Proactive Risk Management through Improved Cyber Situational Awareness

Start Date of Project: 2016-09-01

Duration: 36 months

D7.3 - Pilot Evaluation Report

Deliverable Details	
Deliverable Number	**D7.3**
Revision Number	**E**
Author(s)	**SYNYO, UOXF**
Due Date	**09/18**
Delivered Date	**16/10/2018**
Reviewed by	**OXFORD, ITTI, AIT**
Dissemination Level	**PU**
Contact Person EC	**Alina-Maria Bercea**

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under grant agreement no 700071.

Contributing Partners

Contributing Partners	
1.	SYNYO
2.	UOXF
3.	AIT – Review and Feedback
4.	All – input, feedback and the efficient execution of the pilot

Revision History

Revision	By	Date	Changes
E	SYNYO/UOXF	16/10/2018	Version submitted to REA
A3	SYNYO/UOXF	15/10/2018	Finalised document after proofreading.
A2	SYNYO/UOXF	11/10/2018	Added events, evaluations.
A1	SYNYO	10/07/2018	Initial draft and table of content.

Executive Summary

Pilot 1 consisted of two primary phases, Phase 1 and Phase 2, each with their specific purpose and goal. Phase 1 main focus was to enable and validate the data sharing within the PROTECTIVE project, and during which the operators at the three NREN sites deployed:

- Warden, to facilitate sharing within and across the three constituencies,
- Mentat, the storage facility for the collected alerts,
- HAWAT, the user interface for navigating, searching and reading alerts,
- A collection of existing and newly developed connectors, which enable further types of alerts being fed into the system.

During Phase 2 the front-end was replaced with Prot-Dash, the official PROTECTIVE user interface, and the focus shifted towards the evaluation of the new interface as well as validating the individual components developed in PROTECTIVE, namely context awareness, correlation, alert quality and trust, and alert prioritisation, together with the operators. During this phase, specific questions were raised and the methodology of each of the components discussed, in order to have the operators at each of the three pilot sites validate the methodology used in each component and discussing how the assumption correspond with the daily work of the operators.

The data collection for the data used to support this document, consisted of surveys based on questionnaires, focus group discussions, operator interviews and discussions, in order to elicit the feedback and impressions from the operators. Further, selected technical measurement points was evaluated in order to identify how the information sharing benefited the participants.

Contents

Executive Summary 3
1 Introduction 9
1.1 Ethical Considerations and Reflections 9
1.2 Report Overview 10
2 Pilot Overview 10
2.1 Locations 10
2.2 Evaluated Components and Sharing Infrastructure 10
2.3 Data Collection Activities 12
2.4 Timeline 12
2.5 Internal Communication 13
2.6 Key Findings 13
2.6.1 System as a Whole 13
2.6.2 SC1 - System and Sensor Data Statistics and SC3 - Time Series and Trend Monitoring 13
2.6.3 SC2 - Reputation of Malicious Entities 14
2.6.4 SC6 - Sharing of Threat Intelligence 14
2.6.5 SC7 – Context Awareness 15
3 Pilot 1 - Phase 1 16
3.1 Overview of Activities 16
3.2 Feedback Logger Results 17
3.3 Questionnaire Results 19
3.4 Commentary on Phase 1 results 22
4 Pilot 1 - Phase 2 22
4.1 PROT-Dash overview 22
4.1.1 Dashboard 23
4.1.2 Alert Search 23
4.1.3 Statistics 24
4.2 Overview of Activities 24
4.3 Feedback Logger Results 25
4.4 Questionnaire Results 26
4.5 Commentary on Phase 2 Results 30
5 Pilot Findings 30
5.1 Outcomes from Feedback logs, Questionnaires and Pilot Procedures 30
5.2 Technical Aspects and Measurements 32
6 Lessons Learned, Action Points and Identified Requirements 35
6.1 Action Points for Pilot 2 35

6.2 "Nice to Have" Requirements 35
6.3 Documentation of PROTECTIVE as an Open Source Project 36
7 Summary 37
Appendix A: Connectors 38
Appendix B: Operator Guide for Pilot 1 Phase 1 39
Overview 39
Beginning of Pilot (15th Feb - 7th March): 39
Main Part of Phase 1 (7th March - 27th April): 39
Pilot timeline 40
Key purpose of the pilot 41
Key tasks of operators during the pilot 41
Webinars 42
Using the system 43
Operator Tasks 46
Beginning of Pilot (15th Feb - 7th March): 46
Deployment 46
Daily - Mid-Phase (7th March- 27th April): 46
Using the system 46
Feedback logger 46
End of pilot (27th April - 4th May): 47
Appendix C: Operator Guide for Pilot 1 Phase 2 49
Overview 49
Operator Tasks 50
Beginning of Phase 2 (11 July + 13 July): 50
Main Part of Phase 2 (16th July - 24th August): 50
Pilot 1 - Phase 2 timeline 50
Key purpose of Phase 2 51
Key tasks of operators during the pilot 51
Webinars 53
Feedback Logger 54
Appendix D: PROTECTIVE Pilot 1, Phase 2 – Last Questionnaire 56

s

List of Figures

Figure 1: Mapping between Pilot 1 activities and the upcoming Pilot 2 plan 9
Figure 2: Main piloting goals and locations; CESNET, PSNC and RoEduNet 10
Figure 3: Dark green highlighted components reflect the availability to the operators. Light green were directly evaluated with the operators through e.g., detailed discussions of the applied methodology in usability studies and/or interviews 11
Figure 4: PROTECTIVE sharing infrastructure topology of Pilot 1 11
Figure 5: Timeline and main piloting activities 12
Figure 6: Count of Description of Feedback by Comment Type 17
Figure 7: Count of Description by Module 18
Figure 8 Count of Description by Topic 19
Figure 9: Opinions (Likert scale) feedback after the start of Phase 1 20
Figure 10: Opinions (Likert scale) feedback during the middle and end part of Phase 1 20
Figure 11: Screenshot of Dashboard 23
Figure 12: Alert search view 23
Figure 13: Count of Description by Comment Type 25
Figure 14: Count of Description by Module. 25
Figure 15: Count of Description by Topic. 25
Figure 16: Opinions (Likert scale) feedback during middle part of Phase 2 27
Figure 17: Opinions (Likert scale) feedback during the end part of Phase 2 27
Figure 18: CESNET's Likert scale feedback over the pilot 31
Figure 19: PSNC's Likert scale feedback over the pilot 31
Figure 20: RoEduNet's Likert scale feedback over the pilot 32
Figure 21: Distribution of alerts across NRENs and the alert categorization 33
Figure 22: Distribution of alerts across specific connectors and NRENs 33
Figure 23: Distribution of alerts across connector software and NRENs 34
Figure 24: Screenshot of a part of the user manual 37

List of Tables

Table 1: Opinions about the system as a whole 13
Table 2: Opinions about SC1 and SC3 related components 14
Table 3: Opinions about SC2 related components 14
Table 4: Opinions about SC6 related components 15
Table 5: Opinions about SC7 related components 16
Table 6: Pilot 1 Phase 2 activities overview 16
Table 7: Likert Scale Opinions from Start of Phase 1 19
Table 8: Likert Scale Opinions from Mid of Phase 1 20
Table 9: Pilot 1 Phase 2 activities 24
Table 10: Likert Scale Opinions from Mid Phase 2 26
Table 11: Likert Scale Opinions from End of Phase 2 26
Table 12: Specific connector deployments per NREN instance 34
Table 13: Action Points for Pilot 2 35
Table 14: Summary of “Nice to have” features 35
Table 15: Connectors made available during Phase 1 38

List of Abbreviations

Abbreviations	
CTI	CTI
CSIRT	Computer Security Incident Response Team
ENISA	European Union Agency for Network and Information Security
KPI	Key Performance Indicator
MSSP	Managed Security Service Provider
MSP	Managed Service Provider
NREN	National Research and Education Network
SME	Small Medium Enterprise
TI	Threat Intelligence

1 Introduction

In this report we provide an overview of the lessons learned during the two phases of Pilot 1 of the PROTECTIVE project. The purpose is to document the activities that have been performed, the involved parties and to document the outcomes based on the feedback that has been collected during the pilot. This report builds upon D7.1, but provides an updated, detailed view of the activities while answering the initially raised questions for the different scenarios that have been evaluated.

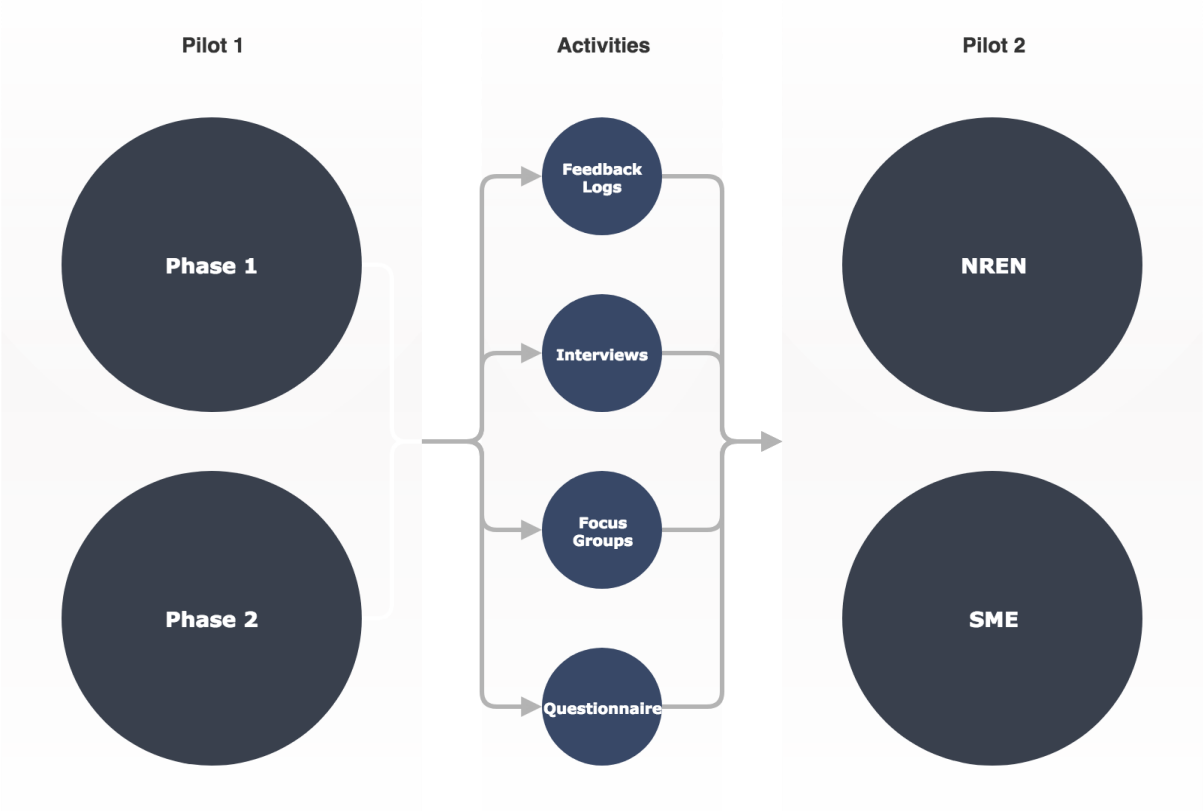


Figure 1: Mapping between Pilot 1 activities and the upcoming Pilot 2 plan

The output and lessons learned from this report feeds into the design and execution of the second pilot, in which external parties will be invited to participate and use the developed system.

Pilot 1 has two types of evaluation in mind:

- **A usability/user-centred evaluation** in which we identify whether the operators perceive PROTECTIVE as useful, and what can be done to improve the tool's usability features, ranging from: deployment, maintenance and day-to-day usage.
- **A technical evaluation** in which we identify whether the tool is able to execute the functionality we set out to achieve, in live environments at reasonable performance.

1.1 Ethical Considerations and Reflections

Ethical considerations for the pilot execution have been prioritised throughout the pilot planning. Additional ethical considerations on a per-system-component level have been dealt with in the majority of the project deliverables. D7.1 overview the ethical considerations specific to Pilot 1, and D7.2 will list ethical considerations for Pilot 2. Ethical considerations and the data management plan can be found in the deliverable D2.5.

1.2 Report Overview

The report is structured as follows; Section 2 provides an overview of the two phases of Pilot 1, documenting the activities performed, the locations as well as the key findings. Section 3 and Section 4 provide a detailed discussion of each of the two phases, respectively, documenting what was evaluated and a detailed description of the outcomes during each of the phases. Section 5 provides an overview of the piloting outcomes, based on the detailed outputs of the two phases. Section 6 focuses on lessons learned which directly impact the second pilot that is scheduled for February 2019. Finally, Section 7 concludes the report.

2 Pilot Overview

In this section, we provide an overview of the piloting activities, providing a summarised view of the activities initially outlined in D7.1, and summarize the overall structure of the pilots regarding data collection, evaluation and methodology. This consists of providing an overview of the piloting activities and the overall goal, it also includes a description of what has been achieved as well as a summary of the main findings across the pilot sites.

2.1 Locations

The piloting activities were performed across three different European countries, namely at the PROTECTIVE pilot partners in Czech Republic, Poland and Romania.

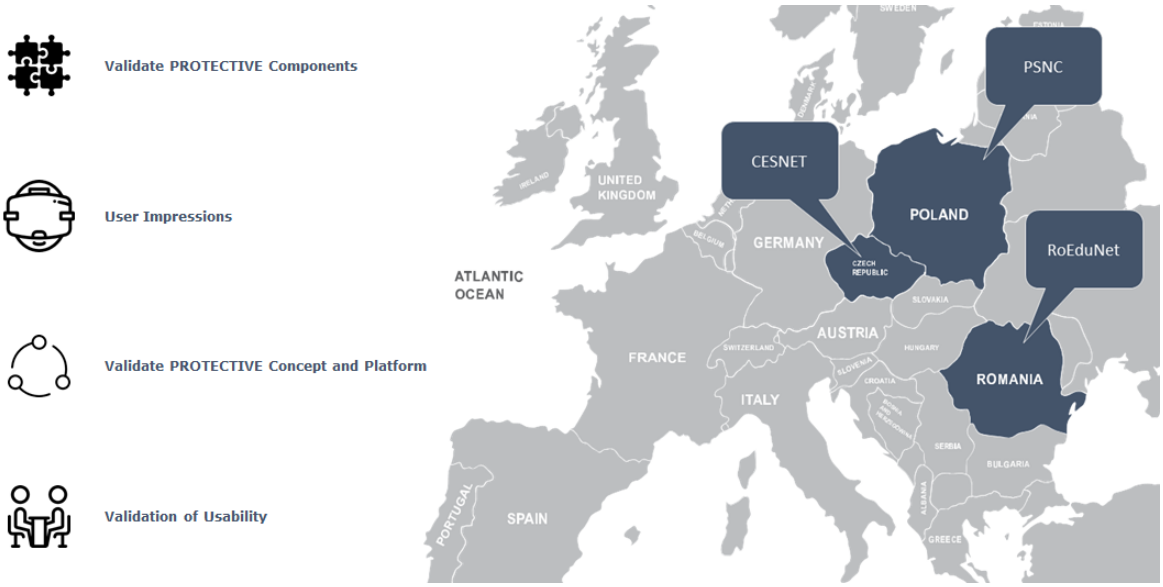


Figure 2: Main piloting goals and locations; CESNET, PSNC and RoEduNet

It is worth noting that GMV ran a separate instance of PROTECTIVE. GMV's engagement did not generate or share events, but was a listening node to facilitate support. Each time there was a support concern, GMV would be able to check whether the issue affected everyone or only a particular node.

2.2 Evaluated Components and Sharing Infrastructure

Operators at PROTECTIVE NRENs were exposed to the core parts of the PROTECTIVE system as illustrated in Figure 3. Components highlighted in dark green were made available throughout the pilot, whereas components highlighted with light green were directly evaluated with the operators through e.g., detailed discussions of the applied methodology in usability studies and/or interviews.

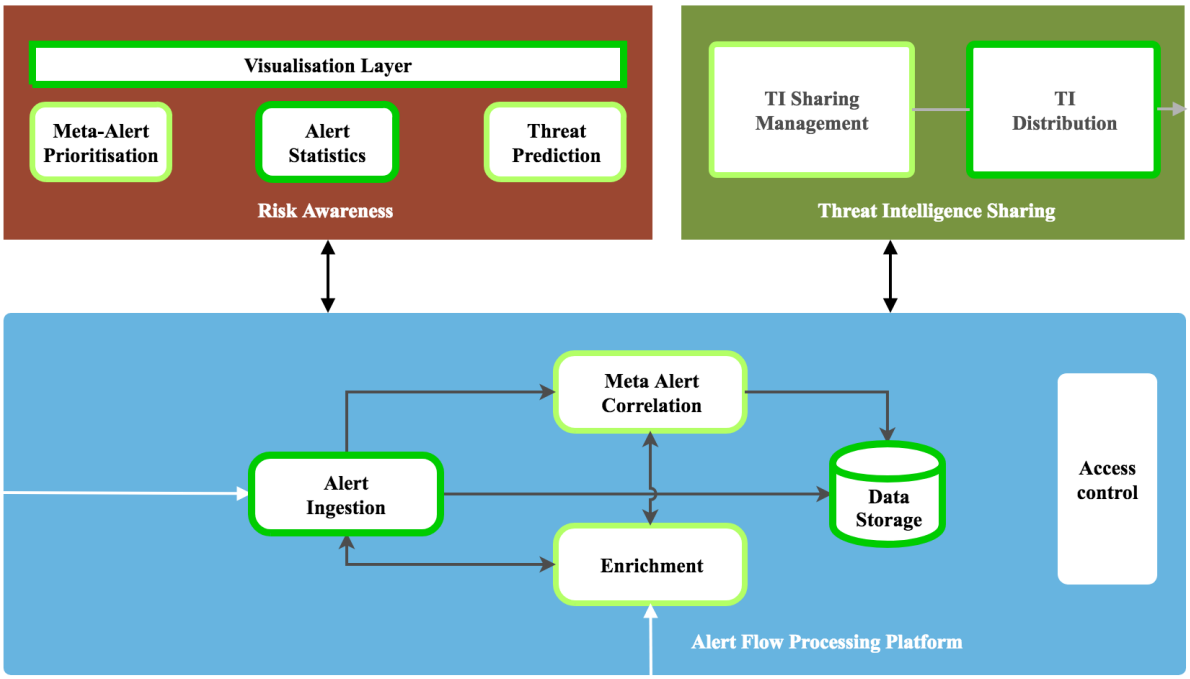


Figure 3: Dark green highlighted components reflect the availability to the operators. Light green were directly evaluated with the operators through e.g., detailed discussions of the applied methodology in usability studies and/or interviews

During the pilot, the three NRENs were connected in a peer-2-peer architecture, enabling each of the NRENs to decide **what** information is shared and with **whom**, directly from their own instance of the PROTECTIVE system and independent of any central authority. This also made it possible for the NRENs to install and use the PROTECTIVE system as a tool for internal monitoring, without having to disclose all the information being collected to the other NRENs. Figure 4 shows the sharing model applied during the first pilot.

