserting its (INT) “RTT” number which represents the ICMP round trip from the scheduler and back. The insertion is frequently updated. When the Scheduler needs to dispatch new malware it simply choose the lower (INT) RTT number -> Support remote analysis machines
- Wake on lan + PXE boot of real machines when malware refuses to run in a VM
- The central database is updated when a new machine registers itself : it just plug a new element using DHCP and DNS name to specify what it does in the architecture. O(1) to add any new device. Each device has a minimal instance of the scheduler -> Scheduler election in case of failure of the central unit.

QUALYS®

ON DEMAND SECURITY

Plugins

- Malware samples are analyzed by applications that will be referenced as ‘plugins’
- The architechture was developed to allow an easy insertion and removal of plugins
- It is very easy to write plugins:

```
#include <stdio.h>
int main(int argc, char **argv) {
    printf(“My plugin result.”);
    return 1;
}
```

QUALYS®

ON DEMAND SECURITY

Community Support

- Plugins are easily created in the platform
- Plugin might choose to run in ALL samples (receive less priority), and will be executed in the new ones
- The failure of a plugin does not affect the whole architecture: Great for academic research
  - Do you need malware samples? Why don't you send us your plugin and get the results of it executed against millions of analyzed samples?

QUALYS®

ON DEMAND SECURITY

Scalability

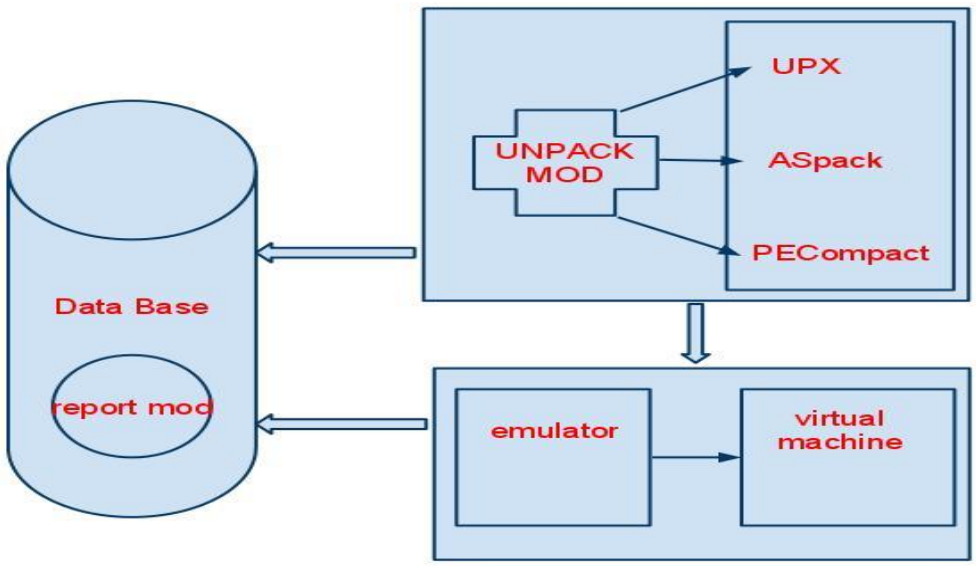
- It is supported in the actual architecture 255 machines running 60 VMs each
  - Tested with 8 machines with average of 60 VMs each
- Network communication is the upper limit
  - Maybe, 10 Gbps networks allows the usage of more machines
- Syslog is used by everything
  - Multiple debugging options

QUALYS®

ON DEMAND SECURITY

Unpackers

- We have an emulator to collect information from binaries and to automatically unpack some specific packers
- Unpackers are just normal plugins in the architecture



QUALYS®

ON DEMAND SECURITY

Dissectors

- Malwares do have specific network traffic associated to them, which can be used to further identify the specimen

- Network traffic analysis takes time:
  - Collect the traffic
  - Open in wireshark
  - See the sessions
  - Get other links, download the samples, re-do the analysis...

