Title	European network of Cybersecurity centres and competence Hub for innovation and Operations
Acronym	ECHO
Number	830943
Type of instrument	Research and Innovation Action
Topic	SU-ICT-03-2018
Starting date	01/02/2019
Duration	48
Website	www.echonetwork.eu

D4.3 INTER-SECTOR CYBERSECURITY TECHNOLOGY ROADMAP

Work package	WP4 Inter-sector Technology Roadmap
Lead author	Peter Kirkov (TBS)
Contributors	Please see table “List of contributors”
Peer reviewers	Andrew James Roberts (TUT)
Version	V1.0
Due date	31/07/2020
Submission date	31/08/2020

Dissemination level

X	PU: Public
	CO: Confidential, only for members of the consortium (including the Commission)
	EU-RES. Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)
	EU-CON. Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)
	EU-SEC. Classified Information: SECRET UE (Commission Decision 2005/444/EC)

The work described in this document has been conducted within the ECHO project. This project has received funding by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 830943

Version history

Revision	Date	Editor	Comments
0.1	14/07/2019	Peter Kirkov (TBS)	Initial approach
0.2	15.04.2020	Peter Kirkov (TBS)	Introduction, Roadmap development methodology chapters, EPIC 3, chapter integration
0.3	28.05.2020	Peter Kirkov (TBS)	Refactoring of Epics and User Stories. Integration of contributions
0.4	03.08.2020	Peter Kirkov (TBS), Krasimir Ivanov (TBS), Matteo Merialdo (RHEA)	Integration of contributions, formatting, citation and reference fixing, rewording. Final deliverable preparation.
0.5	20.08.2020	Peter Kirkov (TBS), Krasimir Ivanov (TBS), Matteo Merialdo (RHEA)	User Story rework. Document structure consolidation. Meta information.
0.7	22.08.2020	Peter Kirkov (TBS)	Reworked User Stories, reorganized document, removal of irrelevant contributions as per initial feedback from stakeholders.
0.7	24.08.2020	Peter Kirkov (TBS)	Removed comments, track changes, fixed pagination. General clean up before QA
0.8	29.08.2020	Peter Kirkov (TBS)	Rework after peer review feedback
1.0	31.08.2020	Ewa Konieczna (VST)	Minor updates and formatting

List of contributors

The list of contributors to this deliverable are presented in the following table:

Section	Contributor
Document information	Krasimir Ivanov (TBS)
1, 2	Peter Kirkov (TBS)
3.1.1	Peter Kirkov (TBS), Marco Pappalardo (CIRM), Riccardo Delpopolo Carciopolo (CIRM), Luis Galindo (TME)
3.1.2	Giuseppe Chechile (FNC), Riccardo Feletto (FNC)
3.1.3	Cagatay Yucel (BU), Daniel McNeill (EXP)
3.1.4	Marco Pappalardo (CIRM), Gabriella Trombino (CIRM)
3.1.5	Antonis Voulgaridis (CERTH), Nikos Oikonomou (CERTH), Minas Spanopoulos-Karalexidis (CERTH), Marcin Niemiec (AGH), Bartlomiej Szpila (AGH) and Michal Ciborowski (AGH)
3.1.6	Peter Kirkov (TBS)
3.1.7	Monica Costantini (LCU)
3.2.1	Luis Galindo (TME), Notis Mengidis (CERTH), Roberto Martinez (TME), Ramón Cebrián (TME)
3.2.2	Matteo Merialdo (RHEA), Andrea Rossi (RHEA) and Pierluigi Cara (RHEA)
3.3.1	Antonis Voulgaridis (CERTH), Ioannis Chalkias (BU), Daniele Cristofori (Z&P)
3.3.2	Peter Kirkov (TBS)
3.4.1	Notis Mengidis (CERTH), Peter Kirkov (TBS)
3.4.2	Valery Yakmadzhiev (TBS)
3.4.3	Matteo Merialdo (RHEA)
3.4.4	Paloma de la Vallée (RMA)
3.4.5	Antonis Voulgaridis (CERTH), Mike Anastasiadis (CERTH), Giorgos Aivatoglou (CERTH)
4.1.1	Oleg Illiashenko (KhAI), Vyacheslav Kharchenko (KhAI), Maryna Kolisnyk (KhAI), Herman Fesenko (KhAI), Marco Pappalardo (CIRM), Gabriella Trombino (CIRM)
4.2.1	Giuseppe Chechile (FNC), Riccardo Feletto (FNC)
4.2.2	Pencho Vasilev (BDI), Veselin Dobrev (BDI), Tsvetelin Tsonev (BDI)
4.2.3	Giuseppe Chechile (FNC), Riccardo Feletto (FNC), Peter Kirkov (TBS)
4.2.4	Fabrizio De Vecchis (RHEA), Peter Kirkov (TBS)
4.3.1	Oleg Illiashenko (KhAI), Vyacheslav Kharchenko (KhAI), Maryna Kolisnyk (KhAI), Herman Fesenko (KhAI)
4.3.2	Daniel McNeill (EXP)
4.3.3	Fabrizio De Vecchis (RHEA), Peter Kirkov (TBS)
5	Peter Kirkov (TBS)
Annex 1	Monica Costantini (LCU)
Annex 2	Luis Galindo (TME), Notis Mengidis (CERTH), Roberto Martinez (TME), Ramón Cebrián (TME)

Table 1: List of contributors

Keywords

FEDERATED CYBER-RANGE, EARLY WARNING SYSTEM, INTER-SECTOR, MULTI-SECTOR, TECHNOLOGICAL, CHALLENGES AND OPPORTUNITIES, ROADMAPS.

Disclaimer

Executive summary

One of the main objective of Work Package 4(WP4) is the development of cybersecurity technology roadmaps as a result of the analysis related to current emerging cybersecurity challenges and associated technologies. These roadmaps will create the foundations for new industrial capabilities and assist in the development of innovative technologies that will aim to address the identified cybersecurity challenges. Based on the roadmaps, the early prototypes research and development in T4.3 will address some of the identified opportunities. To support these objectives and the objectives of ECHO project, at least six inter-sector technology roadmaps will be developed following the challenges identified in D4.1 Transversal technical cybersecurity challenges report.

As from the GA:

ECHO includes delivery of at least 6 cybersecurity technology roadmaps including:

- *ECHO Early Warning System delivered as part of the project;*
- *ECHO Federated Cyber Range delivered as part of the project;*
- *At least 2 additional technology innovations to be completed as part of the project;*
- *At least 2 additional technology innovations to be addressed by the future Cybersecurity Competence Network (CCN)*

Of these six technology roadmaps to be developed, the GA suggests that two of these technology roadmaps describe future (inter sector) opportunities for the development of ECHO Federated Cyber Range (E-FCR) and ECHO Early Warning System (E-EWS) following the development plan of those two platforms in WP5 and WP6, respectively. The purpose of the E-FCR and E-EWS technology roadmaps is to present future development options aiming for the continuous improvement, adaptation and evolution of the ECHO platforms after the end of the project and when the .CCN will be active.

Roadmaps outline innovative and horizontal capabilities that could not have been captured otherwise. The Consortium hence decided to dedicate this initial version of the deliverable to the E-FCR and E-EWS future roadmaps (since they will come useful for the actual software development activities), while D4.10 will include the other four roadmaps.

Explored topics were identified using brainstorming sessions with participation of a broad spectrum of specialist, researchers and experts within the ECHO consortium. These topics were consolidated and developed into specific technology lines using the SCRUM methodology adopted for the roadmap creation in the environment of ECHO consortium.

Due to the number of potential development opportunities, the roadmaps have been divided into sub-sections according to major technological areas described, called EPICs within this document: five for E-FCR and three for E-EWS. Those EPICs were further divided into sub-topics called UserStories (US), each describing specific technology drivers and their targets, technology alternatives and their timelines. In the two roadmaps, total of 29 user stories were identified and developed, some of them including multiple development opportunities for the ECHO platforms.

Table of contents

List of figures

List of tables

1. Introduction

1.1 ECHO Project

The ECHO project (European network of Cybersecurity centres and competence Hub for innovation and Operations) is one of four Pilot projects, launched by the European Commission, to establish and operate a Cybersecurity Competence Network. ECHO delivers an organized and coordinated approach to strengthen the proactive cyber defence of the European Union, through effective and efficient multi-sector collaboration. The project includes 9 Work Packages (WP) from which Work Package 4 (WP4) is focused on the development of cybersecurity technology roadmaps as a result of the analysis related to current and emerging cybersecurity challenges and associated technologies. These roadmaps will create the foundations for new industrial capabilities and assist in the development of innovative technologies that will aim to address these cybersecurity challenges. Consequent research and development of early prototypes will target specific, high-priority opportunities identified in these roadmaps.

The specific objectives that relate to this document are:

- Demonstration of a network of cyber research and competence centres with a central competence hub, having a mandate for increasing participation through a new partner engagement model, including collaboration with other networks funded under the same call.
- Address current cybersecurity EU gaps. Development of an adaptive model for information sharing and collaboration among the network of cybersecurity centres supported by an early warning system and a framework for improved cyber skills development and technology roadmap delivery, in a multiple-sector context.

The ECHO network, at the time of writing of this deliverable, consists of 30 partners, spread across Europe. At the time of writing this document, onboarding activities are progressing to identify and include new partners in the ECHO network.

1.2 Purpose and scope of the document

The focus of task T4.2 *Inter-sector roadmap development* is, as the name suggests, the development of the inter-sector roadmaps, starting from the analysis of challenges and opportunities identified in T4.1 “Detailed analysis of transversal technical cybersecurity challenges” and T2.1 “Sector scenario and use case analysis”. The roadmaps provide a set of tangible opportunities, encouraging partners to engage together in common research, development and innovation of the next generation cybersecurity technologies, applications and services. Achievement of those opportunities would demonstrate mastery of these relevant cybersecurity technologies, including horizontal inter-sector technologies, cybersecurity disciplines and transversal factors. In this deliverable, two of the roadmaps of the ECHO specific technology innovations are documented - ECHO Early Warning System (E-EWS), the ECHO Federated Cyber Range (E-FCR). The roadmaps will support the strategic and long-range planning by matching short-term and long-term goals with; specific technologies, methodologies and solutions. The roadmaps will also explore the evolution of external factors that affect the environment and context of both ECHO platforms.

The deliverable outlines the objectives and features that might be chosen to be developed in the future, and it can be used to develop selected features into series of action plans to guide the implementation of specific versions or capabilities after development planned under ECHO Project is completed. The timeframe of the roadmap is targeted after planned software development processes in ECHO Work Package (WP) 5 and 6 respectively are completed as a way to map future development. These roadmaps will be updated in January 2023, the M48 of ECHO project, as D4.10 “Update – Inter-sector cybersecurity technology roadmap” to reflect

better development options and capabilities of the tools. The update of this deliverable will incorporate additional technology roadmaps to include up to two more specific technology innovations developed under ECHO and up to two technology innovations targeted to meet specific industrial challenges having a horizontal impact across the sectors as per ECHO proposal.

1.3 Structure of the document

The document is divided into five major sections as follows:

Section 1 structures the document description and meta information. It describes the relation of the Technology Roadmap to ECHO project and interlinkage with other ECHO activities and deliverables.

Section 2 contains a description of the roadmap development methodology

Section 3 consists of the technology roadmap of E-FCR, split into sections according to the major technology domains, affecting the platform (called EPICs):

- User Experience
- Connectivity
- Scalability
- Platform
- Exploitation

Section 4 consists of the technology roadmap of E-EWS, split into sections according to the major technology domains affecting the platform:

- User Experience
- Platform
- Exploitation

Conclusions are summarised in the Section 5 of this document.

Each section consists of discrete articles dealing with a particular aspect or technology that influences ECHO Platforms.

The Annex comprises:

- Annex 1 – Customer feedback form for E-FCR
- Annex 2 – 5G Technology concepts

1.4 Relation to other work in the project

The work on this deliverable takes as input the following ECHO deliverables:

- D2.2 “Echo Multi-Sector Assessment Framework” developed as part of T2.2 “Derivation of ECHO Multi-sector Assessment Framework” was consulted in lieu of E-MAF still in development while compiling this roadmap. Pointers for risk assessment, and vocabulary were used in order to stay as true as possible to the E-MAF concept and minimize rework once E-MAF is finalized.
- D4.1 “Transversal Cybersecurity Needs Analysis” which in turn used D2.1 from T2.1 “Sector scenario use case analysis” as an input in order to derive challenges from cybersecurity scenarios and also to identify threats based on known cyberattacks and cybersecurity threat trends and D2.4 from T2.4 “Technological challenges and opportunities” to identify the specifics of each sector, and see whether technical-based approach was required. Appropriate, relevant to E-EWS and E-FCR challenges identified in D4.1 were addressed in the development of the specific technology roadmap UserStories.

This roadmap contains a set of tangible development opportunities that are going to be the basis for further advancement of the ECHO platforms (E-EWS and E-FCR). These will ensure sustainability, adaptability and evolution of the two solutions.

