A CIP-PSP funded pilot action

Grant agreement n°325188

Deliverable	**D1.1.2 – Overall Software Description**
Work package	WP1 Requirements & Specifications
Due date	M30
Submission date	02/12/2015
Revision	2
Status of revision	Final
Responsible partner	ECO
Contributors	FKIE, INCIBE, Telecom Italia, TID, Montimage, ATOS, BDigital (EURECAT) XLAB, UL, ISCTI/GARR, DE-CIX, GDATA, IF(IS), CyberDefcon
Project Number	CIP-ICT PSP-2012-6 / 325188
Project Acronym	ACDC
Project Title	Advanced Cyber Defence Centre
Start Date of Project	01/02/2013

Dissemination Level	
PU: Public	X
PP: Restricted to other programme participants (including the Commission)	
RE: Restricted to a group specified by the consortium (including the Commission)	
CO: Confidential, only for members of the consortium (including the Commission)	

Version History

Rev	Date	Author(s)	Notes
V0.1	21/04/2015	Michael Weirich (eco)	First Draft
V0.2	22/04/2015	Michael Weirich (eco) Thorsten Kraft (eco) Peter Meyer (eco)	Input on CCH, Feng Office and German NSC
V0.4	17/06/2015	Michael Weirich (eco)	Revision of document
V0.5	24/06/2015	Thomas King (DE-CIX)	Contribution for DE-CIX
V0.6	29/06/2015	Ales Cernivec (XLAB)	Contribution
V0.7	05/07/2015	Edgardo Montes de Oca (MI) Christian Nordlohne (ifis) Christos Dimou (BDigital) Antonio Pastor (TID) Katia Velikova (BGPost) Andreas Fobian (GDATA)	Input on sensors and developments
V0.7a	13/07/2015	Tigran Avanesov (UL) Darko Perhoc (Carnet) Roberto Cecchini (ISCTI/GARR)	Contributions on developments
V0.8	15/07/2015	Tiziano Inzerilli (GARR) Roberto Cecchini (GARR)	Contributions on developments
V0.9	16/07/2015	Paolo de Lutiis (TIIT)	Contributions on developments
V1.0	30/07/2015	Michael Weirich (eco) Georg Roßrucker (eco)	Finalisation
V1.1	27/11/2015	Michael Weirich (eco) Georg Roßrucker (eco)	Improvements
V2.0	30/11/2015	Michael Weirich (eco) Georg Roßrucker (eco)	Formatting and layout, corrections

Executive Summary

In this document we provide an architectural overview to the Building Blocks and the interaction of the building blocks and other components within the ACDC context. A brief description of the Tool is delivered along with the implementation of the given tool by the providing partners.

A universal communication protocol is used to ensure communication between the different Tools and components on the one side and the Centralized Clearing House on the other side. we designed a simple protocol for exchanging data between tools through the Centralised Clearing House. This approach reduces the amount of configuration required throughout the solution, decouples the individual tools and their deployments and leaves a minimal attack surface while also being firewall friendly. The protocol uses message exchanges between a tool and the CCH to send and retrieve data and notifications for new or updated data sets.

Even though the initial concept defined having no limitations on data (format) submissions, it has been agreed on across the project participants, that a basic standardisation of the submitted data fields and basic requirement on mandatory fields simplifies the data submission and retrieval. These specifications have been defined as the ACDC - “Schemata”. These have been outlined and defined in the Deliverables D1.7.1/2 “Data Formats Specification”.

Acronyms used in the document

ACDC	Advanced Cyber Defence Centre
API	Application Programming Interface
CCH	Centralized Clearing House
CP	Community Platform
CERT	Computer Emergency Response Team
C&C	Command-and-Control
DAM	Data Access Manager / Part of the Community Portal
DB	Data Base
DDoS	Distributed Denial of Service
DNS	Domain Name System
DPI	Deep Packet Inspection
HTTP	Hyper Text Transfer Protocol
HTTPS	Hyper Text Transfer Protocol Secure
IAM	Identity and Access Management
JSON	JavaScript Object Notation
NSC	National Support Centre
REST	Representational State Transfer
SDN	Software Defined Network
SIEM	Security Information Event Management
SMTP	Simple Mail Transfer Protocol
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
STIX	Structured Threat Information eXpression
SW	Soft Ware
TBD	To Be Defined
UI	User Interface
WS	Web Service
XML	eXtensible Markup Language

Table of Content

List of Tables

List of Figures

1 The ACDC Approach

The objective of the ACDC project is to set up a European Advanced Cyber Defense Centre (ACDC) to mitigate the threat of botnets. ACDC's approach is to

- foster an extensive sharing of information across borders to improve the early detection of botnets
- provide an extensive set of ACDC Tools and Services accessible online for mitigating ongoing attacks
- use the pool of knowledge to create best practices that support organisations in raising their cyber-protection level
- create a European wide network of cyber-defence centres

The goal of the infrastructure is to provide users with solutions to fight botnets, and through data collection to build up an analysis capability of botnet occurrence and behaviour to also provide early detection of emerging botnets. ACDC therefore aims to improve prevention, detection and mitigation of botnets

Through this networked approach, ACDC paves the way for a consolidated approach to protect organisations from cyber-threats and support mitigation of on-going attacks through easy access to an increasing pool of ACDC Tools and Services.

The pillars for this model are clearly shown in D2.3 (Technology Development Framework):

The ACDC integrated process for fighting botnets is depicted in Figure 1 (DoW):

Figure 1: ACDC integrated process for fighting botnets

The ACDC Solution is the software and infrastructures that support the operation of the ACDC integrated process for fighting botnets. The ACDC Solution complies with the ACDC Model and thus is a set of different tools that interact only by sharing data through a Centralized Data Clearing House component, and working together in this way they contribute to the common objective of improving prevention, detection and mitigation of botnets.

ACDC has initially identified 5 types of services, with defined goals and oriented toward clearly identified target audiences:

Sensors and detection tools for networks

- Goal: detect malicious traffic

Systems infections/infected website analysis

- Goal: detect and analyse the malicious behaviour of infected websites and/or systems

- Target audience: owners of websites, SMEs, end-users

Device detection and mitigation - multipurpose tools for users

- Goal: detection of malicious activities in personal infrastructures and end-points (e.g. mobile phones)
- Target audience: SMEs, end-users

Information Sharing Platform / Central Clearing House (CCH)

- Goals: collect and analyse data from different data feeds, generate single EU common reporting picture
- Target audience: CERTs, International collaboration bodies, European agencies, industrial users, telecom operators, LEAs

Support centre

- Goals: provide help and support to infected users, based on the results of the detection through the CCH
- Target audience: owners of websites, SMEs, end-users

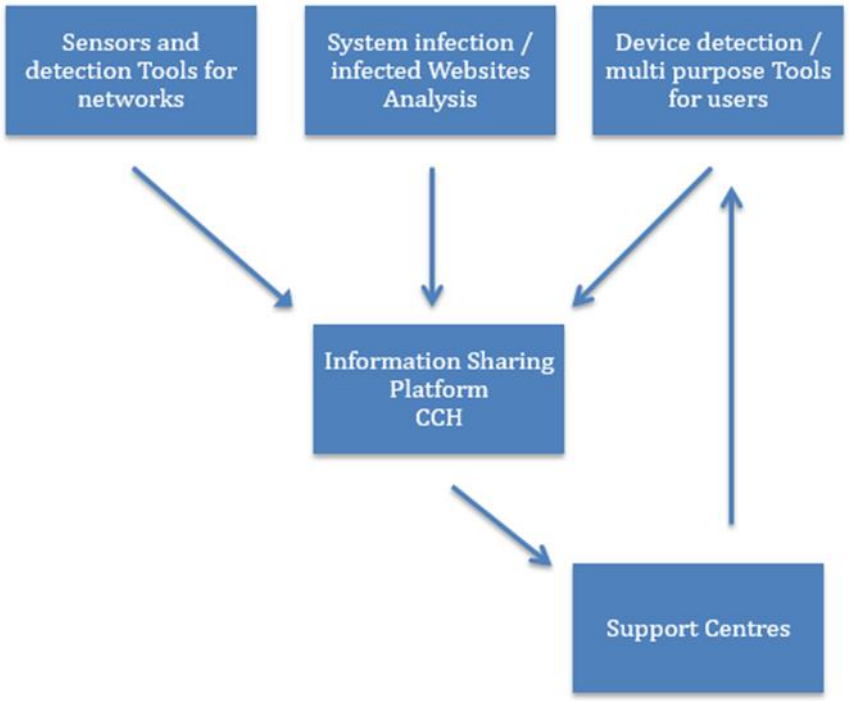

Figure 2:The ACDC Solution –overview

2 Classification of the tools

The ACDC domain tools, the tools that provide a functionality or service in the domain of ACDC, that is fighting botnets, can be subcategorized into smaller building blocks according to different criteria. We distinguish between those building blocks with the functional view of ACDC in mind that bases on how to contribute to the different phases of the ACDC integrated process on fighting botnets.

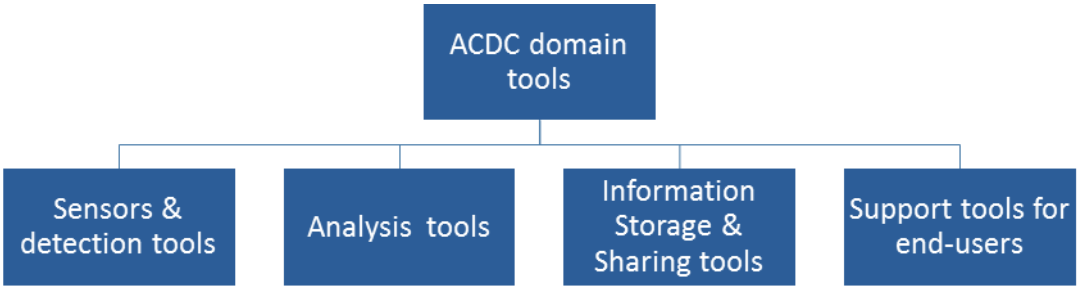

Figure 3: ACDC building blocks - high level classification (D2.3)

Each of this building blocks, or tool groups, consists of the contributions of the ACDC partners that are described in this document. A reference to the State of the Art Catalogue of Tools, developed or adapted for ACDC, can be found in the Community Platform, along with a short description of every available tool.

3 Building Blocks

In Work Package 1, Building Blocks – also called Tool Groups, are defined to specify the requirements and specifications of the different tool and their use within the ACDC Project.

Often a toolset provided by a partner consists of tools that are placed in several building blocks. This document gives an overview on the building blocks and high-level categorisation of the tools, along with their presentation under the providing partner's name.

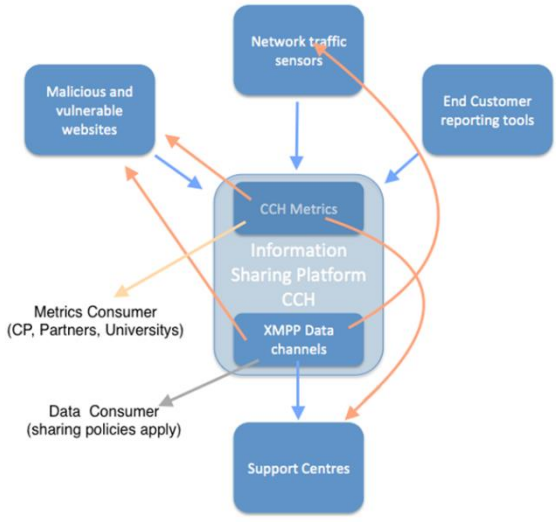

Figure 4: Building blocks and their interaction with the CCH

3.1 Building Block : “Centralized Data Clearing House (CCH)”

3.1.1 Description:

This Tool Group describes the functionality of the CCH, the Central Clearing House or centralized data clearing house.

The CCH plays the central role in the European Advanced Cyber Defence Centre as it provides the database to store incidents and botnet findings.

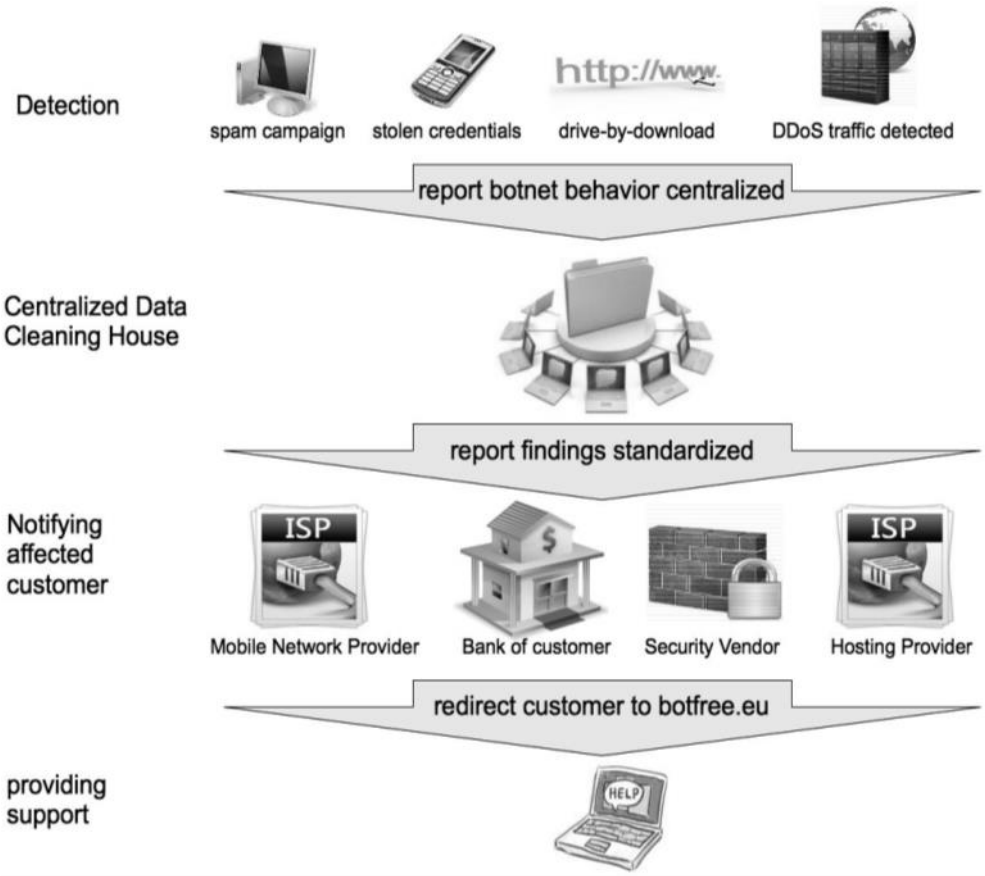

Figure 5: ACDC overview with the stages "detection->storage->reporting->Support"

Furthermore, Incidents and Reports have been provided to interested partners like ISPs, financial institutes, IT-Security vendors or research facilities.

3.1.2 Implementation

The Central Clearing House provides the database in which incidents and botnet findings are stored. Information on given IP or incidents can be enriched with additional information, originating from different sensors which provide a data feed to the CCH.

The Central Clearing House feeds a Redis database. Redis databases typically keep the whole dataset in memory. The in-memory nature of Redis allows it to perform extremely well compared to database systems that write every change to disk before considering a committed transaction. The Redis database keeps the data in memory for a defined timeframe. This setup enables an efficient performance of the database. Although Redis does not provide long-time storage of data, this functionality is handled by the MongoDB database, which is attached to the Redis cluster. This

database supports the storage of the data for long-term use like statistics, metrics or for entire botnet pictures.

Like the overlying Redis database, MongoDB can easily be clustered and its performance increases linearly whenever a new device is added. This is achieved with no downtime or interruption to applications or operation.

The ACDC database consists of the following infrastructure:

- Redis Database
- mongoDB database

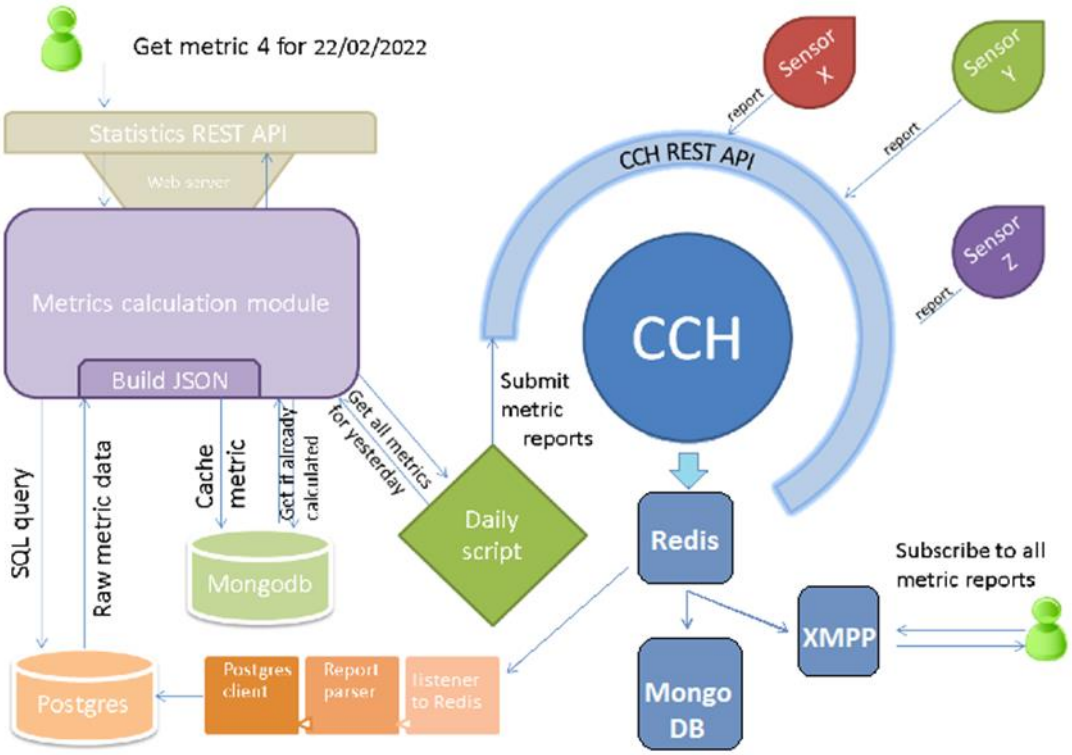

Figure 6: CCH with metrics module, databases and XMPP output

API keys handle access to the infrastructure. These keys manage the access of sensors, tools or users to the Database and are maintained and distributed by the ACDC Community Portal. The Community Portal is the first point of contact for every entity which is interested in sharing data with the Central Clearing House and the participants. The Community Portal manages the access control policies and the relationship between the ACDC Partners.

The API concept is based on a hierarchical model of a three-stage API-Key distribution (CCH-Manager/CCH-Key-Manager/CCH-Key-User). Besides CCH-Manager-Keys, the Community Portal is able to assign individual keys to subscribers in the ACDC community. With such a key, each partner (CCH Key Manager) is capable of creating and assigning individual API-Keys (CCH Key User) to their own sensors, nodes or even sensors running on end-customer devices through the Community Portal, where each key is registered.

Based on the defined settings in the Community Portal, a limitation of the keys provided per organisations can be set, as can the expiry date or the revocation of a key.

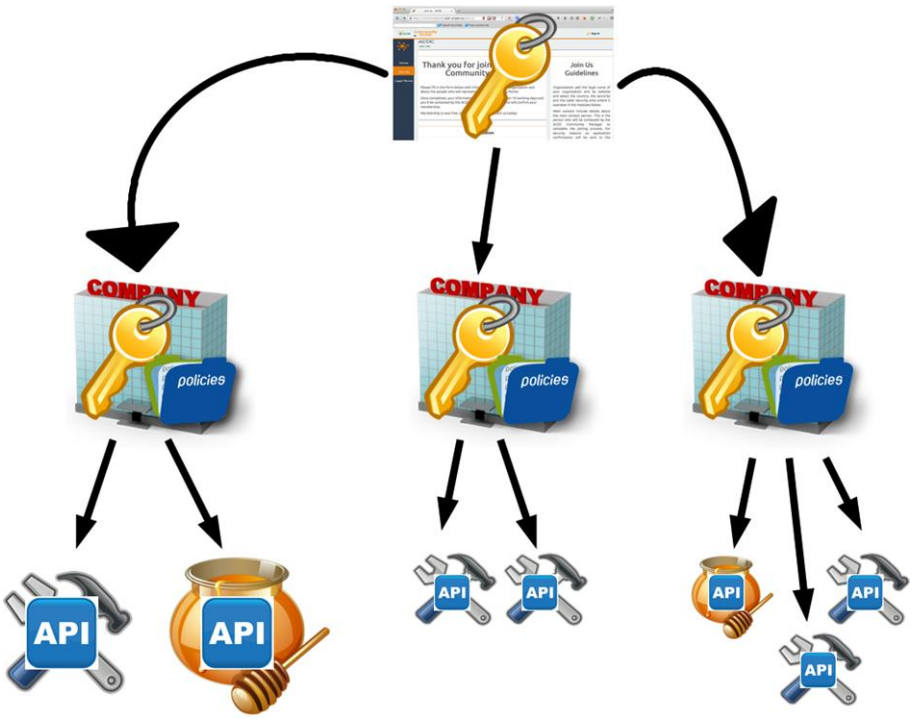

Figure 7: API -Key generation process

Further information on the Key Management is part of deliverable **D6.2.1/2 ACDC Social Platform** and can also be found on the deliverable **D 1.2.2 Centralized Clearing House**.