- The dissectors do everything automatically:
  - Supporting IRC, P2P, HTTP, DNS and other protocols
  - Automatically downloads and inserts in the queue other parts of the malware
  - SSL Inspection (pre-loaded keys)

QUALYS®

ON DEMAND SECURITY

Kernel Driver

- Responsible for collecting everything inside the VM (or the real machine, if needed to boot one)
- Intercept the behaviour of the analyzed sample (function calls, entry points, memory dump)
- Uses a proprietary RPC channel to communicate it findings (i.e. to say the binary refused to run)
- Implements inotify-like interface for Windows

QUALYS®

ON DEMAND SECURITY

Innovations ?

- We are testing the performance benefits of SSD
  - Read-only VM image loaded from an SSD disk
  - Read-write temp VM image in normal SATA 3 disks
- We already discarded 10 G NICs
  - Trying to use RDMA to read malware memory (no need for the driver)
  - Easier to detect such NICs than to detect our driver

QUALYS®

ON DEMAND SECURITY

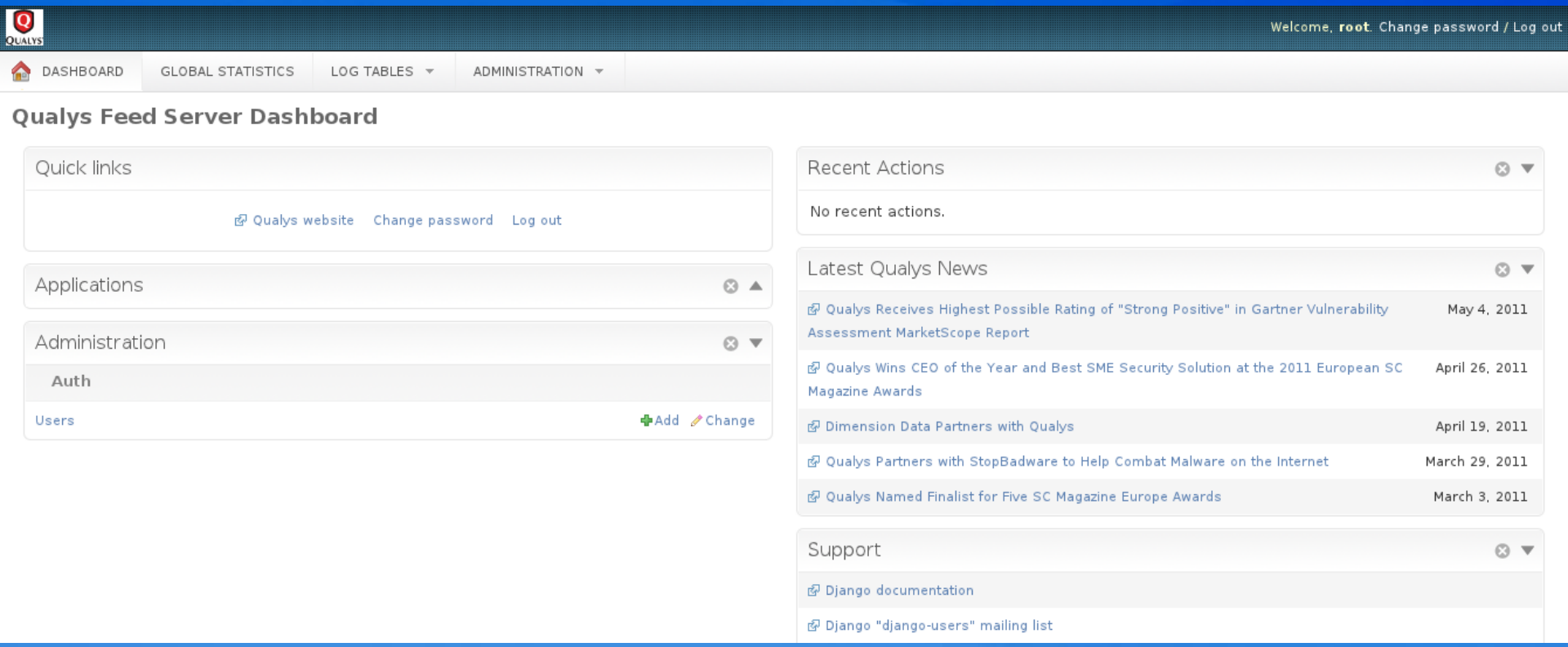
Interface and Results

- We analyzed more than 30 millions of malicious binaries, received from different sources and collected from different web collections, in 8 different real machines running up to 60 virtual machines each with a total of more than 50 cores.
- From these, 68% were actually packed.
- From that we conclude that most of the analyzed malwares are simple variations to avoid detection by actual anti-virus software (more than 90%). Also, most of the detected variations are detected by the heuristics engines of the anti-virus, other than specific signatures.

