UNIVERSITY OF SANTIAGO DE COMPOSTELA

ESCOLA TÉCNICA SUPERIOR DE ENXEÑARÍA

Improvements in IDS: adding functionality to Wazuh

Autor:
Andrés Santiago Gómez Vidal

Directores:
Purificación Cariñena Amigo
Andrés Tarascó Acuña

Computer Engineering Degree

February 2019

Final degree project presented at the Escola Técnica Superior de Enxeñaría of the University of Santiago de Compostela to obtain the Degree in Computer Engineering

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Ms. Purificación Cariñena Amigo, Professor Computing Science and Artificial Intelligence at the University of Santiago de Compostela and **Mr. Andrés Tarascó Acuña**, Managing Director at Tarlogic Security S.L.

STATE:

That the present report entitled *Improvements in IDS: adding functionality to Wazuh* written by **Andrés Santiago Gómez Vidal** in order to obtain the ECTS corresponding to the final degree project of the Computer Engineering degree was conducted under our direction in the department of Computer Science and Artificial Intelligence of the University of Santiago de Compostela.

For the purpose to be duly recorded, this document was signed in Santiago de Compostela on February TODO, 2019:

The director,	The codirector,	The student,
(Purificación Cariñena Amigo)	(Andrés Tarascó Acuña)	(Andrés Gómez Vidal)