3.2 Building Block: Support Centres

3.2.1 Description

Deliverable D1.3.2 describes the specifications for the Tool Group "Support Centres". The support centres are the first point of contact for victims of cybercrime and the main resource of information and knowledge towards prevention, awareness and dissemination of infected electronic devices. The support centres represent the initiative to the broad public by interacting directly with end-users and the project. During the project span of ACDC, 10 National Support Centres (NSC) have been launched. These will continue with their operation after the project end under the umbrella of European Anti-Botnet Support Centre Alliance.

End-users are redirected by their ISP or other organisations to an NSC and are provided assistance, support and free tools in case of malware detection in their infrastructure.

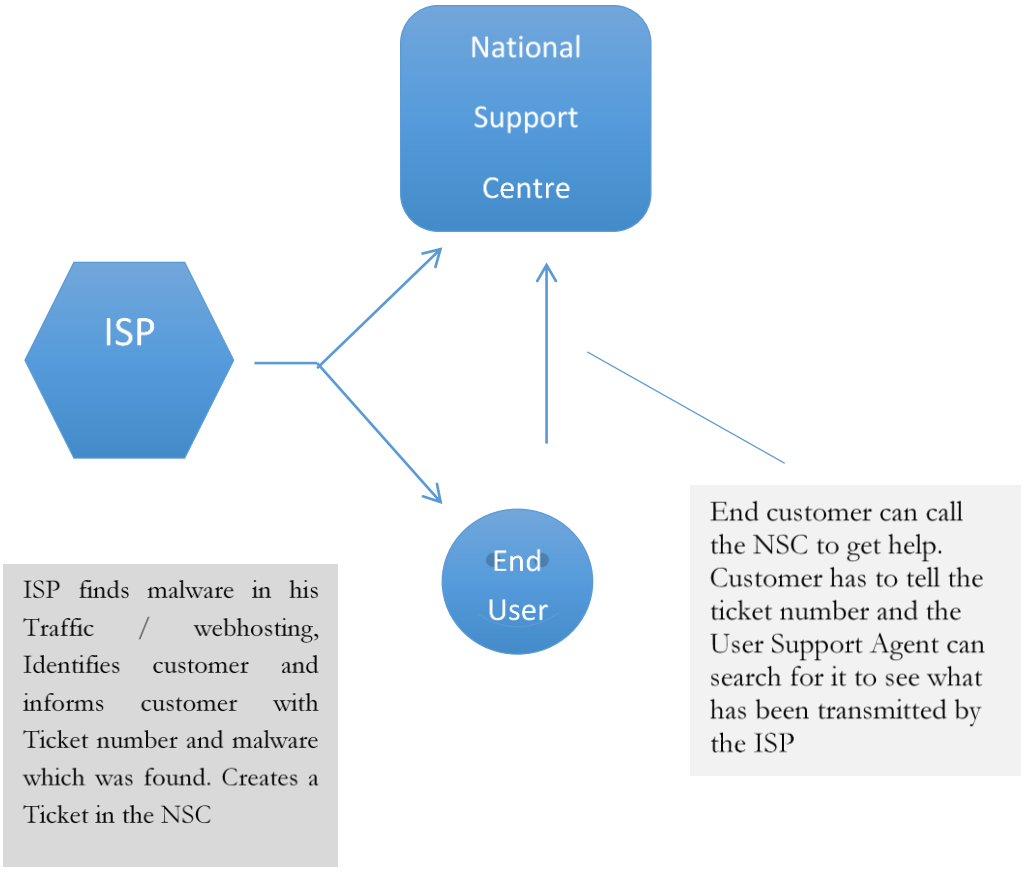

Figure 8: Workflow of a National Support Centre

3.2.2 Implementation Examples

Anti Botnet Advisory Centre - Germany

The German National Support Centre was launched in 2010. It provides various levels of service and can be seen as a pilot model for other national support centres. ECO provides the following services within the German Support Centre:

The "Anti-Botnet-Beratungszentrum" named "Botfrei" is the German national support centre, consisting of a website and a user help desk with phone and email-support. The service includes a forum, a blog and links to activities on social networks. The webpage includes

additional pages with information on given threats and it offers free removal tools like the "EU-cleaners" in collaboration with anti-virus vendors. The blog is powered by a WordPress installation, which is enhanced by CMSSECURITY – a self-developed security solution. End users / customers can contact the "Botfrei" experts by phone or email; the customer handling is supported by an OTRS Ticket-system.

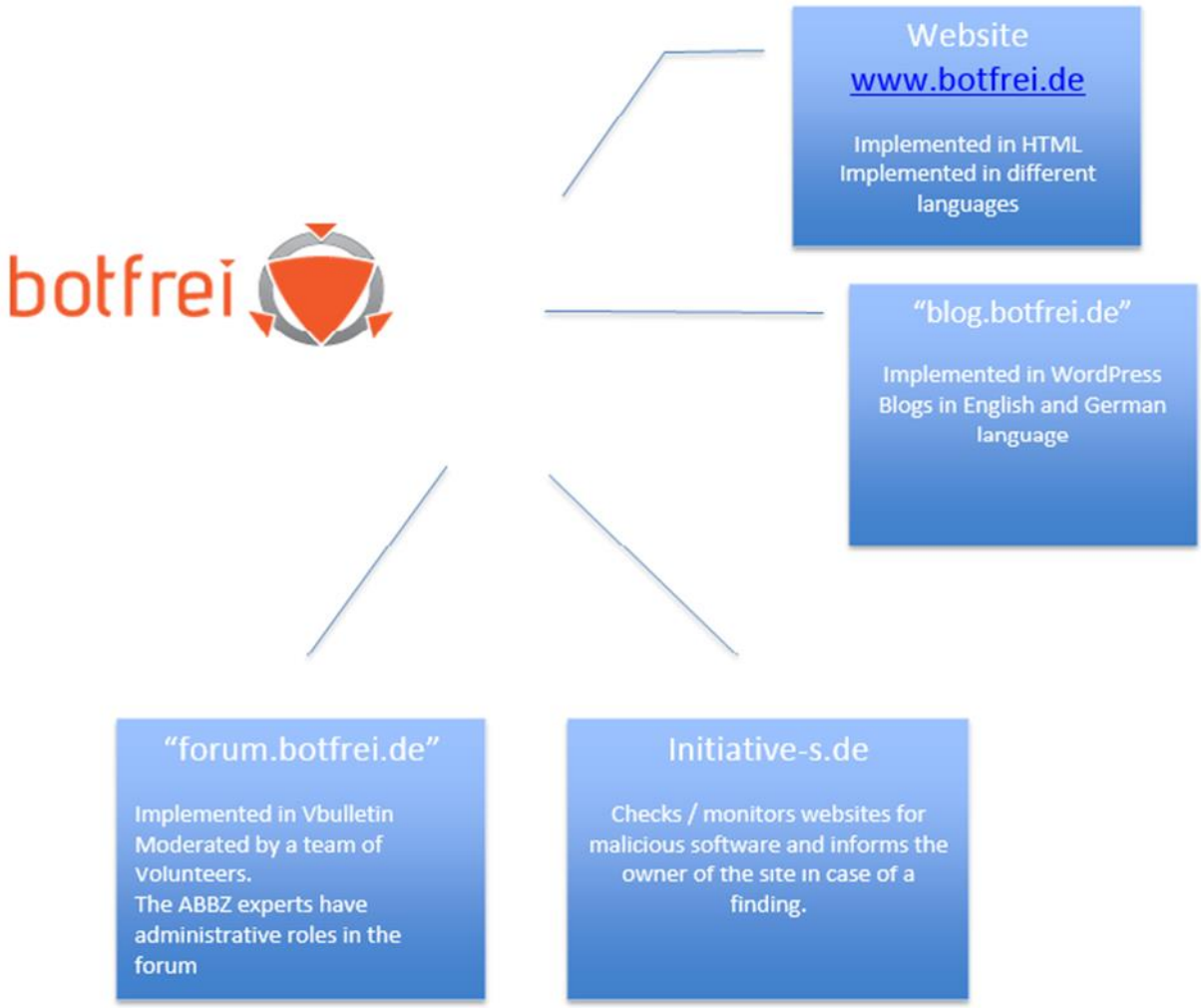

Figure 9: Components of the German NSC (botfrei.de)

Software in the German NSC:

- WordPress for the Website "Botfrei.de" and the Blog "blog.Botfrei.de" are managed internally by eco. The statistics are tracked by the analytics system PIWIK
- Vbulletin for the forum on "forum.Botfrei.de"
- OTRS is a ticketing system in which ISPs and other partners can deliver tickets for end-customers.
- Mail services to get in contact with the users and feed tickets and incidents to the OTRS Ticket system

Italian NSC Workflow Example:

GARR shows the integration of the Italian NSC within the ACDC project and their HORGA Sensor system.

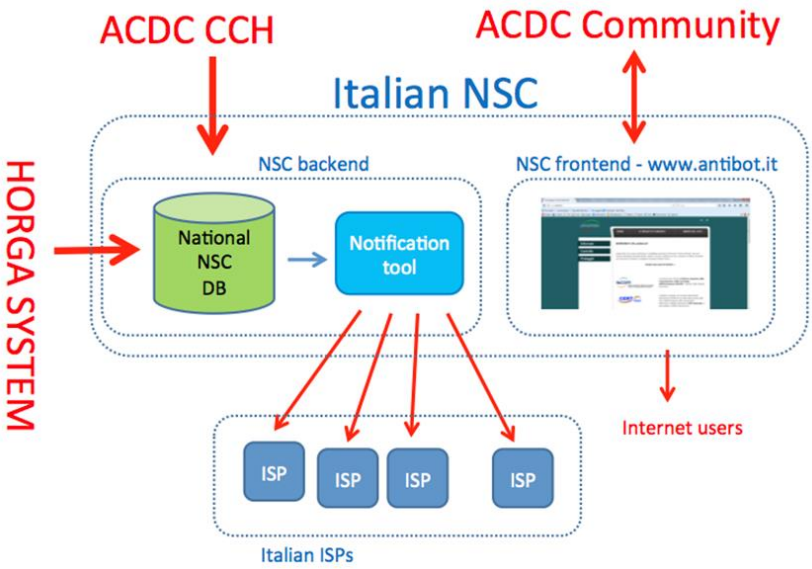

Figure 10: Italian NSC, operating in the ACDC environment

Spanish NSC, Workflow Example

The ACDC partner INCIBE has a long tradition on supporting end-users and running a CERT, this picture shows the integration of the Spanish NSC into INCIBE´s toolset and the ACDC infrastructure.

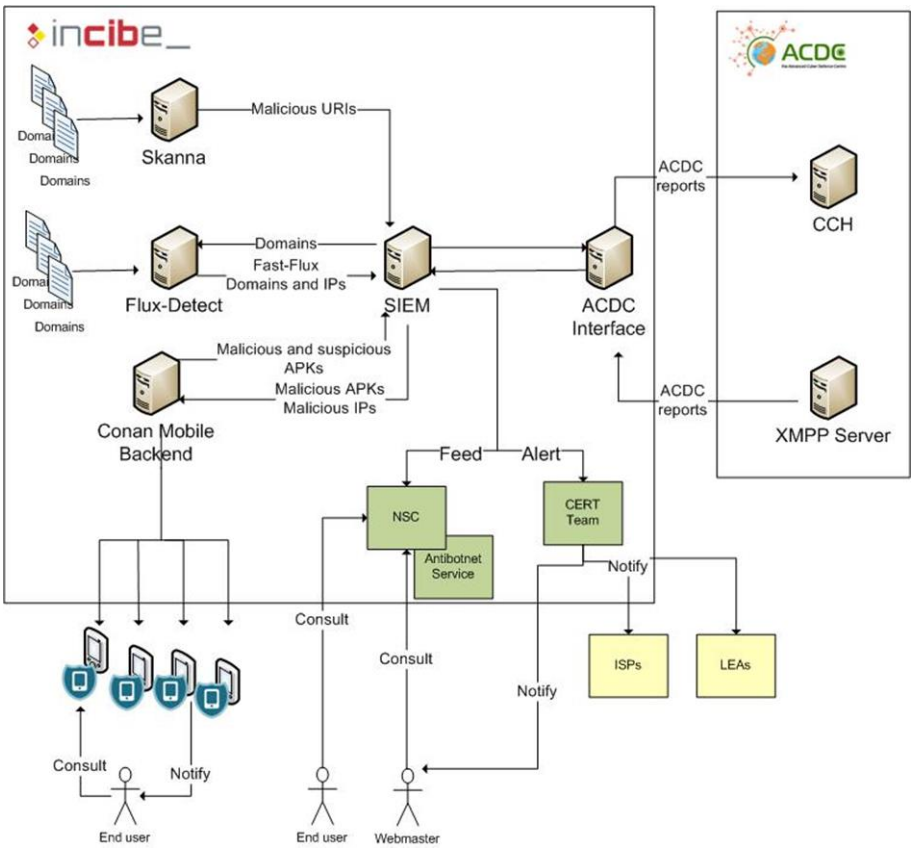

Figure 11:INCIBE - overview on Toolset and ACDC integration

Active Support Centres

Country	Operator	Website
Germany	Eco e.V.	http://www.botfrei.de
Belgium	LSEC	http://www.botvrij.be
Spain	INCIBE	http://www.osi.es/servicio-antibotnet
Croatia	CARNet	http://antibot.hr
Romania	CERT-RO	http://www.botfree.ro
Italy	ISCTI	http://www.antibot.it
France	CECyF	http://www.antibot.fr
Portugal	FCN/FCCT	http://www.antibot.pt
Luxembourg	CIRCL	http://www.botfree.lu
Bulgaria	CERT.bg	http://www.antibot.bg

Table 1: active National Support Centres in the ACDC project

Deliverable 1.3.2 covers the National Support Centres and provides further information on the implementation and operation of the NSC.

3.3 Building Block: Malicious or Vulnerable Websites

A detailed description of this Tool Group, its components and definitions are further described in the deliverable D1.4.2. The following examples should provide a basic overview of some of the tools for the detection of malicious or vulnerable websites.

3.3.1 Eco / Initiative-s

Initiative-S offers early warning protection for the website as well as a confidence building seal indicating that this site is constantly monitored and checked for malicious activity.

The services that are visible to the public are described on the project's webpage:

https://www.initiative-s.de/en/index.html.

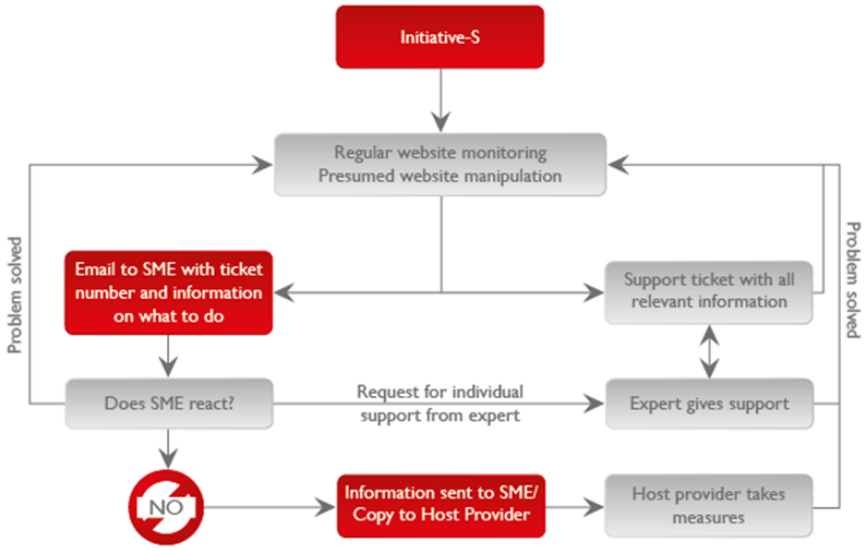

Figure 12 : Initiative-s - Workflow and intergation in the german NSC

Initiative-S is hosted by eco, source code and documentation is not available, as this tool is provided to ACDC as a service and is proprietary of eco. It is deployed and operated internally by eco.

3.3.2 ATOS Software components integrated in ACDC

Atos has deployed and integrated four software components into the ACDC infrastructure:

Sensors

- DNS traffic sensor and analysis tool
- Netflow traffic sensors and behaviour analysis tools for botnet detection

Analysis and correlation tools

- SL-SIEM (correlator)
- HPPD service (phishing analysis)

3.3.2.1 DNS Traffic sensor and analysis Tools

3.3.2.1.1 Overview

The DNS traffic analysis component consists of a set of modules that analyse the DNS traffic of a monitored network looking for certain patterns and features that lead to identification of domains and IPs that could potentially belong to a fast-flux network, used to support botnet activities and DDoS attacks.

The fast-flux features analysed are grouped into four analysis groups:

- Time-based analysis: statistical analysis looks for certain patterns regarding timestamp of the different queries and responses to the servers.
- TTL-based analysis: looks for suspicious values and behaviour regarding the TTL (Time to Live) assigned to domains queried.
- Domain name-based analysis: based on the use of blacklists, whitelists and machine learning techniques, looks for suspicious domain names in the traffic.
- DNS answer-based analysis: looks for suspicious patterns in the response returned by the DNS in terms of number and variation of the response IPs, geographical distribution of response IPs and reverse DNS lookup analysis to detect use of bogon IPs.

Besides the identification of fast-flux network behaviour in the DNS traffic, the sensor also detects amplification DDoS attacks of type DNS amplification. By analysing the DNS traffic captured within the monitored network, looking for UDP packets (DNS requests sent to the monitored DNS servers) with specific characteristics:

- much larger response than query
- use of ANY in the DNS query
- DNS query source IPs from outside the monitored network (suspicious of being spoofed IPs)
- volume of DNS requests

3.3.2.1.2 Integration with ACDC: Architectural view

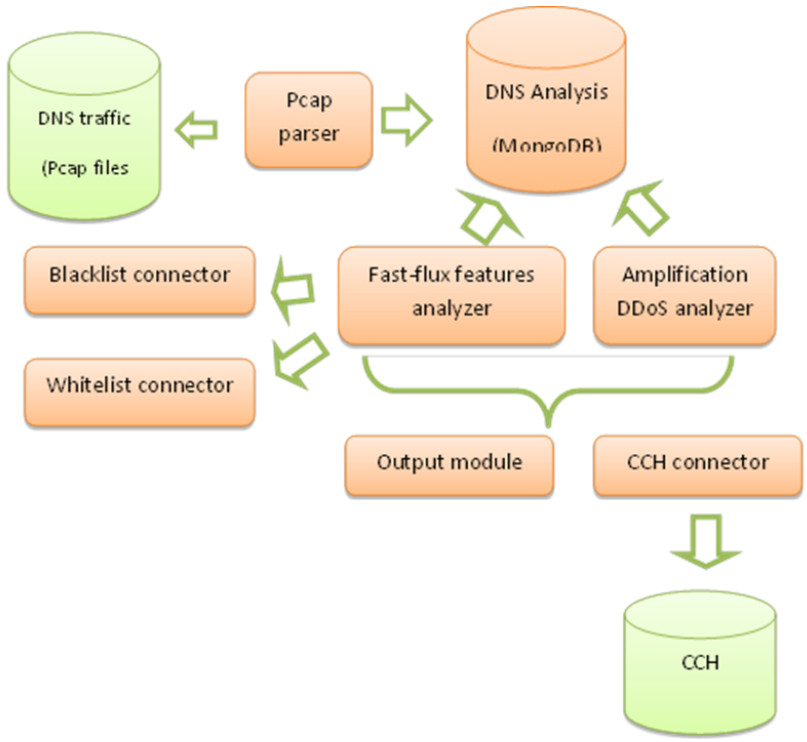

Figure 13: Integration and interactions of the ATOS Toolset in ACDC

This component requires continuous monitoring of DNS traffic for long periods of time (especially for the time-based and DNS-answer based analysis) to achieve acceptable results of accuracy. Also the tool is considered a proof of concept and could be assessed to be between TRL 3 and 4, thus, the reports sent to the CCH of category eu.acdc.fast_flux and eu.acdc.bot (subcategory fast_flux) have confidence level of 0.5. The value of the information reported is to be used for further analysis by other tools participating in the experiments and in correlation processes, for instance in rules that take into account multiple low-to-medium confidence level reports of the same bot IP, fast-flux domain or attacker IP reported by different tools within a specific time-frame.