1.5 Applicable and reference documents

The following documents contain requirements applicable to the generation of this document:

Reference	Document Title	Document Reference	Version	Date
[GA]	Grant Agreement 830943 – ECHO		N/A	02/04/2019
[PH]	D1.1 Project Handbook	ECHO_D1.1_v1.42	1.42	20/10/2019
[PQP]	D1.3 Project Quality Plan	ECHO_D1.3_v1.2	1.2	19/05/2020

Table 2: Applicable documents

The following documents have been consulted for the generation of this document:

Reference	Document Title	Document Reference	Version	Date
[1]	CONSULTA PÚBLICA del MINETAD Plan Nacional de 5G	Retrieved July 30 2020 from: https://avancedigital.gob.es/es-es/Participacion/RespuestasPlan5G/40_Telefonica_NO_confidencial.pdf		
[2]	Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges.	Ordonez-Lucena, J., Ameigeiras, P., Lopez, D., Ramos-Munoz, J.J., Lorca, J. and Folgueira, J., 2017. Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges. *IEEE Communications Magazine*, *55*(5), pp.80-87.		2017
[3]	SDN Architecture Overview	TR, O., 2016. 521,“. *SDN Architecture Overview*, (1.1).	1.1	2016
[4]	Telefónica’s UNICA architecture strategy for network virtualisation	Cooperson, D. and Chappell, C., 2017. Telefónica’s UNICA architecture strategy for network virtualisation. *White paper*.		2017
[5]	Applying SDN Architecture to 5G Slicing	TR, O., 2016. 526,“. Applying SDN Architecture to 5G Slicing.		2016
[6]	Network Function Virtualisation (NFV)-Architectural Framework	ETSI, N., 2013. GS NFV 002-V1.1.1-Network Function Virtualisation (NFV)-Architectural Framework. (2014)	1.1.1	2014
[7]	Who owns the road? The IoT-connected car of today-and tomorrow	Ninan, S., Gangula, B., von Alten, M. and Snidermann, B., 2015. Who owns the road? The IoT-connected car of today—and tomorrow. *Deloitte University Press, August*, *18*, p.2015.		2015
[8]	Open Access and Edge Computing Whitepaper	Telefónica, 2019. Open Access and Edge Computing Whitepaper. Retrieved July 30 2020 from: https://www.telefonica.com/documents/737979/144981357/whitepaper-telefonica-opa-mec-feb-2019.pdf		2019

Reference	Document Title	Document Reference	Version	Date
[9]	Virtualisation essentials	Portnoy, M., 2012. *Virtualisation essentials* (Vol. 19). John Wiley & Sons.		2012
[10]	Formal requirements for virtualisable third generation architectures	Popek, G.J. and Goldberg, R.P., 1974. Formal requirements for virtualisable third generation architectures. *Communications of the ACM*, *17*(7), pp.412-421.		1974
[11]	Virtualising I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor	Sugerman, J., Venkitachalam, G. and Lim, B.H., 2001, June. Virtualising I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor. In *USENIX Annual Technical Conference, General Track* (pp. 1-14).		2001
[12]	The 2020 State of Virtualisation Technology	Spiceworks. The 2020 State of Virtualisation Technology. Retrieved July 23, 2020 from: https://www.spiceworks.com/marketing/reports/state-of-virtualisation/		2020
[13]	Intent based network operations	Campanella, A., 2019, March. Intent based network operations. In 2019 Optical Fiber Communications Conference and Exhibition (OFC) (pp. 1-3). IEEE.		2019
[14]	Unikernels: Rise of the virtual library operating system	Madhavapeddy, A. and Scott, D.J., 2013. Unikernels: Rise of the virtual library operating system. Queue, 11(11), pp.30-44.		2013
[15]	A performance benchmarking analysis of hypervisors containers and unikernels on ARMv8 and x86 CPUs.	Acharya, A., Fanguède, J., Paolino, M. and Raho, D., 2018, June. A performance benchmarking analysis of hypervisors containers and unikernels on ARMv8 and x86 CPUs. In 2018 European Conference on Networks and Communications (EuCNC) (pp. 282-9). IEEE.		2018
[16]	Detecting Network Threats using OSINT Knowledge-based IDS	Vacas, Ivo; Medeiros, Iberia; Neves, Nuno - Detecting Network Threats using OSINT Knowledge-based IDS, 2018		2018
[17]	CONSULTA PÚBLICA del MINETAD Plan Nacional de 5G	Retrieved July 30 2020 from: https://avancedigital.gob.es/es-es/Participacion/RespuestasPlan5G/40_Telefonica_NO_confidencial.pdf		
[18]	Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges.	Ordonez-Lucena, J., Ameigeiras, P., Lopez, D., Ramos-Munoz, J.J., Lorca, J. and Folgueira, J., 2017. Network slicing for 5G with SDN/NFV: Concepts,		2017

Reference	Document Title	Document Reference	Version	Date
		architectures, and challenges. *IEEE Communications Magazine*, *55*(5), pp.80-87.		
[19]	SDN Architecture Overview	TR, O., 2016. 521,". *SDN Architecture Overview*, (1.1).	1.1	2016
[20]	Telefónica's UNICA architecture strategy for network virtualisation	Cooperson, D. and Chappell, C., 2017. Telefónica's UNICA architecture strategy for network virtualisation. *White paper*.		2017
[21]	Applying SDN Architecture to 5G Slicing	TR, O., 2016. 526,". Applying SDN Architecture to 5G Slicing.		2016
[22]	Network Function Virtualisation (NFV)-Architectural Framework	ETSI, N., 2013. GS NFV 002-V1.1.1-Network Function Virtualisation (NFV)-Architectural Framework. (2014)	1.1.1	2014
[23]	Who owns the road? The IoT-connected car of today-and tomorrow	Ninan, S., Gangula, B., von Alten, M. and Snidermann, B., 2015. Who owns the road? The IoT-connected car of today—and tomorrow. *Deloitte University Press, August*, *18*, p.2015.		2015
[24]	Open Access and Edge Computing Whitepaper	Telefónica, 2019. Open Access and Edge Computing Whitepaper. Retrieved July 30 2020 from: https://www.telefonica.com/documents/737979/144981357/whitepaper-telefonica-opa-mec-feb-2019.pdf		2019
[25]	Virtualisation essentials	Portnoy, M., 2012. *Virtualisation essentials* (Vol. 19). John Wiley & Sons.		2012
[26]	Formal requirements for virtualisable third generation architectures	Popek, G.J. and Goldberg, R.P., 1974. Formal requirements for virtualisable third generation architectures. *Communications of the ACM*, *17*(7), pp.412-421.		1974
[27]	Virtualising I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor	Sugerman, J., Venkitachalam, G. and Lim, B.H., 2001, June. Virtualising I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor. In *USENIX Annual Technical Conference, General Track* (pp. 1-14).		2001
[28]	The 2020 State of Virtualisation Technology	Spiceworks. The 2020 State of Virtualisation Technology. Retrieved July 23, 2020 from: https://www.spiceworks.com/marketing/reports/state-of-virtualisation/		2020
[29]	Intent based network operations	Campanella, A., 2019, March. Intent based network operations. In 2019 Optical Fiber Communications Conference and Exhibition (OFC) (pp. 1-3). IEEE.		2019

Reference	Document Title	Document Reference	Version	Date
[30]	Unikernels: Rise of the virtual library operating system	Madhavapeddy, A. and Scott, D.J., 2013. Unikernels: Rise of the virtual library operating system. Queue, 11(11), pp.30-44.		2013
[31]	A performance benchmarking analysis of hypervisors containers and unikernels on ARMv8 and x86 CPUs.	Acharya, A., Fanguède, J., Paolino, M. and Raho, D., 2018, June. A performance benchmarking analysis of hypervisors containers and unikernels on ARMv8 and x86 CPUs. In 2018 European Conference on Networks and Communications (EuCNC) (pp. 282-9). IEEE.		2018
[32]	Detecting Network Threats using OSINT Knowledge-based IDS	Vacas, Ivo; Medeiros, Iberia; Neves, Nuno - Detecting Network Threats using OSINT Knowledge-based IDS, 2018		2018
[33]	Definition: Threat Intelligence	Rob McMillan, Retrieved July 30 from: https://www.gartner.com/en/documents/2487216/definition-threat-intelligence		16 May 2013
[34]	Transparency in Algorithmic Decision Making	Andreas Rauber (TU Wien and SBA), Roberto Trasarti and Fosca Giannotti (ISTI-CNR), ERCIM News 116 January 2019, available at https://ercim-news.ercim.eu/en116/special/indicaeting-automatically-detecting-extracting-and-correlating-cyber-threat-intelligence-from-raw-computer-log-data		January 2019
[35]		Structured Threat Information eXpression and Trusted Automated eXchange of Intelligence Information standards		
[36]	RFC 5070: The Incident Object Description Exchange Format	https://tools.ietf.org/html/rfc5070	Obsoleted by: 7970 Updated by: 6685	December 2007
[37]	Gibb, Will: OpenIOC: Back to the Basics	https://www.fireeye.com/blog/threat-research/2013/10/openioc-basics.html		October 01, 2013
[38]	Purple Teaming for Success	https://blog.secureideas.com/2014/04/purple-teaming-for-success.html		April 29, 2014
[39]	Adversary Simulation Becomes a Thing…	https://blog.cobaltstrike.com/2014/11/12/adversary-simulation-becomes-a-thing/		November 12, 2014
[40]	APT3 Adversary Emulation Plan	Christopher A. Korban Douglas P. Miller Adam Pennington Cody B. Thomas MTR170446 MITRE TECHNICAL REPORT Retrieved July 30 2020 from:		September 2017

Reference	Document Title	Document Reference	Version	Date
		https://attack.mitre.org/docs/APT3_Adversary_Emulation_Plan.pdf		
[41]	CVE-2017-0199: In the Wild Attacks Leveraging HTA Handler	Genwei Jiang, Rahul Mohandas, Jonathan Leathery, Alex Berry, Lennard Galang, Retrieved July 30 from https://www.fireeye.com/blog/threat-research/2017/04/cve-2017-0199-hta-handler.html		April 11, 2017
[42]	Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) License	Retrieved July 30 2020 from: https://creativecommons.org/licenses/by-sa/4.0/	N/A	
[43]	NIST cyber ranges brochure	Retrieved on July 30, 2020 from https://www.nist.gov/system/files/documents/2018/02/13/cyber_ranges.pdf	N/A	N/A
[44]	OECD (2019), The Role of Digital Platforms in the Collection of VAT/GST on Online Sales, OECD, Paris.	Retrieved July 30 2020 from: www.oecd.org/tax/consumption/the-role-of-digital-platforms-in-the-collection-of-vat-gst-on-online-sales.pdf	N/A	2019
[45]	The Conditions of Learning (4th Ed.).	Gagne, R. New York: Holt, Rinehart & Winston.	4th ed	1985
[46]	Principles of Instructional Design	Gagne, R., Briggs, L. & Wager, W; Fort Worth, TX: HBJ College Publishers.	4th ed	1992
[47]	. Instructional Design: International Perspectives. Theory, research, and models. Vol. 1	Tennyson, Robert; Dijkstra, S.; Schott, Frank; Seel, Norbert (1997), . Mahwah, NJ: Lawrence Erlbaum Associates, Inc. p. 42.	N/A	1997
[48]	An integrated model for automating instructional design and delivery.	Merrill, M. D. In J. M. Spector, M. C. Polson, & D. J. Muraida (Eds.), Automating instructional design: Concepts and issues (pp. 147–190). Englewood Cliffs, NJ: Educational Technology.	N/A	1993
[49]	First principles of instruction.	Merrill, M. D. (2001), Journal of Structural Learning and Intelligent Systems, 14(4), 459–466.	N/A	2001
[50]	Framework for Evaluating Cybersecurity-Related Skills Based on Computer Simulations	Sten Mäses, Liina Randmann, Olaf Maennel, Birgy Lorenz, Stenmap, Tallinn University of Technology, Tallinn, Estonia	N/A	2018
[51]	From Simple Scoring Towards a Meaningful Interpretation of Learning in Cybersecurity Exercises	Margus Ernits, Kaie Maennel, Sten Mäses, Toomas Lepik, and Olaf Maennel, , School of Information Technologies, Tallinn University of Technology, Tallinn, Estonia, 2020	N/A	2020
[52]	An Introduction to Cyber Modeling and Simulation	Jerry M Couretas, John Wiley & Sons, Inc.	N/A	2019
[53]	Webinar - Workforce Ready Cybersecurity Theory and Practice in Academic Studies	Cyberbit, Retrieved on July 30 from https://www.youtube.com/watch?v=FRs6XdQSlUo	N/A	March 11th, 2020