Index

List of Figures

List of Tables

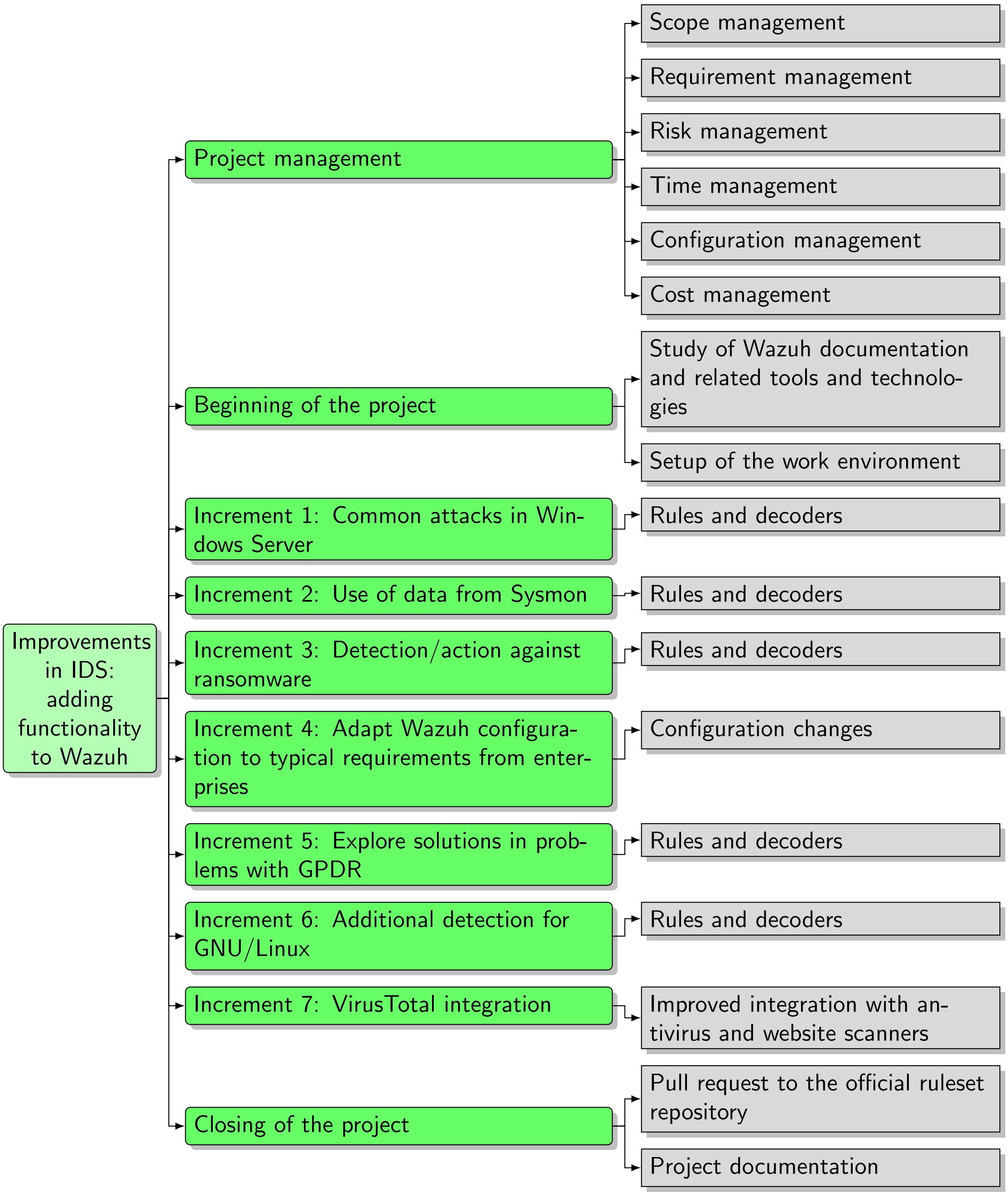
Chapter 1

Introdución

Chapter 2

Planning

2.1 Initial WBS


Improvements in IDS: adding functionality to Wazuh
Project management
Scope management
Requirement management
Risk management
Time management
Configuration management
Cost management
Beginning of the project
Study of Wazuh documentation and related tools and technologies
Setup of the work environment
Increment 1: Common attacks in Windows Server
Rules and decoders
Increment 2: Use of data from Sysmon
Rules and decoders
Increment 3: Detection/action against ransomware
Rules and decoders
Increment 4: Adapt Wazuh configuration to typical requirements from enterprises
Configuration changes
Increment 5: Explore solutions in problems with GPDR
Rules and decoders
Increment 6: Additional detection for GNU/Linux
Rules and decoders
Increment 7: VirusTotal integration
Improved integration with antivirus and website scanners
Closing of the project
Pull request to the official ruleset repository
Project documentation

2.2 Initial planning

2.3 Final planning

Chapter 3

Requirements

Chapter 4

Design

Chapter 5

Conclusions and additions

5.1 Risk management

5.1.1 Risk metrics

Chances of the risk happening	Probability
≥80%	High
Between 30% and 80%	Medium
≤30%	Low

Table 5.1: Probability classification of risks

Resource in Place / Effort / Cost	Impact
≥20%	High
Between 10% and 20%	Medium
≤10%	Low

Table 5.2: Impact classification of risks

Exposition		Probability		
		High	**Medium**	**Low**
Impact	**High**	High	High	Medium
	Medium	High	Medium	Low
	Low	Medium	Low	Low

Table 5.3: Method of calculation of Exposition based of Probability and Impact

5.1.2 Risk types

5.1.3 Risk identification

Table 5.4: Project risks

Identifier	Name
R-001	Optimist planning, “best case” (instead of a realistic “expected case”)
R-002	Bad requirement specification
R-003	Design errors
R-004	Lack of key information from sources
R-005	Lack of feedback or support from the security consultants of Tarlogic
R-006	The learning curve of some technologies is larger than expected
R-007	The unexplained parts of the project take more time than expected
R-008	Cannot access source material
R-009	Unexpected changes to any of the APIs used in the project
R-010	Loss of work
R-011	Wrong management of the project’s configuration
R-012	A delay in one task leads to cascading delays in the dependent tasks
R-013	Unnecessary work
R-014	The quality of the product is not enough
R-015	Sickness or overwork
R-016	Performance issues

5.1.4 Risk analysis

Identifier	**R-000**
Name	Bla
Description	Bla bla bla bla bla bla bla bla bla bla bla bla bla Bla bla bla bla bla bla bla bla bla bla bla bla bla Bla bla bla bla bla bla bla bla bla bla bla bla bla bla
Probability	Low , Medium , High
Impact	Low , Medium , High
Exposition	Low , Medium , High
Indicator	Bla bla bla bla bla bla bla bla bla bla bla bla bla

Identifier	**R-001**
Name	Optimist planning, “best case” (instead of a realistic “expected case”)
Description	An optimistic planning at the start of the project does not take into account problems or delays, and so it does not allocate time for them, leading to cascading delays if they happen.
Probability	Medium
Impact	High
Exposition	High
Indicator	There are 3 consecutive delays, after the beginning of the project.

Identifier	**R-002**
Name	Bad requirement specification
Description	The requirements specified at the beginning of the project are not specific enough, are not needed or there are new requirements after the beginning of the project.
Probability	High
Impact	High
Exposition	High
Indicator	There are 3 changes in the requirements specification.

Identifier	**R-003**
Name	Design errors
Description	A design is not enough or is incorrect, needing a re-design and probably changes in the next steps it was used.
Probability	Low
Impact	Medium
Exposition	Medium
Indicator	There are 3 designs that need rework.

Identifier	**R-004**
Name	Lack of key information from sources
Description	Not having key information from articles, documentation or manuals can result in unexpected delays, added difficulty or the need to rework completely the functionality.
Probability	Medium
Impact	High
Exposition	High
Indicator	The duration of the study of the attack and the needed tools takes 50% than expected.

Identifier	**R-005**
Name	Lack of feedback or support from the security consultants of Tarlogic
Description	Because I do not know enough of some technical aspects of cibersecurity to solve all the problems in this by myself in time, Tarlogic has promised to help (in a tutoring way) if a problem arises. This help could be critical to solve or get around some of the most complex problems, which probably happen to be critical points, needing to be dealt with to continue working on that stage.
Probability	Medium
Impact	High
Exposition	High
Indicator	A simple technical question takes more than 2 working days to be answered or a complex question takes more than 7 working days.