The sensor uses different sources of input data, besides the raw DNS traffic data, such as known online blacklists (e.g. Google Safe Browsing for known Malware and Phishing sites) and whitelists (e.g. Alexa Top 1000 sites), but also uses the reports provided by other WP3 experiments participants, through the CCH XMPP channel service. The reports used by the sensor and obtained from the CCH are of category eu.acdc.fast_flux, eu.acdc.malicious_uri confirmed domains. These reports contain information about confirmed malicious fast-flux DNS servers and domains, and malicious websites, and these are used to enhance the analysis algorithms (in the same way other blacklists are used).

3.3.2.2 Service-Level-SIEM (SL-SIEM)_AHPHS

3.3.2.2.1 Overview

The Atos High Performance Security (AHPS) is a commercial service provided by Atos that is composed of many security services. Atos cannot provide the source code nor the executables of the services composing the AHPS to ACDC, but only a commercial agreement for running and operating the tool (in-house) during a pre-defined period of time and under a pre-defined environment conditions. However, there is a version of the SIEM part of the AHPS service offering, called "Atos Service Level-SIEM" (SL-SIEM), which we use for research in the Atos Lab and that contributed to the FI-WARE project

Security Monitoring Generic Enabler. This version is based on an open source SIEM that Atos has extended with various plug-ins and modules to add further correlation capabilities and make it more scalable, and to adapt to ACDC model. The limitations of current SIEM systems are mainly related with performance and scalability, leading to the inability to process vast amounts of diverse data in a short amount of time. The next generation of SIEM solutions should overcome the performance limitations of its predecessors, allowing in this way the monitoring of more systems, the processing of more complex rules or even the correlation of events at different layers. To achieve the above goals, the Service Level SIEM (SLS) component delivered by Atos for the implementation of the Service Level SIEM component included in the FI-WARE project Security Monitoring GE integrates a standard distribution of the open source OSSIM SIEM with other processes running in a topology deployed in a Storm cluster to provide high-performance correlation capabilities.

The Service Level SIEM server is the component which receives the events already normalized, coming from the different slave nodes (i.e. Security Probes), stores them in a MySQL database and performs its correlation to show the user through the Security Dashboard interface the relevant security events and incidents detected.

The Security Probes are SIEM agents that can be distributed and installed remotely to collect events generated by different data sources at the monitored target infrastructure, depending on the plugins activated (e.g. Nagios, snort, syslog, STIX, etc.), and that send those events to the SL-SIEM server.

More detailed information about the Service Level SIEM architecture can be found in the FI-WARE Service Level SIEM Open Specifications

To deal with the high flow and frequency of reports received in the SL-SIEM component, the following improvements have been implemented:

Instead of a single-node configuration for the SL-SIEM component, it has been distributed in a Storm cluster with three nodes:

- a master node: running the Storm Nimbus process as well as the OSSIM server and database
- two worker nodes: running the Storm Supervisor processes to do the processing and correlation of the incoming reports

Usage of a limited and configurable pool of connections to the database in the pro-cesses running in the Storm cluster to avoid an overflow in the number of database connections.

Capability to generate CCH reports also with the added directives of second level (re-ports with a higher confidence level e.g. because a certain directive has been detected a predefined number of times)

The SL-SIEM perform continuous correlation of events sent to the CCH by other participants in the experiments and received into our XMPP channel, as per the data sharing policies put in place in the Community Portal DAM.

The CCH reports received are input into the correlation engine and evaluated against some correlation directives designed for the purpose of each of the experiments.

These directives evaluate the occurrence of certain conditions in the reports received in or-der to:

1. Increase the quality (i.e. the confidence level) of the information in the CCH
2. Assert new (or confirm the existing) knowledge derived from the information al-ready stored in the CCH

The following cases are evaluated in each of the WP3 experiments:

DDoS:

- Multiple reports of attacks with low to medium confidence level and the same origin IPs in a predefined period of time.

  Reported by the same tool
  Reported by different tools

- Distributed DDoS attack based on multiple DDoS attacks of the same type (with the same target and different origin IPs) in a predefined period of time.

  A suspicious bot (not confirmed) which is the origin of a confirmed at-tack

Fast-flux:

- Multiple reports of fast-flux domains with low to medium confidence level in a de-fined period of time.

  Reported by the same tool
  Reported by different tools

- Multiple reports of fast flux bots with low to medium confidence level in a defined period of time.

  Reported by the same tool
  Reported by different tools

- FF bots used in other malicious activities such as attacks, communication with C&C servers, hosting malware distribution sites, hosting malicious websites, etc.

- FF domains that are used for other malicious activities such as spam campaigns, malware distribution sites, malicious websites, as DNS of a confirmed C&C server.

Mobile:

- Multiple DDoS attack reports with low-to-medium confidence level performed by confirmed mobile bots

  Reported by the same tool
  Reported by different tools

- Multiple malware attack reports with low-to-medium confidence level performed by confirmed mobile bots

Reported by the same tool
Reported by different tools

- Multiple reports with low-to-medium confidence level with the same C&C server used by mobile bots

Reported by the same tool
Reported by different tools

Spam:

- Multiple reports of spam campaigns with the same URIs with low to medium confidence level in a defined period of time.

Reported by the same tool
Reported by different tools

- Multiple reports of spam bots with the same URIs with low to medium confidence level in a defined period of time.

Reported by the same tool
Reported by different tools

- Suspicious URI used in spam confirmed as malicious URI
- Suspicious spambot detected in confirmed abuse attack

Websites:

- Multiple reports of malicious URIs with low-to-medium confidence level in a defined period of time.

Reported by the same tool
Reported by different tools

- Multiple reports of vulnerable URIs with low-to-medium confidence level in a defined period of time.

Reported by the same tool
Reported by different tools

- Multiple reports of malware with low-to-medium confidence level in a defined period of time.

Reported by the same tool
Reported by different tools

- Suspicious website involved in confirmed attack of type Abuse and Spam
- Suspicious website supporting a confirmed attack of type Compromise
- Suspicious website involved in confirmed malware attack

3.3.2.2.2 Integration with ACDC: Architectural view

The SL-SIEM has been adapted to the ACDC model in order to work with the CCH (sending and receiving data reports, using the CCH REST API and the XMPP service respectively) and with the STIX aggregator.

In order to do that, various plugins have been developed:

- STIX plugin: The ACDC STIX plugin is a component developed by Atos in the context of WP2 that allows feeding the Atos Service Level SIEM (SL-SIEM) with the cyber-threat observations aggregated by the STIX aggregator component. The cyber-threat observations, produced by the different sensor and analyser components in ACDC, are represented using the STIX format (stix.mitre.org) and aggregated by the STIX aggregator. In order to be used by the Atos SL-SIEM, it is necessary that these cyber-threat observations are ex-pressed in the SIEM normalized event format. The ACDC STIX plugin parses the STIX data and generates its representation in the SIEM normalized event format. Once in the correct format, the events are fed into the Atos SL-SIEM.
    - CCH plugin: The ACDC CCH plugin is a component developed by Atos to collect data from the ACDC CCH component, parse it and generate one (or more) Atos SL-SIEM normalized event(s). The events are fed into the SL-SIEM for processing and correlation. The ACDC CCH plugin is composed of various modules:
    - CCH XMPP client: this module connects to the CCH XMPP channel associated to the SL-SIEM read key created for this purpose by means of the Data Access Management (DAM) service of the ACDC Community Portal. The ACDC DAM service permits the creation of data sharing policies associated to the SLS read key, regulating the access of the SLS to the data sent to the CCH by other members of the ACDC consortium. The CCH XMPP client allows the SLS to receive messages (reports) from the CCH, encapsulated in JSON data format as described in deliverable D1.7.2.
    - CCH parser: this module parses the CCH JSON reports, and translates them into SLS event format. There are various data formats supported by the CCH (as it is defined in deliverable D1.7) and therefore the parse must implement all the supported formats.
    - CCH REST API connector: The SLS correlator engine raises SL-SIEM correlation alarms whenever a rule matches a series of SL-SIEM event occurrences. SLS alarms have the same data format as SLS events and actually are fed into the SLS server for consideration and further analysis. The ACDC CCH connector script is used by the SL-SIEM to submit alarms raised by the SLSIEM server as a result of a correlation process to the ACDC CCH. Not all the alarms are submitted to the CCH, because some of the alarms are used internally for second-level correlation, or for other internal purposes (i.e. displaying in the SL-SIEM graphic dashboard, sending emails to system/network administrator, statistical purposes, etc.). Defined policies and actions determine which alarm types must be submitted and under which circumstances.

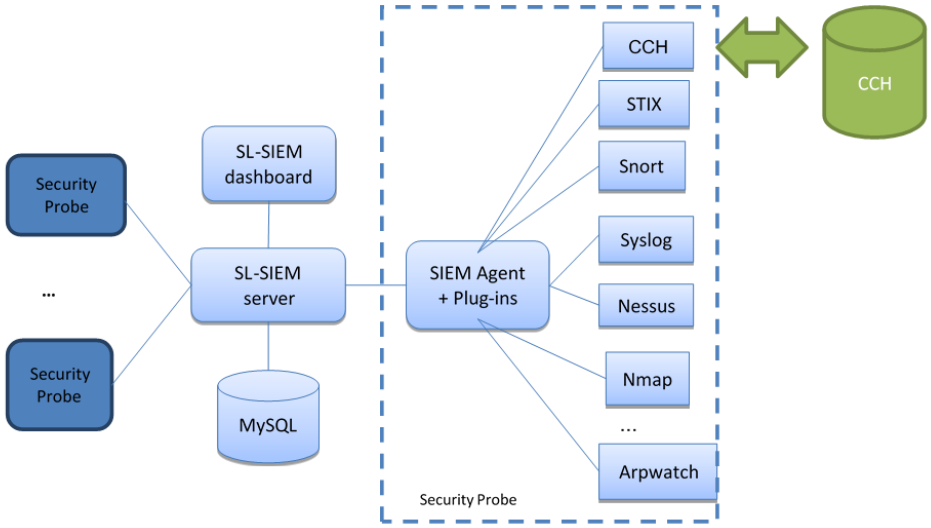

Figure 14: ATOS - SL_SIEM in the ACDC environment

3.3.2.3 High Precision Phishing Detection Service

3.3.2.3.1 Overview

The High Precision Phishing Detection module (HPPD) is a machine learning based module that performs comparison of certain features of host names with known phishing and benign web sites. It is a stand-alone, fully automated module capable of real-time/stream learning, meant to serve as a high-performance pre-filter placed before other modules that require larger time frames to perform analyses and classification. The phishing analysis service, called High Precision Phishing Detection (HPPD for short), was developed in the context of project NECOMA and contributes to the AHPS offering with the analysis of phishing websites. In ACDC, the HPPD module is used through the REST API service offered with two purposes, and the input from ACDC is different in each case:

- Training: the confirmed malicious_URI reports of subcategory phishing received by the SLS component through the CCH XMPP channel are used as another source of train data for the HPPD.
- Analysis: The SL-SIEM component invokes the analysis API with suspicious URIs with low to medium confidence levels received from the CCH by means of the XMPP client connection. The HPPD analysis service classifies the URIs into malicious or legitimate with a probability value assigned to them

3.3.2.3.2 Implementation

For the participation in the WP3 experiment WEBSITE, the service is deployed in the Atos environment, as part of the deployment of the SL-SIEM. The invocation of the HPPD REST API for analysis with suspicious URIs received from the CCH through the XMPP channel produces a classification of the URI into malicious or legitimate. This classification has a probability value assigned which determines the reliability of the results, and thus is used to derive the confidence level of the report sent to the CCH. No report will be sent to the CCH unless the probability is higher than 80%.

The HPPD component requires continuous training for enhancing the classification capabilities of the machine-learning algorithm. This module relies on two external datasets: Alexa and PhishTank . These are used as training data sets and provide patterns of features for the module, where Alexa serves as

a benign web sites source and Phish Tank determines malicious patterns. As another source of training data, the HPPD module uses the CCH re-ports of malicious URIs subcategory phishing, which have a confidence level of 1.0 (i.e. con-firmed).

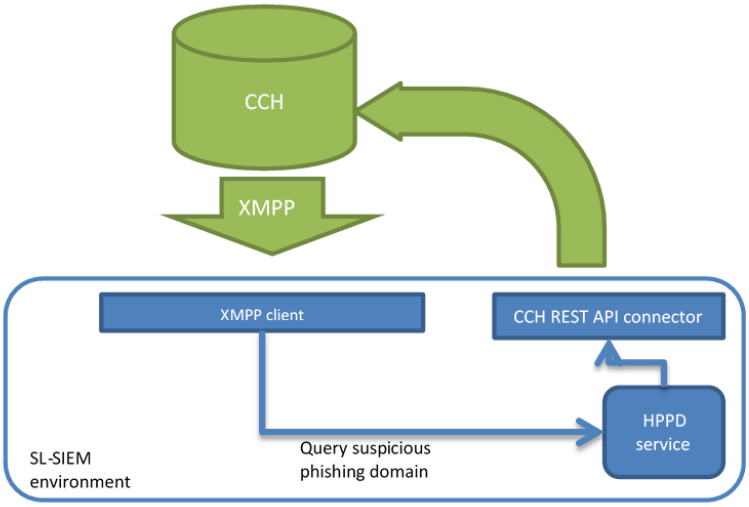

Figure 15: SL_SIEM connection to the CCH

3.3.2.4 Network traffic sensors and behaviour analysis tools for botnet detection

3.3.2.4.1 Overview

This type of sensor analyses, primarily, Netflow traffic data generated by routing and switching devices that are Netflow-capable (e.g. CISCO, Adtran, NEC, etc). But software capture tools, such as softflow or nProbe, are also able to sniff the network traffic and produce an output in Netflow format that can be analysed by these sensors. The analysis of Netflow data aims at identifying botnets by discovering anomalous behaviour in the network traffic. These observations may lead, for instance, to the identification of the hosts in the network that are part of a botnet, but also to the identification of a compromised network device and the C&C server that is sending commands to it. Botnets detected by these sensors normally compromise a vulnerable router or switch device (usually not properly configured), giving the C&C server control over the network to recruit all the hosts in the corresponding subnet to per-form malicious activities. An example of this type of botnet is the Chuck Norris botnet. The sensor analyses the Netflow traffic based on the methods described in the literature. Other botnet types can be detected by observing http headers in the Netflow data, allowing the identification of malware distribution content web servers. The analysis of Netflow data over a period of time can be used for the identification of clusters of hosts with unusually high rates of inter-connections that simulate the behaviour of regular peer-to-peer networks but are actually an active botnet in disguise. When botnets are in an idle state, they tend to reflect P2P behaviour as they try to infect the rest of the computers in the network and wait instructions from the Command and Control server (C&C). The sensor analyses the Netflow traffic to look for suspicious clusters of IPs (i.e. suspicious P2P botnets) based on the work described in the paper by J. François et al. , where a method is proposed to search for these malicious P2P connection patterns, based on a mathematical calculation to find clusters of IPs of highly dense traffic between a small number of computers.

The tools are considered a proof of concept and could be assessed to be between TRL 3 and 4.

3.3.2.4.2 Implementation:

For the participation in WP3 experiments, the sensor has been deployed in the Atos research network environment. The sensor captures Netflow traffic and analyses it to look for suspicious patterns that may lead to the identification of hosts in the corporate network that may belong to a botnet. Although the sensor does not send any report to the CCH, the sensor retrieves reports from the CCH about C&C servers to look for connections to these servers in the monitored traffic, supporting the analysis algorithms (as a blacklist of C&C servers). The output of the analysis, that is clusters of IPs in the Atos network suspected of being part of a botnet, are not reported to the CCH because the IPs belong to the Atos corporate network and are internal IP addresses of no use outside Atos environment. However, the results of the analysis are very useful for Atos system and network administrators to ban the detected connections and to perform mitigation actions such as removing them from the corporate network and analysing the affected hosts looking for any virus or malware installed or assessing potential existing vulnerabilities, for example.

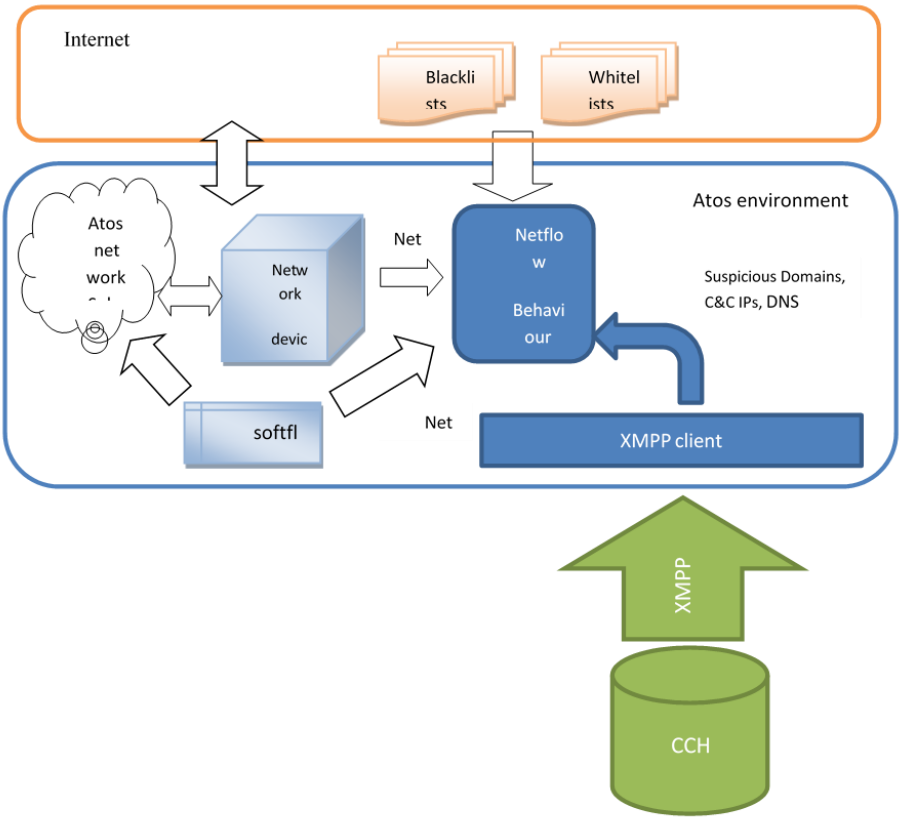

Figure 16: receiving data from the CCH

3.3.3 Fraunhofer FKIE components in ACDC

3.3.3.1 HoneyUnit

The HoneyUnit, provided by Fraunhofer FKIE, is a generic security tool for analysing the runtime behaviour of web pages and SVG documents. It can detect attempts to exploit the client's web browser both through static signature matching as well as through dynamic tests analysing the runtime behaviour of a simulated web browser rendering the web page and a JavaScript engine executing JavaScript code embedded in that page.

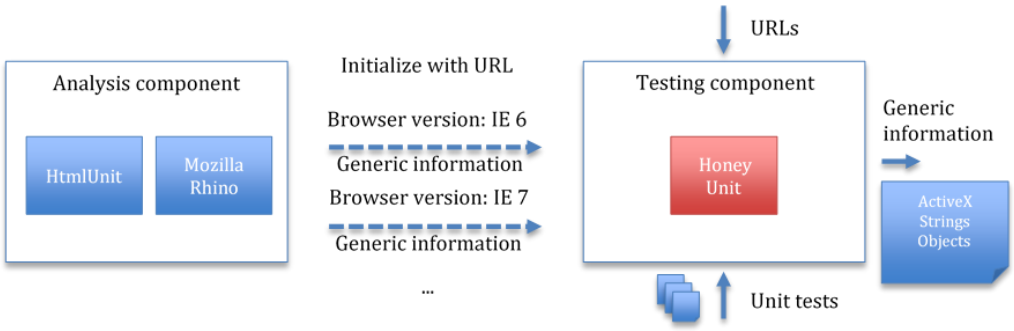

Figure 17: Overview to the HoneyUnit analysis process

The HoneyUnit does not receive any data from the ACDC solution. Since it is designed as a library, it can be wrapped in a solution that retrieves URLs from other parts of the ACDC solution, e.g. the Centralised Clearing House.