Reference	Document Title	Document Reference	Version	Date
[54]	"From Game Design Elements to Gamefulness: Defining Gamification"	Deterding Sebastian, Dixon Dan, Khaled Rilla, Nacke, Lennart		September 2018
[55]	Cyber-security training: A comparative analysis of cyber ranges and emerging trends,	Evangelos C. Chaskos, MSc Thesis, NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS, SCHOOL OF SCIENCE, DEPARTMENT OF INFORMATICS AND TELECOMMUNICATION Athens	N/A	March 2019
[56]	8 methodologies that every 21st century teacher should know	Retrieved on July 30 from https://www.realinfluencers.es/en/2019/05/09/8-21st-century-methodologies/	N/A	9th May 2019 • 23rd, redacted on October 2019
[57]	The three dimensions of learning.	Illeris, Knud, Malabar, Fla: Krieger Pub. Co. ISBN 9781575242583.	N/A	2004
[58]	Human learning	Ormrod, Jeanne. Boston: Pearson. ISBN 9780132595186.	6th ed	2012
[59]	The Life and Times of Cybersecurity Professionals.	Enterprise Strategy Group, Research Report. Retrieved July 30 from https://cdn.ymaws.com/www.members.issa.org/resource/resmgr/surveys/ESG-ISSA-2017-full.pdf	N/A	2017
[60]	Perspectives on Learning,	Phillips, D., Soltis, J. F. (2015). 5th Edition. Thinking about education series. Teachers College Press, 2015, p. 22, ISBN 0807771201, 9780807771204	5th ed	2015
[61]	e-Learning, online learning, and distance learning environments: Are they the same?	Moore, J., Dickson-Deane, C. and Galyen, K.. Internet and Higher Education, 14(2011):129-135.		2011
[62]	Alsawaier, R. (2018) "The Effect of Gamification on Motivation and Engagement",	The International Journal of Information and Learning Technology, Vol 35, No. 1, pp 56-79.	N/A	December, 2019
[63]	*Dicheva, D. et al (2015) "Gamification in Education: A Systematic Mapping Study",	Journal of Educational Technology & Society, Vol 18, No. 3, pp 75.	N/A	December, 2019
[64]	Zichermann, G., and Cunningham, C. (2011) Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps,	O'Reilly Media, Sebastopol.	N/A	December, 2019
[65]	Bell, K. (2017) Game On!: Gamification, Gameful Design, and the Rise of the Gamer Educator,	Johns Hopkins University Press, Baltimore.	N/A	December, 2019

Reference	Document Title	Document Reference	Version	Date
[66]	Bollini, L. (2017) Beautiful interfaces. From user experience to user interface design.	The Design Journal - An International Journal for All Aspects of Design, 20:sup1, S89-S101.	N/A	2017
[67]	Ivan Burmistrov, T. Z. Flat Designv s Traditional Design: Comparative Experimental Study.	15th Human-Computer Interaction (INTERACT) (pp. 106-114). Bamberg, Germany: 978-3-319-22668-2_10 . hal-01599895 .	N/A	September, 2015
[68]	Konstantinos Spiliotopoulos, M. R. (2018). A Comparative Study of Skeuomorphic and Flat Design from a UX Perspective.	Multimodal Technologies and Interaction, 2,31.	N/A	2018
[69]	SAKAMOTO, M. (2017). Relationship between operation and texture and stereoscopic feeling of screen.	International Conference on Biometrics and Kansei Engineering, (pp. 29-32).	N/A	2017
[70]	Suleman Shahid, J. t. (2016). Skeuomorphic, Flat or Material Design:Requirements for Designing Mobile Planning Applications for Students with Autism Spectrum Disorder.	MobileHCI '16 Adjunct. Florence, Italy.	N/A	2016
[71]	Urbano, I. C. (2018). From Skeuomorphism to Flat design:Investigating Older Adults Experience.	Lisboa, Portugal: Tecnico Lisboa.	N/A	2018
[72]	Taj Eldin, Suliman M. Ali , Hany H. Ammar Lane : Pricing Models for Cloud Computing Services, a Survey	International Journal of Computer Applications Technology and Research Volume 5– Issue 3, 126 - 131, 2016, ISSN:- 2319–8656 www.ijcat.com 126	N/A	2016
[73]	Scrum (software development)	Krasimir Ivanov- Own work; Part of the image is based on public domain graphics from Open Clip Art Library (openclipart.org). CC BY-SA 4.0 File:Scrum process.svg published in https://en.wikipedia.org/wiki/Scrum_(software_development)	N/A	2020
[74]	Integrated Learning Environment Course Development and Life-Cycle Maintenance. Available online (retreived on 21.07.2020) at: https://www.public.navy.mil/netc/ile/documents/NAVEDTRA136/NAVEDTRA_136.pdf	USA Navy, Naval Education and Training Command	NAVED TRA 136	2010
[75]	Rapid Prototyping: An Alternative Instructional Design Strategy	Tripp, Steven D. & Bichelmeyer, Barbara	N/A	1990

Reference	Document Title	Document Reference	Version	Date
[76]	Securing maritime logistics and supply chain: The Medusa and MITIGATE approaches. 1st NMIOTC conference on cyber security in the maritime environment, NATO Maritime Interdiction Operational Training Centre, Chania, Crete.	Papastergiou, S., & Polemi, D.	N/A	4–5 Oct 2016
[77]	ECEL 2018 17th European Conference on e-Learning proceedings	Rabah et al. page number 490 from the conference proceedings	N/A	2018
[78]	Enterprise gamification: Will it drive better business performance?	Dion Hinchcliffe retrieved online from https://www.zdnet.com/article/enterprise-gamification-will-it-drive-better-business-performance/	N/A	28 April 2012
[79]	Artificial Intelligence for Cybersecurity	Matteo Bonfanti, Kevin Kohler, CSS Analyses in Security Policy, issue 265, Center for Security Studies (CSS), ETH Zurich	N/A	June 2020
[80]	Adversary Simulation Becomes a Thing…	Raphael Mudge, Cobalt Strike blog, Retrieved online from: https://blog.cobaltstrike.com/2014/11/12/adversary-simulation-becomes-a-thing/	N/A	12 November 2014
[81]	Purple Teaming for Success	Published on Secure Ideas blog. Retrieved online from: https://blog.secureideas.com/2014/04/purple-teaming-for-success.html	N/A	29 April 2014
[82]	The Scrum Guide™	Developed and sustained by Scrum creators: Ken Schwaber and Jeff Sutherland		November, 2017
[D5.1]	ECHO D5.1 E-EWS High Level Design	ECHO_D5.1_v1.2	V1.2	31/10/2019
[D4.1]	ECHO Transversal technical cybersecurity challenges report	ECHO_D4.1_v1.0	V1.0	18/06/2020
[D6.5]	ECHO D6.5 UPDATE - E-FCR HIGH-LEVEL DESIGN	ECHO_D6.5_v1.0	V1.0	31.07.2020

Table 3: References

1.6 Intellectual Property Rights

Based on the legal framework provided in the ECHO Grant Agreement and the Consortium Agreement, ECHO specific Intellectual Property Rights(IPR) procedures have been established to protect the innovations and knowledge developed within this deliverable.

1.7 Glossary of acronyms

Acronym	Description
AAA	Authentication, Authorization and Accounting
AI	Artificial Intelligence
APIs	Application Programming Interface
APV	Attribute Pair Value
AR	Augmented Reality
AWS	Amazon Web Services
CEM	Common Methodoloy for Information Technology Security Evaluation
CERT	Computer Emergency Response Teams
CLOSINT	Close Source Intelligece
CPU	Central Processing Unit
CRIS	Cyber Range Interoperability Standard
CR	Cyber Range
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Specification System
E-EWS	Early Warning System
E-FCR	ECHO Federated Cyber Range
E-MSAF	ECHO Multi-Sector Assessment Framework
ECHO	European network of Cybersecurity centres and competence Hub for innovation and Operations
eMBB	Enhanced Mobile Broadband
EOL	End-Of-Life
ESDL	ECHO – Service Description Language
EU	European Union
EUCC	European Union Cybersecurity Certification
GDPR	General Data Protection Regulation
HMI	Human-Machine Interface
HW	Hardware
IaaS	Infrastructure-as-a-Service
ICS	Industrial Control Systems
ICT	Information Communication Technology
IDS	Intrusion Detection Systems
IoT	Internet of Things
IPR	Intellectual Property Rights
ISP	Internet Service Providers
IT	Information Technology
ITSEF	Information Technology Security Evaluation Facility

Acronym	Description
ITU	International Telecommunication Union
JVN	Japan Vulnerability Notes
LBS	Location Based Services
LTE	Long Term Evolution
ML	Machine Learning
Cat-M	Machine Type Communications
mMTC	Machine Type Mass Communications
MANO	Management and Orchestration
M2M	Massive machine-to-machine
mmWave	Millimetre Wave Band
MEC	Multi-Access Edge Computing
MPLS	Multiprotocol Label Switching
NaaS	Network as a Service
NB-IoT	Narrowband Communications for the Internet of Things
NFV	Network Function Virtualisation
NICE	National Initiative for Cybersecurity Education
NVD	National Vulnerability Database
NIS	Network and Information Security
NFV	Network Function Virtualisation
NFV	Network Function Virtualisation
NF	Network functions
NFVI	Network Functions Virtualisation Infrastructure
NaaS	Networks as a Service
ONF	Open Network Foundation
OSS/BSS	Operation/Business Support System
OT	Operations Technology
PaaS	Platform-as-a-Service
PAM	Pluggable Authentication Module
PASTA	Process for Attack Simulation and Threat Analysis
QA	Quality Audit
QoS	Quality of Service
RADIUS	Remote Authentication Dial-In User Service
RBAC	Role-based access control systems
SOC	Security Operations Centre
SW	Software
SD-WAN	Software Defined WAN
SaaS	Software-as-a-Service
SDN	Software-Defined Networking
SCTP	Stream Control Transmission Protocol
TCP	Transmission Control Protocol
TLS	Transport Layer Security
uRLLC	Ultra Reliable Low-Latency Communications
UMTS	Universal Mobile Telecomunications System
UDP	User Datagram Protocol

Acronym	Description
UX	User Experience
US	UserStories
V2X	Vehicle-to-Infrastructure
VM	Virtual Machine
VMM	Virtual Machine Monitor
VNF	Virtual Network Function
VPN	Virtual Private Network
VR	Virtual Reality
VIM	Virtualized Infrastructure Manager
VNFM	VNF Manager
WBA	Wireless Broadband Access
WP	Work Package

2. Roadmap development methodology

2.1 Mapping SCRUM to Roadmap development

SCRUM based methodology was selected for development of the roadmap document. SCRUM is an agile process framework used for managing complex knowledge work, facilitating effective team collaboration, with an initial emphasis on software development, which makes it suitable to be explored for complex work, research and advanced technologies such as roadmap developing for cybersecurity domain. Figure 1 below depicts the general SCRUM process [73] that the Roadmap development will go through.

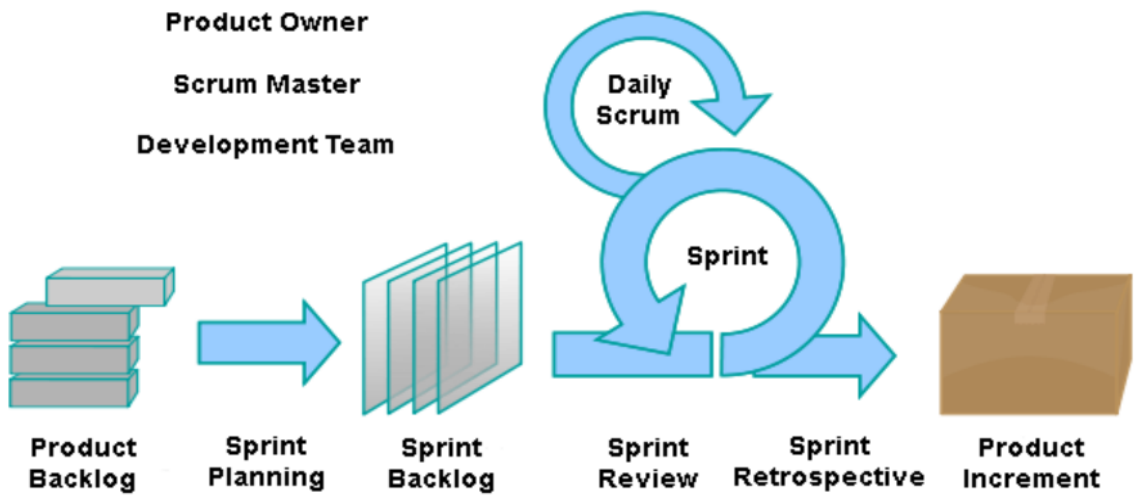

Figure 1: SCRUM Process Overview

SCRUM is one of the most popular methodologies used for software development and revolves around iterative work of self-organizing teams working together towards a common goal. The framework is also popular in other domains such as marketing and research, where the final product of the team's labour is somewhat clear, but not always the path of development. Main component facilitating the iterations is the Scrum Sprint, defined by Ken Schwaber and Jeff Sutherlandas as [82] a time-box of one month or less during which a "Done", useable, and potentially releasable product Increment is created. Sprint iterations are going to be the basic time-boxed development unit utilized to measure and track the progress. It is also going to be used to facilitate the interaction, collaboration, and synchronization between the contributing partners. Moving away from the above mentioned Schwaber and Sutherlandas' time-boxing length of one month or less, the length is agreed and fixed in advance for each sprint and for the roadmap development is initially being set at one month for all partners to get comfortable with the nature of the roadmap development process, the iterations and foster frequent interactions. After the 6th iteration the length of each Sprint is going to be prolonged to two months.