QUALYS®

ON DEMAND SECURITY

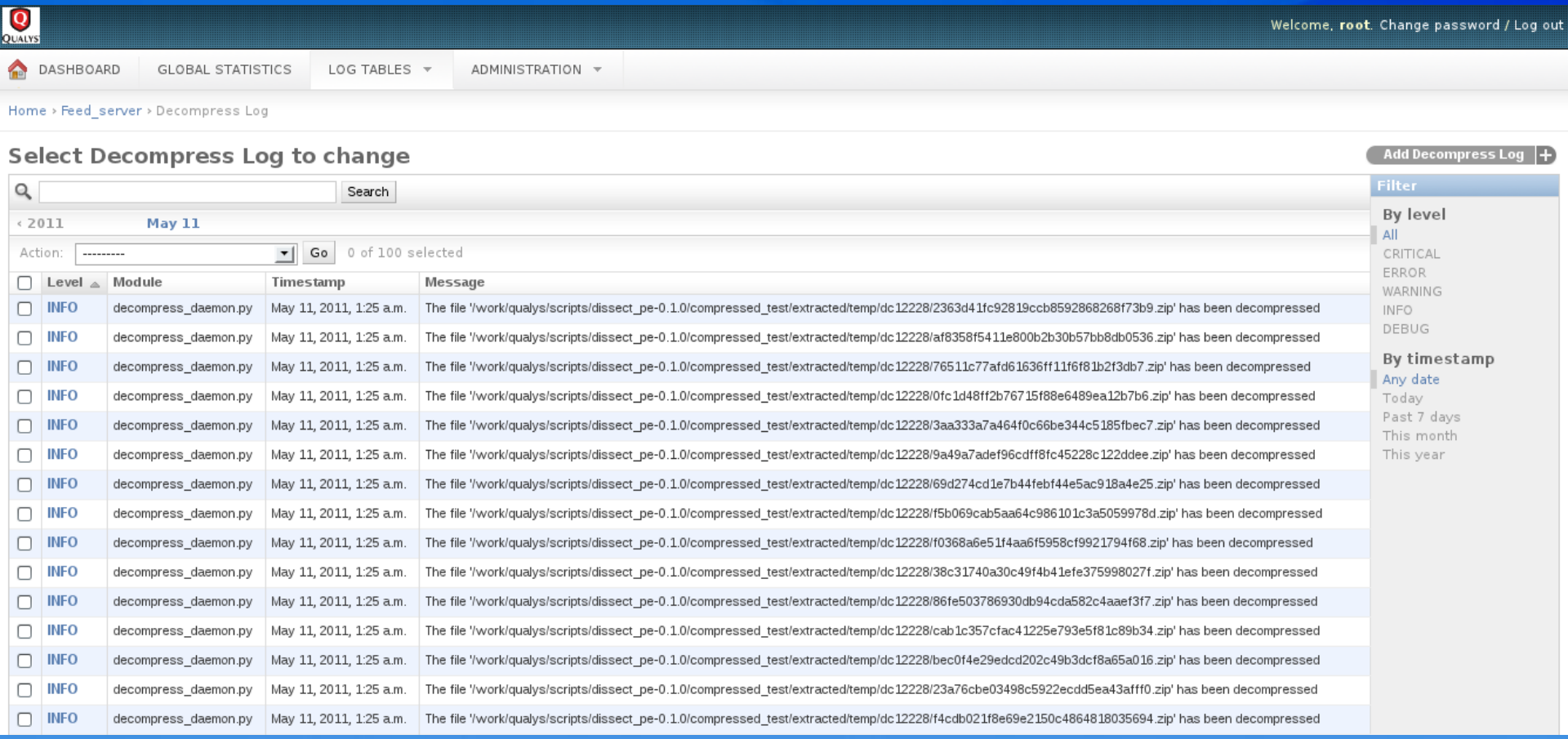
Main Screen



QUALYS®

ON DEMAND SECURITY