3.3.3.2 HoneyAgent

The HoneyAgent, provided by Fraunhofer FKIE, performs dynamic analysis to detect attempts of malicious Java applets to break out of the Java sandbox. It combines a set of dynamically applied signatures as well as heuristics that can detect successful exploitation attempts using unknown vulnerabilities in the Java VM. In most of these cases, the additional sandboxing layer provided by the HoneyAgent protects the host system against further damage. The sandboxing layer can emulate the effect of some vulnerabilities, allowing the malicious code to continue execution and download additional stages of the attack, which the HoneyAgent can submit to the CCH for further analysis.

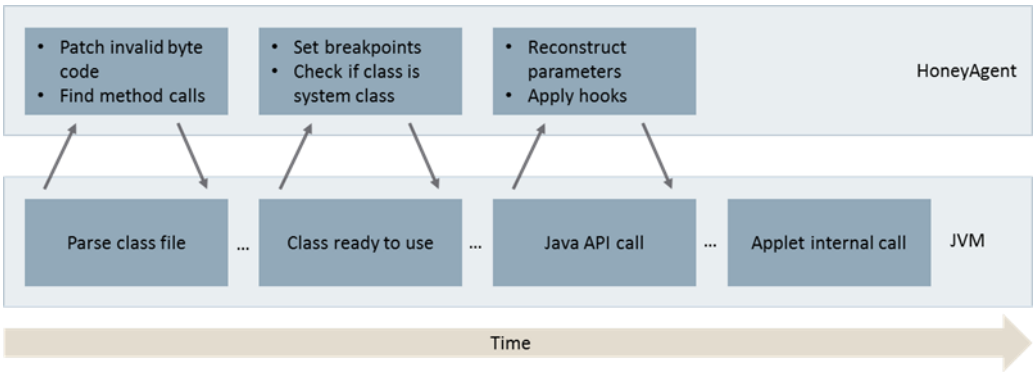

Figure 18: overview on the HoneyAgent analysis process

3.3.3.3 PDF Scrutinizer

Content delivered by malicious websites is not necessarily limited to HTML documents and JavaScript. Such websites can also deliver malicious PDF documents in order to attack the victim's browser or its PDF rendering plugin.

The PDF Scrutinizer, provided by Fraunhofer FKIE, dynamically analyses the content of a PDF document in order to identify malicious patterns or behaviour in it. This tool can be used for example in conjunction with the HoneyUnit to expand its detection capabilities allowing a more comprehensive and thus more accurate analysis of malicious websites.

The PDF Scrutinizer does not receive data from the ACDC solution. However, since it is de-signed as library, it can be wrapped in a tool that retrieves URLs from another part of the solution, e.g. the Centralised Clearing House.

3.3.4 Site Vet Webservice by Cyberdefcon

SiteVet is a web service that provides data on malicious activity hosted worldwide. Data is combined from multiple sources – community partners as well as CyberDefcon's own re-search data – and processed using unique algorithms to provide meaningful results. The focus is on Autonomous Systems and the "reputation" of hosts.

SiteVet is configured to retrieve data from the ACDC Central Clearing House that will aid in the reputational analysis of Autonomous Systems, IP addresses, and domain names.

Specifically, SiteVet can receive the following data types from the CCH:

- eu.acdc.bot
- eu.acdc.c2_server
- eu.acdc.fast_flux
- eu.acdc.malicious_uri
- eu.acdc.vulnerable_uri

The data is used from these sources to complement data from partners and CyberDefcon's research data.

3.3.5 WebCheck by Cyberdefcon

WebCheck is a server plugin for webmasters that identifies and remediates malware and vulnerabilities hosted from the server. It focuses both on cleaning websites and on forging trust by guaranteeing a website is safe.

WebCheck utilises data from CyberDefcon's other ACDC tool, SiteVet. The data provided includes lists of malware, badware, botnets, spam and vulnerabilities. These lists are in the form of blacklisting, recorded instances and static signatures. Also it retrieves data through the SiteVet tool in the acdc.eu.malicious_uri category.

WebCheck processing is carried out by both the centralised WebCheck server and the local server. If a full installation has not been carried out (such as when used as a trial or in situations where a local installation is not possible e.g. due to lack of administrative access), the processing occurs only by the centralised server.

The centralised server is limited to analysing websites externally and checking this information against data retrieved from the external sources (SiteVet and the ACDC Central Clearing House). This external analysis can occur in two ways: the first is from a quick scan (e.g. when a new user wishes to check their website immediately), the second is from a full scan, which is intended for full customers and requires a site crawl to have occurred. The resulting information to the end user includes blacklisting data, reputational data, basic vulnerabilities and basic performance issues.

The local installation also has access to the file system and configuration files, which enables the tool to provide more advanced vulnerability information, malware information, and advanced performance issues, as well as facilities to clean and mitigate against vulnerabilities and malware.

For the purposes of ACDC, WebCheck aims to discover malicious "intermediary" URLs that are not hosted from the end user's website. Primarily, this occurs as a result of the local scanning

functionalities of WebCheck discovering some malware or vulnerabilities that call back to a malicious third-party URL.

3.3.6 Skanna – INCIBE

Skanna is a system that, for a given set of domains, analyses websites served under that do-main in order to create an inventory of technologies used. It uses sandboxing, dynamic and static analysis to identify whether a given website has been compromised and whether it engages in any malicious activities. It discovers potentially vulnerable websites by comparing the software used against an inventory based on well-known vulnerability databases. This contributes to a faster discovery of possibly compromised websites due to the exploitation of known vulnerabilities.

Skanna provides reports to the CCH indicating that a given domain or URI is considered malicious or suspicious (with a high likelihood of being malicious) and the reasons for reaching that conclusion.

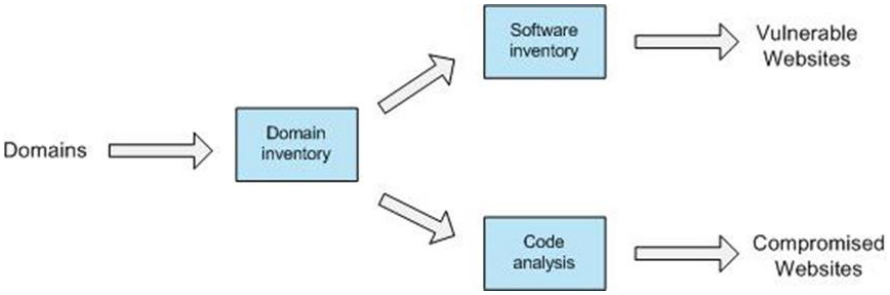

Figure 19: Skanna process description

3.3.7 Website Analysis Component by GDATA

The website analysis component is an interface to G Data's internal website analysis components. The component can identify malicious websites based on G Data's internal analysis systems.

The interface feeds the analysis system via the input stream from the CCH and, after a re-port is produced, the output is send back to the CCH for further processing.

The generated output can be used by the experiments to notify CERTs and website owners.

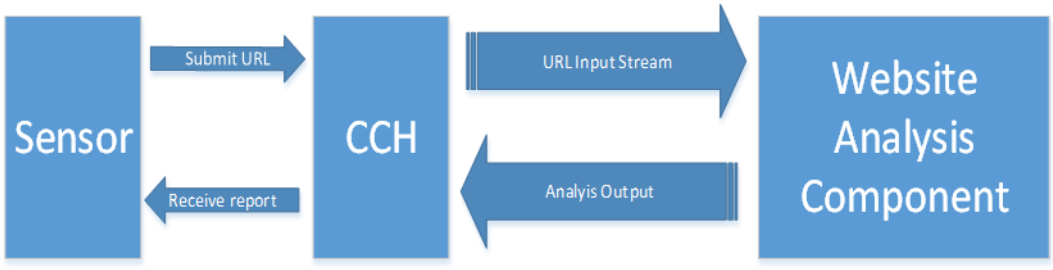

Figure 20: Website Analysis component service lifecycle workflow

3.4 Building Block: Network Traffic Sensors

The Network Traffic Sensors are responsible for collecting and providing data on infected systems (bots) for ACDC. They are one of the (primary) sources of data for the ACDC Centralized Clearing House, providing information related to infected systems on the Internet that are used for malicious purposes.

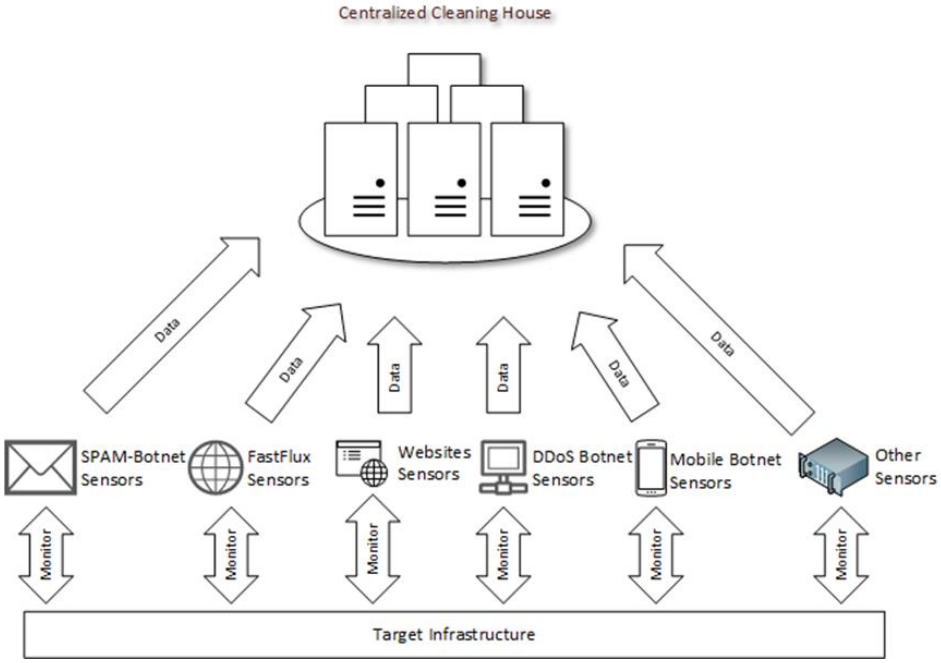

Figure 21: ACDC network Sensors - General Architecture

3.4.1 Spam Botnet Sensors

The Spam-Botnet sensors are focused on gathering data related to Spam botnets used primarily for Spam message distribution

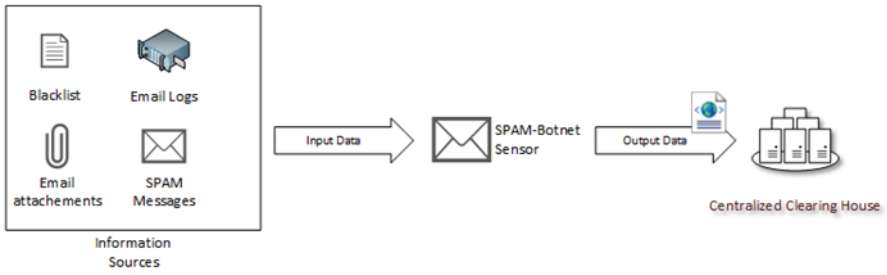

Figure 22: Spam Botnet Sensor - General Architecture

Depending on the specific type of sensor, the sensor receives input data from specific sources, such as logs from email servers, email messages to be analysed or already market as spam by anti-spam filter engines, email attachments, etc.

The sensor should then process these data according to its specification and, when evidence of botnet related activity is detected, send it to the Centralized Clearing House, in a standardized form and using the Clearing House's API.

3.4.2 Fast Flux Botnet Sensors

Fast-Flux Botnet Sensor are focused on targeting systems and domain names used in Fast-Flux activities on the Internet, and provide this information to the Centralized Clearing House.

Usually, the IP address behind a webpage is static. In contrast to this, the Fast-Flux method uses a specific domain (e.g., www.example.com) and assigns new IP addresses to it within a short time interval (approximately every three minutes). The bulk of IP addresses used usually points to infected computers which are part of the same botnet, and all these machines (i.e., the bots operating on them) host the same website. In other words, a user who thinks he connects to the benign service of www.example.com is frequently redirected to another server without noticing it, as the visible content never changes.

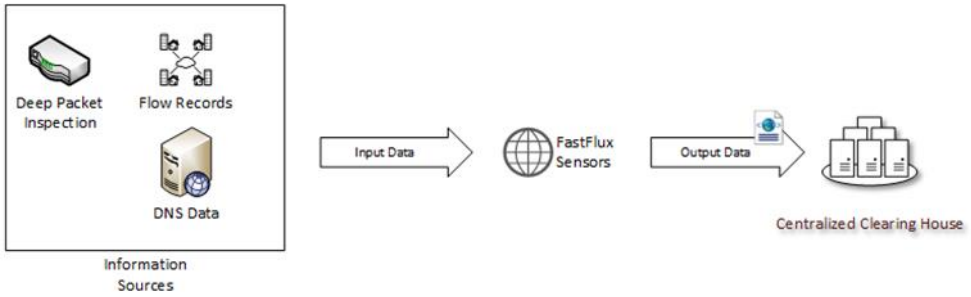

Figure 23: Fast-Flux Botnet Sensors - General Architecture

Depending on the specific type of sensor, it receives input data from specific sources, such as DNS zones or servers, network flow records, packet inspection mechanisms, etc. In terms of spatial analysis DNS-information could be used to extract geographical information of IP addresses that are or have been associated with a specific domain. Here, not only DNS-information about a domain itself but also information about their responding DNS-servers should be evaluated.

The sensor should then process the data according to its specification and, when evidence of Fast-Flux botnet related activity is detected, send it to the Centralized Clearing House, in a standardized form and using the Clearing House's API.

3.4.3 Malicious and vulnerable Website Sensors

Vulnerable web sites are very often target of the attacks done by hackers manually or these attacks are performed from compromised bots. The attacks performed by compromised bots to port 80 are performed automatically and are usually related to remote file inclusion at-tack types or attacks which do not require assistance of other compromised systems. In this sense the most interesting attack type is remote file inclusion, since it includes in the attack another system hosting malware. Such attacks could exploit vulnerabilities in web sites thus turning web site for example into php bot or do other types of attacks like cross site scripting etc. Such attack turns regular web site into malicious one.

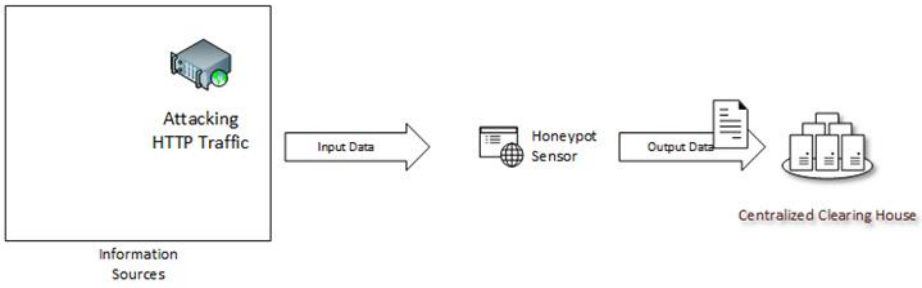

Figure 24: Website Sensors - General Architecture

Through the use of passive sensors that simulate given vulnerabilities – Honeypots – which will be set on a given network, one can identify malicious or vulnerable websites, on the internet, used for malicious proposes.

The sensor should then process these data according to its specification and, when evidence of botnet related activity is detected, send it to the Centralized Clearing House, in a standardized form and using the Clearing House's API.

3.4.4 DDoS (Distributed Denial of Service) Botnet Sensors

The Distributed Denial of Service (DDoS) Botnet Sensors are focused on targeting systems and networks used in DDoS activities on the Internet, and provide this information to the Centralized Clearing House.

DDoS attacks imply a massive amount of requests being done to a specific target. The success of an attack is directly related to the amount of traffic generated, something that can be specially accomplished by using botnets. When a specific target has been chosen, botmasters contact their bots and initiate the attack, which is nothing more than accessing the targets service as often as possible.

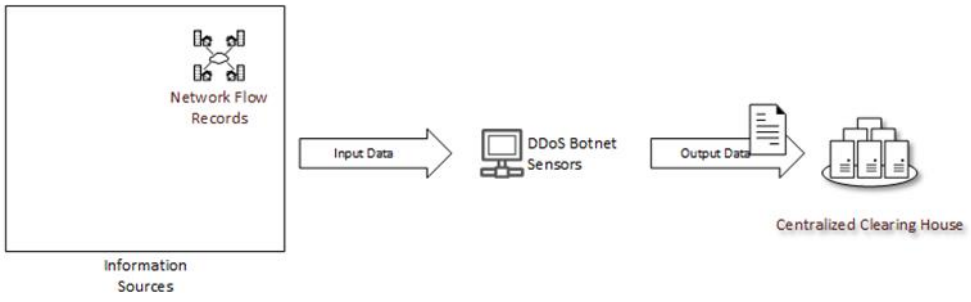

Figure 25: DDoS Botnet Sensor - General Architecture

Depending on the specific type of sensor, it receives input data from specific sources, such as network flow records.

The sensor then processes these data according to its specification and, when evidence of DDoS botnet related activity is detected, send it to the Centralized Clearing House, in a standardized form and using the Clearing House's API.

3.4.5 Mobile Botnet Sensors

The Mobile Botnet Sensors will be focused on targeting mobile systems infected with mal-ware and controlled by a botmaster for specific purposes, and provide this information to the Centralized Clearing House.

Mobile phones are today nothing less than pocket size computers and their use cases com-prise much more than making telephone calls and writing text messages. Smartphones, i.e., mobile phones with sophisticated capabilities, advanced mobile computing competencies and broad band connectivity, are employed to connect to and make use of a wide range of different services. Many of these services (e.g., email, banking, shopping, or social communities) require the indication of personal user credentials, which in turn are often saved on the device for convenience reasons. Because of this, attacking modern phones is a promising endeavour.

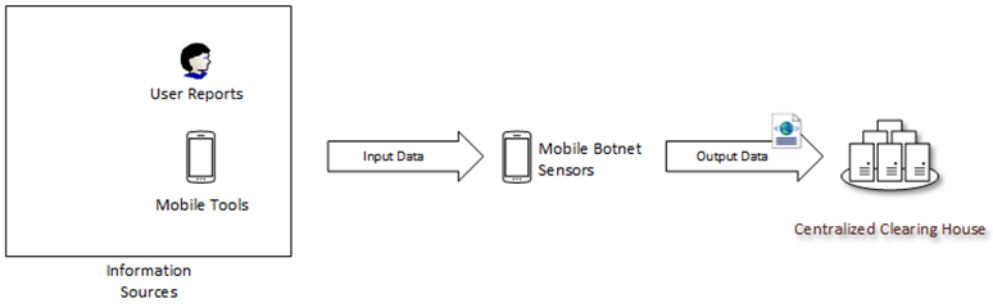

Figure 26: Mobile Botnet Sensors - General Architecture

Depending on the specific type of sensor, it receives input data from specific sources, such as the data collected by the mobile tools or the reports from the users.

The sensor should then process these data according to its specification and, when evidence of a mobile bot is detected, send it to the Centralized Clearing House, in a standardized form and using the Clearing House's API.

3.4.6 Other Network Sensors

TI-IT HoneyNet tool is a collection of sensors based on low-interaction honeypots (Source code of honeypot sensors are publicly available) from which different types of events are collected.

The Dionaea honeypot sensor is mainly used for the malware collection (binary file) and for capturing session information about potential anomaly events in terms of connection parameters, source of the incident and its properties. The Glastopf web application honeypot is instead deployed to gather data from attacks targeting web applications. Finally the others honeypot (i.e. Kippo) sensors have been deployed to monitor brute force attempts against the SSH service.

The honeypot sensors collect traffic on public network; the IP addresses assigned to the sensors reflect the geographical distribution of Telecom Italia (TI) Company's geographical ad-dress plan across Italian regions. The distributed network of honeypot sensors deployed in-crease the capacity of ACDC in terms of incident detection, event correlation and trend analysis. The HoneyNet tool is not publicized and, since there is no valid reason to try to connect to the fake services offered by the sensors, the captured data are only related to scanning activities or to malicious infection attempts towards TI infrastructure. No user's traffic is monitored.

Considering the architecture of the ACDC Framework and the services classification, TIIT HoneyNet Tool is mapped as an ACDC service that offers data related to events in the do-main of cyber security.

The data acquired by the HoneyNet have the objective to detect anomalous activities and in this way, TIIT HoneyNet contributes to the detection phase in the ACDC process. The collected data are specifically related to suspicious IPs, domains and binaries connected to mal-ware.

Figure below reports how the HoneyNet tool has been integrated in the overall ACDC architecture and the interaction with the other ACC components.