In the beginning of each Sprint, a Sprint planning meeting is held where the upcoming goals are being defined, agreed upon from all partners and translated into a Sprint backlog, an ordered list of everything that is known to be needed in the product[82], with a simple work breakdown structure and estimated effort for each goal to be achieved. Once the backlog is defined each partner commits to delivering the accepted task. On a weekly basis huddles/ progress meeting are going to be held for the contributors to report progress, any impediments in the process, discuss ideas and inform the others regarding future plans. The huddles are also going to be used to discuss possibilities for collaboration that were not present at earlier stages. At the end of each Sprint, Review and Retrospective meetings are going to be held simultaneously to review process, identify lessons learned and propose improvement in the work for the next sprint.

Work units are organized in major technology domains, affecting the E-EWS or E-FCR platform, called EPICs in this methodology. One such domain can be “User Experience” or “Technological Platform” for example. Those EPICs are further sprit into User Stories (US). Each US is focusing on specific technology driver for the domain, concrete technology target or alternatives that might affect future development of the ECHO platform in scope of the roadmap. Each US can contain one or multiple opportunities for future development.

The daily huddles are going to be replaced with weekly, however the requirement for them are going to remain the same[73]:

- All members of the development team come prepared.
- The huddle starts precisely on time even if some development team members are missing.
- Should happen at the same weekday and time.
- is limited (timeboxed) to thirty minutes.
- Anyone is welcome, though only development team members should contribute.

During the weekly huddle, each team member would answer the typical for the daily huddle questions, just modified for the length between meetings[73]:

- What did I complete last week that contributed to the team meeting our sprint goal?
- What do I plan to complete next week to contribute to the team meeting our sprint goal?
- Do I see any impediment that could prevent me or the team from meeting our sprint goal?

Any impediment (e.g., stumbling block, risk, issue, delayed dependency, assumption proved unfounded) identified in the huddle should be captured by the scrum master and designated to a team member to be working toward a resolution.[73]

After the sixth consecutive sprint, the huddles were held at bi-weekly recurrence in order to facilitate adopted approach and to allow more effective contributions.

2.2 SCRUM Roles in Roadmap development

There are four major roles in the Scrum framework adapted to Roadmap development:

- **Product Owner** - The product owner defines the product in customer-centric terms (typically user stories), adds them to the Product Backlog, and prioritizes them based on importance and dependencies
- **Development team** - The development team has from three to nine members who carry out all tasks required to build increments of valuable output every sprint. The team is self-organizing.
- **Scrum Master** - Scrum is facilitated by a scrum master, who is accountable for removing impediments to the ability of the team to deliver the product goals and deliverables. The scrum master is not a traditional team lead or project manager but acts as a buffer between the team and any distracting influences. The scrum master ensures that the scrum framework is followed. The scrum master helps to ensure the team follows the agreed processes in the Scrum framework, often facilitates key sessions, and encourages the team to improve.
- **Scrum Governance Body** - auxiliary team of ECHO experts that will provide guidance to avoid tunnel vision from the Core team

For E-EWS Development the following role assignments were selected at task meeting in Sofia on 30.09.2019:

- Core team:

    - Product Owner- Fabrizio De Vecchis for E-EWS roadmap and Matteo Merialdo for E-FCR roadmap
    - SCRUM Master- Krasimir Ivanov
    - Dev Team- Partners with more than 4 CP-M committed to task 4.2
- Non-Core Team:
    - Stakeholder- Wim Mees
    - SCRUM Governance Body- Partners with 3 or less CP-M committed to task 4.2

2.3 Development phases

Development process for the Inter-Sector Technology Roadmaps consist of the following phases:

- **Phase 1: Setup,** with the following tasks:
  - Define scope, vision and boundaries for each individual roadmap.
  - Determine Product owners and Development teams.

For the purpose of development of ECHO E-EWS and E-FCR Roadmaps naturally the definition of the scope, vision and boundaries was defined of the activities already in progress within ECHO project – Work Packages 5 and 6. In those work packages product owners were already identified and also solid software architectures and development target were determined. Roadmap development teams were selected on base of specific expertise and resources available for the task.

- **Phase 2: Development**, with the following tasks:
  - Identification of the critical system requirements and their targets.
  - Identification of major technology areas.
  - Specification of the technology drivers and their targets.
  - Identification of technology alternatives and their timelines if applicable
  - Recommendations of technology alternatives that should be pursued, if applicable

For the purpose of identification of the critical system requirements, D5.1 E-EWS High Level Design and D6.1 E-FCR High Level Design were analysed. Major technology areas, drivers and targets were identified in brainstorming sessions with broad panel of experts within the ECHO network, to capture a wider range of opinions and reflect sector and technology diversity that the network brings . Intentionally the audience of the brainstorming sessions was extended, including by using of technological tools to enable remote participation in order to ensure maximization of number and in result – the quality of ideas and identification of major technology trends, as well as specific technologies an d drivers that will affect the future environment ECHO platforms will operate in. As result virtually all ECHO Partners participated in the development of the current document. The results from the brainstorming sessions were consolidated, correlated and established as UserStories under the Epics, containing the major technology areas. Those UserStories were assigned to specific development teams, based on their expertise and later consolidated in single document and distributed in the Sprints.

- **Phase 3: Follow up** with specific tasks:
  - Quality Audit (QA) process, Stakeholder review.
  - Update of roadmaps with finalized inputs at M48.

D4.3 follows ECHO standard Quality Assurance process, as outlined in D1.3 Project Quality Plan and is reviewed by peer reviewers, as well as stakeholders, or the Product Owners under SCRUM terminology. By utilizing this process, it is ensured that roadmaps fulfil their purpose to outline concrete opportunities, related to the ECHO platforms.

Inter-Sector technology roadmaps will be updated at M48 of the ECHO project with the finalized inputs identified in section 1.4 of this document and with roadmaps of the T4.3 prototypes.

3. ECHO Federated Cyber Range Roadmap

The goal of the E-FCR developed in the ECHO project is to interconnect existing cyber-range capabilities through a convenient portal operating as a broker between user requirements and a pool of available cyber range capabilities. The objective of the E-FCR is to solve the problem of simulation of the complex realities and inter-sector dependencies of an inter-sector scenario by establishing a mechanism by which the independent cyber-range capabilities can be interconnected and accessed via a convenient portal for configuration and management. As cybersecurity is dynamic there is a need to identify future development options. Those opportunities can be utilized to achieve sustainability of the platform in the dynamic environment through: continuous improvement, adaptation and evolution of the platforms using innovative and horizontal capabilities.

The specific concept of the E-FCR includes a portal/dashboard where users can develop their single and multisector cyber scenarios and request the portal to start the intended scenario from among a pool of interconnected cyber-ranges operated by partners of the ECHO Network. As part of this process, the portal will also serve as a means to validate whether the complete scenario can be simulated or highlighting the case where some aspects of the scenario may require additional simulation capabilities, whether due to complexity of scenarios, missing of specific virtual images or specific cyber physical assets.

Using the above as starting point, the roadmap development process followed the approach, defined in Chapter 2 of this document and produced the consolidated major technological areas (EPICs), outlined below. Those were identified during the brainstorming session of extended ECHO experts panel, with their respective supportive “user stories” – technology drivers, targets, alternatives that affect future development of E-FCR after the period of ECHO Project. The five major domains, outcome from the development process, capture different aspects of technology, environment and requirements evolution:

- **User experience** domain deals with how users of the Cyber Ranges and E-FCR interact with the tools
- **Connectivity** domain explores effect of future connectivity development
- **Scalability** discusses ways to scale the E-FCR platform
- **Platform** domain deals with Cyber Ranges platforms and integration of new tools
- **Exploitation** domain focuses on novel uses of the E-FCR platform

One of the main benefits of implementation of E-FCR would be addressing of the identified in T4.1 challenge “Lack of cyber situational awareness in national critical infrastructure and gaps in defense-in-depth architecture hacking “ discussed in section 4.2.2 of [D4.1]

3.1 User Experience of E-FCR

3.1.1 Gamification in Cyber Ranges

Theory of gamification

As per [77] It is important to establish differentiation between 'gaming' and 'playing'. While playing is a freeform, creative and open-ended process, gaming is a highly structured process oriented toward discrete, clearly defined goals. Gamification, then is defined as "the use of game design elements characteristic for games (rather than play or playfulness) in non-game contexts". This differentiates it from serious games and design for playful interactions.

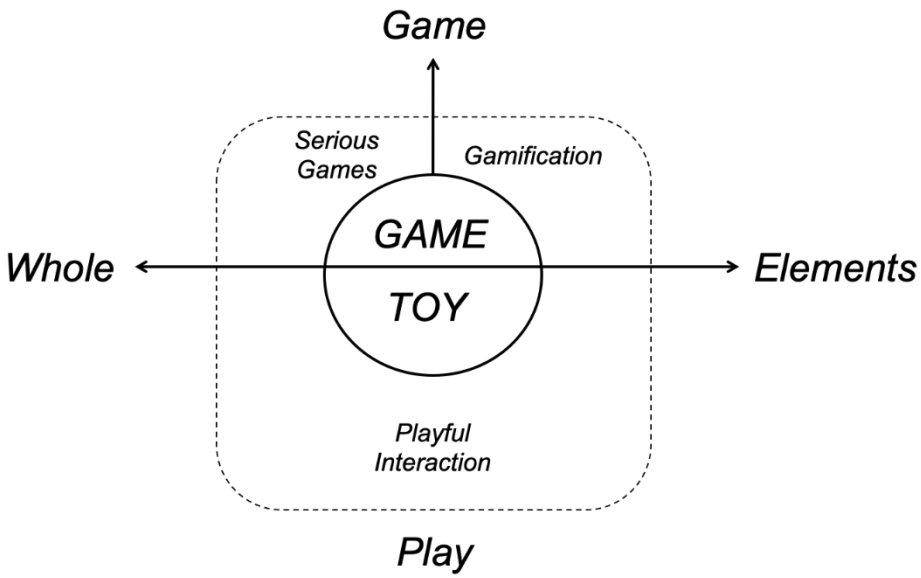

Figure 2: Difference between gaming and playing

We can find principles and design elements in the synthesis by [63] of the more prominent frameworks found in the literature (e.g [64]):

As per [62] gamification design principles are:

- Goals and challenges
- Visible status – informs participants about a task's completion status or else shows students how they are progressing
- Social engagement - feeds into purported needs for competition against individuals or teams and may include team projects and group learning opportunities
    - Competition
    - Cooperation
    - Collaboration
- Freedom of choice - participants are free to choose whichever task(s) they want to complete from a multitude of tasks
- Freedom to fail - exemplified in contexts where participants were given the chance to submit assignments again and to revise their work without a penalty

- Rapid feedback
- Access, unlock content

According to [62] mechanics of gamification include

- Badges
- Points
- Levels
- Rewards
- Leaderboards
- Progress bars
- Currency
- Avatars
- Countdown clock

Gamification is not only about using game mechanics in courses, rather, using them to overcome challenges in education and meeting objectives by increasing participants motivation. This leads to positive psychological outcomes that motivate the participants behaviour.

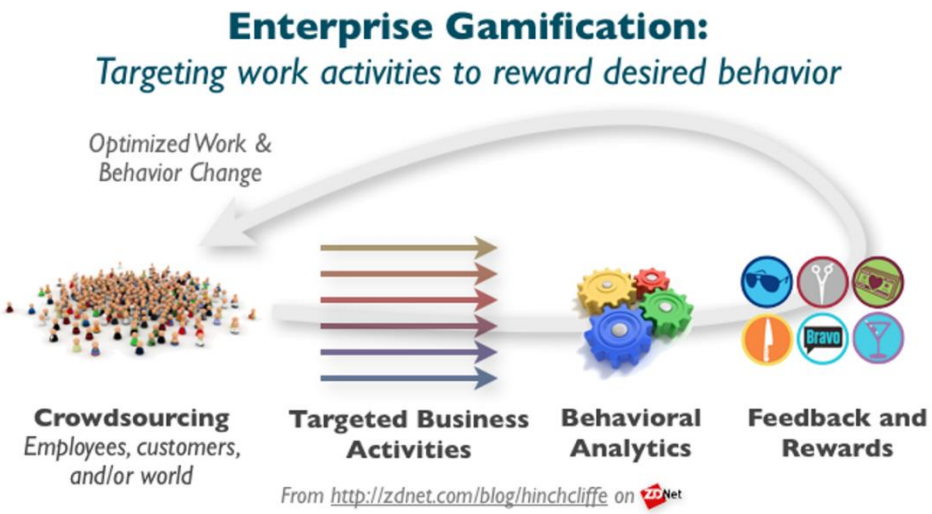

Figure 3: Gamification in Enterprise[78]

For example the social network, Foursquare, uses gamification to motivate users to “check in” wherever they are, thus, promoting the use of Location Based Services (LBS) to grant the virtual reward of being the major of a location were that person had the highest number of check-ins.

Gamification application to FCR

Application of design principles to cyber range (CR) exercises can be established in several domains:

- Gamification for participants:
  - Goals and challenges associated with the different trainings. Each goal and challenge can be time based and awarded with points

  - Global leaderboard for participants
  - Designing the exercises in a way that provides multiple tasks for a team (i.e. multiple targets to attack for Red Team member) so the participant can have freedom of choice
  - Status or level of the knowledge of the participant, calculated from the goal points, expressed visually with a badge.
  - Exercises that are available to participants above certain level (unlock content)
- Gamification for content providers
  - Visual badges for the cyber ranges based on the feedback from Capacity/Capability Map, QoS manager and Service Broker concerning categories as “uptime”, “performance”, “helpdesk reaction time”, etc
  - Content providers leader boards in different categories (i.e. number of exercises uploaded to FCR, number/hours of services provided, available resources, etc)
- Gamification for CR Technology providers
  - To provide recognition in the PEF community of those who contribute certain technology that, for example, provides faster processing of cyber-attacks.