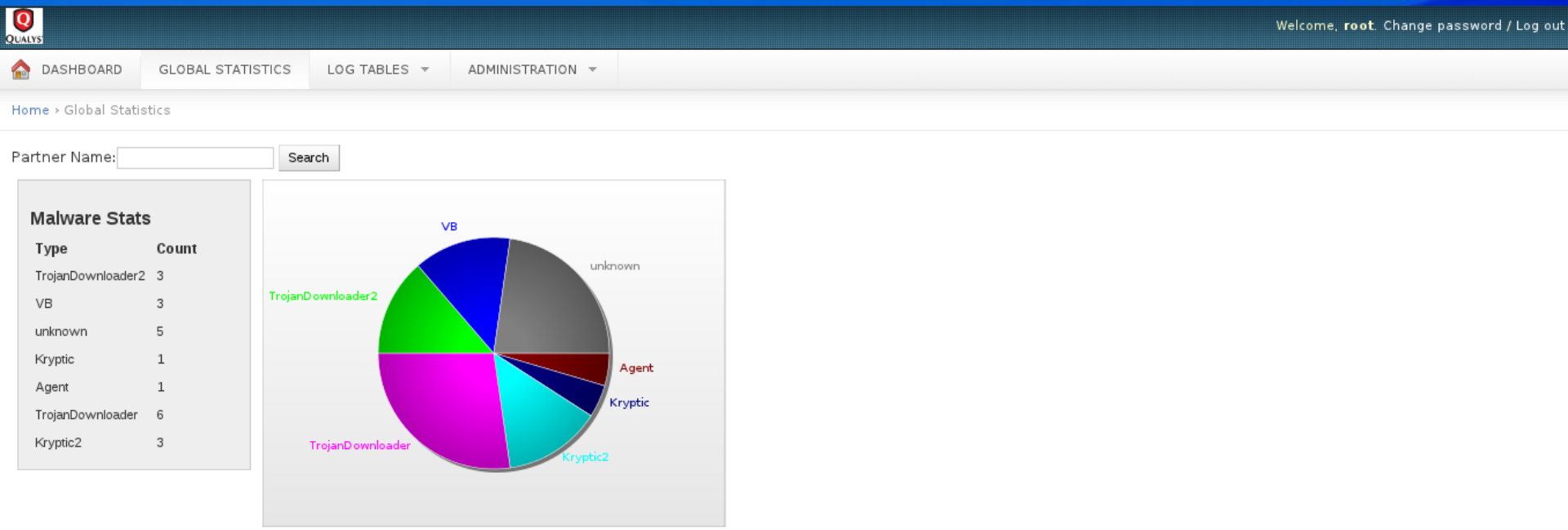
Archive Decompression



QUALYS®

ON DEMAND SECURITY

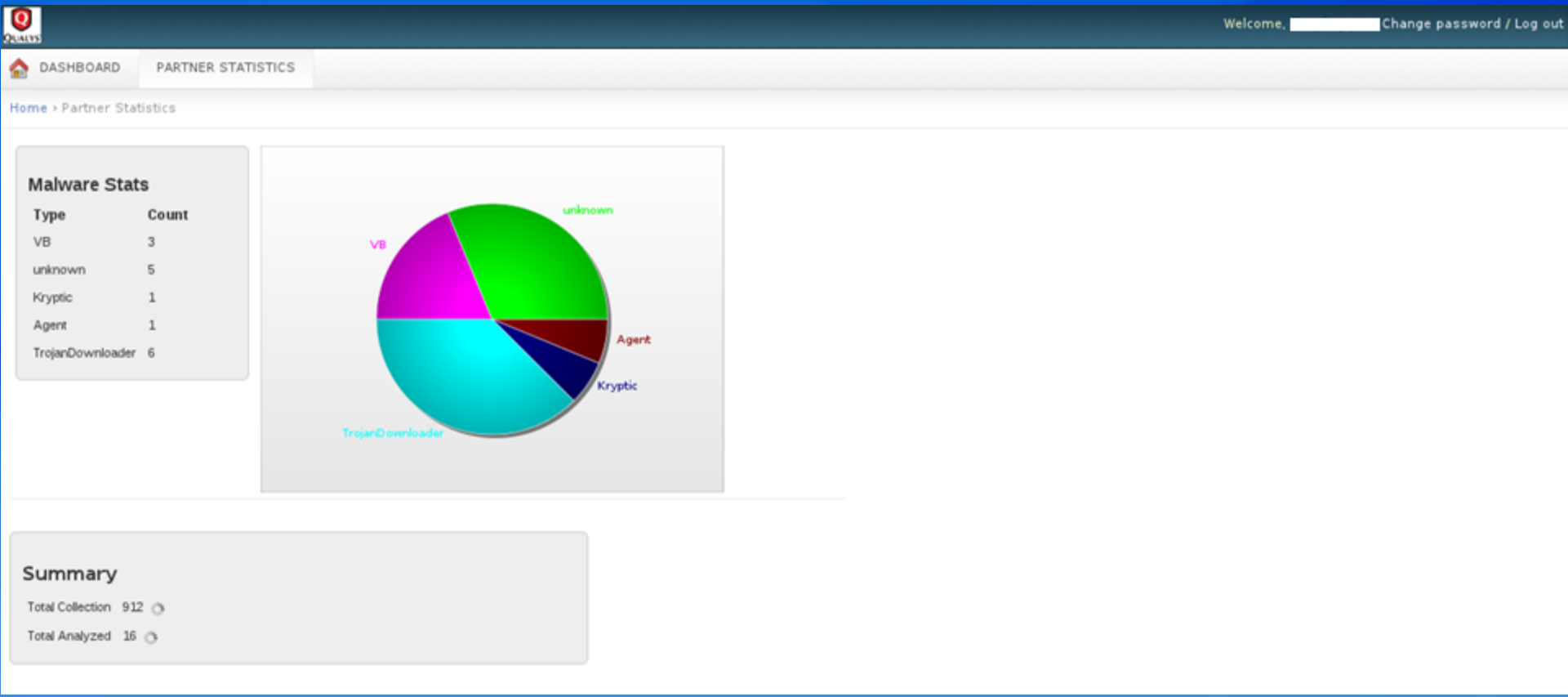