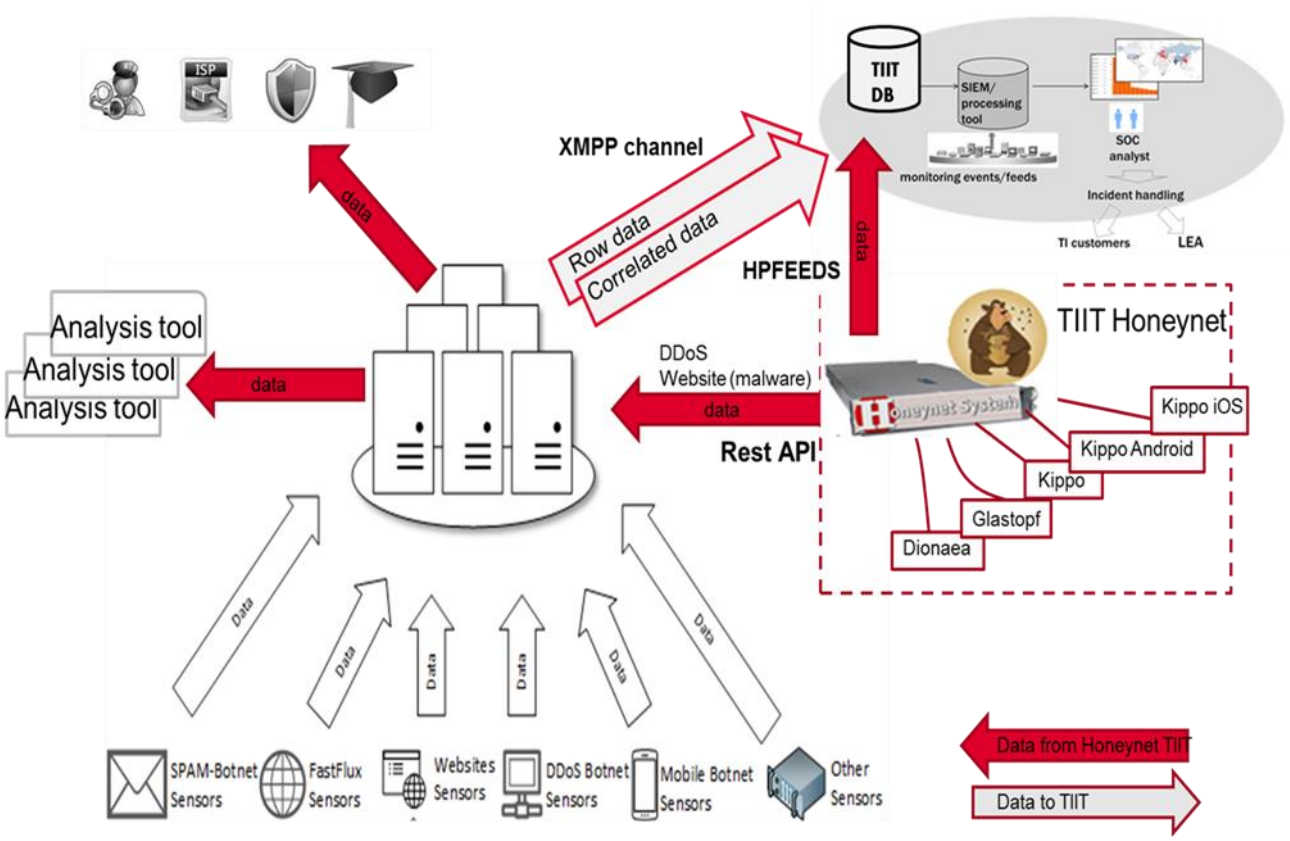

Figure 27: Integration of TIIT HoneyNet Tool with the ACDC architecture

The HoneyNet tool interacts principally with the CCH (the core element in the picture) by reporting in real time the collected data. This type of interaction with the CCH is realized via REST API. The data are sent according to data exchange format defined by WP1.

The data provided by TIIT HoneyNet extends the CCH information base and can be used by ISP/CERT connected to the ACDC infrastructure to detect suspicious activities carried on against the TI infrastructures by IP addresses under their constituencies.

Through the CCH, TIIT HoneyNet also interacts with the other ACDC tools, in particular with the analysis tools by providing row data useful to perform a more in-depth analysis and correlation. For example the reports of the malware type sent to the CCH are useful to perform analysis of the attached code

On the other hand the HoneyNet tool interacts with the CCH to receive all the events, both row and correlated data, associated to Telecom Italia ASNs. In this case the interaction is realized through the XMPP channel. Specific write API keys connected to the XMPP channel have been associated to specific report category together with read API keys.

An internal tool is in charge to receive and process the data generated by CCH through the XMPP channel. The received datasets are then stored in the internal TIIT DB and used to per-form statistics and, depending on the confidence level associated on each reports, alerts can be forwarded to the internal Telecom Italia Security Operation Center (SOC), in charge to manage the possible incidents.

It is still under development the integration of the data received from the CCH into an alert correlation engine, e.g. a SIEM (Security Information Event Management), in order to dis-criminate and produce alarms only for the most critical events, e.g. to an unknown malware or new kind of attacks.

3.5 End Customer Reporting Tools

3.5.1 Conan Mobile – INCIBE

3.5.1.1 Overview

Conan Mobile is a "End user tool" for Android devices. It helps users to check the security state of the device configuration and installed apps. To reach this objective it realizes three main activities:

- Configuration devices analysis:
  It evaluates the device configuration and gives the user recommendations to improve the security level of his device.
- Analysis of installed applications:
  It classifies the apps regarding their dangerous level.
  It classifies the apps permission regarding their risk.
- Proactive service:
  Real time check of the events generated in the device. Some of these events will generate a notification:
  Connection to an unsecure WIFI network.
  Changes done to the Hosts file.
  Dangerous package installation.
  Monitoring of SMS and calls.

In addition to the features described above, Conan Mobile uses the GCM service to notify the users about any threat or alert discovered.

3.5.1.2 Implementation

As this tool is installed on the end-user device it is used to directly notify users and warn them about dangerous configuration of their devices, connections done to malicious sites, malicious APKs installed and finally thanks to the GCM service it is also possible to notify them about any warn generated from the NSC or CERT. Taking advantage of this direct interaction with the user, notification and mitigation are done without any further action. Be-sides, Conan Mobile has been integrated with the INCIBE's Anti-Botnet Service, which is provided from the Spanish NSC.

Conan Mobiles direct contribution to ACDC, regarding sharing data, is to send malicious and suspicious APKs detected. For those suspicious APKS which are available on the backend it sends the binary too. This data can help other partners to know which malicious APKs are currently installed and used by users and try to prevent them to install them. It can also increase the intelligent of the model because

it provides malware samples that can be further analysed by other partners and may help in knowing how the malware is distributed and acting through correlation and aggregation actions

The Conan Mobile tool and back-end are described at deliverable **D2.3** and it is also available at Google play store:

https://play.google.com/store/apps/details?id=es.inteco.conanmobile

3.5.2 HitmanPro EU Cleaner powered by Surfright

HitmanPro is designed to work alongside existing security programs without any conflicts. It scans the computer quickly (less than 5 minutes) and does not slow down the computer (except for the few minutes it is scanning). HitmanPro does not need to be installed. It can be run straight from a USB flash drive, a CD/DVD, local or network attached hard drive.

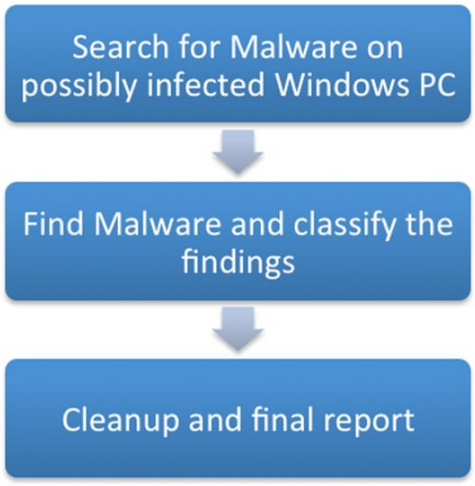

Figure 28: HitmanPro - Architecture Scheme

The “EU-Cleaner provided by Surfright” is able to detect and remove Ransomware and Trojans from Windows machines. It uses the classic list of Malware samples, which is maintained on the local machine and in case of new findings those are listed in the “Surfright-Scan-Cloud” as hash values. The Cleaner also uses an advanced Heuristic to find hidden Malware and can detect new versions of existing Ransomware by using its heuristics.

The Tool is hosted on the ECO webpages:

https://www.Botfrei.de/en/eucleaner.html#surfright

The tool is also available for the other National Support Centres via the ACDC-project website:

http://acdc-project.eu/software/device-detection-and-mitigation-multipurpose-tools-for-users/

The tool detects the language of the operating system at installation time and will install in the selected language.

3.5.3 EU Cleaner powered by Avira

“EU-Cleaner powered by Avira” is a tool to scan infected end user windows machines for malware and to clean infected machines. The tool is provided by AVIRA to be distributed to end users by the National Support Centres.

Avira is providing "Mitigation Tools" for infected customers for free.

The EU-Cleaners are available in the following languages:

DE: http://personal.avira-update.com/package/eucleaner/win32/de/avira-eu-cleaner_de.exe
EN: http://personal.avira-update.com/package/eucleaner/win32/en/avira-eu-cleaner_en.exe
ES: http://personal.avira-update.com/package/eucleaner/win32/es/avira-eu-cleaner_es.exe
FR: http://personal.avira-update.com/package/eucleaner/win32/fr/avira-eu-cleaner_fr.exe
IT: http://personal.avira-update.com/package/eucleaner/win32/it/avira-eu-cleaner_it.exe

3.6 Partner Solutions for the ACDC Environment – overview

3.6.1 DE-CIX DDoS Scanner

3.6.1.1 Overview

To help customers mitigate the effects of Distributed Denial of Service (DDoS) attacks against their networks, DE-CIX introduced customer-triggered blackholing in 2013.

If a customer is attacked by DDoS on a certain IP address, the customer can utilize the DE-CIX Black-holing feature to mitigate the effects inflicted by the attack. The Black-holing feature works in the following way: The customer announces the IP address under attack in a route announcement to DE-CIX operated route servers. This route announcement carries as next hop the IP address of the Black-holing feature. The route server redistributes to all other customers this particular route announcement. All traffic related to this route announcement is forward to the Black-holing IP address where it is then dropped before it can in-crease the load or overload the customer's resources. This is depicted in Figure 29: Architecture of Black-Hole feature.

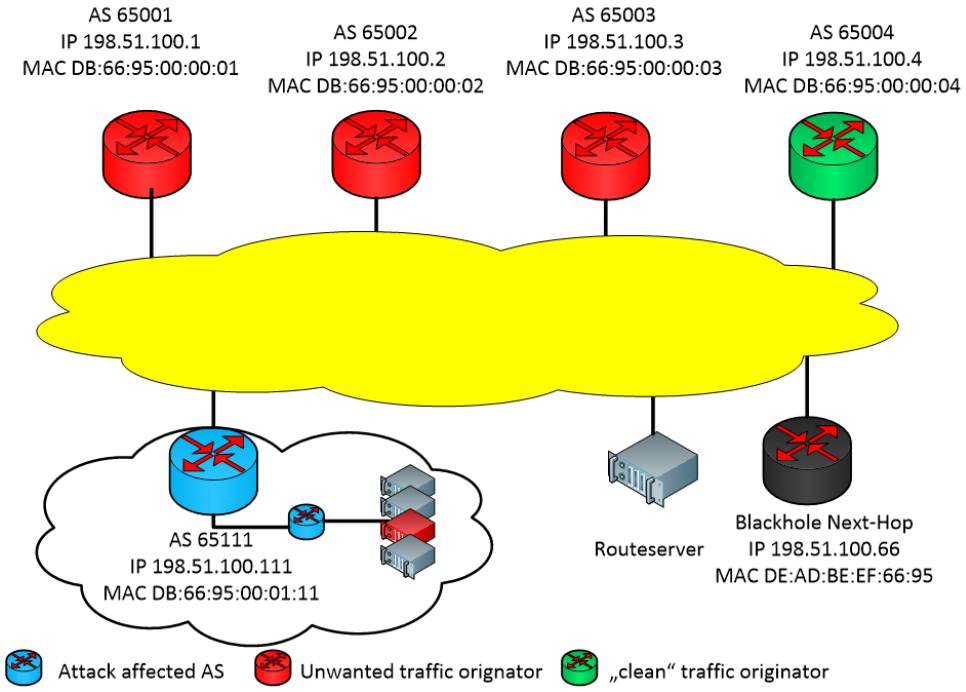

Figure 29: Architecture of Black-Hole feature

Historical data on use of the Black-holing feature is not retained. However, utilization of the feature was measured for a period of seven days in March 2014. During this time the Black-holing IP was set

as next hop for more than 600 routes coming from 28 customers. These numbers show that the Black-holing feature is a simple but effective tool to mitigate DDoS.

3.6.1.2 Implementation

The idea is to use the traffic send to the Black-Hole IP address as input for a DDoS detection sensor. The traffic arriving at the Black-Hole IP address is DDoS attack traffic that customers of DE-CIX want to discard. The DDoS detection sensor collects this traffic and then signals the ACDC Central Clearing House that a DDoS attack is going on. Additional information such as which AS is the source of the DDoS attack, which AS is the destination of the DDoS attack, which AS is sending the DDoS attack traffic to DE-CIX, which AS is forwarding the DDoS at-tack traffic, time duration of the DDoS attack, throughput could be added. For this kind of signalling an appropriate data format must be select. The picture below depicts the architecture of the sensor.

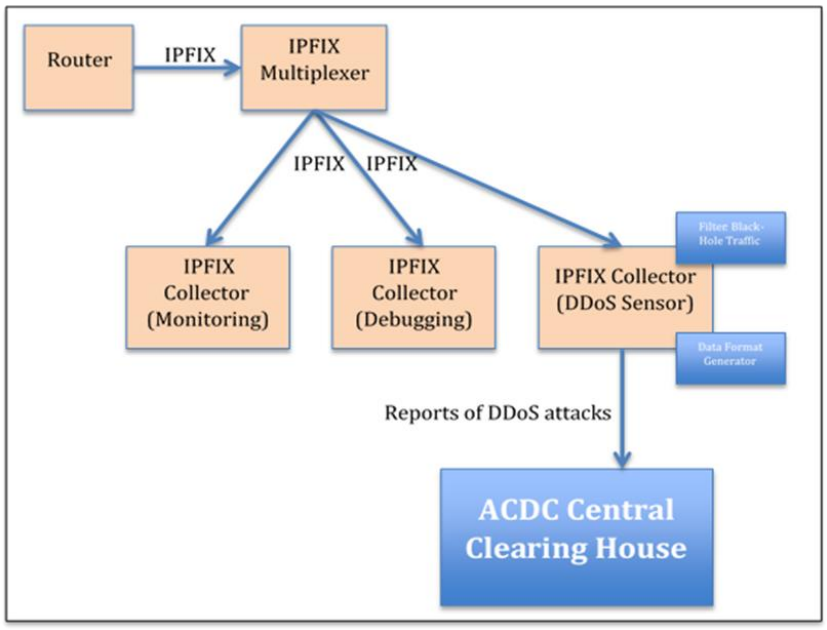

Figure 30: DE-CIX DDoS Sensor

IPFIX is a mechanism that can be used for monitoring and debugging network equipment such as switches and routers. In high throughput environments such as DE-CIX IPFIX can be used in a packet-sampling configuration. Packet sampling means that not every packet is selected for further investigation, instead, every x^{th} packet is selected. So x-1 packets will not be detected by IPFIX if it is used in a packet-sampling configuration. However, on a large scale and if traffic is similar up to a certain point the results generated from IPFIX in the packet-sampling configuration are still valid. The error introduced by this mechanism can be calculated with a certain upper limit.

At DE-CIX every 10000th packet is selected that is transmitted over the DE-CIX switching fabric. With such a high value for selecting packets the IPFIX data stream is still about 50Mbit/s and growing by about 50 percent a year.

The packets that are selected by IPFIX are exported to a so-called IPFIX collector. An IPFIX-collector is an application running on a server in order to aggregate information such as packet size, source and

destination. This information can then be used for monitoring reports (e.g., total traffic on the switching fabric).

IPFIX was not available in the operating system running on the routers DE-CIX is using. So, DE-CIX asked the vendors of its routers (e.g., Alcatel Lucent) to implement IPFIX, which they did. Since the beginning of 2014 IPFIX is available at DE-CIX. DE-CIX activated IPFIX for monitoring and debugging purposes on its routers in March 2014.

The routers used by DE-CIX export an IPFIX data stream to one single location. In order to be able to export this IPFIX data stream for different purposes such as monitoring, debugging, and a DDoS sensor an IPFIX multiplexer is required that copies the IPFIX data stream to various IPFIX collectors. This IPFIX multiplexer does not only multiplex IPFIX data steams but it allows changing the configuration of the IPFIX setup (e.g., adding new IPFIX collectors) with-out the need to change the configuration of the routers used in the production network. The IPFIX multiplexer also provides a layer of security as it separates communication between IPFIX collectors and the routers used for production.

A specialized IPFIX collector is implemented that is capable of signalling on-going DDoS attacks to the ACDC Central Clearing House. For this, a filter has been developed that is able to filter out only IPFIX data that was dropped by the Black-hole feature. Based on this, reports are generated for the DDoS attacks that are detected. These reports must then be formatted according to the data format and transmitted to the ACDC Central Clearing House.

The DATA Format used to notify the CCH is a JSON based format.

{“version”=1, “status”: STATUS, “src-ip”: “xxx.xxx.xxx.xxx", “dest-ip”: “xxx.xxx.xxx.xxx”, “src-port”: xxx, “dest-port”: xxx, “start”: millisecondsSinceEpoche, “end”: millisec-ondsSinceEpoche, “throughput”: xxx }

Meaning of keywords:

- STATUS: can be start, end, update
- throughput: is measure in kbit/s
- version: current data format version
- start: time in milliseconds when DDoS traffic was detected the first time
- end: if the status is start or update the end value is not set
- src-ip: Source IP address
- dest-ip: Destination IP address
- src-port: Source port number
- dest-port: Destination port number

3.6.2 HORGA Tool by GARR

3.6.2.1 Overview

HORGA is a system of low-interaction honeypots deployed in the GARR network. In addition to assigned address, It includes two /24 darknets.

The following pictures show the general architecture of the system.

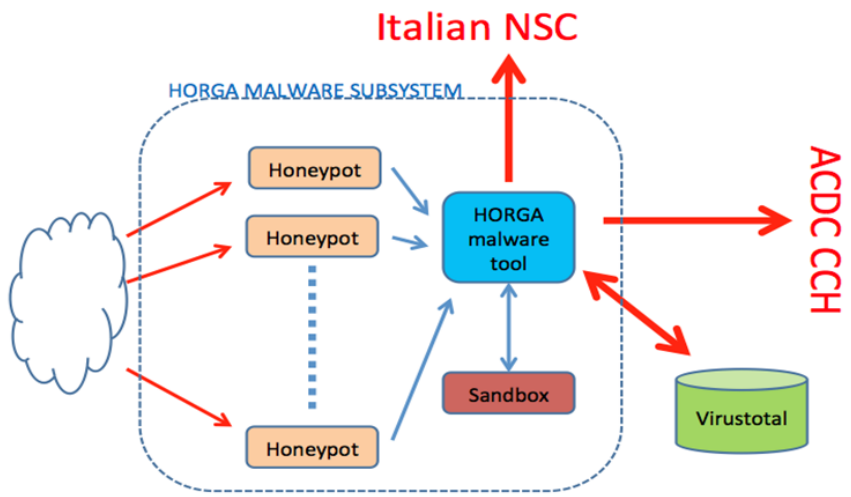

Figure 31: HORGA malware Tool

Different types of events are collected by using different kinds of low-interaction honeypots, e.g. dionaea, amun or thp.

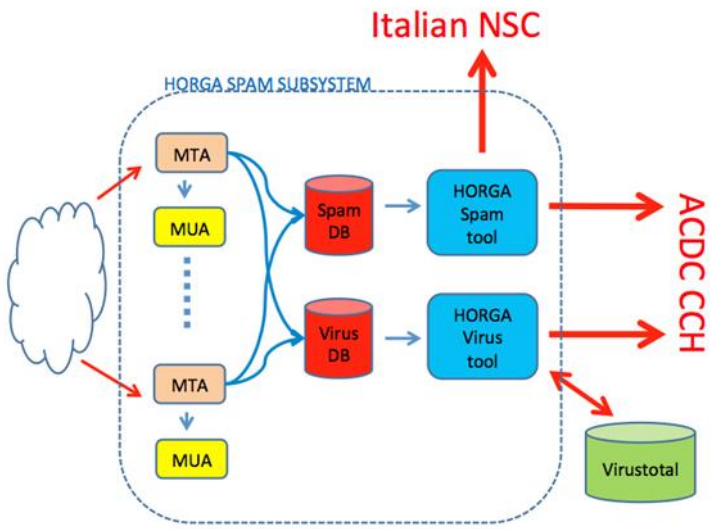

Figure 32: HORGA Spam Tool

3.6.2.2 Implementation

All the traffic in the darknets, which is to be considered malicious, is recorded by tcpdump-like sensors. The logs are collected for further analysis and used to open security incidents.