In creating the mechanics of the gamification, special attention should be paid so the awards and achievements are fair. ,Improperly designed and applied gamification can cause the effect of demotivating and disengaging participants. Rules for how to design and apply rewards should be established and a supervision mechanism should be put in place to approve proposed goals from the content providers.

Another important aspect is the rapid feedback principle that requires immediate provision of achievements and awards to the participant. To be able to provide capability for rapid feedback, the E-FCR should be able to monitor the competition of goals in the CR exercises automatically via API. This also requires monitoring of the achievement to goals in each individual CR which can be especially challenging.

Phases of introducing gamification

In order to validate the gamification approach and results, which techniques work and which don’t in the FCR environment gradual approach can be adopted. Such approach will allow to efficiently introduce the concept without investing significant development resources without real-world feedback. For example the following three-phase implementation can be planned:

Phase 1

In Phase 1, easily accessible metrics can be introduced without need of building additional CR functionality or modifying core FCR components, as the data needed is already present. Such metrics can be:

- For participants: number of exercises participated, cumulative length of time, etc.
- For content providers: number of purchases of the exercises, average vote on exercise from participants, etc.

All metrics should be visualized on a global leaderboard and feedback should be collected on how users of the E-FCR platform perceive the gamification.

Phase 2

In Phase 2, complex metrics can be introduced, that would need more resources to be implemented, as well as some modifications of core E-FCR components and/or CR components.

- For participants: Goals and achievements in exercises that award points to each participant, based on the discretion of the White team decision (via dedicated interface, that could be hosted in the FCR as this will facilitate the rapid feedback). The total number of points accumulate and are visualized on global leaderboard. Levels can be introduced with different avatar or badge visuals to amplify the competition effect. Based on the goals achieved, progress bar for the exercise can also be visualized.
- For content providers: number of goals solved in their exercises

Visualization of metrics in this phase should be integrated in main feedback mechanism of E-FCR platform for reporting simulation status, goal, objectives and results.

Phase 3

In Phase 3 very sophisticated metrics should be introduced after significant modifications and development on both E-FCR and CR platforms. Completely new instrumentation should be introduced in the platform with the sole purpose of communication of real-time feedback to users.

- For participants: Integrated real-time feedback on goals/objectives, including across exercises and CR. Introduction of automatic assessment of difficulty to decide on award points.
- For content providers: Integrated real-time feedback on all exercises, including across different CRs and platforms, to allow real-time adaptation of parameters.

At this phase gamification should be deeply integrated in all aspects of the User Experience of E-FCR, communicated or accessible via multiple channels and used as marketing tool to achieve E-FCR goals and objectives.

3.1.2 Development of FCR Services Catalogue

An CR is a highly complex virtual environment in which organization can test their cyber capabilities, training the personnel and improve their cyber-security process, skills and technologies.

As E-FCR operates as a broker, connecting multiple CRs, one of the main purposes of the platform will be the supply of CRs that best respond to a customer need. For this reason, the design, development and the evolution of the Service Catalogue will be a core functionality of the E-FCR in order to best meet the customer requirements.

In the Service Catalogue there will be available the list of all cyber range services that are part of the E-FCR. For better identification for the needs of services they will be grouped by their properties. The properties of services will include: service specific data, their pricing and the related Quality of Service (QoS).

In the same way of grouping, the customers will also be able to perform specific service search based of the properties of cyber ranges; keywords contained and capacity and capabilities of the service. During the developing of E-FCR and after the release, new more useful properties will be identified and added in Service Catalogue to best fit the customer needs.

The Service Catalogue will also be able to make comparison between two or more different servicesto highlight their difference based on specific needs of the customer. In the initial phase the service will be compared by their capacity and capabilities, but in future evolution of E-FCR, will be useful also to compare service by the pricing to allow all the customers to find the best service that fit their needs and their financial requirements.

The Service Catalogue will also offer the possibility to store and create templates of the service for future reuse, for this purpose, also the possibility of sharing the template will positively increase the user experience of E-FCR.

3.1.3 Exercise workflow

This section discusses the workflow of setting up exercise / experiment / certification or training in E-FCR from point of view of Developer or Partner in the FCR

User Management

When particular a CR is discussed as a training or testing facility , care should be applied to distinguish between the "users," i.e. those undergoing the training or testing (generally persons) and the "customers," i.e. those paying for the training or testing (government offices, enterprises or institutions). Both side's needs should be considered during the planning, development, testing, refinement and advertising of the training or testing environment.

The same should be applied to the E-FCR, but with consideration what makes the E-FCR different – it consists of two components, which are neither totally dependent on each other, nor totally independent.

First, it consists of a marketplace where the services of various associated cyber ranges are offered to clients. This marketplace allows consumers of cyber range services (governments, enterprises, institutions) to have a better view and access to the range of possibilities, allowing better asset utilization and planning for capability development, whether their own or those of adversaries. The marketplace is also intended as a place for partnered or federated cyber ranges to offer their services or upload content, acting as both a virtual store and advertising.

It is important to note then that the "users" of the marketplace, those for whom the user experience must be optimized, consist of the consumers of cyber range services, and, perhaps to a lesser extent, the cyber range service providers themselves.

However, the "customers" of the marketplace, those who pay directly for its services, are a partner of federated cyber ranges themselves. This is quite similar to physical markets which charge rents or virtual markets like eBay or Amazon. The benefits (access to customers) of being a partner in the federation or marketplace must outweigh the costs.

Second, the E-FCR is intended to allow users (that is those needing cyber range services) to purchase services from various providers in a way that these services work together. This will require interoperability and communication between partner cyber ranges both at a technical level and in regulatory and skills standards.

There is a clear benefit to the user here, while, partners of the federation are the "customers" in that they ultimately pay to reach technical, regulatory and skills standards compliance.

Asset Management

When designing exercises, various assets are used – virtualized servers, services, tools or combination of those, in order to build complex environment in which the simulation or exercise to take place. Applying proper asset management in the process of design can positively affect the workflow of the project and provide with a seamless User Experience(UX) both for employees and customers. The correct mapping and acclamation of the assets assigned for the design of the product can assist in building a solution that has effective asset controls, metadata management, system usage and compliance.

With the use of an effective asset management solution, a developer team is offered the ability to not only to effectively organize a project but also to protect assets from being stolen and produce discontinuities in the workflow.

Effective asset management can produce results not only for the developer team. With its properties of storing, sharing and tracking components it can also assist in the cooperation between the whole group of stakeholders; organizations, businesses, departments, and customers which is a major part of a successful UX workflow.

3.1.4 Machine Learning offering suggestion to E-FCR clients

It is a major challenge to ensure that the E-FCR platform is always updated to include latest vulnerabilities and exploits in the simulation as everyday new vulnerabilities are being produced and even more new cybersecurity tools are being developed.Thus, the E-FCRplatform, to be flexible and competitive than competing Cyber Range products in the marketplace, must provide updated courses and training ofnew vulnerabilities. Several EU projects have also added or are currently adding a number of advances on top of their services implementation and/or integration.

The E-FCR platform should search for current technological threats and vulnerabilities on the internet by collecting information of the high priority vulnerabilities. Moreover, it should regularly check the latest available products and capabilities of all Cyber Range providers in the marketplace for further information.

As assessed in D2.2, Section 4, on top of (e.g. their RA and RM) implementations, projects like MITIGATE **[76]** exploited innovative technologies such as big data technologies for threat analysis to develop new approaches for predictive threat attribution..

Usually the list of threats is formulated taking into account both previous incidents experience and threat assessments, as well as available from industry/standardization bodies, national governments, legal bodies, etc threat catalogues. An Machine Learning (ML) could also exploit on top of that information retrieved from social media and other existing sources/repositories that would play a very important role in gathering peripheral information and drawing conclusions. New innovative services for the clients, build on that technology, could be produced by combining those and semantic web technologies (including ontologies) for an advanced, automated, innovative representation of assets and application of risk and assurance models, including the integration of data from open intelligence sources (i.e. social network and crowd-sourcing).

One opportunity how such technology can be implemented in E-FCR can be as follows: Based on the results of the previous searches the E-FCR platform automatically creates various training courses focused on how to handle, mitigate and more importantly, how to prevent specific vulnerabilities and threats. These courses concerning newly-found vulnerabilities can be either from the red or the blue team's point of view, which will help the participants to better understand the structure of any new-found vulnerability. With this capability the E-FCR platform can provide an course about a high priority new threat, explains the threat and teaches how to prevent it, while the competitive Cyber Ranges need manual work and thus time to provide such trainings. This creates major advantage on the ECHO platform against other competitors and will probably attract more participants.

The type of vulnerability encountered is important. In recent years, the most prolific attacks on businesses is ransomware, however, the future trend could see this change to attacks on applications through web-based injections. To meet the dynamic threat environment, the E-FCR platform can be used by business and academia researchers to conduct analysis on persistent and emerging threats to their environment and use the knowledge to inform better decision making.

In conclusion, the continuous changes on the material and the knowledgebase of E-FCR, alongside other European projects will provide to the participants all necessary knowledge on the latest cybersecurity topics which serves as a common denominator on the reinforcement of cybersecurity in the EU and the primary purpose of the E-FCR and ECHO network.

3.1.5 AI for scheduling purposes and increased efficiency

Over the past decades, the exponential increase in complexity, along with the need for optimal resource utilization, has proved a major obstacle for classical planning and scheduling models and methods. Dealing with time and resource constraints, continuous changes, caused AI to become the rule rather than the exception in optimization processes. Planning and scheduling solutions based on AI usually lead to 'trial and error' processes.

Specifically, AI planning and scheduling are application/research areas that have quite a lot of interdependencies, with similar subjects but different optimization approaches. Both rely on certain activities that have to be performed in order to achieve given goals. All the activities need resources and a set of constraints has to be observed in order to get valid solutions. One of the main differences between planning and scheduling is the use of temporal issues. Scheduling in nearly all cases deals with concrete temporal assignments of activities to resources, whereas planning mainly deals with the order in which the activities have to be performed (planning focuses on "what has to be done" whereas scheduling focuses on "when this has to be done").

Planning

The task of planning is to find control algorithms that enable agents of all kinds (humans, machines etc.) to synthesize courses of action that achieve their goals. Generally, in order to define a planning problem one requires a holistic description of the initial environment, a description of the goals (i.e. the environment in one of its possible final states), and a set of possible actions that can be performed and lead to an environmental state shift. The plan as the output of the planning process then shows the sequence of actions to be executed in order to achieve the desired goal/environmental state. Widely used is a representation based on logic and algorithms for efficiently searching the combinatorial search space. Most of the systems developed are domain-independent and therefore allow planning in different application areas. Application areas of planning range from process and project planning.

Scheduling

Scheduling is the process of extraction of activities and assignments, namely schedules, which are due to be fulfilled by a set of resources, while also regard a specific set of constraints and accomplish a number of goals during the current schedule. The order of activities can be an output of a planning system. In order to define a scheduling problem, there is the need to schedule the potential described activities, the definition and provision, in an optimal manner, of the underlying resources to be used by these activities. Furthermore, any constraints (e.g. technical restrictions on the use of resources that must be regarded or economic constraints) that should be fulfilled, e.g. reducing the cost of inventory, minimization of total execution time or avoidance of idle time in bottleneck resources, etc, should also be taken into account during the scheduling process. Additionally, goal functions are often used to compare solutions or contribute to the definition of "optimal" solutions.

Planning and Scheduling Algorithms

In some application areas, planning and scheduling tasks are strongly interconnected, e.g. in E-FCR where the deployment process and the reservation schedule have to be determined for every new customer Service Request. Therefore one of the research tasks is the combination of planning and scheduling to find the best solutions for those kinds of problems. Additionally, plans and schedules usually are created within a dynamic environment where the plan or schedule often becomes invalid due to events from the environment. So there is the need for an adaptation to the actual situation.

To address those issues, a variety of algorithms and techniques have been implemented that can be used for planning, including:

- classical planning (goal-directed planning used to find a goal state with desired property),
- reduction to other problems (obtained by reduction to the propositional satisfiability problem),
- temporal planning (the definition of a state has to include information about the current absolute time and how far the execution of each active action has proceeded, cause of temporally overlapping actions with duration being taken concurrently),
- probabilistic planning (solved with iterative methods such as value iteration and policy iteration),
- preference-based planning (producing a plan that additionally satisfy user-specified preferences),
- and conditional planning (a plan can react to sensor signals which are unknown for the planner).

Some good examples of planning and scheduling based on AI are:

- autonomous space stations and satellites,
- logistics and crisis management,
- video games,
- manufacturing process planning,
- robotics,
- urban facilities (smart cities),
- energy plants,
- driverless cars,
- risk and consequences assessment (medicine),
- conducting clinical trial protocols.