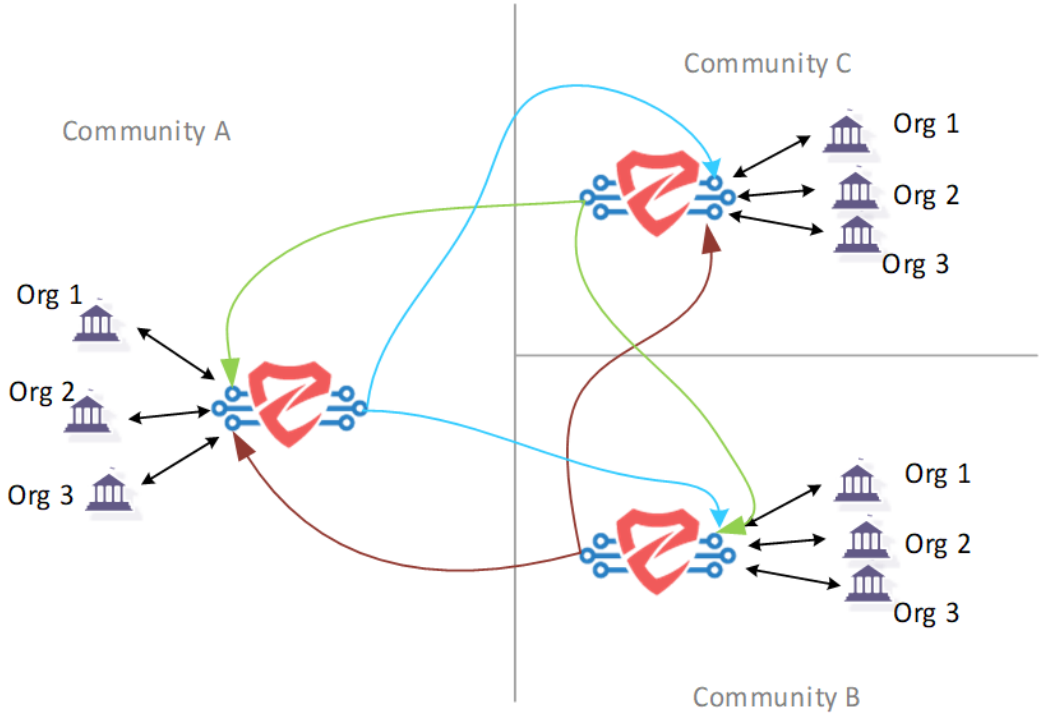


Figure 4: PROTECTIVE sharing infrastructure topology of Pilot 1

2.3 Data Collection Activities

The figure below provides an overview of the overall structure followed in this report, reflecting how the individual data collection activities feed into the overall evaluation of Pilot 1.

Feedback Logs -- Each of the NRENs involved were provided with a reporting tool, through which they could report various events, bug reports and feature requests. While the tool was available throughout the entire duration of Pilot 1, the logs were collected on a weekly basis as to enable monitoring of the progress throughout the pilot itself. For immediate support, e.g., technical support for the installation or configuration or the system, a piloting mailing list was provided.

Questionnaires -- At strategic points during Pilot 1, the operators were asked to fill out questionnaires in which they were asked about their overall experience with the system, technical issues, success stories etc.

Interviews and Focus Group Discussions -- In parallel to the above data collection tools, interviews with the operators were held, discussing both the overall experience with the PROTECTIVE system, as well as individual interviews with operators regarding the methodologies applied in various PROTECTIVE components and to validate initial assumptions. These shared sessions were motivated by having a collaborative exchange between the operators, during which they could potentially exchange their experience within their own domain, while the individual sessions allowed for more in-depth discussions.

Usability Studies -- Dedicated sessions were set up with the operators in order to discuss the usability of the user facing parts of the solution. Here, the operators had the opportunity to provide feedback on specific aspects in the user interface, what information should or could be provided.

Technical Measurements -- Selected metrics and statistics were collected as part of the pilot, covering the amount of exchanged information between the NRENs, the type of alerts being exchange and from where they were generated. These are documented in Section 5.2.

2.4 Timeline

Overall, Pilot 1 consisted of two phases, each with their own software deliveries and specific scopes, a high-level overview is provided in Figure 5. Later sections detail and discuss the specific activities that were performed, and elaborate on their purpose.

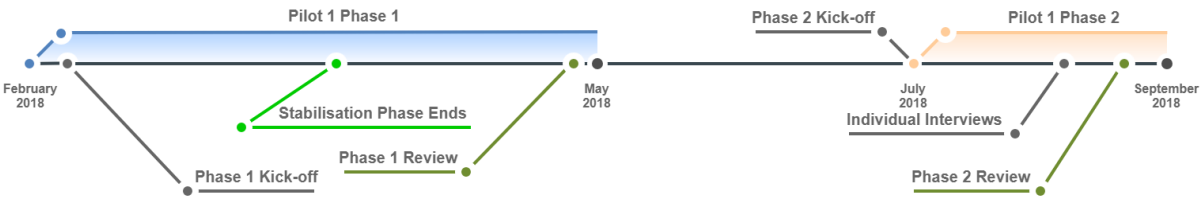


Figure 5: Timeline and main piloting activities

In summary, Phase 1 of Pilot one focused on the deployment and installation of the PROTECTIVE system into the constituency of the NREN, enabling the backbone on which the additional functionality could be added. The second phase focused on specific aspects of the PROTECTIVE concept, and more in-depth discussions of the methodologies and concepts applied.

2.5 Internal Communication

In order to ease the communication between the partners, an internal mailing list was established, where the operators could easily interact with the consortium members, regarding installation of the system and how to use various aspects. The mailing list is being used as one of the sources of input when writing up this document, and, while was perfectly suitable for a limited group, as was the case during the first pilot, the second pilot, and beyond, will consider using collaborative tools provided by GitHub, making the information publicly available, and are more persistent, as to help the community.

2.6 Key Findings

Here we document the key finding, based on the questions defined in "D7.1 - Detailed Pilot Plan v1", in order to provide an overview of the main lessons learned. The answers are a summary of the elaborated analysis of the data collected as part of the piloting activities and documented throughout the rest of this report.

2.6.1 System as a Whole

Table 1: Opinions about the system as a whole

QID:	Question	Methodology
GCE.Q1	What is the operator general opinion of the tool?	Focus Group Discussions
GCE.Q2	How much is e.g., downtime of core services reduced, as a function of the PROTECTIVE system?	Interviews
GCE.Q3	What is the increase in efficiency of the operator - taking the maintenance efforts[1] into account?	Interviews

GCE.Q1 Overall the feedback was positive, with selected success stories. The majority of the discussions, both during the deployment of the PROTECTIVE and as part of the overall feedback, was oriented towards the deployment and installation process itself.

→ Effort is being undertaken to integrate the feedback towards simplifying the installation procedures, both from the point of the process as well as addressing ambiguities in the installation guides.

GCE.Q2 No core system attacks were reflected in the data that was collected during the first pilot.

→ For Pilot 2, reports will be generated that count the number of hits between alerts and their applicability (e.g. make a log of all received alerts that apply to the local constituency) in terms of whether they affect the specific constituency.

GCE.Q3 The focus of the first pilot was on improving the data sharing, and a general deployment of the PROTECTIVE system and connectors.

→ As with QCE.Q1, additional focus will be provided towards simplifying the installation procedure and the documentation of the various components.

2.6.2 SC1 - System and Sensor Data Statistics and SC3 - Time Series and Trend Monitoring

Initially, SC3 was not intended to be evaluated during Pilot 1, however, a combined tool for supporting both Scenario 1 and Scenario 3 was made available to the operator during Phase 2 of Pilot 1. They are both described here, since the questions defined for SC1 also apply for SC3. Use-case 1 is supposed to

[1] Maintenance refers to the administration and management of the individual components, such as; defining and updating contextual information, maintenance of connectors.

enable the operators to easily define the various statistics which they are interested in seeing, either from a system perspective, e.g., from what constituencies alerts have been received or how active certain sensors are, or from a content point of view, e.g., what type of alerts have been received based on their category. On the other hand, Scenario 3 build atop of this information and adds another dimension, namely how the various statistics have behaved over time, which is then used to evaluate overall trends.

Table 2: Opinions about SC1 and SC3 related components

QID:	Question	Methodology
SC1.Q1	What is the benefit of the overview?	Interviews, usability tests
SC1.Q2	How much more efficient does the operator workflow become?	Interviews
SC1.Q3	What is the timeliness and accuracy of the information being provided to the users?	Point measurements
SC1.Q4	Which are the primary features used and how?	Interviews

SC1.Q1 The operators are now capable of defining exactly what they wish to see and how they wish to express it based on their needs and preferences.

SC1.Q2 The operators are provided with additional information which may include additional information that they would not necessarily have had access to without PROTECTIVE.

SC1.Q3 Since the PROTECTIVE installations mainly focused on sharing of alerts, limited processing delay was introduced by the system itself, and thus the main delay would originate from the sensors or connectors that were used.

SC1.Q4 The operators primarily used the search functionality to check if their assets were reflected in the collected information.

→ Improvements will be made on making it more straightforward to find specific assets through the search functionality provided by the system, e.g., to enable the operators to search by IP ranges.

2.6.3 SC2 - Reputation of Malicious Entities

Table 3: Opinions about SC2 related components

QID:	Question	Methodology
SC2.Q1	Would an operator use detailed threat information on a daily basis?	Interviews, Questionnaires and Diaries
SC2.Q2	Accuracy of the provided metrics?	Usability tests

SC2.Q1 Depending on the situation, but in particular, the detailed threat information would be used to retrospectively do forensics work to identify events leading up to a certain event.

SC2.Q2 The operators generally agreed with the methodology applied. While the operators were informed of the general methodology, they were not presented with the values of other NRENs.

2.6.4 SC6 - Sharing of Threat Intelligence

Table 4: Opinions about SC6 related components

QID:	Question	Methodology
SC6.Q1	How does the operator's/user's view improve with the added information?	Technical measurement points, Interviews, Questionnaires and Diaries
SC6.Q2	What technical, human factor and ethical issues emerge during TI sharing?	Interviews, Questionnaires and Diaries
SC6.Q3	What is the increased advantage of cross-constituency sharing over the increased amount of information?	Technical measurement points

SC6.Q1 Operators are being provided with information that their own system would not necessarily have detected, thus providing improved insights and more context to the situational awareness.

SC6.Q2 In order to answer this question, mid-way through Pilot 1, we decided to organise a focus-group workshop, dissemination and networking event, with an audience of 20 participants from 13 different CSIRTs. This event is detailed in D8.8. The audience made up a set of expert stakeholders who have a vested interested in "GDPR compliant threat intelligence". Specifically, the workshop:

- Documented key challenges that CSIRTs have faced in recent months with regards to GDPR as well as present day identifying present day capabilities and limitations in existing systems.
- Had CSIRTs present and peer-review solutions on how CTI-GDPR challenges can be addressed moving forward. These will feed into a CTI-sharing community policy document that PROTECTIVE will use in pilot 2.
- Acted as a networking event between CSIRTs.
- Promoted recruitment for Pilot 2.

The workshop helped us generate use-cases of policies for a CTI-sharing community agreement and peer-review potential rulesets for PROTECTIVE's Information Sharing Compliance module The outcome of this work, will be published in a separate white paper/technical report, estimated to be published January 2019.

SC6.Q3 External parties may detect activities in which an attack may originate from inside another constituency, which is made possible to detect through cross-constituency sharing of threat information. A specific example in which this was validated, is that RoEduNet managed to identify malicious activity coming from their constituency and that was targeted towards CESNET. At the same time, some of the alerts that were shared across constituencies were anonymised, reducing a direct mapping between an alert and a specific constituency.

→ A possible improvement when it comes to sharing of anonymised CTI might be, though communities that share (partially) non-anonymised CTI with the appropriate constituencies.

2.6.5 SC7 – Context Awareness

Context awareness is a primary input for the prioritisation of meta-alerts. While the impact of the context on the prioritisation is validated in SC4, the validation of SC7 focuses on the usability aspects of the management of a given constitution. It is also used to estimate the complexity of maintaining the context awareness mechanisms versus the expected benefits that may be achieved through it in terms of relevant and accurate information. While the former may be rated through short-term interaction with the operators, e.g., interviews and simulations, the latter requires long term monitoring of the situation to capture how the operators tackle potential changes to the configuration over time.

Table 5: Opinions about SC7 related components

QID:	Question	Methodology
SC7.Q1	Accuracy of the modelling and criticality scoring?	Simulations
SC7.Q2	To what extent does the context awareness actually contribute to the information reduction, e.g. comparatively to investing effort into maintaining an accurate view of the constitution models?	Interviews, Interviews, Questionnaires, Diaries, and Estimations

SC7.Q1 -- Given how the model is defined, it is up to the end-users of the system to specify the level of detail that they need to be reflected in the modelling and scoring of the criticality, thus making this question obsolete.

SC7.Q2 -- The operators were not required to model their own assets during Pilot 1, as the feature is not a mandatory component of the processing pipeline.

3 Pilot 1 - Phase 1

The primary focus of Phase 1 was on evaluating the underlying technologies and ensuring the stability of the PROTECTIVE system, the individual components, user experience etc. Specifically;

- To enable CTI sharing across the three NRENs participating in the pilot.
- To deploy existing and newly developed sensors, which plug into the PROTECTIVE solution.
- To gradually introduce the systems to the piloting partners, and learn about installation issues.
- To obtain both usability (including installation) and functionality feedback from end-users to refine the system.

In the first phase, the operators used MENTAT to access, navigate and visualise the alerts that have been collected from both internally deployed sensors as well as from the two other NRENs. An overview of the deployed connectors is provided in Appendix A:.

3.1 Overview of Activities

Pilot 1 Phase 1 lasted from February 12th 2018 to May 2nd 2018. During the pilot the operators were exposed to alerts from other NRENs, thus increasing their situational awareness by providing them with a broader overview of ongoing activities. The first phase, was split up into two stages; a deployment and integration phase, during which the NRENs would deploy the PROTECTIVE system and could provide feedback to the installation procedure, and a second stage, during which the operators would be using the system on a regular basis. Specific events during Phase 1 are listed in the table below. An operator guide of the pilot was given to each operator. This included the timeline shown in Table 6 as well as information about webinars to attend, feedback logger instructions, links to questionnaires to fill in, as well as in-depth information about expectations of the operators and how we will analyse the data, see Appendix B: “Operator Guide for Pilot 1 Phase 1”.

Table 6: Pilot 1 Phase 2 activities overview

When	Activity	Requirements	Comments
Feb 12th	Introduction to PROTECTIVE webinar	Overall PROTECTIVE presentation, NREN key personnel identified	The participating operators were introduced to the PROTECTIVE project as a whole as well as the developed solution that they would be evaluating over the next months.

Feb 12st	Installation and configuration of Warden, connectors and Mentat	Installation guide and method for collecting feedback, a dedicated person who can do the installation	A dedicated session focused on the installation of the PROTECTIVE system and the necessary dependencies. Here the operators could raise any issues or questions regarding the installation itself.
Feb 15th	Initial survey and first impressions	Survey with preliminary questions	The operators were provided with the first questionnaire. The results are documented in Section 3.2 and 4.3.
Feb 12st Feb 28th	Stabilisation and Adaptation	Monitoring of whether the system is running and receiving messages, identification of issues	This phase ensures that PROTECTIVE is running well and is populated with initial data, by the time that the operators are expected to use it.
March 7th	Final collection of feedback on **deployment** experience	Questionnaire is ready	The operators could provide feedback along the entire deployment stage, this only mark when all the data needs to be collected
March 7th	"Introduction to PROTECTIVE" webinar	Elaborated tasks descriptions, specific tasks that can make the operators familiar with the system.	This includes the Mentat tool, how to use the diary (what kind of events are we interested in, escalation, usefulness, bugs etc.)
March 15th	Intermediate survey and interview	Specific questions have been prepared that the operators may address during the discussion.	Discussion with the relevant operators about their experience in order to follow-up on whether they are using the system, and their overall experience until now in order to collect preliminary data.
May 2nd	Final interview	Questions are ready	End of Phase 1

The following sections document and analyse the information and data points that were collected, among others, throughout the above listed activities.