Global Statistics Sample



QUALYS®

ON DEMAND SECURITY

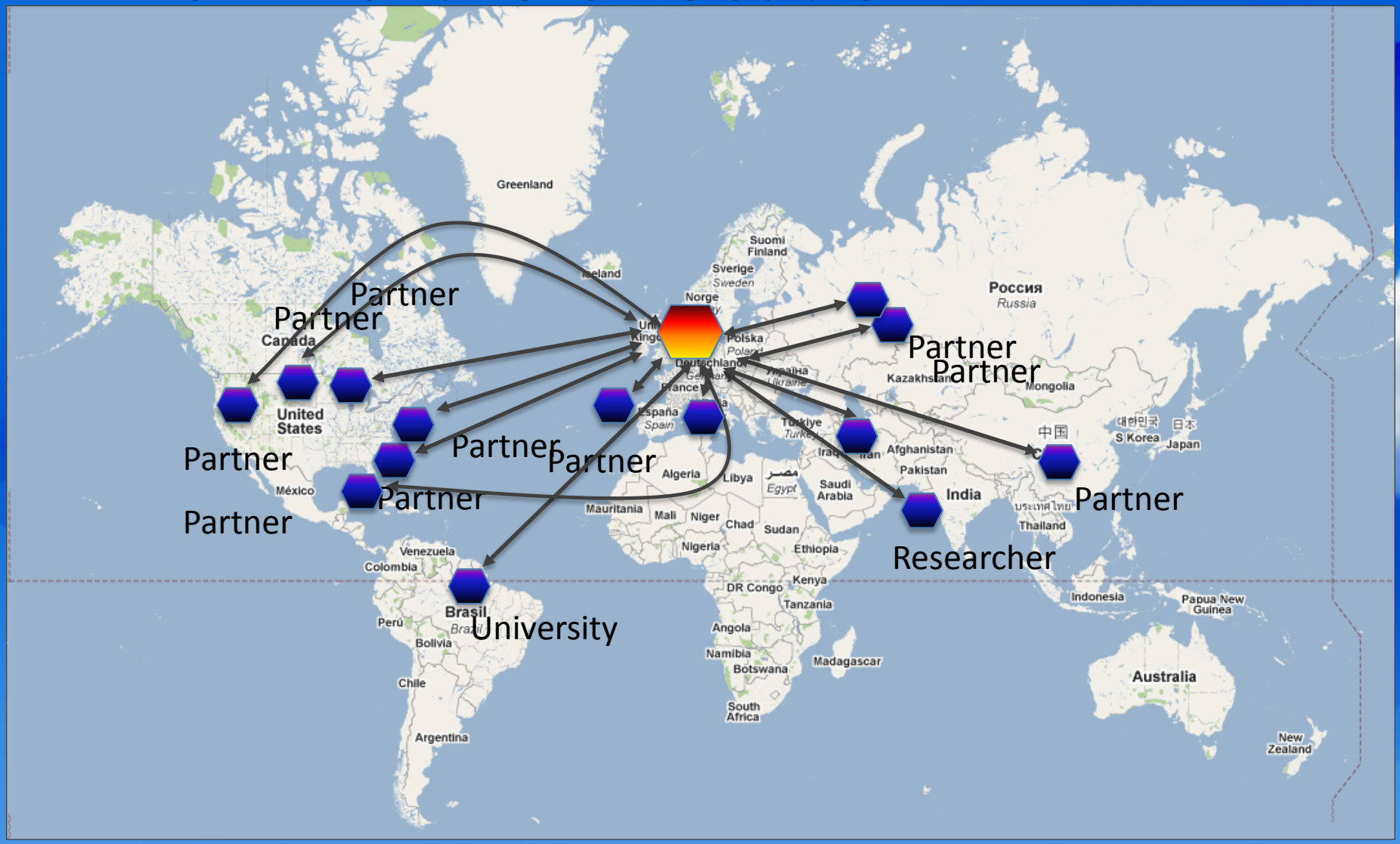
Per-partner Statistics