AI Planning/Scheduling in E-FCR

Based on the aforementioned examples, it is apparent that AI is used in many fields for increased efficiency and scheduling purposes, and therefore the E-FCR provides a great platform where AI can be applied - even in simpler forms - to improve the quality of service.

The E-FCR provides an excellent environment to conduct experiments, exercises, research and prototype testing. It enables the combination of multiple scenarios from different content Providers on multiple Cyber Ranges, with a variety of availabilities and resource allocations. In that context it becomes clear that the scheduling and deployment of customer Service Requests poses an optimization issue for the platform. Regarding the optimal scheduling of the requests, AI techniques and methods offer a significant solution by the use of more complex decision making rules, applying more sophisticated scheduling knowledge, while also tracking constraints and resource requirements.

Moreover, scheduling and planning with AI can assist the E-FCR platform in:

- Rapid resolution of scenario scheduling-conflicts (i.e. scenarios submitted by different content Providers but with common Cyber Range requirements),
- load balancing for efficient server utilization,
- dynamic reallocation of resources in case of Cyber Range outages (scheduled or otherwise),
- customer assistance in the creation/selection of Service Requests (suggestions based on customer input or E-FCR QoS metrics).

Specifically, E-FCR can leverage on AI techniques and monitor the status of the whole ecosystem while creating and processing the potential availability/reservation schedules. It can also offer the potential to dynamically adjust time schedules and resources used for each request in order to maintain proper and optimal execution. Furthermore, it would be able to take into account several QoS metrics from the corresponding Cyber Ranges and decide in a specific time interval the ones that are best to serve a potential Service Request. This entire process contributes to a more efficient handling and resource allocation/utilization from the providers' perspective, thus increasing profit (i.e. providing the ability to accommodate more requests or reduce wages by the increased resource utilization efficiency and reduction of operational costs). From the customers' perspective, the major advantage is the ability to maintain or even to reduce end-user latencies and also, reduce the chance to reject a customer's request due to network congestion or other outages and inefficiencies in the platform.

3.1.6 Billing models

E-FCR billing process is managed by the Billing Manager component of the platform, tasked with determination how much to charge a customer for the service that is to be provided.

As with all billing models we can have static or dynamic billing [72]:

- Static billing has advantages that is more easily understood from clients site, thus lower the entry barrier, lower risks for clients, as they know who much they will pay at the end of the service. This type of billing also reduces the risk of non-paying client and makes profit-estimation easier. However this model of billing might be unfair to either the consumer or CR provides as it relies on assumptions when invoice is created. If the assumptions are not correct, the client will be overcharged or undercharged, depending on the error. Static billing also doesn't provide options for the providers to easily change the invoiced price
- Dynamic billing supports maximisation of profits with each customer without the risk of wrong assumptions. It is also fair to customers, as they can pay just for what they've consumed and is more competitive as allows easy adjustment to the market. Disadvantages are that as billing model can be too complex, it can drive off some customers. The complex billing process is also bigger risk for wrong processing and also there is business risk inherent to this type of billing: the invoice and payment are run after the exercise is completed, which is not the case with the current billing models. Contractual precautions must be taken in order to address this type of risk, as it might create disagreement between ECHO governing body (as E-FCR operator) and CR operators.

E-FCR currently have 2 separate models, both static:

- Standard service: if a client choose one or more of the standard services that can be provided without customization at all or with small customization in terms of requested schedule, number of end-users requested and reachability. The invoice amount is calculated by the predetermined prices of the service components.

- Custom service offering: The invoice amount is manually created together with the offer for the services, based on initial client request. The amount is determined in the proposal and is stating until end of the service.

With the advancement of the proposed services and with ever increasing number of CRs looking to integrate Public Cloud platforms as a way to accommodate much larger and complex simulations (ref. UserStory 3.4.2 "Virtualization evolution", section "Comparative Analysis Of Commercial Cyber Ranges Roadmaps" of the current document), other more dynamic billing methods can be envisioned:

- Pay-as-you-go billing: This is the typical billing model of public cloud usage, where an client is charged on monthly basis after service consumption only for the resources used – i.e. number of CPU cores, amount of RAM, etc. In the case of E-FCR such billing model might be implemented with enhancement of the Billing Manager and feeding data about actual systems usage from the participating CRs. In this way the cumulative resources usage from all the participating CRs can be calculated and invoiced on the end of the exercise. This model can simplify billing process in cases where very complex simulations are performed and help introduce elasticity in the provided services. It also allows for better financial security in case of participating CR is using public cloud – the business risk of wrongly assessing the needed resources and quoting fixed price, while resources are paid by the CR operator per consumption is virtually non-existent.
- Rule based billing: In this billing model several quotes are created for one and the same exercise and a rule or trigger is used to determine which variant of the quote/total invoice amount will be used. The trigger for example might be performance of participants of the exercise – in this case an "challenge", or "competition" might be created with participant's that achieve specific results have their invoice decreased as reward. Such setup might be beneficial in Talent search or Academic research scenarios.
- Competition-based billing method – in this model the prices is determined dynamically in the beginning of the sales cycle – i.e. when the invoice is created and from then on the invoice price is static. Here the price can be determined in several dynamic ways – by automated rule-based engines or by running automated or non-automated internal competition among participating CRs, capable of fulfilling the client's needs. Such approach will ensure competitiveness and can be very attractive for clients, but if not managed carefully can lead to price wars and dumping practices among the CRs participating in the E-FCR platform.

3.1.7 Customer Feedback

A develop team will want to know if their services are helpful and might need help prioritizing what to build next. So, one important question that arises is if the prototype is corresponding what the team had set out to develop and if it's suitable for the purpose for which it has been designed.

During development of the cyber range federated framework, built on the requirements derived from the user research, is important test it often in functionality and in usability. Functional and usability testing require different mindsets from the person testing and ideally should not be carried out by anyone involved in the development of the product.

Of course, in a market, the ultimate feedback comes in the form of profit. This may seem obvious, but it's important to also appropriately define profit. The profit of a particular venture is the measure not just of the earnings minus the expenses, but only the earnings above the standard rate of interest, i.e. what those expenses would have earned in other areas. Of course, this is simply a measure of good or poor performance on the market. Good performance indicating we are serving our customers well, poor performance indicating that some change is necessary.

Indeed, the preferred method would be to use customer (and user) feedback mechanisms which are faster, finer-grained and less severe than simple market measures.

Also worth noting is that our users and customers are not entirely the same set of entities, meaning we don't have direct market feedback from users.

For this reason, user feedback must be gathered, which serves the dual purpose of metrics to plan the improvement of the system, as well as metrics used to advertise to potential customers, i.e. those providing cyber range services on the federated platform.

In that follow, we look at the different types of user feedback, the ways to collect useful feedback, and analyze types of feedback that are most important and most useful for the ECHO context. User feedback is input about their experience and satisfaction levels regarding to a product or a service.

Feedback can come in from a variety of channels (email, social media) or messenger tools, is an helpful information and criticism that pertains to a certain action and is meant at improving, innovating.

Customer feedback is an important information for product development, for improving the user experience and overall user satisfaction levels. Proper analysis provides a better view of what it has to change and improve on to help increase user interest and reduce user support cases.

Collecting customer feedback

The type of customer giving the feedback matters, The advice given by expert users has value, knowing what an expert thinks about what is developing is fundamental to validate functionality and usability but at the same time advice given by a less expert user could help in simplifying the interfaces or in identifying new features.

Gather customer feedback at every opportunity, every customer interaction is an opportunity for feedback. Avoid the trap of "we don't want to bother our customers."

Unprompted feedback matters, Unprompted feedback deserves special attention. The customer issues that aren't known, which may have been ignored, can be the most important things that need to hear.

Focus on continuous improvement, Enlist the aid of users to improve an aspect or features.

Make it easy for customers to provide feedback, Employ multiple input points: in person, email, Web sites, meeting, conferences and the telephone. It is important for continuous improvement have a high volume of feedback and this depend strictly from the capacity to actively solicit good and bad feedback.

The customers possess good ideas, Often the customers can are capable of knowing what really makes a product or service valuable, and therefore customer input is necessary. Asking customers to participate in your problem-solving is a strength and not a weakness as a developer.

Seek real-time feedback and Use feedback to make changes quickly, Customers become loyal to responsive organization, especially ones that keep them in the loop of how their feedback was used (or wasn't).

Leverage technology to aid your efforts, Online survey tools makes it very easy to gather feedback. They are typically fast, efficient, and inexpensive. They automatically tabulate data and don't require a techie to launch.

Collect customer feedback with a Form

As we said, customer feedback is essential to improving product and even to understanding of users.

In that follow, the tips for designing and writing customers feedback forms that get more responses and more useful feedback for improving a framework.

Improve feedback form's usability- The best customer feedback forms are user friendly, that meaning: simple design, straightforward instructions, responsive and easy to fill in. And so:

Leave white space - The form should appear uncluttered and easy on the eye. A cluttered form can scare people away because it looks time-consuming and hard.

Label the form fields well - Placing each label close to its corresponding field reduces the time it takes for a user to complete the form.

No mandatory fields - If the user doesn't want to answer one of the questions, don't stop them from submitting the form.

Set the tabbing order - This is helpful for desktop and mobile users who wish to tab through the form using the keyboard rather than lifting their mouse or finger to move to the next field.

Short but clear sentences - Cut down the number of questions in the form so it only includes those that have the clear goal of gaining a better understanding of your users' experience. The fewer questions, the better.

Try to find out what you don't know - The most useful feedback helps you learn something you didn't already know. The best way to do that is to include a "free text" box.

Create consistent rating scales - If you include more than one question with a rating scale, make sure that the scales are consistent from question to question. If 1 is the best and 5 is the worst in one question, don't change the scale from 1 to 10 on the next question.

Proposed structure of such feedback collection form can be found in Annex 1 of the current document.

3.2 E-FCR Connectivity

3.2.1 Implications of 5G technology

Current development of fifth generation mobile technology

Fifth generation mobile networks are not yet fully standardized, however commercial deployments are already in place with close to complete functionality. This next generation of communication offer vast qualitive and conceptual improvements over existing technologies that will allow not only much faster connectivity, but also will enable wider penetration of technologies as IoT, autonomous systems and real-time communication, including Machine-to-Machine communications. As result the way people and systems communicate will drastically change, which in turn will change the attack surface of both mobile networks and the systems that communicate, utilizing those networks. Those changes will certainly impact the following challenges, identified as part of T4.1:

- Constantly increasing attack surface
- Cross device dependencies
- Access to IoT devices
- Negative effects of complexity and connectivity

Annex 2 of the current document contains further exploration and description of the details and capabilities of 5G technology as standardization candidate.

5G technology application to E-FCR

The use of 5G will provide connectivity to all kinds of devices that thanks to this data network will be able to connect to the Internet, thus opening up the possibilities of network attacks. The creation of network slices for different types of services will make those slices, that are most vulnerable due to the type of device (e.g. IoT), the gateway to other critical slices, such as health or power supply. Obviously, access to the pure 5G data transport network from an attack by one of the network slices is within the realm of possibility, but logically unlikely, since if the 5G network is shut down by an attack, all telecommunications would automatically fail affecting the total activity of the country where that network is located.

With this criticality, the application of 5G to the FCR should be more focused on the simulation of different devices connected by 5G and analyze the possible attacks and solutions available for example simulating of autonomous vehicles, car connected to the 5G network using edge computing for real-time decision making. These vehicles, completely sensorized and connected, will be a clear target for cybercriminals. Being able to replicate in the FCR the characteristics of a stand-alone vehicle would give the system great possibilities to replicate these attacks and find solutions as soon as possible.

Another area that 5G is expected to be applied as communications technology will be industrial SCADA-type systems. At present, these types of systems, with very limited security measures (having been isolated from the Internet, practically the only access was physical) are increasing their exposure to vulnerabilities by making them ‘intelligent’ through the inclusion of an endless number of sensors. In this line, the FCR could simulate SCADA systems with connected IoT devices to replicate attacks.

The ever expanding digitalization of all devices (i.e. smart watches, smart fridges, smart heart pacemakers), require real-time connection to a data network, which will be done massively through 5G, will again increase the exposure of virtually any person, animal or thing on Earth to be cyber-attacked. It seems foreseeable that the number of attacks will skyrocket as there is trend to massively connect any device to the Internet. Again, the simulation of this type of environment in the FCR will allow an researcher to obtain response mechanisms with a certain speed in the face of these cyber-attacks, so replicating common environments (it is totally impossible to simulate all types of devices that exist or will exist in the world) becomes a very relevant piece to include in the E-FCR.

3.2.2 Comparison of CR interconnectivity

Technical interconnection between different cyber ranges is one of the central items of the E-FCR development.

There are multiple options for technical interconnection of cyber ranges:

- Layer 1 physical interconnection
- Layer 2 datalink interconnection
- Layer 3 logical interconnection
- Software Defined WAN (SD-WAN)

Layer 1 physical interconnection

Physical (**Layer 1**) interconnection of cyber ranges into a federation is the most performant solution but of course extremely challenging and expensive, in particular for a commercial solution [Allar-Federation]. Physical interconnection means a whole dedicated physical network (e.g., fibre optics cabling) must be deployed between CRs.