3.2 Feedback Logger Results

The total collected feedback log entries in Phase 1 was 220 (45 + 88 + 87) collected from 13/03/2018 to 28/04/2018. Figures 6-8 present the results collected in terms of: Type of log, Module and Topic.

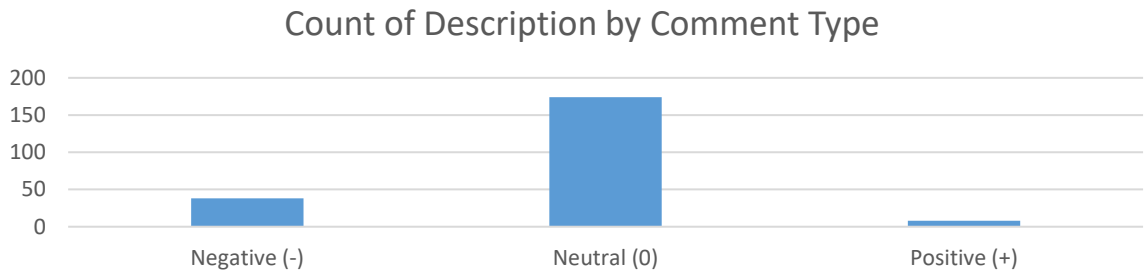


Figure 6: Count of Description of Feedback by Comment Type

Figure 6 shows 38 negative comments, 174 neutral comments and 8 positive comments. A breakdown of reasons given for comment types are the following:

- **Negative (-), 37 logs:**
  - 9 of these related to troubleshooting issues, these were issues with:
    - Connectivity to other NRENs, connectors and using Mentat and Warden.
  - 4 were related to system checks: 1 crash, 2 connectivity (incoming and outgoing issue) to other PROTECTIVE nodes, and 1 issue related to the incoming directory of CTI events.
  - 24 were related to bugs, these were issues with:
    - Saving alert queries.

  - One NREN not receiving data from another NREN – this was due to a misconfiguration.
  - Connectivity.
  - System Exceptions.
- 1 was about other, which was misclassified as a troubleshooting issue of one NREN not receiving CTI from another NREN.

- **Neutral (0), 174 logs:**
  - 173 of these were regular system checks (i.e. nothing noteworthy happened, but PROTECTIVE was used to check for updates).
  - 1 gave a comment on a non-serious python script issue.
- **Positive (+), 8 logs:**
  - 2 about system checks: positive feedback about PROTECTIVE pilot support.
  - 3 about troubleshooting: issues being resolved.
  - 2 about performance: connectivity and ability to receive data at expected performance.
  - 1 about other: troubleshooting issue being resolved.

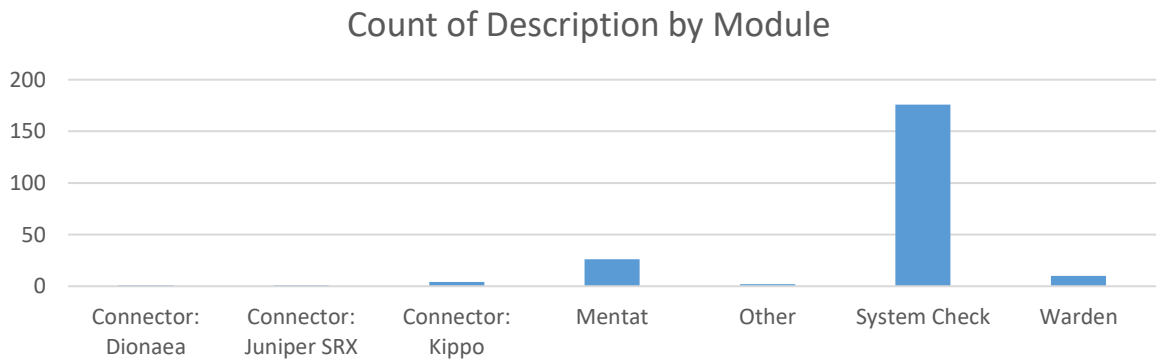


Figure 7: Count of Description by Module

Figure 7 shows 1 comment about Dionaea (a connector), 1 about Juniper SRX (a connector), 4 about Kippo (a connector), 26 about Mentat, 176 System checks, 10 Warden comments and 2 other comments. A breakdown of reasons given for module comments are the following:

- **Connectors** (Dionaea, Juniper SRX and Kippo): 1 performance comment and 5 troubleshooting comments about connectivity issues and pilot support.
- **Mentat**: comments about 22 bugs, 3 troubleshooting issues and 1 system check.
- **System Check**: 176 comments about simply checking the system as a whole and no issue with any module.
- **Warden**: 5 troubleshooting, 2 system check, 2 bugs and 1 performance comments.
- **Other**: 1 positive and 1 negative about connectivity issues, one positive, one negative.

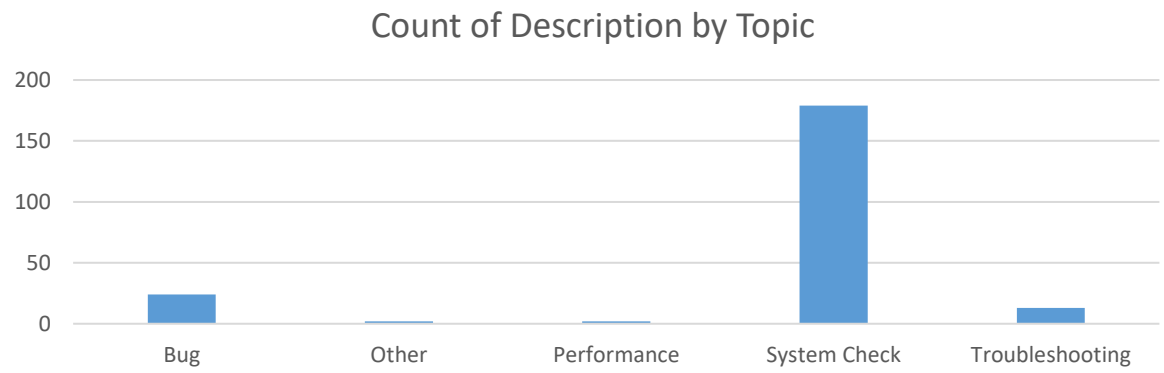


Figure 8 Count of Description by Topic

Figure 8 shows 24 comments about bugs, 2 about other, 2 about performance, 179 about system checks, and 13 about troubleshooting. These issues follow in line with comments from the previous two figures, we therefore will not include a breakdown of these – as that will lead to repeated information. The purpose of this figure is to provide a different perspective of the aforementioned data.

We plan on addressing the negative comments from Phase 1 by improving the installation and usage guide (incl. how to add new connectors guides), enhance the alert querying capabilities, improve system stability, and maintain our existing too support approach. We believe that the volume of positive logs do not necessarily reflect reality as this is highlighted by the questionnaire results.

3.3 Questionnaire Results

The questions used in the questionnaire can be found in Appendix D. Below in Table 7 and 8 follows the Likert-scale results, averages and standard deviations for operator opinions about key aspects of the tool such as their opinion on ease of installation/use, the perceived value of tool aspects including: viewing other NREN alerts, statistics (about the alerts), querying alerts and briefs (see Appendix B for more information about each of these aspects). It is worth noting that the reason "ease of installation" is only added once is because after installation, we believe we cannot ask this again of operators.

Table 7: Likert Scale Opinions from Start of Phase 1

Start Phase 1	CESNET	PSNC	RoEduNet	Avg.	Std.Dev.
Ease of Install?	3	2	4	3	1
Ease of Use?	4	3	5	4	1
Value of viewing other NREN events?	3	4	5	4	1
Value of statistics View?	4	4	4	4	0
Value of Alert Querying?	5	4	5	4.666667	0.57735
Value of Briefs?	4	4	5	4.333333	0.57735
Avg	3.833333	3.5	4.666667		
Std.Dev	0.752773	0.83666	0.516398		

Table 8: Likert Scale Opinions from Mid of Phase 1

Mid Phase 1	CESNET	PSNC	RoEduNet	Avg.	Std.Dev.
Ease of Use?	4	3	5	4	1
Value of viewing other NREN events?	3	5	5	4.333333	1.154701
Value of statistics View?	4	5	5	4.666667	0.57735
Value of Alert Querying?	5	3	5	4.333333	1.154701
Value of Briefs?	4	3	5	4	1
Avg	4	3.8	5		
Std.Dev	0.707107	1.095445	0		

Figures 9 and 10 show the results of Tables 7 and 8 in chart form, comparing each questionnaire having the results of each NREN compared with each other. In section 5, we show these results across all questionnaire cross sections to provide a temporal comparison of these results.

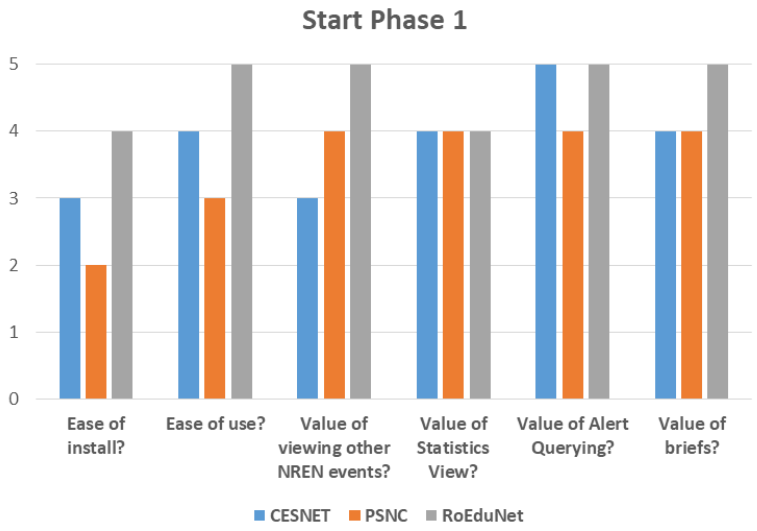


Figure 9: Opinions (Likert scale) feedback after the start of Phase 1

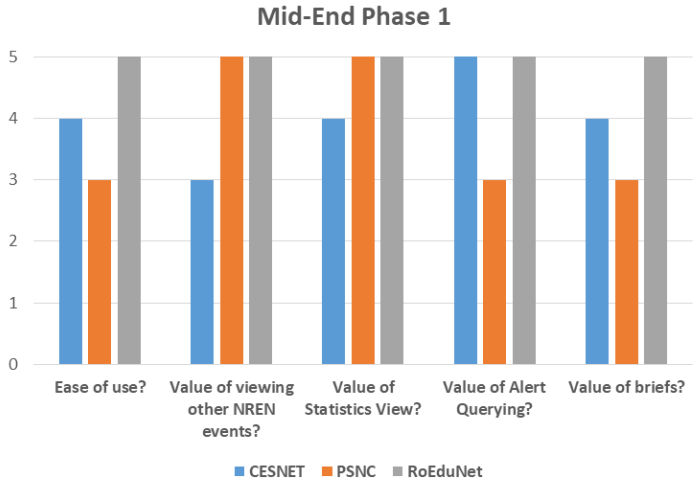


Figure 10: Opinions (Likert scale) feedback during the middle and end part of Phase 1

The feedback logger provides insights into what key issues emerged in the pilot, while the Likert scale provides insight into operators' opinions of different aspects of the tool. The Likert scale values suggests that feedback from the operators is largely positive, but there is room for improvement. Specifically, we make the following key observations:

- The installation process can be improved.
- The value of statistics increased over time. We will therefore look to add new statistics features moving forward.
- There is divided opinions on "ease of use" and "value in viewing other NREN alerts". We will focus on improving these topics moving forward.
- The value of alert querying dropped – the questionnaire results and engagement with operators suggested that more feature need to be implemented for PROTECTIVE to become as fully-fledged as other commercial products.
- The value of briefs (see Appendix B for screenshots) in Mentat is higher than expected. Briefs were not in the PROTECTIVE requirements and specifications. However, as it already existed in the original Mentat GUI, we wanted to assess its value. We are considering implementing it for pilot 2 given its high score.

The likert scale provides insight into operators' opinions of different aspects of the tool, while the remainder of the questionnaire was designed to have operators elaborate on key issues in using the tool. Instead of listing the operator's feedback individually, we will list the key comments made by operators. We categorise them into three main classes: positive, negative and desirables. Below follow the key findings from the questionnaire feedback.

Positive:

- The approaches to presenting or exploring patterns in CTI (e.g. stats, briefs and queries).
- RoEduNet commented: "*we've noticed that one of the faculties had a server compromised which nobody know about*". Specifically, RoEduNet obtained CTI from CESNET which indicated that one of the IP addresses owned by one of RoEduNet's clients, was issuing abnormal requests and trying to scan and brute force CESNET servers. With this information RoEduNet went to the owner of the server and used this information in their incident handling processes. RoEduNet determined that the machine was running a root-kit.
- Despite the issues emerging, the operators still reported that they perceive the tool as useful.
- Level of support from the PROTECTIVE team is appropriate, with comprehensive answers and less than 24h response time.

Negative - needs improvement/Challenges:

- Minor issues appeared w.r.t. installation, configuration and general use. These related to:
  - Installation processes: lack of information about what sudo permissions are required causing system permission issues. There were also confusion about usage of secret keys and certificates in the tool.
  - Installation notes needing improving to remove ambiguity and adding more specificity. For instance, the inclusion of a correct setup of IP tables and sudo permissions. Additional diagrams should be included to provide a better picture of how the tool works conceptually.
  - Time needed to understand the installation and usage – the operators believe the time necessary to understand the system can be reduced by simplify installation and usage guides.
  - How when the system was misconfigured, graphs and briefs were not displayed. This left some confusion in terms of whether the PROTECTIVE tool itself worked or not.
  - How one of the NRENs ran out of storage for alerts on its HDD. The PROTECTIVE tool needs

mechanisms to alert operators when it is about the run out of storage and facilities to help operators migrate stored data and upgrade storage.
    - How more PROTECTIVE logs (including error and warnings) could help operators receive confirmation about correctness of installation and tool usage.
- Docker was perceived as more complex than necessary. It is currently unclear to us whether this concern a learning curve issue or a more general concern (i.e. is Docker generally to complex or is the concern chiefly a learning curve one?).
- Occasionally the operator guide would be out of date.

Desire:

- Single script installation (or install wizard) – there was a strong desire to simplify the installation proves further. Getting the tool up and running, should be a matter of following a script.
- Searching across multiple sources (IP ranges) in the alert database.
- Enable email ticket system integration from PROTECTIVE (clicking some icon near the alert would open desktop e-mail client with filled-in subject and body of the message).

3.4 Commentary on Phase 1 results

Our key take aways from Phase 1 is that it was successful with minor setbacks related to installation and system usage issues (connectivity and crash). We consider such issues to be expected when rolling out a new system.

We plan on addressing the negative comments from Phase 1 by improving the installation and usage guide (incl. how to add new connectors guides), enhance the alert querying capabilities, improve system stability, and maintain our existing too support approach. We believe that the volume of positive logs do not necessarily reflect reality (i.e. the tool was well-received) as this is highlighted by the questionnaire feedback.

We take RoEduNet's comment on how it aided their incident handling as a victory, but acknowledge that more than one such instance would have been desirable to demonstrate success. We expect that more PROTECTIVE nodes would need to more success stories.

The procedures related to the pilot itself were well-received. The response times for support were always less than 24h and comprehensive according to the operators. The operator guide was perceived as complete but occasionally out of date, which we will improve moving forward.

4 Pilot 1 - Phase 2

The scope of the second phase extended the first phase and introduced additional PROTECTIVE components into the deployed system and to the operators. The primary goal was to evaluate specific PROTECTIVE features, namely the PROTECTIVE dashboard (PROT-Dash) user interface, as well as the specific contributions of the PROTECTIVE project. These included the trust, context awareness, correlation and visualisation components. The rest of this section provides an overview of the user-interface that was made available to the operators. It details an overview of the activities performed during the second phase in order to introduce the operators to the system, gather their feedback etc. and finally, an analysis of the feedback is provided.

4.1 PROT-Dash overview

The following section briefly describe the different views that were provided by the PROT-Dash user interfaces, and their primary functionality.

4.1.1 Dashboard

The dashboard view served two main functions during the second phase of the pilot. Firstly, it provided an overall overview of the most recent state of the system, namely what has occurred during the last six hours as well as a detailed view of the most recent events. Secondly, the dashboard serves as a reference guide for how the system may be used and how to actually use it, supported by the user-manual that was developed as well.