QUALYS®

ON DEMAND SECURITY

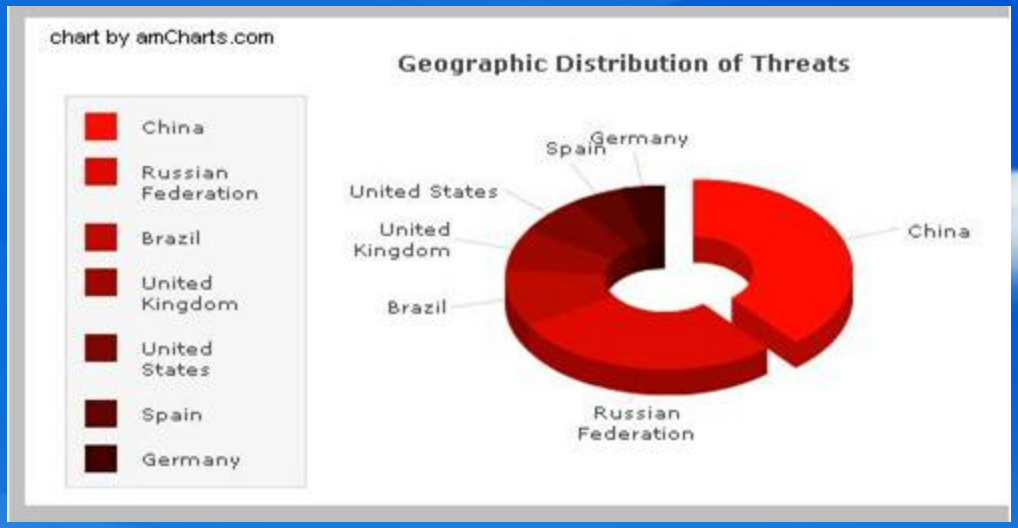
Main Partners Location



QUALYS®

ON DEMAND SECURITY