Benefits of the Layer 1 physical interconnection

Reliable low-latency/high-bandwidth access to remotely shared resources.

Limitations of the Layer 1 physical interconnection solution

Feasibility is extremely dependent on geography.

Real project of federation would be reasonably unsustainable because of economics, coordination and management (ownerships of physical mediums for interconnections and so on).

Layer 2 datalink interconnection

Layer 2 VPN emulates the behaviour of a LAN across an ~~L2 switched~~, IP or Multiprotocol Label Switching (MPLS)-enabled IP network, allowing Ethernet devices to communicate with each other as they would, when connected to a common LAN segment~~. Point-to-point L2 connections are necessary when creating L2VPNs.~~ Layer 2 Ethernet frames will be encapsulated by the chosen VPN technology (e.g., SSL or IPSec tunnel, MPLS L2VPN) in order to stretch the broadcast domain of a common IP subnet between remote sites. Extending an Ethernet-based LAN over a natively unreliable medium as the internet (like the SSL/IPSec-based approach) is an inconsistency that brings inevitable problems. Such issues are due to the high-latency and possible intermittency of connectivity which are against the principle of the Ethernet technology and related protocols (L2VPNs are typically adopted for very specific applications, like offline remote backups with low-level protocols, for instance). To limit this kind of issues as much as possible, at least an extremely reliable WAN solution must be adopted, like MPLS. But building an MPLS-based layer 2 VPN assumes cooperation between an Internet Service Provider (ISP) and the Cyber range Providers with costly investments and contracts. Plus, such configurations are natively static, so not suitable to dynamic reallocation of network resources while instantiating/deinstantiating scenarios for CR purposes.

Benefits of the Layer 2 datalink interconnection

Common broadcast domains (VLANs) shared between remote CRs which can help emulate Layer 2 reachability and conditions across remote sites (ARP resolution, DoS attacks/defence based on broadcast/multicast traffic and so on).

Limitations of the Layer 2 datalink interconnection

The typical latency and reliability of a real local network cannot be emulated. This may bring serious stability issues inside a Scenario's network.

The MPLS-based approach (more reliable than other L2 VPNs) will take work at ISP level and therefore costly investments.

Software Defined WAN

Although it is sometimes referred to as a "technology" by itself, SD-WAN is basically a concept. Like the SDN and NFV, it is more of an abstraction of several complex techniques rather than a specific technology based on a specific set of standards whatsoever (e.g., MPLS is considered an actual technology instead). Therefore, a real analysis of an SD-WAN system for E-FCR cannot be performed without adopting and designing a specific implementation of a commercial SD-WAN solution from any vendor like Cisco, Riverbed, VMware and so on. What can be done, in order to understand applicability and feasibility of such systems inside the E-FCR, is just to provide some comments on their common applications and compare them to the goals of the E-FCR interconnection.

SD-WAN is a cutting-edge breed of solutions meant to manage and optimize interconnection of a large number of multiple remote sites/infrastructures (WAN topologies of datacentres, branch offices, manufacturing/retail sites, cloud-based hosting for any type of business and industry). Such architectures typically include at least a single physical/virtual dedicated appliance for each branch site plus a main appliance and/or a controller (in the cloud or in a central DataCenter).

Companies and organizations with specific business continuity, high-availability and performance optimization requirements can leverage SD-WAN advanced features in order to achieve a mid/long-term ROI from this kind of expensive solutions. The most valuable among those features is for sure the dynamic application-aware routing or, in other words, intelligent and dynamic path selection. That is what distinguishes SD-WAN from classic VPN technologies.

SD-WAN products can be considered low cost solutions if compared to more classic technologies (e.g., MPLS with additional integrated QoS services), but it's still an investment that sounds hard to justify with the scopes of E-FCR interconnection.

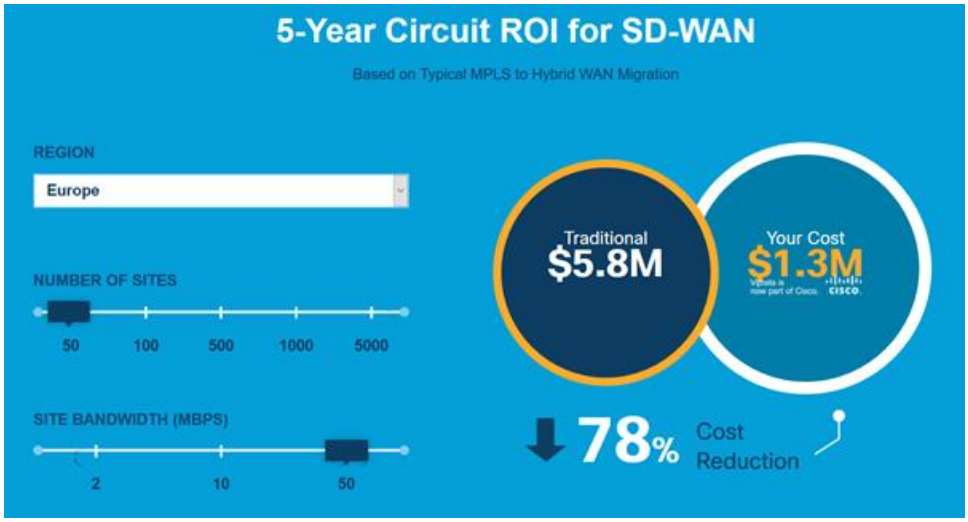

Figure 4: 5-Year Circuit ROI for SD-WAN, taken from https://www.cisco.com/c/en/us/solutions/enterprise-networks/sd-wan/roi-calculator.html

Consider that the above calculator is meant to highlight the cost effectiveness of SD-WAN.

So, reducing the application of such solutions to connecting single CRs just for a limited set of Scenarios and for limited amounts of time (when the Scenario/Training needs to be run) looks quite out of scope and oversized in terms of investments.

Besides, we can suppose that an SD-WAN system could be used to terminate directly the interconnection between CR platforms, but in that case, the IP traffic coming from the Scenarios will be intercepted by the SD-WAN appliance like any other traffic of the infrastructure and therefore will need to be assigned static IP subnets to comply with IP addressing of each site. That considerably limits the flexibility of the IP addressing inside the Scenarios and poses the issue of reusability of those IP subnets.

On the other hand, SD-WAN is something that can still be deployed or not on top of already existing VPN interconnections between CRs if, in the future, the need arises to achieve a far more “strictly binding” network among CRs’ infrastructures. In that case, SD-WAN system will apply its advanced features to manage and optimize inter-CR traffic, like it is supposed to do with any other type of traffic between remote sites joining the SD-WAN.

Benefits of the SD-WAN

As above mentioned, SD-WAN advanced features can achieve a mid/long-term ROI in terms of high-availability and business continuity. The most valuable among those features is for sure the dynamic application-aware routing or, in other words, intelligent and dynamic path selection. That is what distinguishes SD-WAN from classic VPN technologies.

SD-WAN products can be considered low cost solutions if compared to more classic technologies (e.g., MPLS with additional integrated QoS services).

Limitations of the SD-WAN

This approach seems technically very reliable, but did not seem actually applicable in the E-FCR context:

- Long-term agreements with Internet Providers seem complex to be achieved by the actual ECHO project (while, when the reorganization of ECHO into a future network of centres will be achieved, this issue could be mitigated)
- The solution seems expensive and not particularly agile for being proposed to the commercial market
- As explained above, the IP traffic coming from the Scenarios will be intercepted and managed by an SD-WAN appliance and therefore will need to be assigned static IP subnets to comply with IP addressing of each site's infrastructure. That limits the flexibility of the IP addressing inside the Scenarios and poses the issue of reusability of IP subnets.

As a conclusion, this approach is not actually feasible for E-FCR. However, it could be considered if the E-FCR will be successfully transformed into a profitable/financed and sufficiently widely used service towards the end of the ECHO project.

Layer 3 logical interconnection

Due to the configuration complexity and the needed investments for Layer 1 and 2 interconnections, ECHO focused on **Layer 3 VPN**. Also in this case, several options have been analysed.

Fully-meshed network of VPN site-to-site tunnels

For instance, a *fully-meshed network of VPN site-to-site tunnels* among Providers could provide a deep integration of infrastructures [Allar-Federation], but would bring overwhelming efforts in terms of management of routing policies (e.g., possible overlapping networks inside different scenarios, routing configuration complexity with interaction between infrastructures) and governance of cross-provider configurations (e.g., security reasons). In this case, an efficient meshed network of N federated members will need a total of $[N*(N-1) / 2]$ VPNs to be established (e.g., a federation of 3 members will need 3 VPN tunnels while a federation of 7 members will need 21 tunnels with commonly agreed policies for IP addressing, dynamic routing configurations and route filtering for security reasons).

Client-to-server Layer 3 VPN

Within the Layer 3 logical interconnection option, the team studied another possibility which may overcome the above-mentioned challenges could avoid the stretching/merging of virtual infrastructures: the aim would be the sharing of cyber range resources among several Providers, simply allowing each member of the federation to reach other member's resources (not moving resources or centrally managing all Virtualisation infrastructures). Such solution would consist on opening a *client-to-server Layer 3 VPN* between Cyber range Providers each time a scenario within a Service Request includes resources available on other cyber ranges.

As illustrated in the example below, Scenario 1 includes resources from Provider A, B and C while Scenario 2 includes resources from Provider D and B.

In the first case, A and B will provide their own packages for the client VPN connection to C (the package could consist of either a complete Virtual Machine (VM) with its configuration or the VM requirements with configuration). Provider C will take care of its own Domain Name System (DNS) and Network Address Translation (NAT) configurations (to apply on the provided package) in order to guarantee proper access to remote resources from its own part of Scenario 1.

In the second case, B will provide VPN-client package to D that will take care of its own DNS and NAT configurations (to apply on the provided package) in order to guarantee proper access to remote resources from its own part of Scenario 2.

In all cases, the server side of the VPN will need ad hoc Port Address Translation (PAT) rules (on the edge firewall of its infrastructure) in order to properly forward incoming VPN traffic from different remote Cyber ranges (VPN clients) to the correct scenario (VPN server).

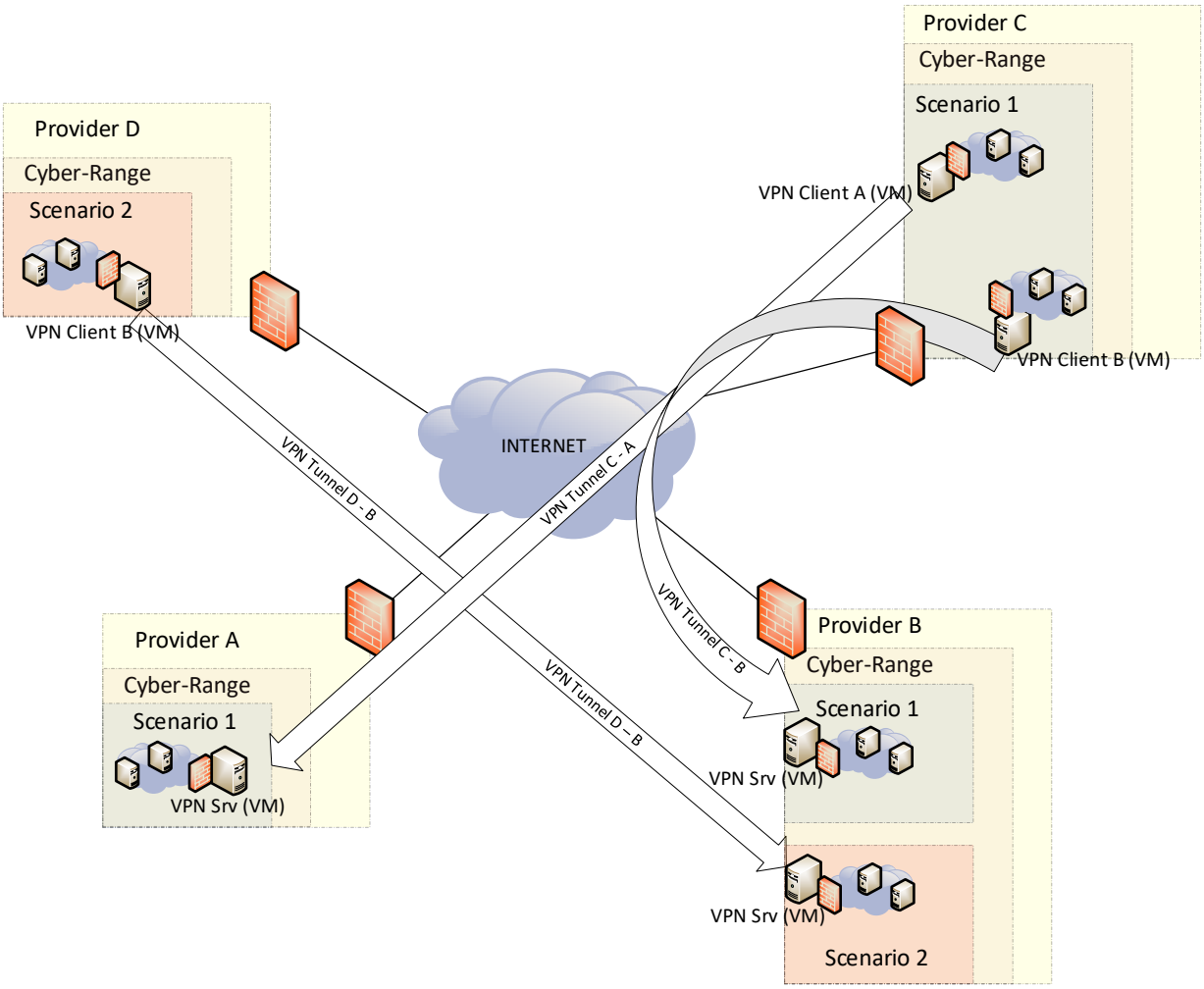

Figure 5: Cyber ranges interconnection with VPNs on different scenarios

The above described approach will not prevent any couple or group of federated members to implement a site-to-site VPN (or leverage an existing one) based on relevant partnerships and they will continue to manage those interconnections and will be responsible of their possible adaptations. For instance, site-to-site L2 VPNs could be used for specific needs of broadcast domain (VLAN) stretching between remote sites.