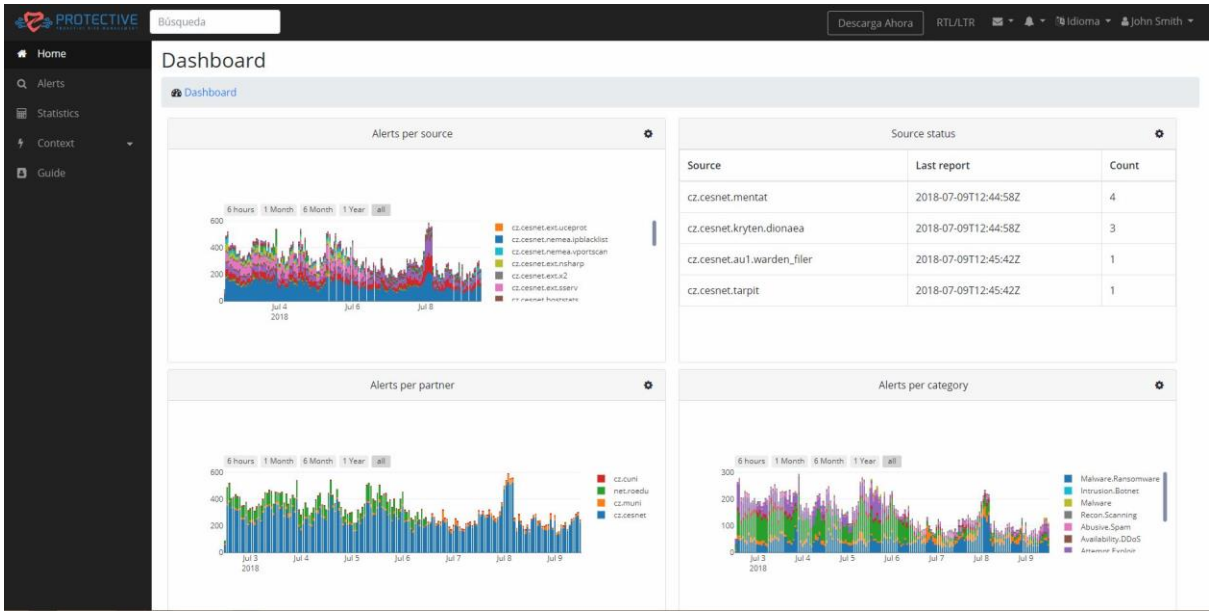


Figure 11: Screenshot of Dashboard

4.1.2 Alert Search

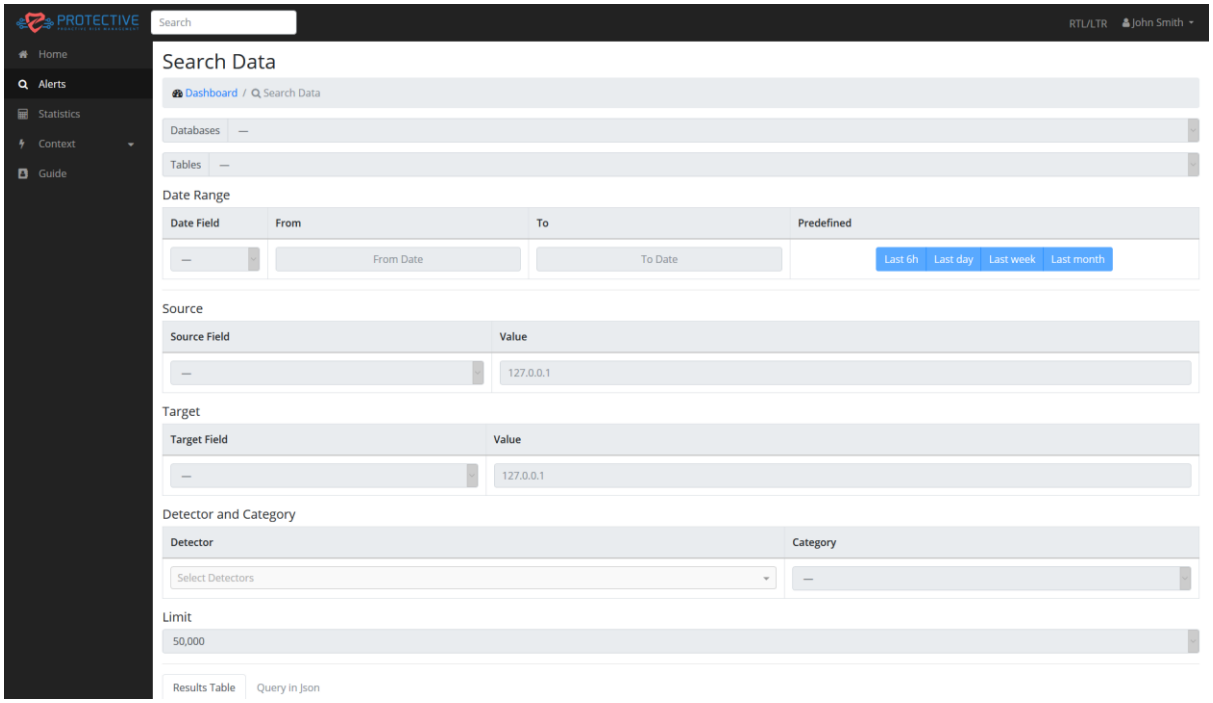


Figure 12: Alert search view

The alerts view enables the operators to query the local alerts based on their defined criteria. It provides access to all the databases tables running locally, e.g., the alerts, meta-alerts and

administrative information. The queries of the operators may be refined to filter the resulting information, as also illustrated in Figure 12, the following, main criteria:

- The time windows for when the events that they are interested in may have occurred.
- According to the source that initiated the event(s) in the alert.
- According to the target that was affected by the event(s) in the alert.
- Detector category, in order to filter by specific activities.

4.1.3 Statistics

The statistics view supports the implementation of Scenario 1 - System and Sensor Data Statistics and Scenario 3 – Time Series and Trend Monitoring, and enables to operator to configure the system as they need it in terms of what kind of statistics they wish to be informed about.

4.2 Overview of Activities

Phase 2 consisted of two main activities, namely the evaluation and deployment the PROTECTIVE user interface, which was made available throughout the second phase of the pilot, and individual sessions with the operators, during which core concepts and usability aspects were evaluated. The purpose of the individual sessions was to get detailed feedback from the operators, for activities that the operators would not be directly exposed to, e.g., how correlation of meta-alerts is performed, in order to have the operators validate the methodology and to map their routines to what PROTECTIVE provides. The usability-oriented sessions had the users discuss the PROTECTIVE user interface in order to gather detailed feedback on the usage of PROTECTIVE.

Table 9: Pilot 1 Phase 2 activities

Date	Activity	Requirements	Comment
July 11th	Kick-off	PROTECTIVE 2.0 is deployed, Pilot partners, GMV, AIT.	The purpose was to kick-off the second phase, prepare the operators for the upcoming activities.
July 13th	Follow up troubleshooting for Phase 2 software	Operators from pilot partners GMV.	Initial follow-up session in order to assure that everything was running and to catch any potential issues.
August 1st	Follow up, troubleshooting, and questionnaire	Operators from pilot partners GMV.	Collection of impressions regarding the running system and to gather technical feedback.
August 6th - 17th	User-Interface Aspects	Search, Statistics and Detailed Views. Operators are available. UOXF, SYNYO.	**Individual** sessions with each of the three pilot partners, discussing the usability aspects of the system and the revised user interface.
August 6th - 17th	Core System Components	Trust, Context Awareness, Correlation and Prioritisation. Operators are available. UOXF, SYNYO.	**Individual** sessions with each of the three pilot partners, discussing the individual components of the PROTECTIVE system.
August 24th	Wrap-up and final data collection (logs carried on until 31/08/2018)	All actions finished. Attendance from operators required.	Final closing session, summarizing the Phase 2 activity and collection of final issues, comments and feedback.

The rest of this section documents the information collected during the activities described in the above table, as well as by using the feedback logger and all other means of communication.

4.3 Feedback Logger Results

The total collected feedback log entries in Phase 1 was 86 (39 + 10 + 37) collected from 01/08/2018 to 31/08/2018. Figures 13-15 present the results collected in terms of: Type of log, module and topic.

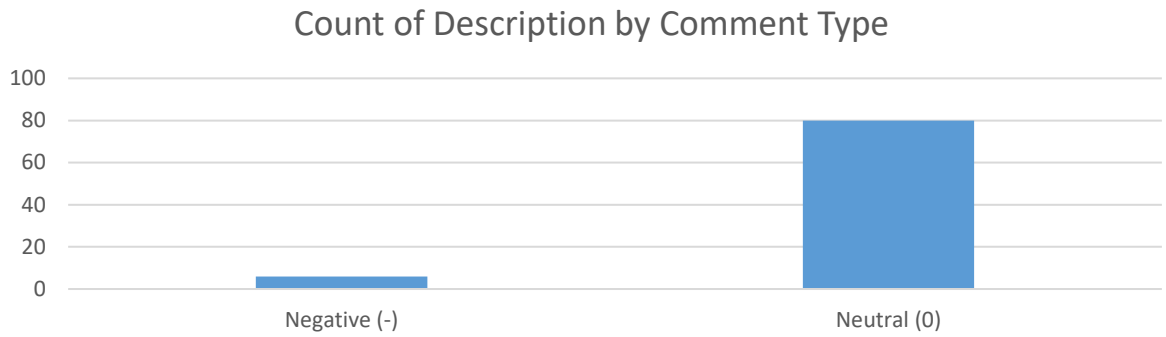


Figure 13: Count of Description by Comment Type.

Figure 13 shows 6 negative comments, 80 neutral comments. No positive logs were recorded in the feedback logger. A breakdown of reasons given for comment types are the following:

- **Negative (-) 6 logs:**
  - 1 feature issue: When exporting PNG screenshots, the same name is used.
  - 1 feature issue: Editing plot is mislabelled, and the feature related to editing in chart studio cannot be done.
  - 4 system check issue: 3 data requests failing (connecting to other PROTECTIVE nodes) and 1 necessary restart due to a system crash.
- **Neutral (0) 80 logs, all System Checks:**
  - 14 system checks related to connectivity checks specifically.
  - 66 system checks related to overall system checks.

Figure 14: Count of Description by Module.

Figure 15: Count of Description by Topic.

Figure 14 and Figure 15 shows 2 comment about the dashboard/features and 84 system checks. The dashboard/feature comments are the same two as the aforementioned feature comments related to Figure 13. We see that no issues with the connectors or install issues emerged, instead comments relate to features in the dashboard and minor stability issues.

While Phase 2 is a shorter timeframe we still observe that during Phase 2, fewer issues emerged overall, even with a tool upgrade taking place. The tool ran more stably, and although similar, minor issues for running the tool still prevailed, including another tool crash and connectivity/connector issues. Despite this, we still consider the stability and usability improved overall. We plan on addressing the negative comments from Phase 1 by further improving the installation and usage guide (incl. how to add new connectors guides), enhancing alert querying capabilities, improving system stability, and maintain our existing too support approach. We believe that lack of positive logs does not necessarily reflect reality as this is highlighted by the questionnaire results.

4.4 Questionnaire Results

Similar to Section 3.3, the questions used in the questionnaire can be found in Appendix D. Below in Table 10 and 11 follows the Likert-scale results, averages and standard deviations for operator opinions about key aspects of the tool such as their opinion on ease of installation/use, the perceived value of tool aspects including: viewing other NREN alerts, statistics (about the alerts) and querying alerts (see Appendix C for more information about each of these aspects). “Ease of install” and “Value of Briefs” are non-existent because in this phase the operators only upgraded the tool, and the upgrade removed Briefs as they were replaced by the dashboard.

Table 10: Likert Scale Opinions from Mid Phase 2

Mid Phase 2	CESNET	PSNC	RoEduNet	Avg.	Std.Dev.
Ease of Use?	4	3	5	4	1
Value of viewing other NREN events?	3	5	3	3.666667	1.154701
Value of statistics View?	4	4	2	3.333333	1.154701
Value of Alert Querying?	5	4	3	4	1
Avg	4	4	3.25		
Std.Dev	0.816497	0.816497	1.258306		

Table 11: Likert Scale Opinions from End of Phase 2

End Phase 2	CESNET	PSNC	RoEduNet	Avg.	Std.Dev.
Ease of Use?	4	3	5	4	1
Value of viewing other NREN events?	4	3	4	3.666667	0.57735
Value of statistics View?	4	4	4	4	0
Value of Alert Querying?	5	2	5	4	1.732051
Avg	4.25	3	4.5		
Std.Dev	0.5	0.816497	0.57735		

Figures 16 and 17 show the results of Tables 7 and 8 in chart form, comparing each questionnaire having the results of each NREN compared with each other.

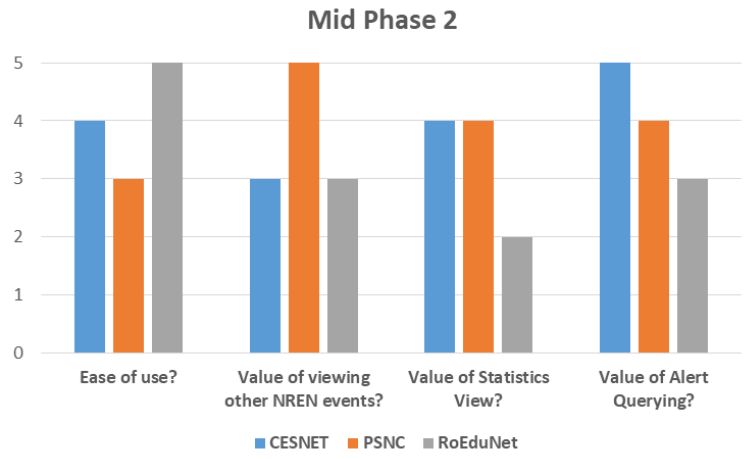


Figure 16: Opinions (Likert scale) feedback during middle part of Phase 2

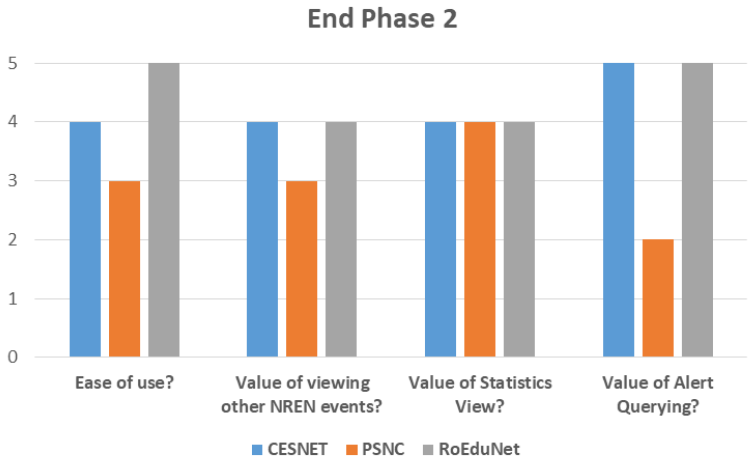


Figure 17: Opinions (Likert scale) feedback during the end part of Phase 2

The Likert scale values suggests that feedback from the operators is again largely positive, less so than Phase 1. Upon further enquiry, we discovered this is due to the fact that the new dashboard, as it is awaiting some of the more advanced features to be integrated, including Trust, Meta Alerts and Context Awareness. There is still room for improvement. Specifically, we make the following key observations:

- The value of statistics increased over time. We will therefore look to add new statistics features moving forward.
- There is still divided opinions on “ease of use” and “value in viewing other NREN alerts”. We will focus on improving these topics moving forward.
- The value of alert querying dropped – the questionnaire results and engagement with operators suggested that more feature need to be implemented for PROTECTIVE to become as fully-fledged as other commercial products.

This questionnaire, unlike the Phase 1 version, included another section focusing on the new features to be integrated. These were reviewed in separate webinars and the answers were covered in the last questionnaire (see the final section in Appendix D: PROTECTIVE Pilot 1, Phase 2 – Last Questionnaire).

The remainder of the questionnaire was designed to have operators elaborate on key issues in using the tool – the last questionnaire had an altogether separate section to also include feedback on trust, meta alerts, context awareness and new features for the dashboard (beyond Phase 2). Again, we categorise them into three main classes: positive, negative and desirables. We also provide key commentary on the new features for the tool in Trust, Meta Alerts, Context Awareness and Visualization separately. Below follow the key findings from the questionnaire feedback:

Positive:

- The approaches to presenting or exploring patterns in CTI (e.g. stats, visualizations and queries).
- The installation instructions have been improved.
- Upgrading went smoothly (esp. compared to the previous installation in Phase 1)
- Despite the issues, the operators still reported that they perceive the tool as useful.
- Level of support from the PROTECTIVE team is appropriate.

Negative - needs improvement/challenges:

- Minor issues appeared w.r.t. general use. These related to:
  - One system crash – stability can be improved further.
  - Connectivity to connectors can be improved.
  - Feature issues (see previous section discussing feedback logs).
  - The dashboard needs more features.
- Occasionally the operator guide would be out of date.
- More IP searching options would be helpful.
- Running the tool in a VM requires additional resources (and not the default resources given as the tool was underprovisioned).
- It would be more convenient for newcomers to have a complete installation guide that includes the docker and docker composer installation steps.
- In this pilot, there were no cases of CTI event being directly helpful to an incident.

Desire:

- More tools to troubleshoot issues through error and warning messages.
- Being able to migrate the tool to other hardware straightforwardly (e.g. through scripts).
- Automate the uninstallation process through scripts, including the ability to delete all dockers[2], create database backups (as part of the uninstallation process).
- The number of rows per page is always 4 regardless of limit, and this should be changed.
- Currently operators are unable to sort by column, and this should be made possible.
- Filter events by "source country" and “partner (peer)” should be made possible.
- There is currently no default statistics views/graphs defined, already predefined views would be helpful to ease the experience.
- User guide: would be helpful to add more working examples/samples based on real use cases.

Trust:

- It is not clear how IP recurrence is computed.
- It would be helpful to process what is more trustworthy first. A nice feature would be a possibility to manually specify trustworthiness of specific detectors or organizations (so it is based on not only type of detector).

[2] This is partially a limitation coming from Docker, which provides limited clean-up capabilities.

- General agreement that the trust scores are reasonable values.
- At the start, all detectors should have the same score and work them out individually and manually set some score influencers based on importance and quality that the detector is expected to give.
- In order to obtain a good understanding of trust scoring it is necessary to relate them to more real world scenarios to operations in the tool.