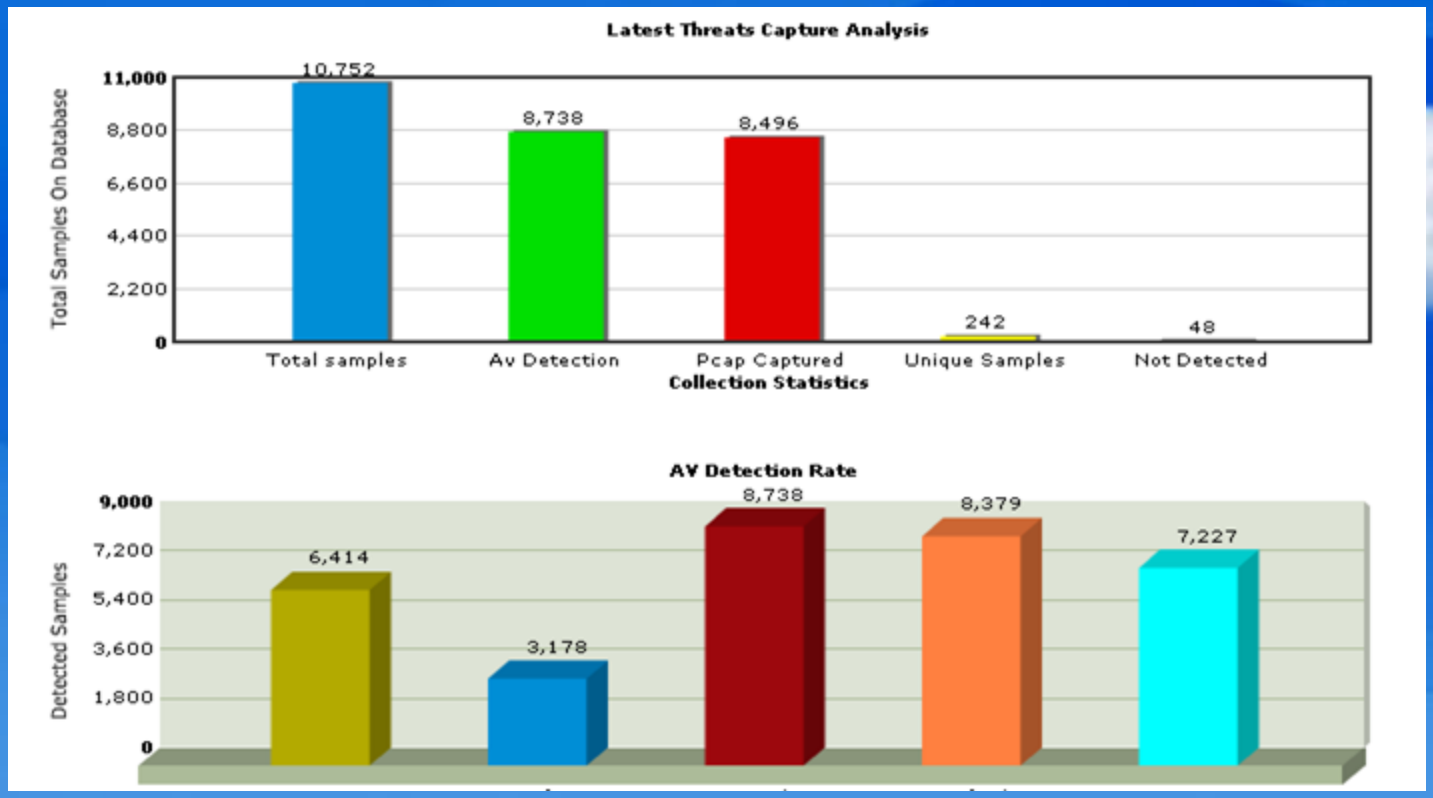
Geographical Distribuition of Threats Not-yet integrated to the portal (backend-only)