The most important features are:

- Detection of automated scans
- Detection of brute force scans
- Collect of binaries of malware

Binaries captured by the sensors are run in sandboxes, public domain and in house, so as to acquire information regarding the nodes, presumably botnet controllers, to which the mal-ware tries to

connect. From the sandboxes we obtain a further list of IP addresses and URLs, presumably much more interesting of the IP´s of the attacking nodes.

Data from malware captured is sent to CCH via the REST APIs and, when relevant, to GARR-CERT or the National Italian Cert.

3.6.2.3 Interaction with ACDC and other Parties

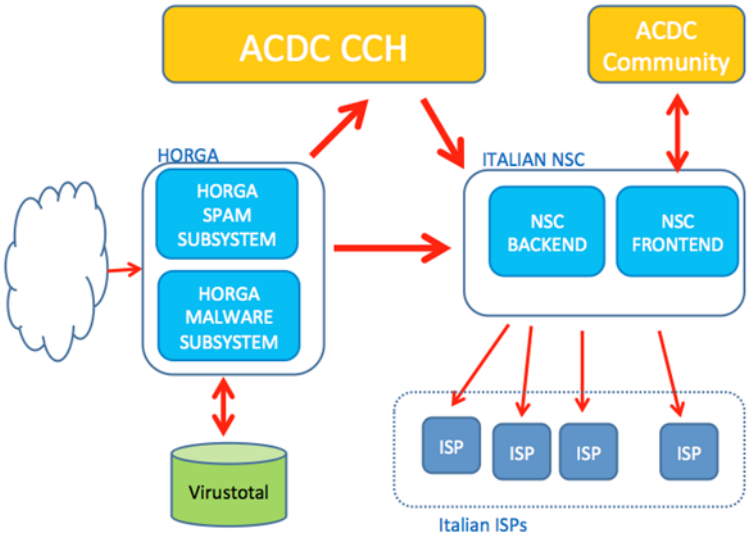

Figure 33:HORGA with links to ACDC and NSC

3.6.3 RelBot by University of Luxembourg

3.6.3.1 Overview

A tool for detecting the infected hosts by monitoring Netflow records was developed by UL and integrated with the CCH. A common scenario would be to deploy the RelBot tool in a network and provide it the access to NetFlow/IPFIX records collected from the routing and switching devices. If the tool detects a suspicious communications (based on the fact that many p2p botnets use keep-alive messages to their neighbours that are sent with a fixed period), a report is generated and can be sent to the CCH.

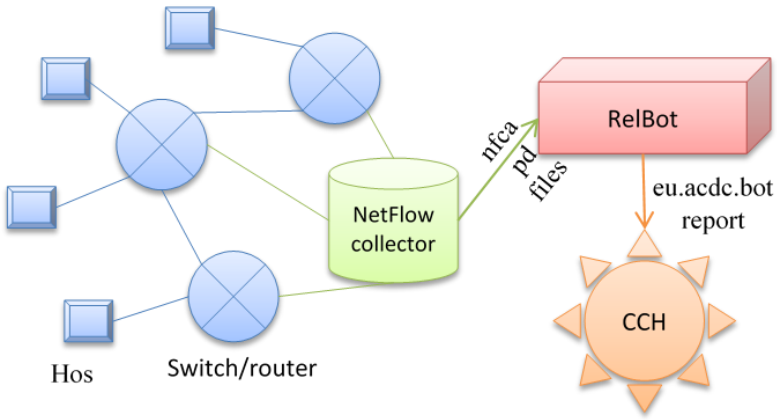

Figure 34: Interaction of RelBot with ACDC

3.6.3.2 Implementation

RelBot is a modular software (see Figure 35) whose all components, including the CCH submitter, are released under Apache 2.0 license and available on GitHub (https://github.com/tigran-a/relbot).

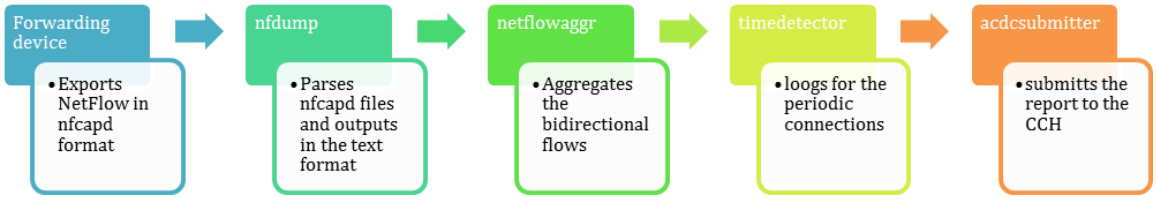

Figure 35: RelBot Toolchain

The code repository contains an installation guide and configuration tips. The functional tests of RelBot were described in Deliverable D2.5.

3.6.4 TID – High level overview of ACDC components

TID role in the general ACDC infrastructure and collaboration model is driven by several components. TID contributes with 2 types of components described in D2.3 Technological Development Framework: Detection tools and Utilities for data transformation. As part of detection tools TID includes a set of sensors based on DPI (Deep Packet Inspection) technology for botnets identification based on traffic behavioural: SPAMBot and DNSBot detector. HP Sentinel, a SDN (Software Defined Networking) based botnet detection, is another sensor tool integrated. Also a more privacy-friendly detection sensor is contributed through a low interaction HoneyNet based on Telefónica infrastructure, which allows collect malware and botnets attacks. Finally in order to support data reporting of previous tools and smoothly integration in ACDC's CCH, an additional utility component has been developed: ISPAdaptor, a network operator oriented tool that allows send and receive data in different formats, which also allows data analysis and data visualization.

Figure 36:TID Tools integration in ACDC represents the high-level integration with ACDC of the different TID's components, being the ISPAdaptor the central point on the integration process. Figure 1 (a) shows how each Sensor tool exposes the relevant information to ISPAdaptor, who is in charge of normalize and send the information to CCH, using the HTTPS REST API. Available information related with Telefónica constituency is also collected from CCH's XMPP channel. Figure 1 (b) summarizes how ISPAdaptor stores and shares with business units the relevant information after internal processes analysis.

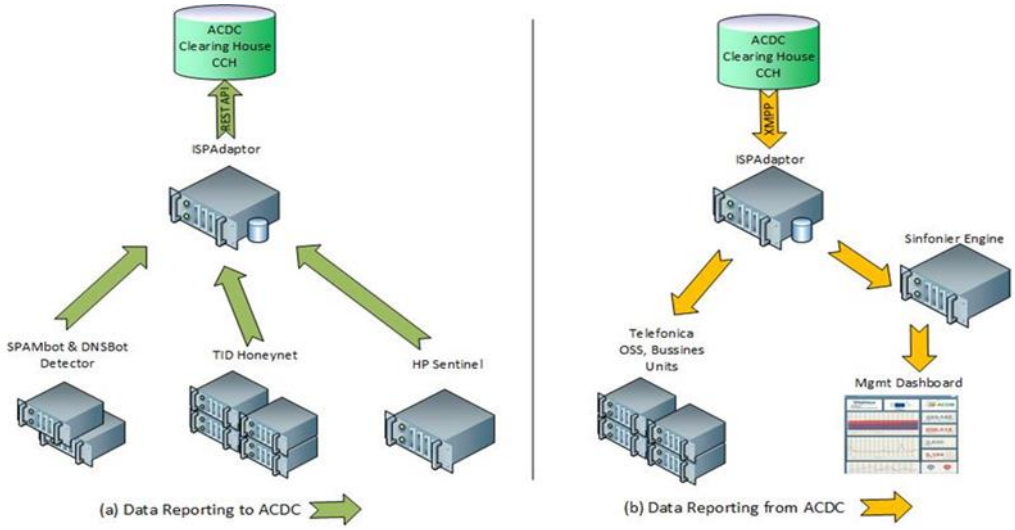

Figure 36:TID Tools integration in ACDC

3.6.4.1 SPAMBot and DNSBot

TID has developed a DPI (Deep Packet Inspection) product based on common Hardware (Commodity off-the-shelf or COTS) but with high performance for inline analysis when a high bandwidth requirement is present like in Network operator infrastructures. This in-house DPI has the potential for detections of different kinds of malware. These detectors are composed of security modules oriented in specific traffic, like SPAM, identifying SMTP traffic and making deep (L2-L7) analysis, or DNS analysing deeply (L2-L7) flows. Both are able to identify infected users by a botnet trying to contact with a C&C, FastFlux or generating spam traffic. This tool is aimed at Identification of ISP residential users or SME clients that generate spam caused by botnets. Identified IP addresses can be reported externally to ACDC sending in-formation through ISPAdaptor. Also is planned to collect relevant malicious domains from ACDC and added in blacklist domains detections

3.6.4.2 HP Sentinel

Hewlett-Packard Sentinel is a Malware detection tool based on Software Defined Networking (SDN) Architecture. This solution Uses network switches with OpenFlow capacities and a Security Openflow controller (Hewlett-Packard Sentinel) in charge of compare each DNS query crossing the switch against a Blacklist of Malware Domains from HP's DVLabs. When a user that is connected to the switch generate a DNS Query (for example a malware specimen trying to contact with his C&C), the DNS packet is redirected by Openflow to the controller and when it match against Black list discarded. HP Sentinel Tool can generate information of infected IP and the domain affected, this information is shared with CCH through ISPAdaptor, offering information of real uses of malicious domains.

3.6.4.3 HoneyNet

It is a tool that is composed of several nodes belonging to Telefónica network Spain that act as honeypots collecting data from attacks. It covers from Datacenters to residential (xDSL/FTTH) access, and with capacities from low cost HW (Raspberry PI) devices to standard servers. Automatically extract and send relevant information associated with suspicious botnets activity and share it with ACDC through ISPAdaptor. This HoneyNet is based in sever-al open source solutions: Glastopf that simulate web application services, Kippo for SSH at-tacks, and Amun that simulate services and binary downloads.

3.6.4.4 ISP Adaptor

Network Operator oriented utility tool that acts as an integration and translation service with the objective to export/import the botnets information from the ISP to a different external systems, (currently the CCH and Business units). The system is composed of several modules in charge of collect the data from different sensors and store in DB (MongoDB). Information is translated and sent to CCH through API interface. Also it allows collecting data through an additional module from CCH with XMPP channel. This data is stored and used to support business and to raise Telefónica awareness. It includes integration with Telefónica BigData analysis and visualization engine (Sinfonier) and report generation, through specific formats.

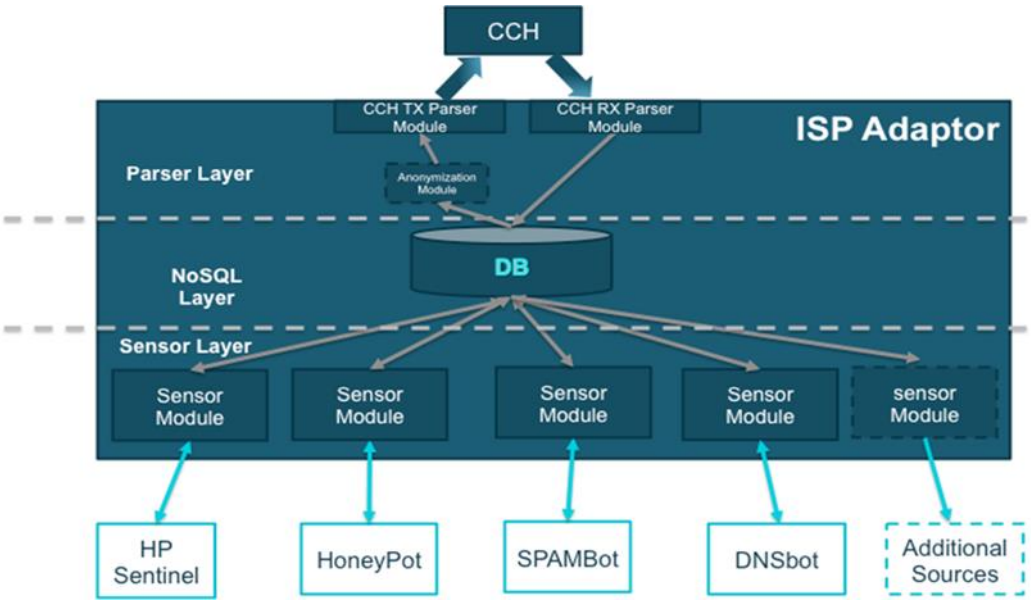

Figure 37: ISP Adaptor by Telefonica

3.6.5 MMT-Tool by Montimage

Montimage is involved in the DDoS experiment doing the actions related to Monitoring tool provider and tool operator roles. The Montimage Monitoring Tool (MMT) was deployed on the BGPOST HoneyNet platform in collaboration with BGPOST. This platform is presented in the next section. This allowed detecting certain types of behaviour that can be related to botnet activity performed by infected or malicious machines.

The activity detected by MMT is related to DDoS activity but this needs to be confirmed by correlating information from other sources. The incidents can be considered unwanted activity but the goal is to determine if they correspond to malicious intent with a high degree of certainty. For this a set of modules have been developed that allow: i) receiving attack related data from any of the CCH data providers; and, ii) correlating this data to determine if the suspicious activity detected by MMT comes from the same machines that have been identified as malicious using other techniques and/or from observations at different locations.

3.6.6 BGPost HoneyNet

The Figure 38 represents the HoneyNet BGPOST platform and shows where MMT was in-stalled.

The BGPOST HoneyNet acts as a trap set to detect, deflect, and counteract attempts at unauthorized use of an information system. The network site appears to be part of a network and seems to contain information or resources of value to attackers. It is actually isolated from the BGPOST's network and is monitored using several tools, including MMT.

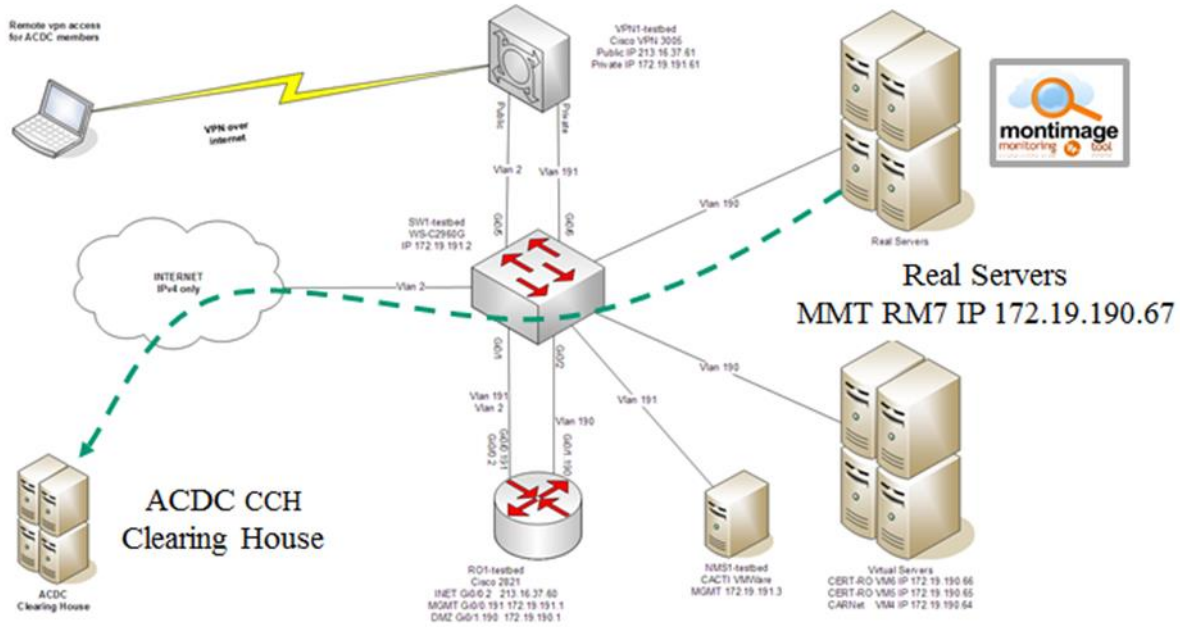

Figure 38: BGPost HoneyNet Platform

3.6.7 Correlation Platform

In ACDC, Montimage adapted its existing tool set and developed new modules to integrate the CCH and data coming from different sources through XMPP channels. The Figure 39 represents the complete platform that correlates data coming from different sources.

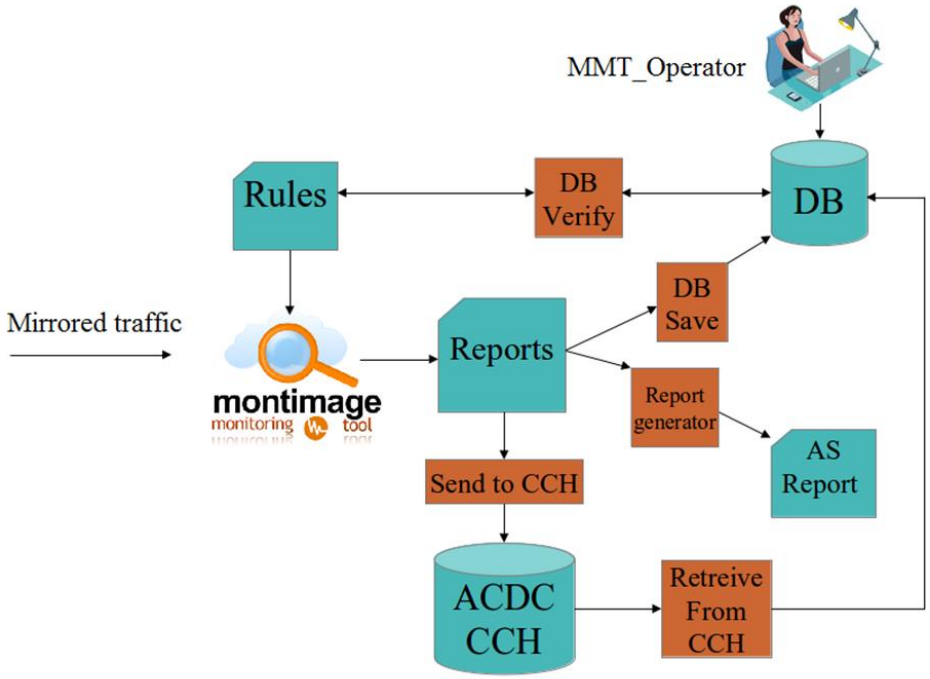

Figure 39:MMT Correlation Platform

The process is as follows:

1. Traffic is analysed by MMT_Probe. This traffic can be mirrored and the probe will not disrupt the traffic; or, the probe can act as a firewall and block attacks detect-ed. The active mode has not yet been finalised and tested.
2. MMT_Probe generates live reports that are sent to the CCH by a python daemon process (Send_to_CCH); used to generate a report identifying the AS (Report_generator); and, saved in the local MongoDB (DB_save). DB_save will generate an event that will be detected by MMT_Operator.
3. Another python daemon process (Retreive_from_CCH) connects to the CCH and recuperates the feeds from other XMPP channels and stores them in the Mon-goDB.
4. MMT_Operator will detect the event generated by MMT_Probe, query the MongoDB to collect all the information already available on the IP address, source of suspicious behaviour, and display an alarm with all this information so that it can be seen by the human operator.
5. Finally, the rules used by MMT_Probe can contain queries to the MongoDB (executed using DB_verify) to determine if other security related information is avail-able from other sources. In this way the local detections by MMT_Security can be refined and made more precise.

Figure 40 shows the present situation. Information from the write keys is received by the read keys created by Montimage. Practically each write key corresponds to a read key. This choice was made to simplify the control and allow selecting exactly what information one wishes to process. Instances of the Retreive_from_CCH module can be created for each read key and the module will store the information received in the MongoDB.