Meta Alerts:

- Key prioritisation criteria should be severity, detection time and MAQuality[3].
- More important meta alerts should be served to operators first.
- It might help to reduce number of events (number of meta-alerts is less then number of alerts). Depending on the, for example, number of alerts in meta-alert priority could be changed. Important missing feature is user friendly GUI to search and visualize Meta-Alerts.
- It would be helpful to automate inclusion of abuse contact email addresses for IP addresses.
- It would be helpful to be able to learn an NREN's prioritization model as every organization or operator may have different priorities.
- There should be some way (in a GUI) to correct the priorities assigned by the current algorithm, e.g. to say "*this alert is important and should have higher priority*" or "*this alert is not as important*".
- A criteria to meta alerts should be that it should belong to the operator's constituency (based on IP addresses).
- It would be helpful to correlate data received from the honeypots with the firewall (UTM) alerts and also with external events received from other NRENs. Using this kind of correlation, even during the pilot phase it would be possible to catch some infected IP addresses that were scanning the services of other NRENs. Using Protective Exchange capability it is possible to find these malicious IP addresses and manually correlate them with our internal range of IP addresses and alerts raised in a network.
- It would be helpful to prioritize alerts that are addressing important clients of the constituency, and the prioritisation should be based on severity and attacker reputation scores. Users should also be able to influence the prioritisation through network blocks or particular IP addresses. Furthermore, it should be possible to import the dataset through the Dashboard and be able to model priorities in an interactive way by assigning user defined values.

Context Awareness:

- Context awareness looks to increase understanding of node relations in the network. It is currently unclear how CA data is connected to alerts/meta-alerts (in GUI) if at all, and this could be made clearer to the operators: for example to show context data for IP addresses in alerts and a link to CA module. Maybe also a possibility to find all alerts related to a given asset.
- Visualizing the network asset hierarchy may help show the impact relationships and pinpoint the source and possibly automate an action, perhaps through object relationships.

Visualization:

- The visualization helps with orientation of incidents and the alert search is also very useful.
- It would be useful to:
  - linking to other external services such as passive DNS, Virus Total, IP reputation looked up on various databases (RBLs), Hostname (DNS lookup), and abuse contact obtained from whois.

[3] Meta-Alert Quality

- extend the search capabilities to be able to find alerts based on country code, ASN, IP source, IP recurrence and category.
- enrich the visualization with lots of information (security contexts) before they get stored, making lookups fast. It is important to have the information accessible as quickly as possible through the dashboard when needed.

4.5 Commentary on Phase 2 Results

Our key take aways from Phase 2 feedback logs is that it was successful with minor setbacks related to installation and system usage issues (connectivity and crash). There is an improvement on Phase 1 in terms of system stability and usage, however, there was no success story in terms of incident handling this time. We plan on addressing the negative comments from Phase 1 by improving the installation and usage guide, enhance the alert querying capabilities (specifically: preventing operators from searching in time periods prior to any events being observed, and allow searching for IP ranges), improve system stability, and maintain our existing tool support approach. We believe that the lack of positive logs do not necessarily reflect reality (i.e. the tool was well-received, also in this iteration) as this is highlighted by the questionnaire feedback. The procedures related to the pilot itself were well-received. The response times for support were usually less than 24h, but due to summer holiday, there were a small number of cases in which emails were not responded to immediately. This will need to improve for Pilot 2.

5 Pilot Findings

5.1 Outcomes from Feedback logs, Questionnaires and Pilot Procedures

Our key take aways from Pilot 1 feedback logs is that it was successful with minor setbacks related to installation and system usage issues (connectivity and crash). There was one success story related to incident handling.

We plan on addressing the negative comments from Pilot 1 by improving the installation and usage guide (incl. how to add new connectors guides), enhance the alert querying capabilities (specifically: preventing operators from searching in time periods prior to any events being observed, and allow searching for IP ranges), improve system stability, add migrating options and single script management, but also maintain our existing tool support approach.

We believe that the small volume of positive logs do not necessarily reflect reality (i.e. the tool was well-received, in both iterations) as this is highlighted by the questionnaire feedback. The procedures related to the pilot itself were well-received. The response times for support were usually less than 24h, although in Phase 2, due to the summer holiday, response times could have improved, there were a small number of cases in which emails were not responded to immediately.

We show the Likert scale results across all questionnaire cross sections to provide a temporal comparison of these results, see Figures 18, 19 and 20 for the results for each NREN. “Ease of install” and “Value of Briefs” are only visible for “P1 Mid[4]” and “P1 Mid and P2 Mid” respectively. In Figure 18, we see a mostly consistent opinion throughout the pilot, with a slight improvement on perceived value of viewing other NREN events at the end of the pilot.

[4] P1 is shorthand for Phase 1. Mid is shorthand for Middle. Both terms are used to indicate timeframe of the cross-sectional results.

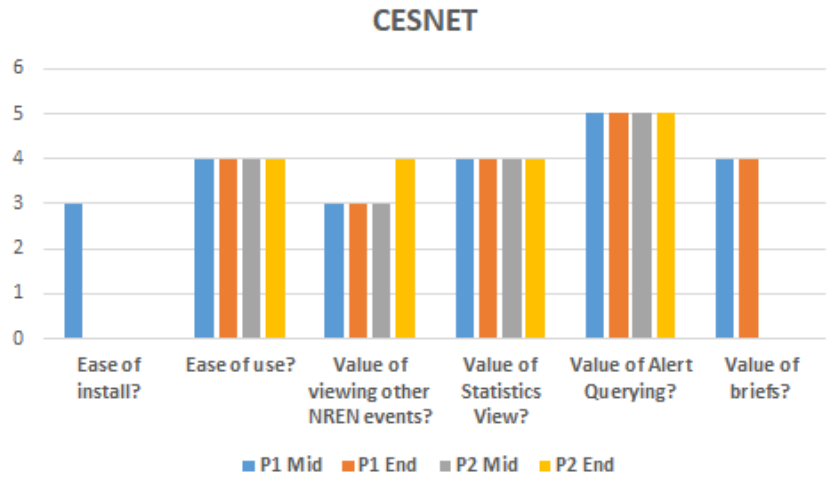


Figure 18: CESNET's Likert scale feedback over the pilot

In Figure 19, we see a more volatile experience with the PROTECTIVE tool. Upon enquiry, this is due to several operators being involved and having different experiences with the tool. The outcome from PSNC is the combined representation across three different operators, whereas Figure 18 shows the experiences from a single operator. Figure 20 shows the feedback from two operators. With PSNC we also note that improving the value of alert querying is the most urgent issue to address. Finally, in Figure 20, we observe the feedback is largely consistent, with the exception of the midpoint of Phase 2. Upon further investigation, this was due to becoming accustom to the system from the upgrade.

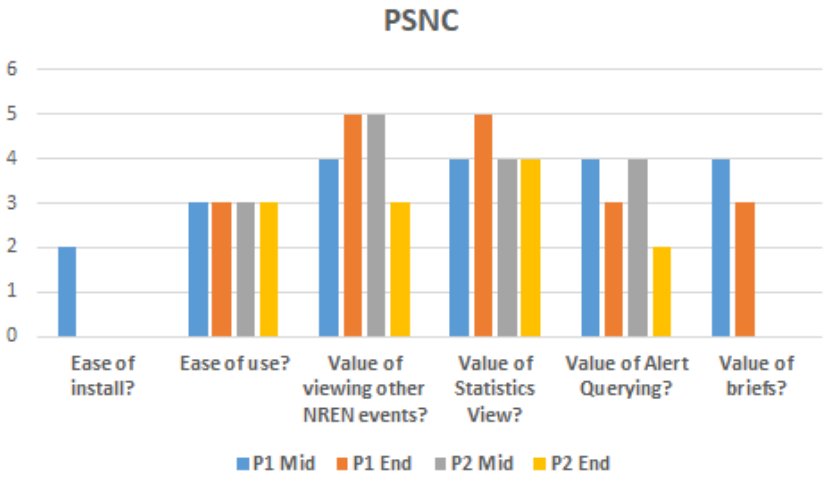


Figure 19: PSNC's Likert scale feedback over the pilot

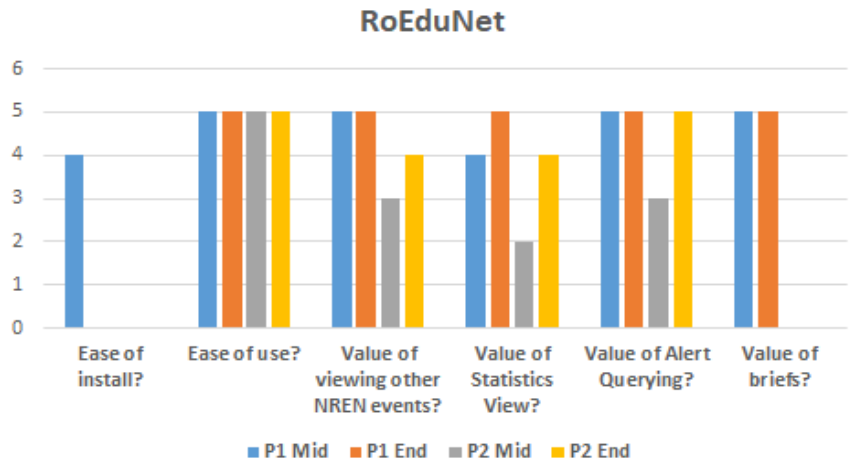


Figure 20: RoEduNet's Likert scale feedback over the pilot

With regards to the pilot feedback: operators feel the procedures are clear and they get the support they need to run and troubleshoot the tool, with support and pilot questions email responses mostly getting fast (less than 24h) responses. Key areas for improvement however include:

- Operator guide needs to be updated more frequently to reflect issues in running the pilot (if any).
- It is reasonable/helpful to have major issue be reported over email, and this approach should be kept.
- There were also minor feedback logger issues:
  - More well-formedness checks to ensure the formats received are comparable to each other.
  - The interface needs to be minimised further to maximise input value and minimise time spent recording feedback – this was corrected for in Phase 2, see Appendix B and C to compare the feedback logger visuals.
  - Add persistence, so if the feedback logger crashes or computer is shut down, it reopens old entries – this was added for Phase 2.
  - Add a 'phone home' button feature so operators do not have to email feedback logs as csv files, but instead only have to click a button to report this week's logs. Adding a phone home button allows operators to say when logs leave the organisation. This feature could be automated, perhaps prompted, but should leave only when operators give a go ahead. A 'phone home' button is planned feature for Pilot 2.
  - It would be helpful to collaborate in a more centralised way – akin to using a google doc sheet, which would entail less work on gathering reports.

5.2 Technical Aspects and Measurements

This section covers the system activities during Pilot 1, focusing on measuring the number of alerts that were made available to the operators, across three different groupings; the category of the alerts, reflecting the frequency of the activities that were recorded during the considered period, the specific connectors which generated the alerts, showing where the majority of the information came from, and finally, the type of software that generated the alerts.

The alerts shared across the three PROTECTIVE instances contain meta information on the category of the alert that has been shared. The below figure, Figure 21, shows the number of alerts being distributed across the three NRENs and the amount in each category as well as the total alerts.

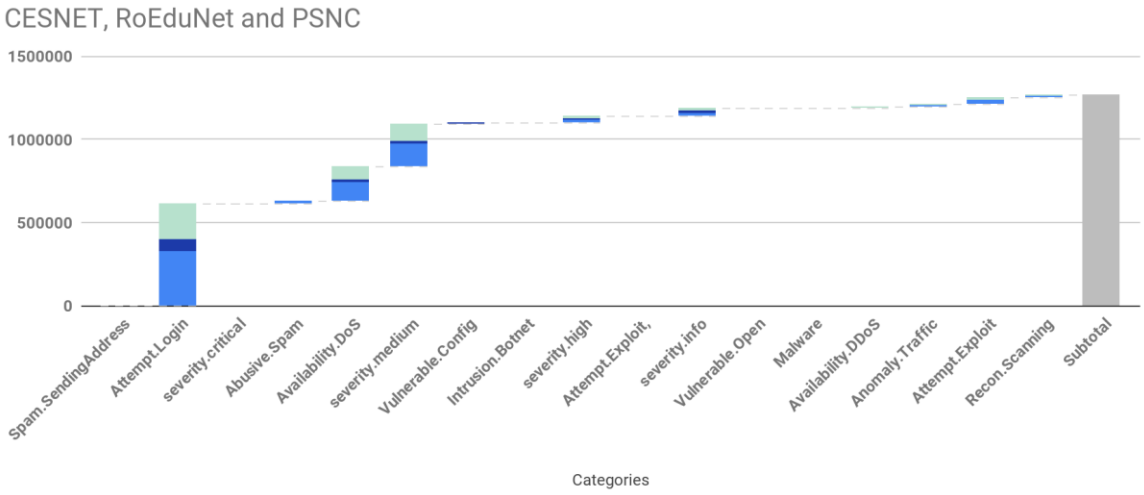


Figure 21: Distribution of alerts across NRENs and the alert categorization

In total, around 1.3 million complete alerts were distributed across the three NRENs, distributed across 18 categories. The main category includes attempted login and Denial of Service attacks.

Each alert, originates from a specific instance of a connector, the connector that detected and therefore created the alert which is then injected into the system. The figure below, Figure 22, shows the distribution of the alerts across the various instances of connectors.

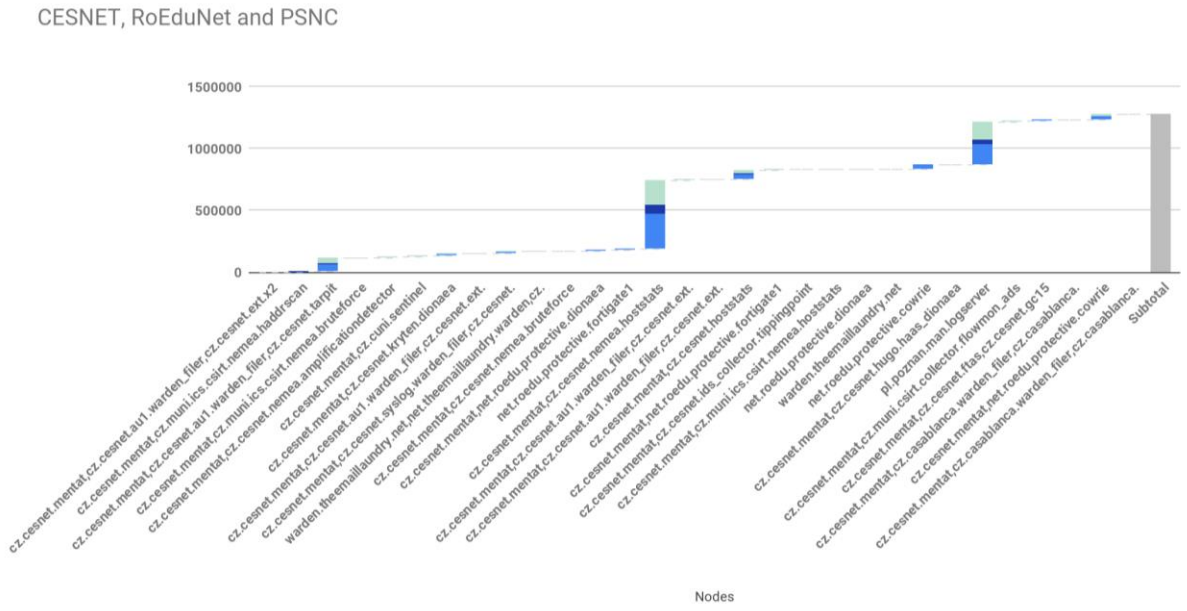


Figure 22: Distribution of alerts across specific connectors and NRENs

Since each connector is an instance of a specific type of software with its own capabilities, we document the frequency of alerts generated across connector **type**, or software, as well as across the different NRENs. The distribution is shown in Figure 23, showing the distribution of alerts across the software implementing the connector and the three NRENs.

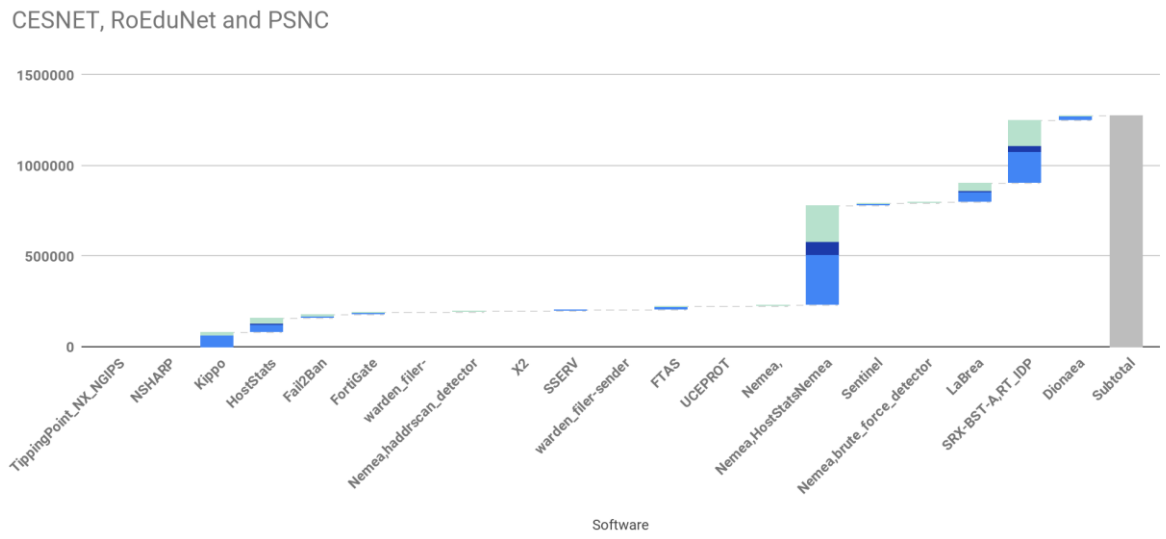


Figure 23: Distribution of alerts across connector software and NRENs

Finally, the table below, Table 12, provides an overview of the connectors that were deployed locally, e.g., installed within the premise of each of the NRENs' constituency. Contrary to the previous statistics shown, which focused on the data in available to the operators, the below table shows which specific connectors were configured and connected to the PROTECTIVE system at each NREN.