QUALYS®

ON DEMAND SECURITY

Capture and Detection Statistics Not-yet integrated to the portal (backend-only)



QUALYS®

ON DEMAND SECURITY

Dropping Points
Not-yet integrated to the portal (backend-only)

Last Seen	Domain	Malware URL	Site Advisor	Google Safe Browsing	Zeus Tracker	Malware Domain	Domain Reg	Country
2009-12-14 12:08:25	hadwares.com	LISTED	---	---	---	LISTED	12-oct-2010	CN
2009-12-14 10:34:53	freehostia.com	---	---	---	---	LISTED	26-may-2014	GB
2009-12-14 15:50:25	shmaltooz.cn	---	---	---	---	LISTED		CN
2009-12-16 12:05:37	mscom-wui.vo.llnwd.net	---	---	---	---	---		NL
2009-12-16 14:05:32	gpdev.net	---	---	---	---	---	13-nov-2010	US
2009-12-16 14:06:30	flap71.com	---	---	---	---	---	27-oct-2010	US
2009-12-16 15:20:57	wapdodoit.ru	LISTED	---	---		LISTED		FR
2009-12-16 15:22:23	estoniashi.ru	---	---	---	---	---		US
2009-12-16 16:05:48	downloadavr15.com	---	---	LISTED	---	---	14-dec-2010	FR
2009-12-16 16:07:11	testavrdown.com	LISTED	---	LISTED	---	LISTED	09-aug-2010	FR
2009-12-16 16:31:08	sunmicro-1.vo.llnwd.net	---	---	---	---	---		NL
2009-12-16 16:31:31	1.vo.llnwd.net	---	---	---	---	---		US
2009-12-16 16:37:18	dealio.com	---	---	---	---	---	08-sep-2010	US
2009-12-16 16:58:24	testavrdown.com	LISTED	---	LISTED	---	LISTED	09-aug-2010	FR
2009-12-16 17:31:27	wapdodoit.ru	LISTED	---	---		LISTED		FR
2009-12-16 18:01:14	orbitdownloader.com	---	---	---	---	---	17-oct-2011	None

QUALYS®

ON DEMAND SECURITY

Binary Information
Not-yet integrated to the portal (backend-only)

Time	Binary						Results URL
2010-06-22 15:24:27	exe.exe	None	None	None	None	None	Link To Result
2010-06-21 11:35:37	file.exe	None	None	Trojan.Win32.Agent	Win32/Oficla.GN trojan", action	None	Link To Result
2010-05-17 10:57:31	MalvRem_34.exe	None	None	None	None	None	Link To Result
2010-05-13 14:54:08	Malware.exe	None	Win32.Parite.B	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result
2010-05-13 14:49:55	Foxit Reader.exe	None	Win32.Parite.B	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result
2010-05-13 14:40:04	Parite.B.exe	None	Win32.Parite.B	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result
2010-05-13 14:34:48	setup103.exe	None	Trojan.Peed.Gen	None	a variant of Win32/Bamital	None	Link To Result
2010-05-13 14:26:37	D.exe	None	Backdoor.Bot.37560	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result
2010-05-13 14:20:17	C.exe	None	Backdoor.Bot.67157	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result
2010-05-13 14:16:52	B.exe	None	Win32.Parite.B	Virus.Win32.Parite	Win32/Parite.B virus", action	W32/Pate	Link To Result

AV vendor names deleted

QUALYS®

ON DEMAND SECURITY

Analysis Output

- Web output
  - We are going to use a local directory that I downloaded from the web interface (Link to the Result in the previous image)
- Flash Player 10 analysis (not all the analyzed binaries are malwares, right? The laboratory can be used to better understand what a binary is doing after all)
  - Output directory
  - Output files
  - Dissection

QUALYS®

ON DEMAND SECURITY

END! Really is!?

Rodrigo Rubira Branco

Founder Dissect || PE – Now the Qualys Vulnerability & Malware Research Lab

rodrigo *noSPAM* kernelhacking.com

http://twitter.com/bsdaemon