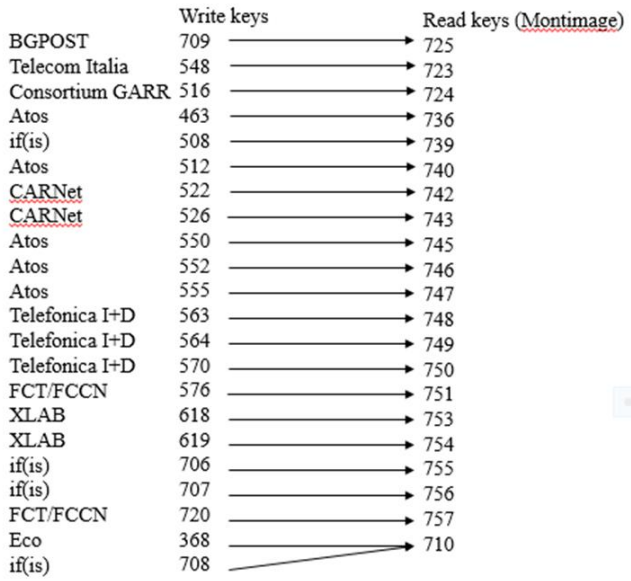

Figure 40:Read Keys that receive data

3.6.8 SEC Incident (SECurity Incident) BDigital/Eurecat

3.6.8.1 Overview:

Barcelona Digital's contribution to the ACDC infrastructure is SEC-Incident (SECurity Incident), a ticketing and security-related event management system to be used in a CSIRT (Computer Security Incident Response Team) or SOC (Security Operations Centre) environments. In the context of the ACDC project, SEC-Incident is capable of managing the status of malware and phishing related issues. More specifically, relevant partner sensors and tools gather malware/phishing related information and store it to the CCH. This information plays an important role in the protection of end users from this kind of attacks, as it is being consumed by end-user tools developed within the project.

However, there is no guarantee that the collected phishing information is up-to-date and still relevant. It is possible that part of it might have been collected earlier and at the time of retrieval it might have been rendered obsolete.

SEC-Incident's objectives include the intelligent, automated management of the status of this kind of incidents. A user of SEC-Incident subscribes to phishing related incidents stored in the CCH. SEC-Incident creates a new issue that enters in the predefined workflow. SEC-Incident handles the issue in an automated way and sends email notifications to the involved users about the status of the incident. Therefore, the added value of integrating SEC-Incident to the project is to validate the actual status of each incident and make only the valid ones available to the consumer applications.

3.6.8.2 Integration and implementation

The following figure shows the SEC-Incident's integration with the ACDC infrastructure, and more specifically the Centralized Clearing House (CCH).

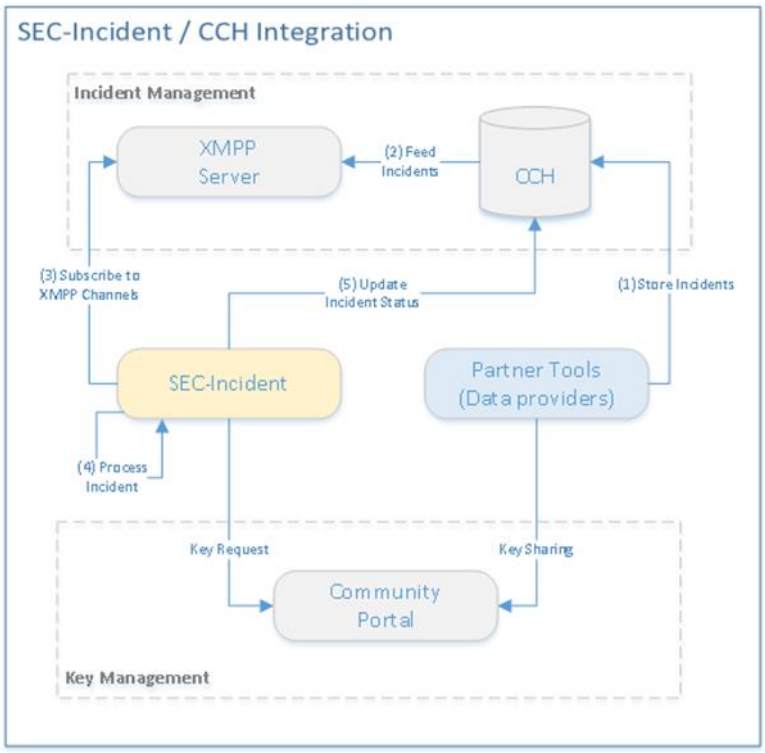

Figure 41: SEC-Incident and CCH Integration

A typical workflow contains the following steps:

1. Partner tools and sensors feed the CCH with malware/phishing incidents
2. The newly added incidents are available to interested and authorized parties through channels of the XMPP service
3. SEC-Incident subscribes to the corresponding XMPP channels and receives security incidents
4. SEC-Incident processes internally the status of the received incidents
5. SEC-Incident updates the status of incidents on the CCH

For steps 3 and 5, SEC-Incident needs to be authorized and authenticated by the CCH. This requirement is satisfied through the Key sharing process, as defined and implemented by the ACDC consortium.

The internal workflow of SEC-Incident is shown in the following figure. The CCH Integration module undertakes user interface tasks, such as management of API Key, subscription to XMPP channels and sending status updates to the CCH. The Incident Management Workflow is based on the JIRA workflow and decides in a semi-automatic way the current status and therefore the validity of each incident. The conclusion of this workflow is forwarded to the CCH Integration, which in turn updates the incident status at the CCH

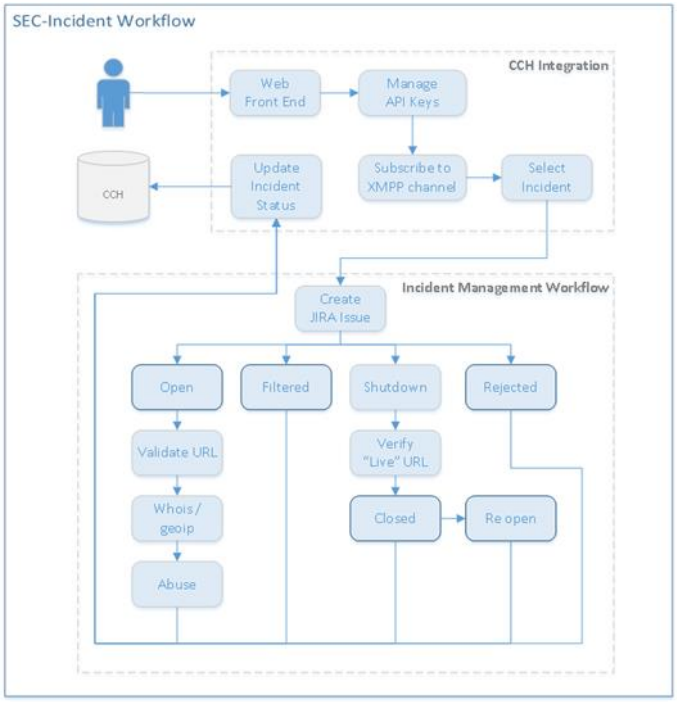

Figure 42: SEC-Incident internal Workflow

Within the context of the ACDC Technology Framework, SEC-Incident has a TLR 9 (actual system proven in operational environment).

3.6.9 If(is) Components – Overview

3.6.9.1 Sandnet

if(is) runs a dynamic malware analysis system for network traffic called Sandnet. Sandnet stores PCAP files of executed binaries (in Sandpuppets) and these are evaluated within a database. We apply signatures (e.g. blacklists) to this data and mark network connections as malicious.

3.6.9.2 DDoS Monitor

The if(is) DDoS Monitoring tool gathers information about DDoS Botnets and the victims of DDoS attacks. This tool was not part of the original list of tools and not part of the original Description of Work, but was added, as it was contributing to a missing component in the detection architecture. Other tools also capture partially information about DDoS attacks, but with this tool additional monitoring can be done on the targets of DDoS attacks.

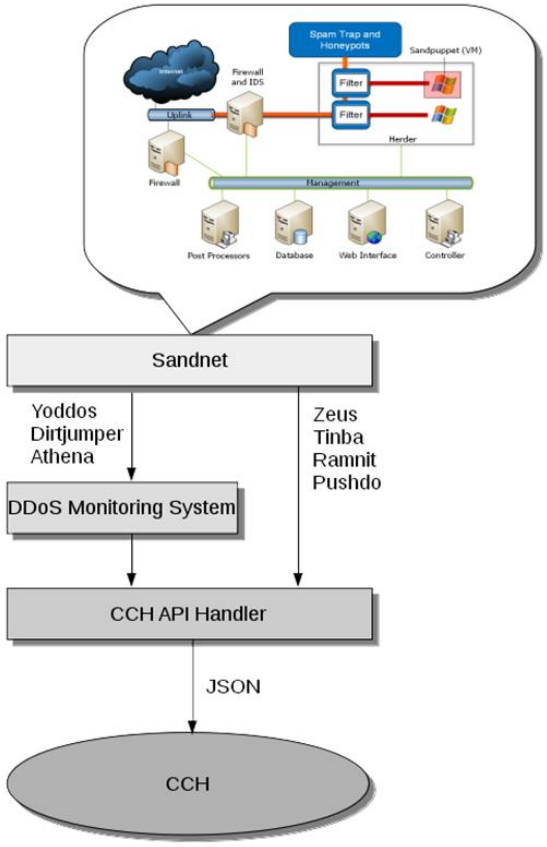

Figure 43:If(is) Tools in the ACDC Environment

The following data is stored for each flow:

- IP address
- URI
- Botnet Family name
- Timestamp

3.6.10 XLAB Toolset

XLAB's tools are reading and sharing data from/to CCH. Information sent to CCH is shared with other ACDC partners as defined with CCH's sharing policies. The main XLAB's sensors tool that participates in MOBILE EXPERIMENT is Device Monitor. Suricata IDS and EventCor-relator are meant to be used as independent refinements of the main mobile sensors. Their development is finished but the integration process was delayed since mobile sensor and data feed was a main priority.

List of sensors/tools:

- Device Monitor
- GCMServer
- Suricata IDS
- EventCorrelator

Figure 44 provides a schema depicting integration and deployment of the XLAB tools

3.6.10.1 Device Monitor

Device Monitor is a mobile application/sensor tool, which monitors mobile devices for malicious mobile events. The detection methods are based on known mobile attacks (SMS hijacks, visits of malicious URLs, detecting of different application exploits and presence of installed known malicious or suspicious applications).

Device Monitor (mobile application – sensor tools) reports directly to GCMServer, and the reports are aggregated within GCMServer and forwarded towards ACDC's CCH.This data is then aggregated and reported to the CCH in appropriate format. Please, refer also to the infrastructure that is depicted in 44.

3.6.10.2 GCM Server

GCMServer acts as Device Monitor's broker meaning it interacts with the external system (CCH) in order to obtain information of malicious activities and report events. Interface to-wards external system can be administered using configuration files. It extends functionality of the Device Monitor tool:

- it connects to CCH and/or STIX Aggregator (broker between CCH and Device Monitor)
- it fetches information from the CCH regarding malicious domains, URLs, and malicious applications
- it provides input to the CCH about malicious activities detected with Device Moni-tor
- is a tool that can be used to submit malicious files (e.g. analysed off-line or with some external tool) manually via its user interface

GCMServer creates rules from the CCH's reports and distributes the rules between Device Monitor instances. GCMServer also aggregates and refines the reports from the Device Monitor applications, converting them into appropriate CCH format. It also provides additional modules to connect to other resources than CCH (e.g. Google Safebrowsing API, Cyscon's API). These external resources act as an additional feed towards GCMServer, and consequently to CCH. With this new knowledge it also refines CCH with new detected APK samples or events from mobile devices.

3.6.10.3 ScuritaIDS

Suricata is the OISF IDP engine, the open source Intrusion Detection and Prevention Engine. High performance on standard x86/x64 based hardware is achieved by multithreaded engine and/or capturing traffic with supported network capture cards. Detection of suspicions traffic is rule based. It is possible to add new rules at runtime which is provided by the extensions. For this purpose we provided extensions interfacing with CCH (generating new Suricata rules from CCH reports). XLAB's Suricata IDS is deployed on XLAB's premises. It provides additional layer of prevention between mobile devices and external resources. It also reports the detections of events towards CCH

3.6.10.4 Event Correlator

EventCorrelator correlates network events from Suricata's fast.log recorded malicious events. The correlation engine is designed to operate with strings and, as such, is able to correlate any string data provided as input from the reports. The correlator checks for PCAP/fast.log, listens for reports originating from GCMServer, and reports new correlations found from IDS and GCMServer's reports. This identifies potentially malicious mobile devices on the network being observed.

3.6.10.5 Infrastructure Overview

The following graph provides high-level view on the components provided by XLAB. Internal message bus connects all components together providing faster and independent communication channel between CCH and the tools. This way in case of a problem on a connection towards the CCH the tools still operate normally and independently.

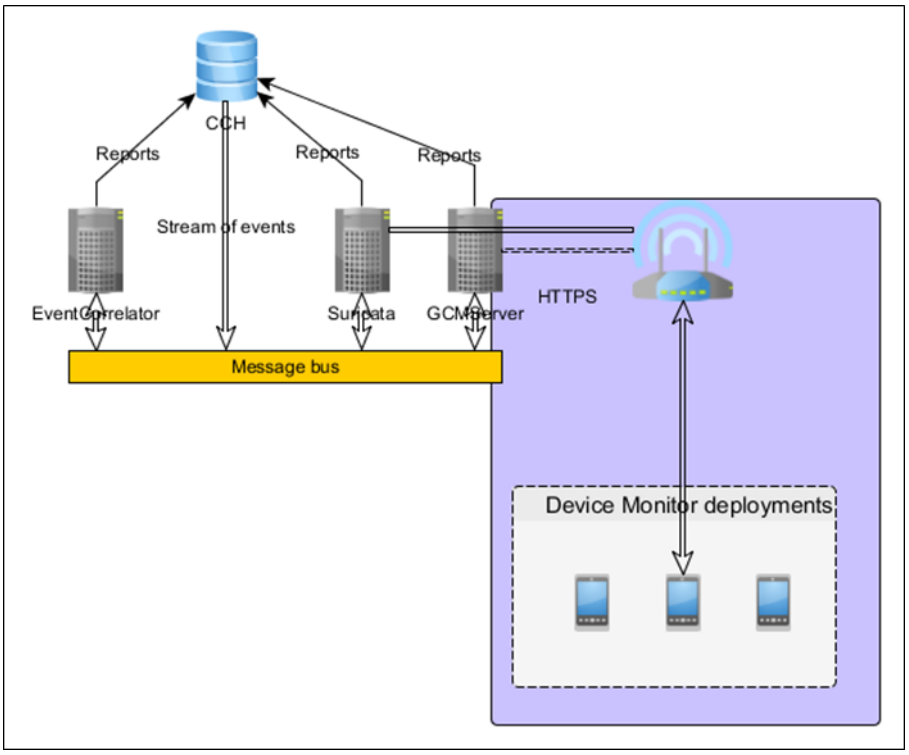

Figure 44:Infrastructure of XLAB´s Tools

4 Interaction of Building Blocks and other Components:

4.1 BGPost

The BGPost Testbed environment is created especially for the needs of ACDC project.

Already implemented partner tools:

- CARNet - Spamtrap and low interaction honeypot(s) multipurpose appliance
- CERT-RO – HoneyNetRO Dionaea and Kippo sensors
- Montimage Monitoring Tool – MMT sensor
- GARR – Horga sensor

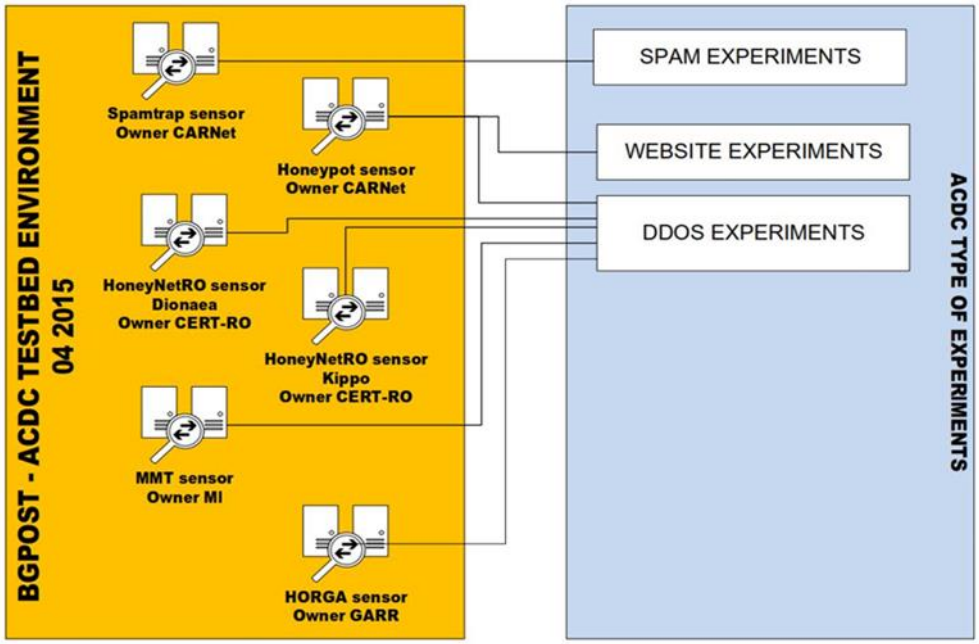

Figure 45: BGPost - Sensor categories

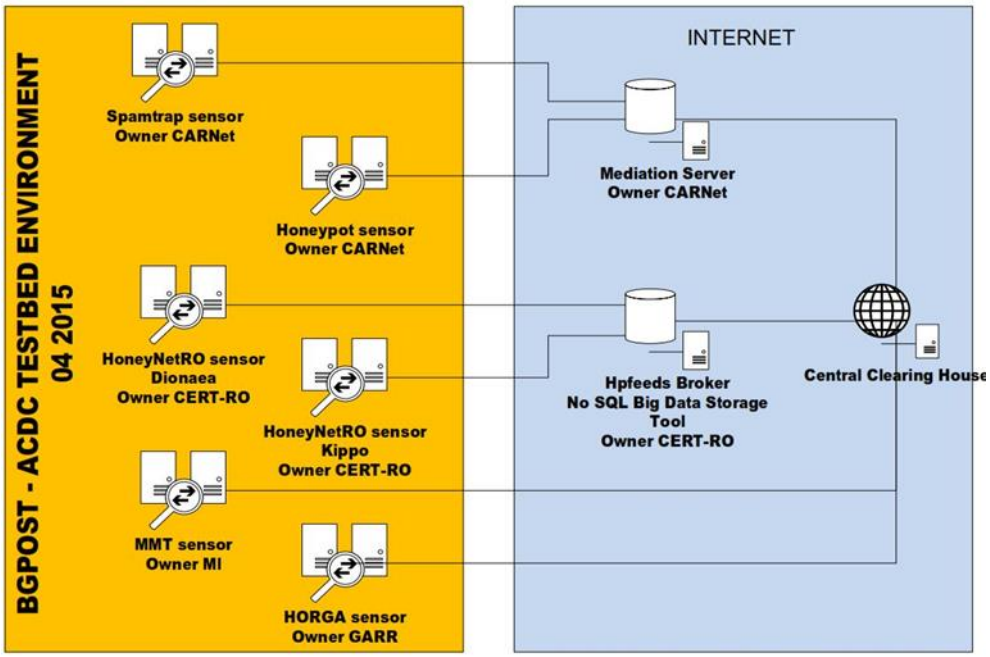

Figure 46: BGPost - Sensor integration with the CCH

BGPost uses physical and virtual servers, to deploy the tools and execute the experiments.

All the traffic between the laboratory area and the Internet is managed from a single point, through specific filtering devices that allow a strict monitoring of the incoming/outgoing traffic.