Table 12: Specific connector deployments per NREN instance

Connector	CESNET	PSNC	RoEduNet
org.protective.warden.rx.internal.eg	■		
cz.cesnet.mentat	■		
net.roedu.protective.node1	■		
pl.poznan.man.protective	■	■	■
cz.cesnet.protective		■	■
kippo.man.poznan.pl		■	
pl.poznan.man.cowrie		■	
pl.poznan.man.dionea		■	
pl.poznan.man.cowrie2		■	
pl.poznan.man.cowrie6		■	
theemaillaundry.dev.protective2	■		■
gmv.development.testbed.node	■		■
net.roedu.protective.dionaea			■
net.roedu.protective.pub.cowrie			■
net.roedu.protective.pub.dionaea			■
net.roedu.protective.cowrie			■

gmv.testbed.node			
pl.poznan.man.logserver			
net.roedu.protective.node1			
net.roedu.protective.warden.rx.internal			
net.roedu.protective.fortigate1			

6 Lessons Learned, Action Points and Identified Requirements

This section documents two aspects of the lessons learned across all the data collection activities performed during the entirety of the first pilot, namely what is essential to include, in order to prepare the PROTECTIVE project to efficiently facilitate Pilot 2, together with both the existing NRENs as well as external parties, that is external NRENs and SME. Further, this section provides an overview of features and functionality that are relevant beyond the pilot and the project. For a holistic view of the piloting activities this section includes all generalised feedback that was collected, even though some aspects may have already been followed-up on.

6.1 Action Points for Pilot 2

The below table, Table 13, provides an overview of actions that will be picked up on and implemented before the PROTECTIVE system is provided to external pilot participants during Pilot 2.

Table 13: Action Points for Pilot 2

Category	Description
Installation procedure	• The readme of the protective-node-installer has been updated to clarify which version of Docker to install when using Linux based systems • It will be necessary to have the installation guide as a document that can be shared with the pilot operators, rather than doing exports from Confluence.
Documentation	• Updated installation guide, independent of PROTECTIVE specific information • Updated user-manual, covering newly integrated features.
System Features	• To be able to search by IP address ranges in the alert/meta-alert search interface
Other	• Editing plot in Plotly chart studio in causes chart never to be imported. This button will be removed in Pilot 2 release

6.2 “Nice to Have” Requirements

Besides the specific aspects that are essential for the deployment of Pilot 2, this reports also summarises aspects that are not necessarily part of the core PROTECTIVE system, but would certainly add value to the solution. The table below lists features and requirements which should be strongly considered, either to be implemented for Pilot 2, if resources are available, or alternatively, beyond of the PROTECTIVE project.

Table 14: Summary of “Nice to have” features

Requirement	Description
Exported visualisation (graphs, plots, etc) should reflect the content in the naming of the exported files	Plotly use the name “newplot.png” for all downloads. It would be good to have a new file name each time. (May be considered if time allows)

The system shall be able to completely uninstall itself	It is planned to add options to the Node automating such activities as: removing the entire installation (including dll of the dockers), creating a backup (on a local machine, and external machine) and moving all the installation (migration). This will be considered for Pilot 2 release
The system shall be able to block whole networks	When the operator detects malicious activity originating from certain networks, there should be a functionality that allows the operator to block an entire IP range, and the blocking is automatically applied on e.g., a connected firewall.
The system shall be able to create notification addressed to external parties.	If an operator identifies a malicious activity from a specific IP address, and e.g., the owner of said IP is not part of the PROTECTIVE ecosystem, the system should be able to create a notification, e.g., an email, that is forwarded to the responsible authority.
The system shall be able to self-diagnose and report the current status	The operators should be able to easily identify the current status of the system and the network, e.g., to be able to verify if all connectors are currently online, or that the connections to other instance of the PROTECTIVE system are active.

6.3 Documentation of PROTECTIVE as an Open Source Project

As of writing of this document, one of the upcoming milestones of the PROTECTIVE project is the open sourcing and the public release of the of PROTECTIVE software stack. Before the actual release can take place, the main feedback from the piloting activities need to be accommodated. However, the documentation of the software stack has been prepared and is also being refined, based on user feedback. These consist of;

- **Installation Guide** - An elaborate guide on how to get started with the PROTECTIVE system, how to install the various components, and how to add additional connectors.
- **User Manual** - An overview of how to use the PROTECTIVE system, once that it is installed, by the operators. The user guide elaborates on what the individual views of the PROTECTIVE system provide, in terms of functionality, as well as document how it can be used.

HOME SCREEN

In this screen you will be presented with 4 hardcoded Widgets:

- Alerts per source: Stacked graph of number of alerts per source per hour over the last 7 days
- Source status: Table with following columns: alert source, time and date of last report seen, number of reports in last five minutes
- Alerts per partner: Stacked graph of number of alerts per partner per hour over the last 7 days
- Alerts per category: Stacked graph of number of alerts per category per hour over the last 7 days

These graphs are refreshed every time that you refresh your browser screen or every time that you go to the home screen.

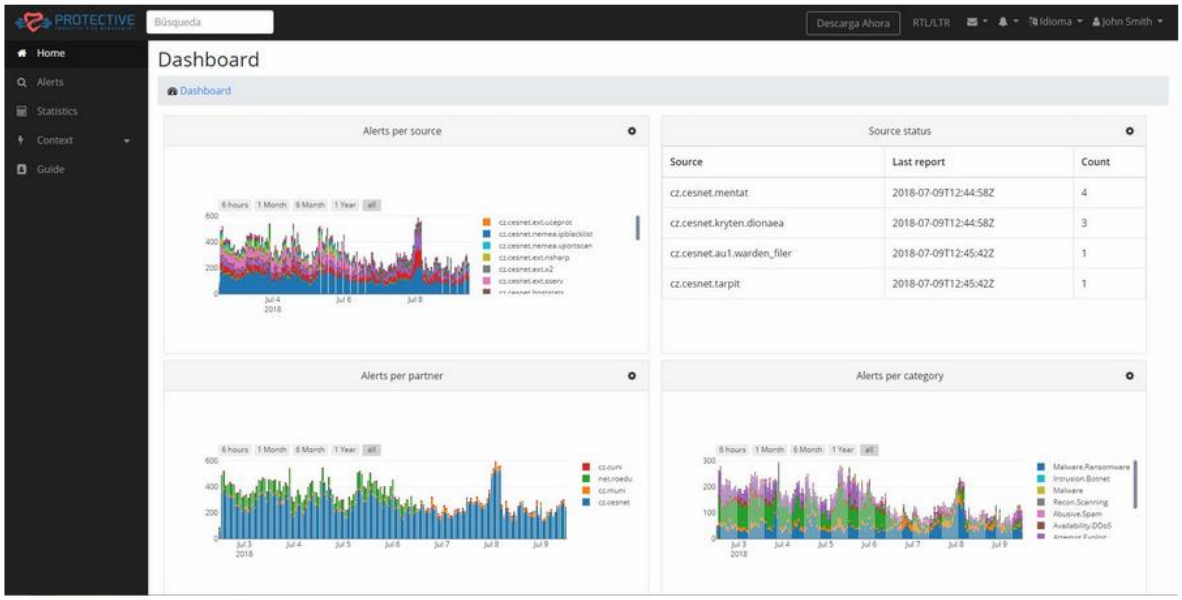


Figure 24: Screenshot of a part of the user manual

The intentions are to have both components available for the second piloting activity of the PROTECTIVE project, as to ensure easy accessibility to the PROTECTIVE application stack at third party participants, that are interested in participating in the second pilot. In particular, the goal is to make the deployment straightforward, in order to support further NRENs installing it and participating in the data sharing and evaluation of the PROTECTIVE project.

7 Summary

In this report we have presented the overall goal and purpose of Pilot 1 as well each of the two phases that made up Pilot 1, with regards to the activities that were performed by the NRENs, in order to get the PROTECTIVE system, the connectors that it depends on and the communication channels across the three NRENs deployed and configured properly. Further, the data collection activities have been documented, that were used to collect feedback from the operators and to validate the assumptions that were made in the development of the PROTECTIVE system. The feedback has and will be used to make improvements in the deployment of Pilot 2, during which external parties, both NRENs and SMEs will be invited to participate in.

The main feedback received from the operators who were evaluating the system targets, was perhaps unsurprisingly, the usability of various aspects of the system, but by pointing out specific aspects, it enables the further development of the PROTECTIVE system to focus on improving specific areas, such as improving the installation procedure to improve documentation in specific areas that reflect the needs to operators, such as how to execute specific tasks. Also, detailed feedback derived from various questions arising on how to achieve certain tasks makes it possible to make those tasks more intuitive.

Appendix A: Connectors

The below table, Table 15, list the connectors that were made available during Pilot 1

Table 15: Connectors made available during Phase 1

Connector	Description
FusionInventory Agent installers	Installs the FusionInventory Agent on Centos6, Centos7, Ubuntu 14.04 and Ubuntu 16.04. Installers are at version 1.0.0-beta.1
Ingestion/Extraction (Juniper SRX Connector)	A custom connector for ingestion phase of the workflow, that will allow to collect data from Juniper SRX Next Generation Firewall together with the additional alert class descriptions provided by the vendor, and to put these data into the IDEA format
Ingestion/Extraction (ForitGate SRX Connector)	A custom connector for ingestion phase of the workflow, that will allow to collect data from FortiGate Unified Threat Management Systems together with alert class descriptions provided by the vendor and mapped into the IDEA format.
Ingestion/Extraction (McAfee SIEM connector)	A custom connector for ingestion phase of the workflow, that will allow to collect data from McAfee SIEM and to put these data into the IDEA format
Ingestion/Extraction (Dionaea honeypot connector)	Warden connector for Dionaea honeypot (from official Warden-contrib repository)
Ingestion/Extraction (Kippo honeypot connector)	Warden connector for Kippo and Cowrie honeypots (from official Warden-contrib repository)
Ingestion/Extraction (LaBrea honeypot connector)	Warden connector for LaBrea (from official Warden-contrib repository)
Ingestion/Extraction (fail2ban connector)	Warden connectors for fail2ban system (from official Warden-contrib repository). Contains scripts for reporting spam- and ssh-related events. Can be easily adapted to other types of events that can be detected by fail2ban.
Ingestion/Extraction (IntelMQ Connector)	This SW reads from 3 different sources of malware info, converts some of their fields into IDEA format and outputs them in a folder of the host machine. Raw data is saved in another folder too.
Ingestion/Extraction (WARDEN parser)	This SW currently implements ingestion from a SQL database and converts data into the IDEA file format. Files are output on the host machine that can then be shared using the sending client for TI sharing. The SW is extensible to incorporate any data source. A class to import from any source can be implemented using the DAO interface as per future requirement.

Appendix B: Operator Guide for Pilot 1 Phase 1[5]

Overview

PROTECTIVE system: PROTECTIVE is a system that aims to help operators and analysts **share events between different public CISRTs**. We are **testing its usability and functionality** in increments. In this pilot phase you will be trying out **basic event aggregation and event sharing functionality** in the system. This "backbone" consists of two tools: **Warden**(an event sharing tool), **Mentat** (a SIEM), but in Phase 2 and beyond, we will append more functionality, including "meta-alerts" (summaries of many events), computation of trust, information sharing compliance, more visualizations, and correlation and prioritisation engines.

Purpose of pilot: Operators/analysts should identify benefits and challenges in using PROTECTIVE to aid threat detection through threat intelligence sharing.

High-level Objectives: Operators should deploy and run PROTECTIVE as a supplementary tool to their existing systems. For most of the time, it should run in the background collecting and processing data automatically, but two or three times per day we expect operators to check the system, and use the tool to help them **identify threats in their own network** OR help **improve their threat awareness** outside their own networks. At the start and end of Phase 1, we will send a **questionnaire to fill in**, we may have **follow-up interviews over skype** or in person at the end, and finally, we hope operators can use a **feedback logger to log issues operators have observed** (akin to a bug tracker), every time operators experience them. Any very serious, system-breaking bugs, please report these as well over email to: pilot@protective-h2020.eu (format template in section 4). Operators should use the statistics screen, and query events using Mentat, see section 6.

Documents to read:

- Operator's Guide (this document): link
- PROTECTIVE Node Installation guide: link
- Feedback Logger tool and README – download tool (available 12/03/2018): link

Please note: we may update the two documents (but not dramatically!), if you need to refer back to this document, please always refer to the online version.

Beginning of Pilot (15th Feb - 7th March):

- **Attend a webinar** on the 15th Feb
- **Installing and try out the tool** using the installation guide.
- **Fill in an initial-impressions questionnaire** will be sent out on 26th Feb to obtain initial feedback. Should be sent back *before webinar 2* on the 7th March.
- **Email major (system breaking) bug issues** if any emerge to: pilot@protective-h2020.eu using the format in section 4 in this document.

Main Part of Phase 1 (7th March - 27th April):

- **Attend a webinar** on the 7th March.
- **Leave the tool running in the SOC**, make sure it is up and running at all times.
- **Check the tool at least 2-3 times per day** (assuming no major incident is happening), consider the system to be a supplementary tool to obtain threat intelligence.
- **Use the feedback logger to report comments**, these can be suggestions, minor bugs, reports on when the tool has been useful, etc., see section 7.2.1.
    - **Once per week, send across the latest CSV files** for Pilot analysis

[5] Please note that all sensitive links, intended for operators, have been taken out of all appendices.

- **Email major (system breaking) bug issues** if any emerge to: pilot@protective-h2020.eu using the format in section 4 in this document.
- **Fill in questionnaires** sent on the 15th March and 27th April.
- **Attend a webinar** on the 27th April.

Pilot timeline

When	Activity	Requirements	Comments
Feb 15th	Introduction to PROTECTIVE and installation webinar	Overall PROTECTIVE presentation, NREN key personnel identified	Person who will install the system locally, person that will use the system when running.
Feb 15st	Installation and configuration of Warden, connectors and Mentat	Installation guide and method for collecting feedback, a dedicated person who can do the installation	
Feb 26th	Initial survey and first impressions	Survey with preliminary questions	Collection of information about what are the expectations? General challenges?
Feb 15th Feb 28th	Stabilisation and Adaptation	Monitoring of whether the system is running and receiving messages, identification of issues	This phase ensures that PROTECTIVE is running well and is populated with initial data, by the time that the operators are expected to use it.
March 7th	Final collection of feedback on deployment experience	Questionnaire	The operators may provide feedback along the entire, this only mark when all the data needs to be collected
March 7th	"Introduction to PROTECTIVE" webinar	Elaborated tasks descriptions, specific tasks that can make the operators familiar with the system.	This includes the Mentat tool, how to use the diary (what kind of events are we interested in, escalation, usefulness, bugs etc)
March 28th	Intermediate survey and interview	Questions	Short discussion with the relevant operators about their experience in order to follow-up on whether they are using the system, and their overall experience until now in order to collect preliminary data.
April 27th	Final questionnaire and interview	Questionnaire + possible follow-up interviews (skype)	
May	Data collection and extraction	Scripts to collect long term statistics from MongoDB queries	We need to monitor how many alerts have been receive a) from local sensor and b) from external

	Obtain a snapshot of PROTECTIVE state. Get data from DB.		sources, per NREN, do post processing etc., collect feedback logs, surveys etc.
June 4th	Installation of **Phase 2** software	Repetition of the above documentation	

Key purpose of the pilot

The pilot is being conducted for the PROTECTIVE consortium to obtain feedback about threat intelligence sharing. Without feedback, we cannot improve or validate the system. It is therefore of importance that we get feedback directly from operators so we can improve the tool accordingly.

Specifically, in this phase we wish to:

- identify the **benefits and challenges in sharing threat intelligence between NRENs**, e.g.:
  - Has data received from other NRENs had positive effects or caused any issues, if so, what?
  - Are you receiving new data types not presently available in your CSIRT?
  - Are you obtaining new information about compromised machines in your network which does not exist in your alerting systems? Are any of these of particular interest?

- identify the **practical issues** that emerge **while deploying and using** the system:
  - Was there a lot of effort involved to get the tool sets up and running?
  - Was the tool stable and were any bugs detected while running the tool?
  - What is the usability of the tools like?
  - Have you had any problems with the database backend?
  - Have you had any problems with dockerization?
  - How can the deployment instructions be improved?

Key tasks of operators during the pilot

During the first phase of the Pilot 1 the main focus is on the evaluation of the deployment and underlying key event sharing functionality and stability of the system as well as the user experience. The purpose is to obtain first indicators of whether we are going in the right direction with the system at all, how stable it is in a live environment (not just testbed scenarios) and identify the real-added value is to the users (operators). Specifically, the key tasks of operators are:

- **Deploy the tool 15th February –** critique the PROTECTIVE setup and deployment instructions provided: link
- **Attend a webinar (15th February, 7th March, 27th April) –** learn how to deploy and use the tool and ask questions about usage, and provide feedback about deployment of the tool.
- **Run the tool in the background and maintain it if necessary (7th March-27th April) –** Keep the tool alive and running in the background in the CSIRT.
- **Check the tool** (ideally two or three times per day, but at least once per day?) – to make sure you are up to date on latest events that are shared.
- **Report feedback logs –** use the feedback logging tool to report how you have used the tool. This includes troubleshooting, (minor) bugs, feature requests, and incident handling (if PROTECTIVE has helped – please state how and what happened? E.g. "an event helped with the takedown of an IP address that was otherwise invisible on our network", see Section 7.2.2 for more info).
- **Report major bugs - In the interest of clarity: send bug reports if the issue is major and need addressing ASAP. Send feedback logs if the bug is not a major issue.** Any very serious bugs, please report these as well over email to: pilot@protective-h2020.eu

If sending in a major bug please provide the following information in the email:

Timestamp: <Time the bug occurred for the first time and (if applicable) how often it occurs.>

Summary: <Should be a short and precise description of the bug. It should be possible to tell what this bug is about, and where it occurs from reading the summary alone.>

Components: <Which part of the system is affected?>

Environment:<Provide specs of the machine. Enter the environment under which you found the bug, make sure to include any information which might be relevant to the bug (e.g. your screen resolution, special plugins, etc.)>

Description: <This is the main description of the bug>. Make sure to include at least:

- A description of the bug.
- A very clear step by step description of how to reproduce the bug (if possible).
- A description of what you expected, and a description of what actually happens.
- Screenshots or other attachments are always helpful, but optional.