Identifier	**R-006**
Name	The learning curve of some technologies is larger than expected
Description	This is a critical need because not having enough knowledge can result in unexpected delays, added difficulty or the need to rework completely the functionality.
Probability	Medium
Impact	Medium
Exposition	Medium
Indicator	The duration of the study of the technologies takes 50% than expected.

Identifier	**R-007**
Name	The unexplained parts of the project take more time than expected
Description	There is not enough specification on what a tasks implies or not enough planning. This means that a part of the project is not understood as it should, and the work done is not what was expected or is not enough, needing more time to finish.
Probability	Low
Impact	High
Exposition	Medium
Indicator	A task takes 15% more time than expected and when the causes are investigated it is revealed that there were ambiguous descriptions or planning.

Identifier	**R-008**
Name	Cannot access source material
Description	All or part of the source material can not be accessed, probably because the only host of the resource is down. In some cases this could mean a delay in a critical task, cascading in other delays and delaying the project for a period unknown.
Probability	Low
Impact	High
Exposition	Medium
Indicator	There have been at least 10 failed attempts to download the source material, at least 5 with a computer A in a network X and at least 5 with a computer B in a network Y.

Identifier	**R-009**
Name	Unexpected changes to any of the APIs used in the project
Description	Changes to an API could affect this project directly or indirectly. Programs could fail or not work as expected. In a project that does not work in a bleeding edge environment, like this, this should be very rare and even if it were to happen it would have to interfere with the part of the API this project uses, which (as this is not bleeding edge) normally would be backwards compatible.
Probability	Low
Impact	Low
Exposition	Low
Indicator	There are 3 failures due to changes in APIs.

Identifier	**R-010**
Name	Loss of work
Description	Due to a bad configuration management or something else, there is a loss of work related to this project.
Probability	Low
Impact	High
Exposition	Medium
Indicator	The need to replicate already done work is greater than 30 minutes.

Identifier	**R-011**
Name	Wrong management of the project's configuration
Description	The project's configuration is inefficient or lacks work. Some of the problems could be: • Wrong baselines • Wrong identification of the configuration elements • It takes more time than expected to manage the project • Wrong use of the tools • Too much time between commits • Changes are unclear This means the project suffer delays because the need to redo management work and/or planned tasks.
Probability	Medium
Impact	High
Exposition	High
Indicator	There are 3 delays because of the configuration of the project.

Identifier	**R-012**
Name	A delay in one task leads to cascading delays in the dependent tasks
Description	A task gets delayed and one or more tasks depends on its completion to start, so they get delayed too.
Probability	Medium
Impact	Medium
Exposition	Medium
Indicator	At least 2 tasks are delayed, due to only one of them needing more time.

Identifier	**R-013**
Name	Unnecessary work
Description	Resources are wasted in work that latter is not used. This could happen because multiple reasons, like wrong assumptions or balancing of the remaining time of the project.
Probability	Low
Impact	Low
Exposition	Low
Indicator	At least 3 commits are reverted.

Identifier	**R-014**
Name	The quality of the product is not enough
Description	The final result is does not comply the quality standard set for this project. This could mean the need to redo work in a later stage or the incorporation to the official repository being rejected.
Probability	Low
Impact	High
Exposition	Medium
Indicator	Getting 10 suggestions to rework functionality.

Identifier	**R-015**
Name	Sickness or overwork
Description	The health of the student deteriorates to the point it affects the project, and it is caused by sickness or overwork.
Probability	Medium
Impact	Medium
Exposition	Medium
Indicator	There is an unexpected delay because the functionality is not done but there has not been any important issues that could explain it.

Identifier	**R-016**
Name	Performance issues
Description	The program is too heavy for the environment and takes too much resources, because there are not good enough optimizations or the problems are poorly approached.
Probability	Low
Impact	Low
Exposition	Low
Indicator	The program takes 30% more resources that at the beginning of the project.

5.1.5 Risk planning

5.1.6 Risk supervision

Appendix A

Manuais técnicos

Manuais técnicos: en función do tipo de Traballo e metodoloxía empregada, o contido poderase dividir en varios documentos. En todo caso, neles incluirase toda a información precisa para aquelas persoas que se vaian a encargar do desenvolvemento e/ou modificación do Sistema (por exemplo código fonte, recursos necesarios, operacións necesarias para modificacións e probas, posibles problemas, etc.). O código fonte poderase entregar en soporte informático en formatos PDF ou postscript.

Appendix B

Manuais de usuario

Manuais de usuario: incluirán toda a información precisa para aquelas persoas que utilicen o Sistema: instalación, utilización, configuración, mensaxes de erro, etc. A documentación do usuario debe ser autocontida, é dicir, para o seu entendemento o usuario final non debe precisar da lectura de outro manual técnico.

Appendix C

Licenza

Se se quere pór unha licenza (GNU GPL, Creative Commons, etc), o texto da licenza vai aquí.