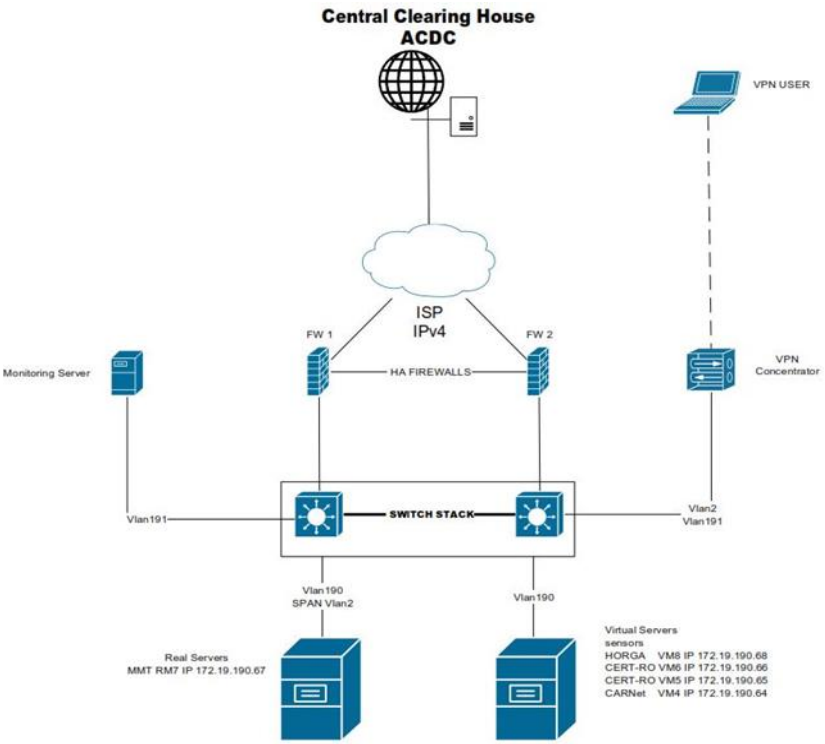

Figure 47:: BGPost Testbed environment – 2014

To include the partners' sensors, and connect with the CCH, BGPost has changed the above topology in 2015

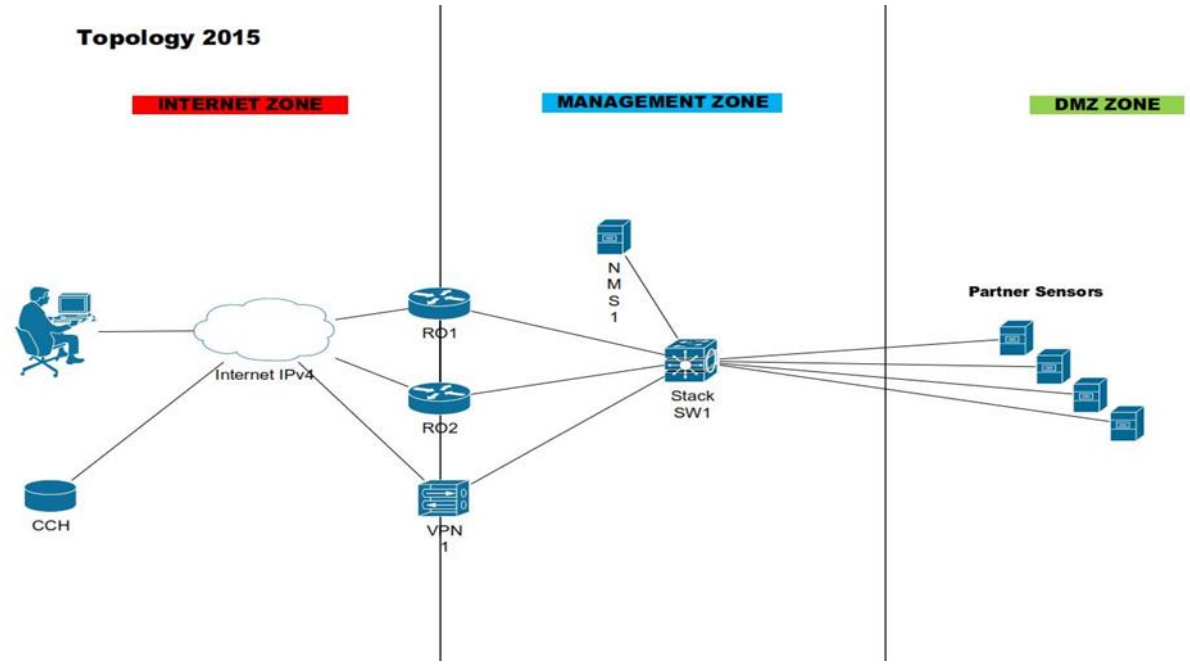

Figure 48:BGPost, new testbed environment 2015

4.2 INCIBEs Toolset integration

4.2.1 CONAN mobile

The Conan Mobile tool and back-end are described at deliverable D2.3 and it is also available at google play:

https://play.google.com/store/apps/details?id=es.inteco.conanmobile

It is also published on the Spanish National Support Centre:

http://www.osi.es/es/conan-mobile

As this tool is installed on the end-user device it is used to directly notify users and warn them about dangerous configuration of their devices, connections done to malicious sites, malicious APKs installed and finally thanks to the GCM service it is also possible to notify them about any warn generated from the NSC or CERT. Taking advantage of this direct interaction with the user, notification and mitigation are done without any further action. Be-sides, Conan Mobile has been integrated with the INCIBE's Anti-Botnet Service, which is provided from the Spanish NSC.

Direct contribution regarding sharing data, is to send malicious and suspicious APKs detect-ed. For those suspicious APKS which are available on the backend it is sent the binary too. This data can help other partners to know which malicious APKs are currently installed and used by users and try to prevent them to install them. It can also increase the intelligent of the model because it provides malware samples that can be further analysed by other par-ners and may help in knowing how the malware is distributed and acting through correlation and aggregation actions.

4.2.2 Skanna

Skanna is described at deliverable D2.3 and it is also described at the Community Portal on the Tools & Services section.

Currently, Skanna can analyse any domain, but main purpose of Skanna is to analyse and detect vulnerable and malicious websites hosted in Spain: .ES websites and .COM, .net websites hosted in Spain. Mitigation of these incidents detected by SKANNA are reported directly to the entity in Spain responsible of .ES TLD and also is notified to websites owners and hosting ISPs by INCIBE's CERT team. This makes not necessary to send to the CCH all the malicious events detected by SKANNA because they are treated and notified internally in Spain.

Direct contribution regarding sharing data, is sharing those URIs which have malicious behaviours and can be potentially useful to other partners. In concrete, it shares URIs that are distributing malware and any other URI that have malicious JavaScript patterns. Only those events with a high reliability are sent to the CCH. With this type of info partners can do blacklists to prevent their users to be in a risky situation and become themselves infected. It also might help in the creation and improvement of the intelligence model of the partners, given them more data or evidences to include in their models. Other types of detections such as defacements

4.2.3 INUC

INUC is the name given to the collaboration established between CERTs within the project. It is sent through the CCH, URIs with any incident detected related to the TLD of each of the CERTs involved. The CERTs are from: Croatia, Romania, Germany, Italy and Portugal. The URIs are obtained from

internal and external sources and it is only legally possible for IN-CIBE to share them with national CERTs responsible of the national TLDs affected for actions related with notification and mitigation.

4.2.4 Flux-Detect

Flux-Detect is described at deliverable D2.3 and it is also described at the Community Portal on the Tools & Services section.

Flux-Detect is fed by an external list of domains and analyse them obtaining domains and IPs involved in Fast-Flux activities. For each domain detected as doing fast-flux activities it is monitored until it ends its malicious activity. In the scope of Flux-Detect a notification is formed by an IP and the domain associated, it implies that the same IP may be notified more than once if it is resolved by different Fast-Flux domains. There is also a time window in which it is not allow notifying the same pair IP/Domain that has already been notified.

Direct contribution regarding sharing data, is sending to the CCH all the IPs detected and the domain associated to them. In addition, those IPs belonging to ASNs of the INCIBE's constituency are directly notified by the CERT team and for the domains it is established an operation flow to notify all of them hosted in Spain.

4.2.5 Whois

INCIBE-Whois is described at deliverable D2.3 and it is also described at the Community Potal on the Tools & Services section.

INCIBE-Whois is offered as a service and offers abuse contact search with great efficiency. It does not share data directly with the CCH, instead it can be used by any partner to obtain abuse contact information. It is already actively used by the INCIBE's CERT team and the Anti-Botnet Service of the Spanish NSC. It is also open to other CERTs and stakeholders within the project. It must be completed a formal request and sign the terms and conditions of use to grant the company access to the service.

4.2.6 Evidence Seeker

Evidence Seeker tool is described at deliverable D2.3 and it is also described at the Community Portal on the Tools & Services section.

Evidence Seeker analyses Apache logs or any log with the same structure as an Apache log to extract and group by abuse contact the evidences found. It is packed with all the components needed to its execution and the documentation associate. It does not share data with the CCH, it is a stand-alone tool provided as a free software under the GNU General Public License available to all the partners within the consortium.

4.2.7 Spanish NSC

Spanish NSC is described at deliverable D2.3 and it is also described at the Community Portal on the Tools & Services section.

NSC global section and pages below it can be found at http://www.osi.es/es/servicio-antibotnet. NSC notify end-users if their public IP connection is involved on botnet activity. It is done through the use of an online check IP service or directly installing a browser plugin. To achieve this objective NSC is fed by internal and external sources, such as ACDC, in order to maintain a botnet and IPs evidence database. Users can also download cleaners or consult detailed information about botnets included in

the service. Besides, it is established a procedure in collaboration with Telefonica to notify botnet affected end-users.

4.2.8 Integration on ACDC (CCH Interface)

To integrate the tools provided into ACDC it was necessary to develop an interface to communicate with the CCH. It includes both, the part for sending data and the client to receive data. Internally all data flows pass through a SIEM solution.

The global infrastructure can be seen in the following picture:

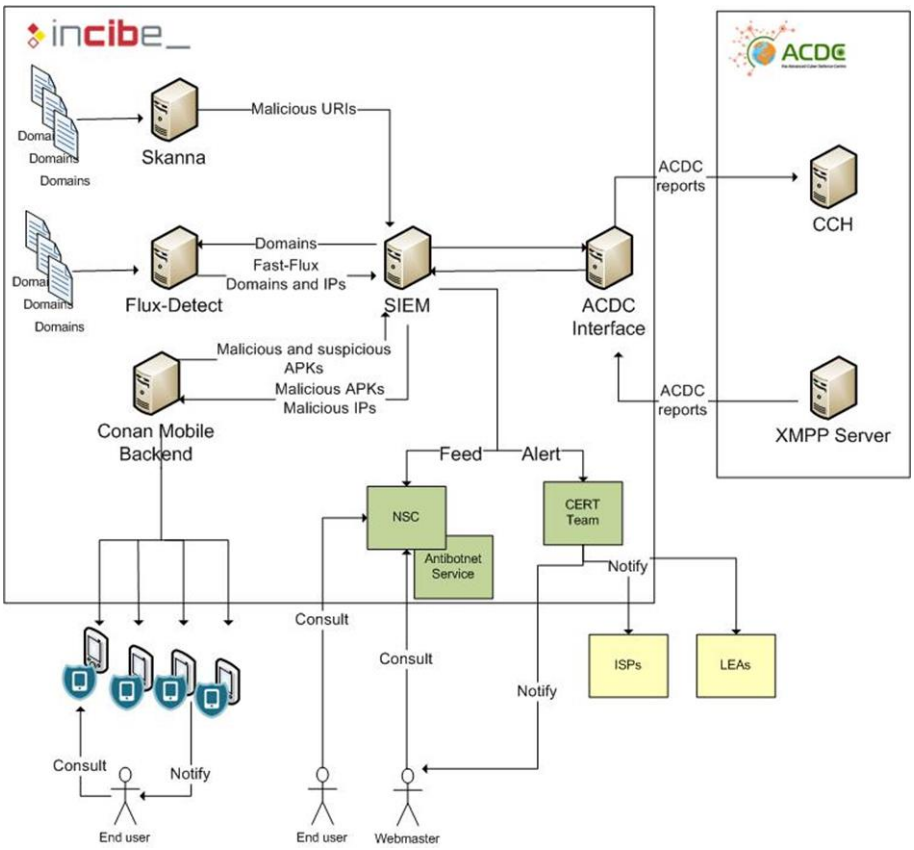

Figure 49: Infrastructure with CCH and End User Support

5 Support Tools / Collaborative Tools

A set of collaborative tools is provided by the partner TEC to facilitate cooperation within the project and to assist in the coordination work.

These tools are:

A versioning system for keeping track of documents, and a mailing list system for information exchange. Both are described in “D 7.1 Project internal IT communication infrastructure”. The versioning system for documents has been replaced by an Office Workspace.

ECO provides a web based team-collaboration software on the ACDC Project Page. It allows the partners to manage project tasks, schedule meetings and events, and share every kind of electronic information. All features are easily accessible within a single web page.

We have chosen Feng office (www.fengoffice.com), hosted on the ACDCs Project Website.

The login is password protected and accessible via: https://workspace.acdc-project.eu (SSL secured).

Once logged in, the User is presented with the main view:

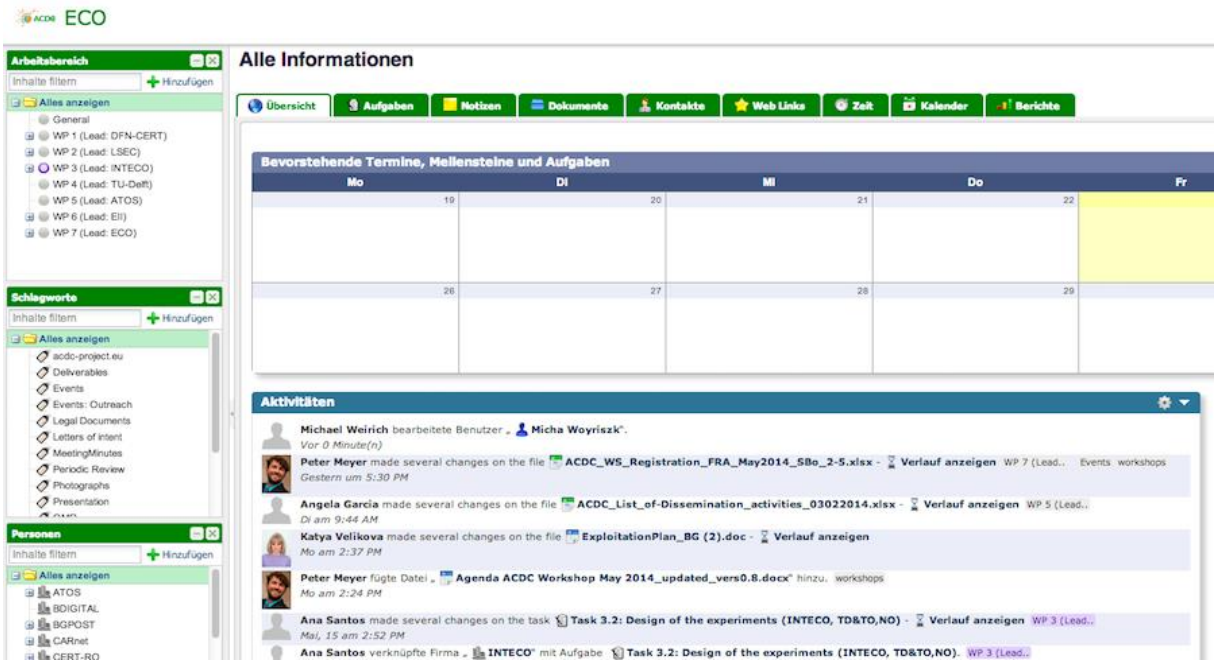

Figure 50: ACDC Workspace – Feng

6 Protocols

6.1 Overview:

In accordance with the DoW for Task 1.1.4 Malicious or Vulnerable Website Analysis, the communication protocol defined in D1.4 is used throughout the ACDC solution. This section only provides a brief overview of that protocol, for more details please refer to deliverableD1.4.

The protocol defined in D1.4 allows ACDC tools to send and receive data to and from the CCH. Tool owners are required to implement this protocol to communicate with the CCH. There might be cases when it´s not feasible to send data to the CCH for example when a Tool wants to transmit large volume data or wants to exchange data with another tool directly. In this case, the tools' maintainers may define and implement an appropriate communication protocol which suits their particular requirements.

To deliver new or updated data sets to the CCH, tools engage in a publish message exchange using the respective connection. Note that publish does not imply that the data will be made publicly available but is a technical term that refers to "the publish-subscribe" pattern implemented by the protocol design. The response by the CCH will include the ID it assigned to the data set, so that tools can easily extend or update the data they delivered as new analysis results become available.

To ensure the confidentiality and integrity of the data exchanged as well as to authenticate the communication peers, connections are wrapped in Transport Layer Security (TLS) sessions. By avoiding third party public key infrastructures, ACDC ensures that it will not be affected by rogue or compromised third party certificate authorities. Tool operators will be able to receive certificates from the partner operating the Community Platform (CP), issued only after verifying the tool operator's identity through at least two independent communication channels. To verify the identity of the CCH, certificate pinning is used, i.e. the hash value of the certificate presented by the CCH is compared against a pre-shared value. Again, this value has to be received and verified using at least two independent communication channels. Deliverable 1.4 lays out additional details with regard to TLS, including a preferred cipher suite.

In summary, the protocol described in D1.4 implements the publish-subscribe pattern, decoupling data generators, i.e. ACDC Tools acting as sensors, and consumers, i.e. ACDC Tools that provide additional analysis using data provided through the CCH. When acting in the former role, a tool may transmit data sets or updated data sets while in the latter role, tools receive notifications for data sets matching their subscriptions and can then explicitly retrieve them using the ID provided. To ensure confidentiality and mutual authentication, TLS is used with certificates issued only after parties have been authenticated through at least two independent channels.

6.2 Data Access Management

The DAM (Data Access Management) component is the part of the ACDC community portal (CP) that provides user-interfaces enabling the ACDC community to share information about incidents and botnet findings through the CCH. Users working for an organization member of the ACDC community can use the DAM to:

- Manage the API-keys allowing CCH-clients to provide and retrieve data from the CCH

- Set the sharing policies that control the sharing of the data provided with other community members.

Additionally, the DAM provides the CCH managers with ways to constraint the CCH functionalities assigned to group of organizations within the ACDC community.

The data exchange between the DAM module and the CCH will be done through REST technology (Representational State Transfer) over an SSL connection. HTTP requests are protected by authentication based on an access token (API-Key) shared between the CCH and the Community portal at deployment time.

The DAM and the community portal provide an interface to require and manage keys provided from the CCH. The CCH provides the service layer (based on REST notation), used by the DAM as a backend. Only additional information about access tokens and connection parameters are stored in the community portal (DAM Entities in the picture), while the rest of data (including the access tokens itself) are provided by the CCH at runtime.

The complete details on how this interface is implemented, the roles involved and the policies that govern the sharing of data stored in the CCH are described in deliverable D6.2.1.

6.3 Data input:

If a user intends to send data to the CCH or wants to receive a data feed from the CCH, he first needs to be identified and it must be checked whether he's allowed to access the CCH or not.

The CP which stores all user specific data and relationships will perform this task.

A data source, e.q. sensor has to identify itself on the CCH API server with the credentials he received by the DAM. Then the sensor can send its information to the CCH, as a JSON.

We have chosen JSON even though the initial concept defined having no limitations on data (format) submissions, it has been agreed on across the project participants, that a basic standardisation of the submitted data fields and basic requirement on mandatory fields simplifies the data submission and retrieval. These specifications have been defined as the ACDC - “Schemata”. These have been outlined and defined in the Deliverables D1.7.1/2 “Data Formats Specification”.

6.4 Data distribution:

Besides the collection and processing of data, the CCH is also designed to distribute data to stakeholders like ISP´s, government agencies, law enforcement, research groups or industry partners.

Due to security concerns, but also with the project aiming to support an open community of stakeholders, the access management has been integrated into the Community Website / Community Portal of the project. The end user policy enforcement is part of the CCH.

One approach of ACDC is mutual data sharing across international borders and an advanced capability of interaction, permissions and restrictions between the involved stakeholders. The Community Portal manages the stakeholder / User database and maintains their relationships. This portal and further settings are described in the deliverable D6.2.1 ACDC Social Platform.

Data sharing and data distribution is handled by the Centralized Clearing House via its integrated XMPP server. Every (read) API Key connects to his own channel where the relevant data – the datasets this key is allowed to see – is streamed into in real time.

Datasets from Keys, which this Channels owner is allowed to see, are also streamed into the XMPP channel.

The data is streamed in the XMPP channel by following rules:

- If an organization has declared IPs, a IP range or a ASN Range in the Community Portal, then every incident that falls within that IP Range is automatically sent to the XMPP channel.
- a read-key will get all reports from write-keys that are connected to this particular read key e.g. from all Keys that have accepted to share data with the specified read key.

To identify which Key sent the dataset and to include additional Information, collected by the CCH to a given report, first a set of "Meta Data" is send. After the Metadata, the original report is streamed. The report is delivered in the schema it was submitted to the CCH.

7 State of the Art Catalogue of Tools

- The table in Deliverable 2.3 "Technology Framework" summarizes the State of the Art (SotA) Catalogue of ACDC tools that the different partners of ACDC bring into the project. These tools provide different types of features and interoperate, by means of the data exchanged through the CCH, to constitute the ACDC Solutions that will be tested in the context of WP3.
- The table which can be found in Deliverable 2.3 classifies the tools according to the categories defined in the taxonomy, gives a brief summary of the tool and provides links to the section in the Annex 9 of the D2.3 document in which each of the tools is described in detail.
- The SotA Catalogue, can also be found as a list of present tools on the ACDC CP (Community Platform), which is maintained regularly and kept updated by the tool owners.