Webinars

Two webinars will be held early in the Pilot and one at the end of the phase. Who should attend the webinars: ideally all NREN points of contact (Gerard, Andrea and Mihai), the person who will be installing the system, and as many operators from CESNET, PSNC and RoEduNet as possible. The meeting will be screen recorded, so should any operators not be able to join, they can review the video again at any time afterwards by request. Should you have any questions about the pilot or using the tool that cannot be resolved easily (and need to be), please email: pilot@protective-h2020.eu

Where: link , see also Appendix A.

First webinar: "Introduction to PROTECTIVE: deployment and testing the tool in an operational environment, and data sharing procedures.**"**

Time: **15th February @ 12:00 GMT – expected duration, up to two hours.**

Agenda:

- Overview the pilot timeline and purpose – setting the scene.
- The operator's pilot guide (this document).
- Installation instructions and deployment tasks: link

Second webinar: "Deployment Feedback and hands-on Mentat tool"

Time: **7th March @ 12:00 GMT**

Agenda:

- Address questions and overall feedback from the deployment period.
- Review how to use the tool (statistics and event queries), showing example high level questions and queries operators may ask.
- Review key tasks in the pilot and what we as the researchers aim to obtain.
- Review how to use the feedback logger.
- Answer any questions operators may have about usage or the pilot.
- Questionnaire to follow shortly after.

Third webinar: "Pilot 1, Phase 1 Review"

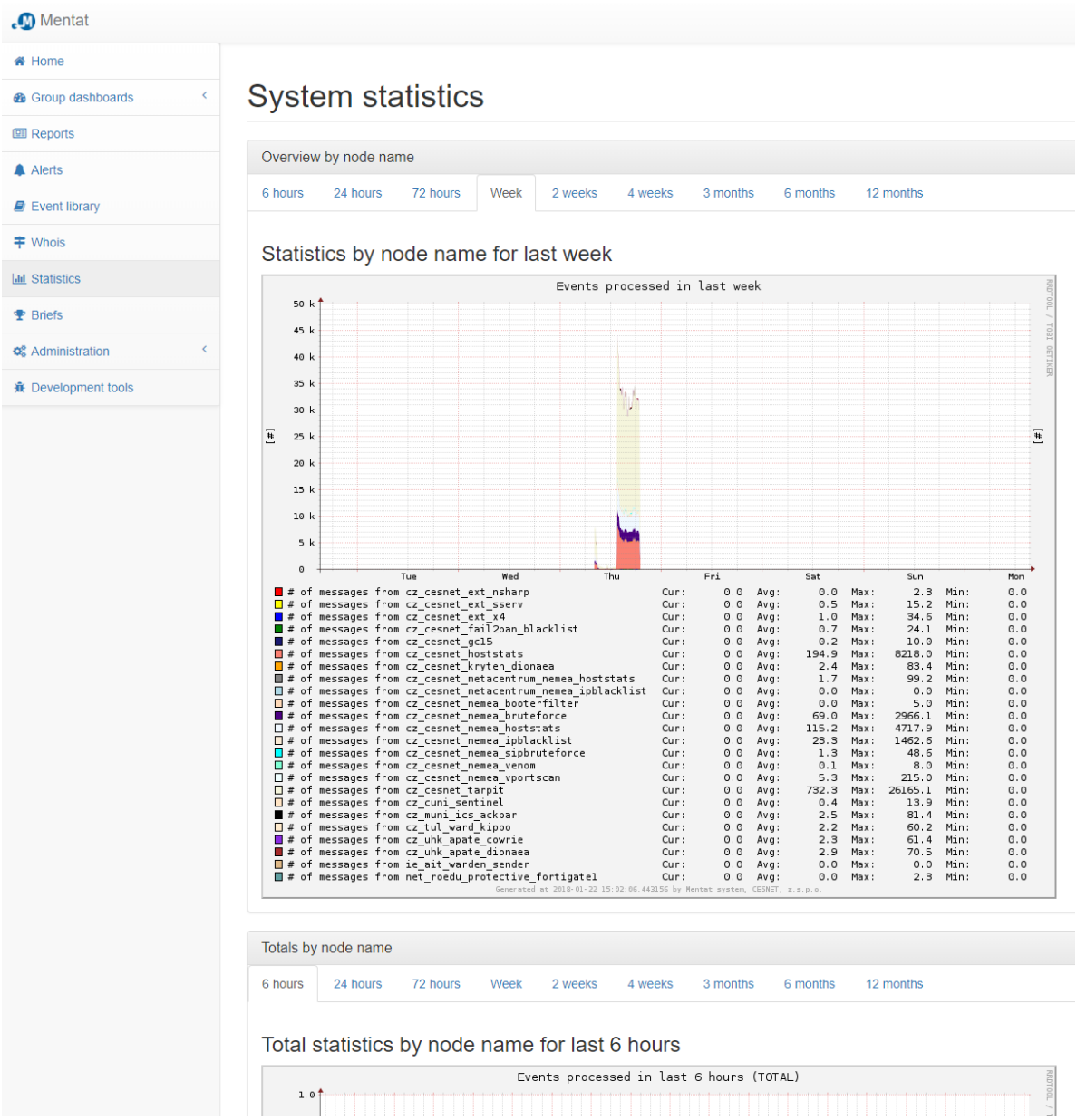
Time: **2nd May @ 12:00 GMT**

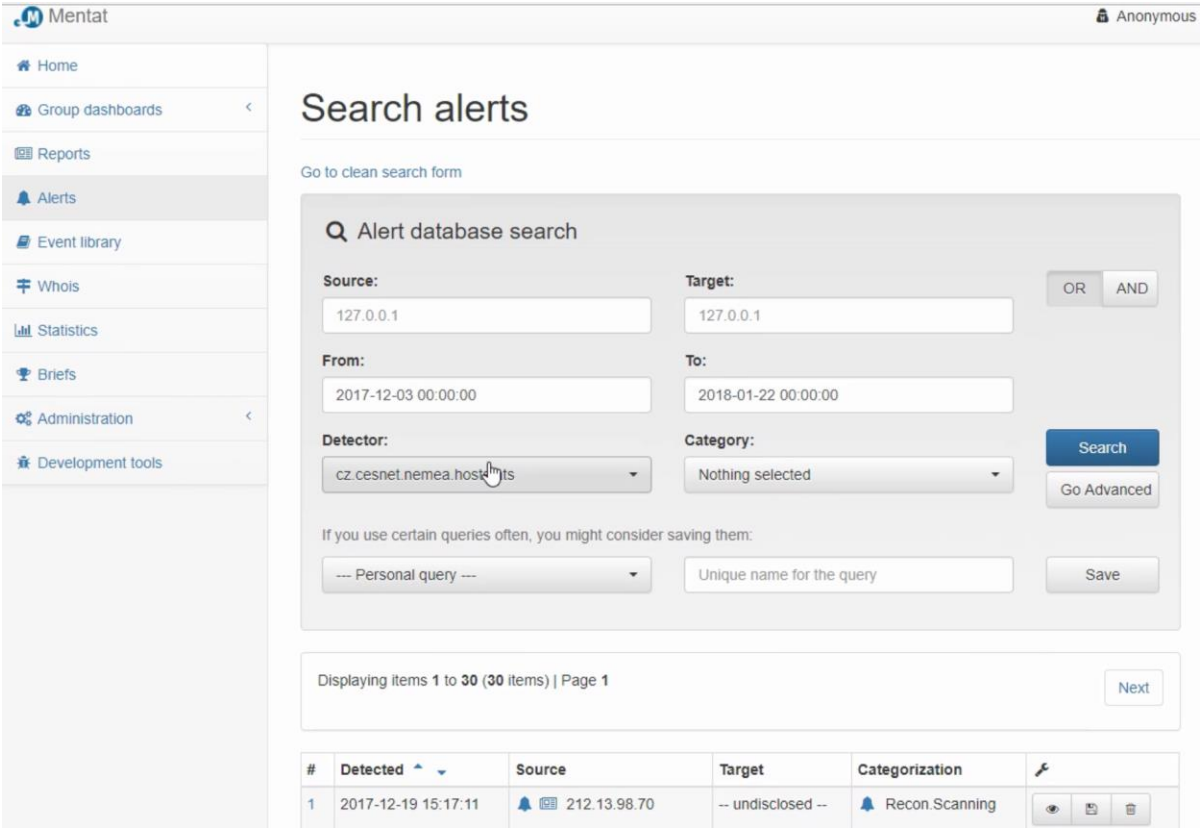
Agenda:

- Review feedback from system usage during pilot
- Review the data collection procedures and how it can be improved for Phase 2
- Questionnaire to follow shortly after + possible interviews.

Using the system

When using the tools, good starting point are: system statistics and event queries:





Repeat periodically (several times a day): (covered in webinar 2)

- Go to "Alerts" page, set "From" to the time last checked, fill the IP prefix of your organization to "Source" field and click "Search". The results show alerts about your IP addresses doing something bad, i.e. they are probably compromised/misused. Assess each alert and decide whether it's a real (new) incident, false alert, or an alert related to the same event/incident as some already being handled (single malicious IP is often reported many times, until the problem is fixed).
- The same as 1), but fill the IP prefix into Target. The results show detected attacks against your network. Most of them are probably just a common noise such as port scans or login attempts to honeypots. If the number of alerts allows it, try to assess each alert as unimportant, false alert, something already seen or a real incident. It may be useful to filter only some attack categories (e.g. DDoS attacks) or detectors (e.g. filter out ones that only report port scans or select only those monitoring important parts of your infrastructure) - it depends on your particular interests and detectors deployed in your network.
- Go to "Statistics" page and check that some data are being processed, and that all detectors report "normal" number of alerts, i.e. whether some detector hasn't stopped working or started to generate much more alerts than before. If such anomaly happens, investigate if something's wrong with the detector, it was a planned reconfiguration, or there's indeed much more/less attacks of given type. Note: If your organization's IP range can't be written as a single IP prefix, the steps 1 and 2 must be repeated for each prefix, as Mentat currently doesn't support multiple prefixes in Source and Target fields.
- Also, you may want to ask participants to fill an entry in the diary every time an alert is classified as an incident to be handled (escalated) or an interesting anomaly appears in statistics graphs.

Operators are encouraged to experiment with alert queries and report back to us (via the feedback logger) which queries have been helpful for them. The feedback logger does not connect to a database so it will not execute the queries in the feedback logger.

Peer Node DNS names:
RoEduNet: link
CESNET: link
PSNC: link

Notes with regards to statistics view:
Please note that there is a discrepancy between ingested timestamps and detection timestamps.

Some questions that may be of interest in asking during tool usage:

- Did you receive an event in PROTECTIVE that you did not receive in your existing system?
- Unusual patterns of data in your statistics view?
- Were any insight missing in PROTECTIVE that you have received in your existing system (that you would expect to see in PROTECTIVE)?
- For events of similar nature, received in both PROTECTIVE and your existing system, what was the time difference between receipt of the event and which system provided the event information first?

Useful searches:
To follow - please send in ones you have found particularly helpful in the feedback logger.

Operator Tasks

Beginning of Pilot (15th Feb - 7th March):

Deployment

Aim: For operators to be able to independently setup and configure the PROTECTIVE tool.

How: Follow the instructions in the installation document: link
Operators will deploy and use the tool. Questionnaires will then be sent at the time specified. These questionnaires aim to capture your perception of PROTECTIVE at **the start and end of the phase**. **This will be covered in webinar 1.** During this phase, there will be no logging of opinions other than one questionnaire. If you have any major issues with the tool, please report them via email in the style specified in section 4 to: pilot@protective-h2020.eu

Daily - Mid-Phase (7th March- 27th April):

Using the system

Aim: to identify issues and feedback with events being received independently by PROTECTIVE.

As a guide, operators should leave the system running the background and check that the system once every 2-3 times per day, and report to the feedback logger if anything has changed and that needs reporting. Operators should use the system as a supplementary tool to support their workflow, not be tied to using it. Operators are expected to check the statistics view AND run a query to check all events that have arrived within the last six hours. **This will be covered in webinar 2.**

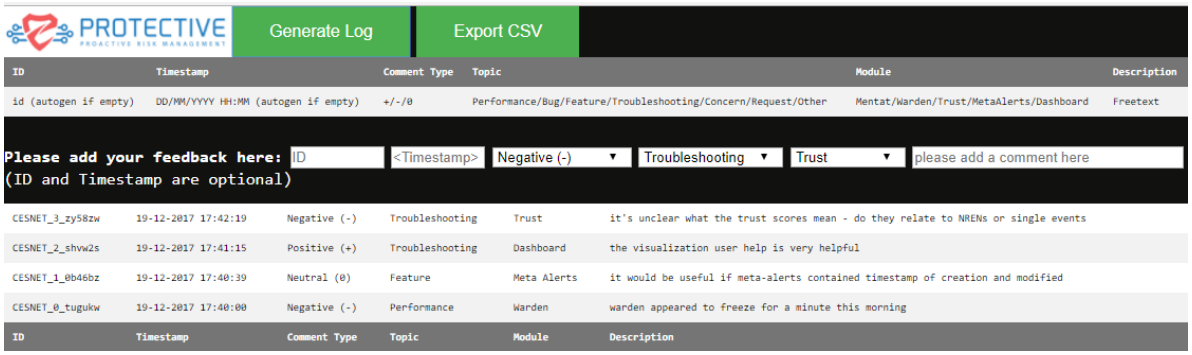
Basic checklist:

- Check statistics view.
- Use the common queries you believe would be of use to you.
- Identify whether you (as an operator) would like to run any additional queries.
- Use the information provided as an external source of information and use it if something meaningful emerges that you can use in your environment.
- If issues emerge or any thoughts to using the system pops up, then make a note of them in the feedback logger.

Feedback logger

Aim: To obtain feedback from using the system.

Feedback logs are a means for operators to self-log emerging feedback using bug tracking-like templates, and follows a CSV-like format. The purpose of the feedback logs is to document key information about emerging usability issues "*as they are happening*". Logs aim to be as short as possible to minimise disruption of the operator, but also to maximise data collected. They are akin to a bug tracker, but differ in that their aim is not solely to identify bugs and troubleshooting concerns, but also opinions and a record of what has happened. The feedback logger will allow us to conduct data analytics on patterns related to usability. The logs would also aid during the evaluation to inform it about which questions are likely to be the most pertinent at the end of each phase of the pilot. **They do not connect to a database - instead only export CSV files directly from JavaScript in a single HTML file. This is done to minimise workload in setting it up and maintaining it.** The log files are stored as a spreadsheet. From the keywords, it is possible to examine the frequency patterns, trends in comments and compare feedback across NRENs. Please keep an eye that the CSVs are being output correctly, and **send weekly CSV files to pilot@protective-h2020.eu** The feedback logger and README installation instructions link is at the start of this document. **This will be covered in webinar 1 and webinar 2.**



We anticipate that **operators would fill in a row up to a few times per day,** although this is subject to emerging issues with the PROTECTIVE system. Fields are filled in next to the "Please add your feedback here" text. When each of the fields in a row are complete, pressing the "Generate Log" button will generate a log. Each x number of lines created will export a CSV file, but the operator can also export at any time using the "Export CSV" button. **Ideally the diary tool should be run on separate hardware from PROTECTIVE, and always on, e.g. on a laptop, if available.** Fields:

- **NUM** – an auto-generated string. (NUM was previously ID, as seen in figure above)
- **Timestamp** – an auto-generated string.
- **Comment Type** *<enum>* - gives an indication about the nature of the feedback in the log, allowing users to specify whether the log is intended to be negative, positive or neutral feedback. Enum options are: "Positive", "Neutral", "Negative".
- **Topic** *<enum>* - describes what the log is concerning using a pre-defined list of words. The list is setup to avoid overlaps in topics. The optional tag column allows for users to specify other tags not covered by topic. Enum options are: "Bug", "Feature", "Performance", "Incident Handling", "Troubleshooting", "Request", "Other"
- **Module** *<enum>* - signifies which technical PROTECTIVE subsystem the log pertains to. Enum options are: <list of all connectors>, "Mentat", "Warden", "Other"
- **Description** *<string>* - free text that an end-user can use to describe their comment in depth. Users are asked to be as brief and informative as possible.

When you check protective and you have no comment, please add:
Comment type: Neutral
Topic: System Check
Module: System Check
and leave the other fields empty. That way we can record that the system is working as expected and you've recorded it for us. If you have any issue or comment you'd like to add, please do.

End of pilot (27th April - 4th May):

Aim: to evaluate the perceived value of CTI sharing using PROTECTIVE, and improve data collection procedures for the next phases.

How: We will use **questionnaires** with closed-ended questions to capture both quantitative and qualitative information about expectations and reflections of the system in use. As mentioned, these will be sent out at the **start and end of the phase**. In the questionnaires we will ask specific questions about:

1) **perceived reduction** of: need for internal events and false positive rates,
2) **perceived increase/improvement** in: confidence in received data, incident response, accuracy of events prioritised in line with analyst preferences, and finally,

3) **perceived value** of CTI sharing.

We may also follow-up these with **interviews (skype and/or in person)** to help with in-depth qualitative understanding of end-user opinions in which they elaborate on issues with regards to using PROTECTIVE. The questions will follow suit from the questionnaire and ask operators to elaborate on answers from the questionnaires. During the interviews, the interviewers present operators with isolated use cases of PROTECTIVE components to allow us to identify usability concerns through an interactive demonstration in which operators also provide feedback as part of the interview questioning the tool is being used with the researchers observing and interviewing the operator. **This will be covered in webinar 3.**

Appendix C: Operator Guide for Pilot 1 Phase 2

Overview

Purpose of phase: Operators/analysts should identify benefits and challenges in using the new PROTECTIVE version. Some aspects of the tool will be mock-ups or using dummy data to illustrate new, upcoming functions.

High-level Objectives: Operators should update their instance of PROTECTIVE from Phase 1. As before, for most of the time, it should run in the background collecting and processing data automatically, but two or three times per day we expect operators to check the system, and use the tool to help them identify threats in their own network OR help improve their threat awareness outside their own networks **using the new dashboard**. At the start and end of Phase 2 (after the first two weeks), we will send a **questionnaire to fill in**, and hope operators will use the **feedback logger to log issues operators have observed** (akin to a bug tracker), every time operators experience them. Any very serious, system-breaking bugs, please report these as well over email to: pilot@protective-h2020.eu . In **August we will run evaluations that will ask operators to try out components of the new systems** (that are in development). We will have a **follow-up questionnaire at the end**, and may follow up with skype interviews.

Important Links:

- Operator's Guide (this document): link - instructions on how this phase of the pilot will run.
- User Guide: link - instructions on how to use the new dashboard.
- PROTECTIVE Node Installation guide: link - installation notes.
- Webinar: link - link to the webinar series: 11/7, 13/7, 1/8, 24/8. Updates about the pilot.
- Past webinar videos and early versions of documents: link
- Feedback Logger tool and README – download and use tool: CESNET , PSNC, RoEduNet
    - Note: only difference is the organization name, all tools are otherwise the same
- Questionnaire to fill in (please do so before 9th August): MID-AUGUST, LATE-AUGUST

All: PowerPoint slides

All: Word doc Questionnaire

- PSNC: Component evaluation (starts with CESNET, PSNC session at 01:48:30), Viz evaluation.
- CESNET: Component evaluation (starts with CESNET, then PSNC session), Viz evaluation.
- RoEduNet: Component evaluation (starts with CESNET, then PSNC session), Viz evaluation (TODO).

All video files are less than 225MB each.

The other video files are: Not necessary to download, just added for reference.

- Phase2_Intro1... - recording of Webinar on 11th August.
- Phase2_Intro2... - recording of Webinar on 13th August.
- Phase2_End... - recording of Webinar on 24th August.

The GMV instance is available here as a reference: http://protective_dev.gmv.com/dashboard

Operator Tasks

Beginning of Phase 2 (11 July + 13 July):

- **Attend a webinar** on the 11th July (introduction) and 13th July (initial troubleshooting).
- **Upgrade the PROTECTIVE tool** between 11th and 13th July. **Identify upgrade issues** for the 13th July meeting.

Main Part of Phase 2 (16th July - 24th August):

- **Attend a webinar** on the 1st August and 17th August.
- **Leave the tool running in the CSIRT**, make sure it is up and running at all times.
- **Check the tool 2-3 times per day** (if no major incident is happening in the CSIRT), consider the system to be a supplementary tool to obtain threat intelligence.
- **Use the feedback logger to report comments (from 1st August onwards)**, these can be suggestions, minor bugs, reports on when the tool has been useful, etc.
    - **Once per week, send across the latest CSV files** for Pilot analysis
- **Email major (system breaking) bug issues** if any emerge to: pilot@protective-h2020.eu using the format in section 4 in this document.
- **Fill in questionnaires** sent on the 1st August and 17th August.

Pilot 1 - Phase 2 timeline

When	Activity	Requirements	Comments
July 11th	Kick-off	Protective 2.0 is been deployed, Pilot partners, GMV, AIT	
July 13th	Follow up troubleshooting for Phase 2 software	Operators are available, pilot partners GMV	
August 1st	Follow up, troubleshooting, and questionnaire	Operators are available, pilot partners, GMV	
August 16th	User-Interfacing Aspects	Search, Statistics and Detailed Views. Operators are available. UOXF, SYNYO.	Visualization of various contexts to alerts, meta-alerts and IP information
August 14th	Core System Components	Trust, Context Awareness, Correlation and Prioritisation. Operators are available. UOXF, SYNYO.	A focused session where the operators are introduced to the various component, what it does and how it does it.
August 24th	Wrap-up and final data collection	All actions finished. Attendance from operators required.	

Key purpose of Phase 2

The pilot is being conducted for the PROTECTIVE consortium to obtain feedback about the updates to the system, and opinions about directions the tool is heading.

Specifically, in this phase we wish to:

- identify the **benefits and issues in the updates of the tool**, e.g.: **New dashboard, Context Awareness, Trust of data and value of Meta Alerts**:
  - Positive effects or caused any issues, if so, what?
  - Are you obtaining new insights about compromised machines in your network which does not exist in your alerting systems? Are any of these of particular interest?
- identify the **practical issues** that emerge **while upgrading and using** the system:
  - Was there a lot of effort involved to get the tool running?
  - Was the tool stable and were any bugs detected while running the tool?
  - What is the usability of the tools like?
  - Have you had any problems with the database backend?
  - Have you had any problems with dockerization?
  - How can the deployment instructions be improved?

Key tasks of operators during the pilot

During the first phase of the Pilot 1 the main focus is on the evaluation of the deployment and underlying key event sharing functionality and stability of the system as well as the user experience. The purpose is to obtain first indicators of whether we are going in the right direction with the system at all, how stable it is in a live environment (not just testbed scenarios) and identify the real-added value is to the users (operators). Specifically, the key tasks of operators are:

- **Upgrade the tool 11th July –** critique the PROTECTIVE setup and deployment instructions provided.
- **Attend webinars (11th July, 13th July, 1st August, 17th August @ 12:00 GMT)** – to cover the progress about the pilot.
- **Run the tool in the background and maintain it if necessary (11th July - 17th August) –** Keep the tool alive and running in the background in the CSIRT.
- **Partake in user evaluations**.
- **Check the tool** (ideally two or three times per day, but at least once per day) – to make sure you are up to date on latest events that are shared.
- **Report feedback logs –** use the feedback logging tool to report how you have used the tool. This includes troubleshooting, (minor) bugs, feature requests, and incident handling (if PROTECTIVE has helped – please state how and what happened? E.g. “an event helped with the takedown of an IP address that was otherwise invisible on our network”).
- **Report major bugs - In the interest of clarity: send bug reports if the issue is major and need addressing ASAP. Send feedback logs if the bug is not a major issue.** Any very serious bugs, please report these as well over email to: pilot@protective-h2020.eu

If sending in a major bug please provide the following information in the email:

Timestamp: <Time the bug occurred for the first time and (if applicable) how often it occurs.>

Summary: <Should be a short and precise description of the bug. It should be possible to tell what this bug is about, and where it occurs from reading the summary alone.>

Components: <Which part of the system is affected?>

Environment:<Provide specs of the machine. Enter the environment under which you found the bug, make sure to include any information which might be relevant to the bug (e.g. your screen resolution, special plugins, etc.)>

Description: <This is the main description of the bug>. Make sure to include at least:

- A description of the bug.
- A very clear step by step description of how to reproduce the bug (if possible).
- A description of what you expected, and a description of what actually happens.
- Screenshots or other attachments are always helpful, but optional.

Webinars

Who should attend the webinars: all operators should attend, ideally all NREN points of contact (Gerard, Andrea and Mihai), the person who will be upgrading the system, and as many operators from CESNET, PSNC and RoEduNet as possible. The meeting will be screen recorded, so should any operators not be able to join, they can review the video again at any time afterwards by request. Should you have any questions about the pilot or using the tool that cannot be resolved easily (and need to be), please email: pilot@protective-h2020.eu

Where: link, see also Appendix A.

First webinar: "Introduction to the upgraded PROTECTIVE and how phase 2 will play out.**"**

Time: **11th July @ 12:00 GMT – expected duration, up to two hours.**

Agenda:

- Overview of phase 2.
  - Purpose and data to be collected.
  - Proposed dates and evaluations.
  - Holiday review (affects dates later in the phase).
- Showing how to do the upgrade.
- How to use the dashboard.
- Questions and AOB.

Second webinar: "Troubleshooting upgrade issues"

Time: **13th July @ 12:00 GMT – expected duration, up to two hours.**

Agenda:

- Troubleshooting and initial impressions.
- Questions and AOB.

Third webinar: "Progress Report"

Time: **1st August @ 12:00 GMT – expected duration, up to two hours.**

Agenda:

- Review of phase 2 progress - raise concerns if any:
  - how have you been using the tool?
  - any concerns we should be aware of?
  - have additional connectors been added?
- Review using the dashboard.
- Feedback logger usage.
- Review mid-phase questionnaires to be sent. Reply due: 9th August
- Update on component evaluations - finalising dates.
- Questions and AOB?

Fourth webinar: "End of Phase 2 Report"

Time: **24th August @ 12:00 GMT – expected duration, up to two hours.**

Agenda:

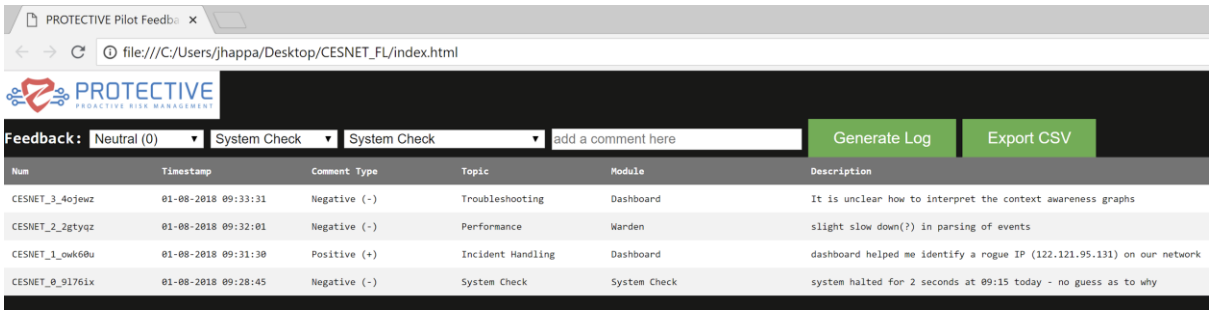
- Review of phase 2 - Reflect on experiences.

- Review end of phase questionnaires to be sent (due before 31st August). Details HERE.
- Suggestion for Pilot 2.

Feedback Logger

Aim: To obtain feedback from using the system.

Feedback logs are a means for operators to self-log emerging feedback using bug tracking-like templates, and follows a CSV-like format. The purpose of the feedback logs is to document key information about emerging usability issues "*as they are happening*". Logs aim to be as short as possible to minimise disruption of the operator, but also to maximise data collected. They are akin to a bug tracker, but differ in that their aim is not solely to identify bugs and troubleshooting concerns, but also opinions and a record of what has happened. The feedback logger will allow us to conduct data analytics on patterns related to usability. The logs would also aid during the evaluation to inform it about which questions are likely to be the most pertinent at the end of each phase of the pilot. **They do not connect to a database - instead only export CSV files directly from JavaScript in a single HTML file. This is done to minimise workload in setting it up and maintaining it.** The log files are stored as a spreadsheet. From the keywords, it is possible to examine the frequency patterns, trends in comments and compare feedback across NRENs. Please keep an eye that the CSVs are being output correctly, and **send weekly CSV files to pilot@protective-h2020.eu** The feedback logger and README installation instructions link is at the start of this document. **This will be covered in webinar 1 and webinar 2.**



We anticipate that **operators would fill in a row up to a few times per day,** although this is subject to emerging issues with the PROTECTIVE system. Fields are filled in next to the "Please add your feedback here" text. When each of the fields in a row are complete, pressing the "Generate Log" button will generate a log. Each x number of lines created will export a CSV file, but the operator can also export at any time using the "Export CSV" button. **Ideally the diary tool should be run on separate hardware from PROTECTIVE, and always on, e.g. on a laptop, if available.** Fields:

- **Comment Type** *<enum>* - gives an indication about the nature of the feedback in the log, allowing users to specify whether the log is intended to be negative, positive or neutral feedback. Enum options are: "Positive", "Neutral", "Negative".
- **Topic** *<enum>* - describes what the log is concerning using a pre-defined list of words. The list is setup to avoid overlaps in topics. The optional tag column allows for users to specify other tags not covered by topic. Enum options are: "Bug", "Feature", "Performance", "Incident Handling", "Troubleshooting", "Request", "Other"
- **Module** *<enum>* - signifies which technical PROTECTIVE subsystem the log pertains to. Enum options are: <list of all connectors>, "Mentat", "Warden", "Other"
- **Description** *<string>* - free text that an end-user can use to describe their comment in depth. Users are asked to be as brief and informative as possible.

Note: NUM and timestamp have been removed since the last version to make each log entry faster to do. Persistence has been added using localStorage in the browser. This means that if the browser crashes or similar, you can simply reload the index file and it will reload the table as you last did it.
If you experience any issues with the local storage of logs, you can clear the table by pressing F12, entering the browser console and type in localStorage.clear()

When you check protective and you have no comment, please add:
Comment type: Neutral
Topic: System Check
Module: System Check
and leave the other fields empty. That way we can record that the system is working as expected and you've recorded it for us. If you have any issue or comment you'd like to add, please do.

Appendix D: PROTECTIVE Pilot 1, Phase 2 – Last Questionnaire[6]

This is a late phase impressions questionnaire. Here we aim to obtain insight about opinions and challenges that operators faced when using the PROTECTIVE system when connected to other NRENs. The questionnaire is aimed at identifying immediate challenges with the latest version tool. Feel free to have multiple operators answer a single sheet, instead of having each operator fill in their own form. We expect the form to take 1-1.5 hours to complete as we ask you to review content from the video in the webinar as well. If you need more space to complete your answer, feel free to make more space. Any questions, please email: pilot@protective-h2020.eu

1. Which NREN are you? CESNET/PSNC/RoEduNet

Section - Ratings questions about the system.

2. On a scale of 1-5 (1 = regarded lowly, 5 = regarded highly), please rate your experience of the system so far:
    a. Ease of Use (once installed)? -
    b. Value of viewing other NRENs events? -
    c. Value of statistics view? -
    d. Value of alert querying? -

For the remaining questions, we are interested in anything else (than what has been covered in the webinars). If your opinion has not changed since the webinar, it is enough to write: see webinar.

Section: Opinion questions about the system.

3. Please outline in bullet points:
    a. Your general opinion of the tool so far, positive and/or negative aspects?
    b. How have you used the system? E.g. how often have you used the tool so far, how have you used the dashboard?
    c. Briefly outline of key experiences so far – were there any challenges were there in upgrading, if so, what? Did you find something interesting or noteworthy in using the tool?
    d. What features and functionality would you like to see in the system going forward?
    e. Have any incidents been handled with the tool to support? If yes, how? Any success stories?
    f. How can the installation guide be improved?
    g. How can the user guide be improved?

Section – Pilot questions.

4. Please outline in bullet points:
    a. Is there anything in running the pilot so far you found problematic so far that has not been mentioned during the webinars?
    b. What is your opinion of the “major-issue” reporting procedure (email: pilot@protective-h2020.eu)? Positive and/or negative aspects.
    c. How can the reporting procedure be improved – if you have any additional thoughts since last time?

[6] Please note: there are only minor differences between the questionnaires at the various stages in the project. The last one is the most comprehensive one (as it includes a subcomponent evaluation specific to correlation, prioritisation, trust, context awareness and visualization. We therefore only include the last one in this report.

d. What is your opinion of the pilot support since last webinar? Would you like more, less or the same amount of involvement from support in the pilot? Are we responding fast enough?
e. How can the pilot guide be improved? E.g. are the guidelines clear, is anything missing?

Section – Pilot User Evaluation questions.

Please take time to study the PowerPoint slides and videos link presented to you last week (slides sent over email).

Correlation (p15):

- Are any of the correlation concept slides unclear to you?
- How would you like to use Correlation as part of your work? Any features missing?
- Are alerts grouped in meta-alerts in a way coherent with your expectations?
- Do meta-alerts contains enough information to react?
- Would you like to add something to meta-alerts?
- Are there any particular correlation scenarios/attack types that should be added?

Prioritisation (p18):

- Are any of the prioritisation slides unclear to you?
- How would you like to use prioritisation as part of your work? Any features missing?
- Which provided prioritisation criteria are relevant for you?
- Which criteria should be added/abandoned?
- Which should be mandatory/voluntary?
- Would you like to use your own data sets to learn prioritisation model? If so, how?

Trust (p35):

- Are any of the Trust concept slides unclear to you?
- How would you like to use Trust as part of your work? Any features missing?
- How much do you agree/disagree with the inputs and outputs of the Trust computation?
- Reviewing an alert with a high-quality score and one with a very low one: do you agree/disagree with the difference in the quality score?
- After viewing the parameters that are considered for the calculation (e.g., IPR, SR, etc.) Is anything missing?
- Critique our approach?

Context Awareness (p56):

- Are any of the context awareness concept slides unclear to you?
- How would you like to use context awareness as part of your work? Any features missing?
- What new insight would you expect to gather from context awareness?
- How can it be improved?

Visualization/Dashboard upgrades (p74):

- Are any of the visualization concept slides unclear to you?

- How would you like to use visualization as part of your work? Any features missing?
- What information from the raw – NERD, IDEA – information should be prioritized?

See: https://nerd.cesnet.cz/

- What additional information/annotations would be relevant/interesting
- What external services could be further linked to?
- What level of interaction would be relevant?