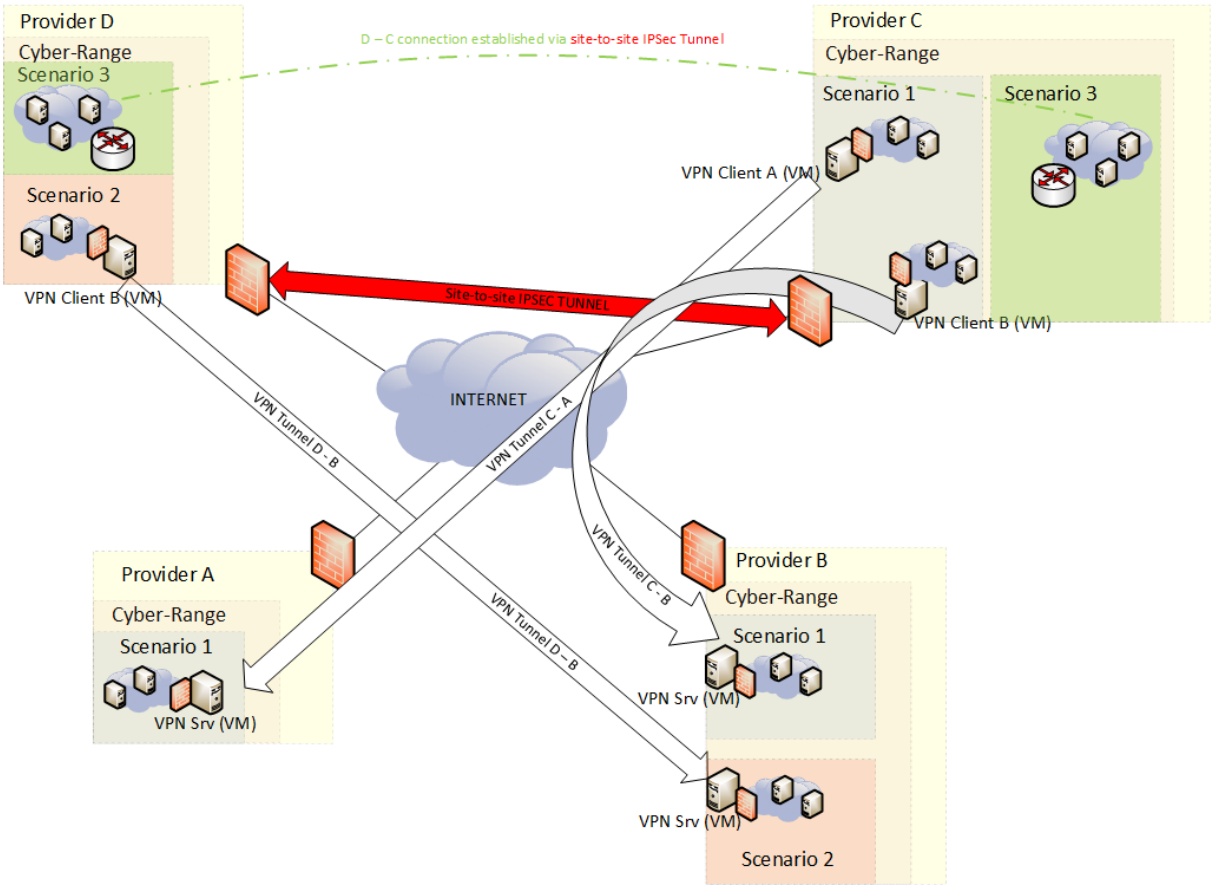

Figure 6: Cyber ranges interconnection with VPNs on different scenarios, existing site-to-site VPN already configured

The VM acting as VPN client/server will need to be properly sized in terms of computational resources in order to bear the routing and encryption/decryption functions for all the interconnection traffic involved in the Scenario. Also Quality of Service (QoS) capabilities and requirements for each Cyber range provider should be defined in order to properly manage such traffic.

The aforementioned solution seems particularly simple to be used within a commercial structure such the E-FCR: a VPN VM will be developed by ECHO and will be configured depending on the characteristics of each onboarding cyber range. In case a Service Request from Customers encompasses one or more scenarios shared with other Cyber range Providers, the VPN VM will be included and deployed in the scenarios, a final configuration will be performed and the connection with other VPN VMs will be established, enabling the sharing of resources.

Tests of the Client-to-server Layer 3 VPN

The approach has been extensively tested during the first months of the ECHO project:

Site X is composed of a set of virtual resources organized on a scenario and instantiated by an OpenNebula cluster, type of Infrastructure-as-a-Service (IaaS) at RHEA Cyber-Security Centre of Excellence premises in Redu (Belgium). RHEA is currently running its own cyber range (CITEF) right on top of this IaaS platform. That includes a virtualized environment consisting of several virtual networks and multiple VMs (including Kali distro).

Site Y is a medium-large cyber range virtual infrastructure based on VMWare and composing an energy sector-specific cyber range hosted by EDP (www.edp.pt), a large energy producer and distributor in Portugal. EDP is RHEA's project partner within H2020 Cyberwiser.eu (www.cyberwiser.eu). EDP cyber range leverages multiple physical SCADA appliances useful to train energy operators with an high degree of realism.

As can be seen, the underlying technologies of the two cyber ranges are quite different (this is a very likely scenario on the commercial world).

In order to test a VPN interconnection between those scenarios, they have been equipped with:

- Linux VM (Ubuntu 16.04.6 Server distro) acting as a VPN server (OpenVPN v2.4.7) inside Scenario X.
- Linux VM (Ubuntu 16.04.3 Server distro) acting as a VPN client (OpenVPN v2.3.10) inside Scenario Y.

During installation of the OpenVPN server, a VPN-client configuration file has been created. That basic configuration includes all VPN negotiation parameters (including a VPN user account) and routes to be added at client side in order to route traffic through the VPN tunnel. That text file can be easily modified to fit different needs. A copy of that file has been transferred to the VM running the OpenVPN client.

At Scenario X, the firewall system of the site's infrastructure has been configured in order to forward incoming VPN protocol traffic from Scenario Y. Namely TCP/UDP port 5190 have been used in this case, but those can be customized if needed.

Finally, the establishment of the VPN has then been successfully tested through the use of a very simple script on the client VPN machine. Such script can avoid human interaction for entering credentials in order to establish the VPN tunnel and can be triggered either manually or automatically (through Linux *cron* tools, for instance). It's also possible to run the OpenVPN client as a service which would probably be the best option for the purposes of Scenarios' interconnections. This operating mode will also be tested soon.

A manual test of client-server VPN connection, with OpenVPN 2.4.6 (and 2.4.7), has also been successfully performed using a Windows 7 machine as a VPN client (automation modes of OpenVPN on Windows client system are yet to be explored though).

The possibility to stretch single VLANs over multiple sites (L2 VPN), with the same OpenVPN technology, is currently under investigation but not yet available.

Once the interconnection was easily established, an interconnected scenario was available for the testers. During the test, some exploitation traffic following a medium complex kill-chain has been run from some Kali machines within Scenario X, targeting systems and physical appliances at Scenario Y.

The tests were very encouraging about the adopted solution, since the interconnection was simple to be achieved and the scenario experience, from a user perspective, was excellent.

Access to remote console of different cyber ranges VMs

Given the positive results of the first tests of the ECHO actual approach for the interconnection, special consideration has been taken into account regarding the access of remote console of the VMs hosted by different Providers (End Users must be able to access to all their assigned VMs, usually via a remote console/GUI). Being part of the final Service delivery, the access to the VM of a scenarios is not provided by the E-FCR, but it is an activity carried out by the Cyber range Provider.

For instance, on Scenario 1 in Figure 6, a user directly connected to Provider C may need remote console access of VMs hosted on Provider A and/or B. In order to achieve this goal, three approaches will be tested:

- Each VM of the remote provider will need to be configured properly to offer a remote connection system to its console (RDP, VNC, SSH etc.) so that they will be reachable through the client-to-site VPN tunnel (this may bring interference to the scenario content, but such conflict can be solved by ruling out the relevant remote connection port/protocols from the scenario activities)
- Each Cyber range Provider will give credentials to access VMs through the regular GUI access of the single cyber range (in this case, the user will use a separate GUI for each Cyber range Provider).
- E-FCR Access Portal will include a Role-Based Access Control RBAC system giving each End-User the remote access to the VMs exposed by each Cyber range Provider for each scenario (this approach will be unlikely adopted in the lifetime of the ECHO project due to its complexity and the available effort, but it could be considered for the future E-FCR roadmap after the ECHO project).

All the above descriptions are not meant to be exhaustive of all the design aspects, but represent valid starting points for further investigations and discussions about implementations.

As a final consideration, further tests of the VPN VM solution will be conducted from M10 to M15, but the initial results, ahead of schedule, are quite positive, even considering some inevitable limitations of the Layer 3 VPNs.

Benefits of the client-to-server Layer 3 VPN

As above explained, client-to-server VPN approach is:

- independent from Virtualisation technology
- independent from physical and IP infrastructure (including network addressing schemes of each CR, that is, no need to coordinate network addresses at infrastructure layer)
- cost effective (no dedicated hardware, no proprietary/licensed solutions).
- based on strong encryption algorithms and secure authentication.

Limitations of the client-to-server Layer 3 VPN

For example, any kind of Layer2-based cyber-attacks cannot be implemented across the remote sides of a L3VPN interconnection. Anyway, this is a limitation that is also implicit in any Scenario with L3 network segmentation (no matter whether it's inside a single CR provider or spanned across multiple CRs).

Interconnectivity conclusions

Despite the aforementioned limitations, the Client-to-server Layer 3 VPN technical solution seems appropriate for the actual E-FCR commercial target due to its agility and extremely cheap cost and configuration simplicity. As detailed in D6.1, this is the solution actually implemented and extensively tested in the E-FCR. However, the team considers the SD-WAN solution as potentially interesting for a future evolution of the E-FCR, when the service will be active and a possibly a bigger budget will be available. In particular, it could be interesting to evaluate an hybrid solution combining the SD-WAN and the client-to-server Layer 3 VPN: some Cyber Range providers (sharing a pre-defined, stable agreement and more performance agreement, likely translating on increased costs for the customer) could be connected with more stable, more performant SD-WAN while others could still rely on the Layer 3 VPN to offer a less premium service still capable to satisfy the business needs.

3.3 E-FCR Scalability

3.3.1 Open Vulnerability DBs Usage

Many uses of cyber ranges involve testing offensive/defensive tools conducted by red/blue teams. In this context, a very important phase of the testing scenario is the generation of malicious traffic conducted by the red team in order to exploit vulnerabilities in the simulated environment. Malicious traffic can look like normal traffic such as normal system administrator traffic with attackers that perform a port scanning, creating accounts, installing new software or changing passwords. The biggest drawback of this approach is the use of the manual operation dedicated to carrying out the cyber-attack by experts.

A framework could be developed to perform automatic cyber-attacks for training purposes within FCR using open vulnerability databases like for example CVE-details and the Metasploit exploit modules. The coordinated generation of automatic cyber-attacks would remove the manual attack operation by experts leading to a faster and more effective cyber-attack simulation.

Several existing Open Source Vulnerability Databases like the Common Vulnerabilities and Exposures (CVE), the National Vulnerability Database (NVD), US-CERT Vulnerability Notes Database, the Chinese NVD (CNNVD), Russia's BDU, Japan Vulnerability Notes (JVN), etc can provide the FCR trainees with a vast range of known vulnerabilities to be used in their exercise scenarios.

To make better use of them, vulnerabilities should be classified according to:

- The public sector they affect (i.e. Maritime, Transportation, Health Care, Energy)
- The type of Infrastructure they target (e.g. critical infrastructure)
- The legal owner of the hardware where the software is installed and running (i.e. On-premises (COTS) and off-premises services, also known as cloud services (IaaS, Platform-as-a-Service(PaaS), SaaS))

The information retrieved from the databases should be cached by the FCR, which should also be in position to track any developments or changes in the contents of the databases and the threat landscape in order for the developers to adopt them in new scenarios and exercises.

A cyber security training scenario usually needs information to set up a virtual environment. Normally, this information includes CVEs, exploits, hosts and network. CVEs and exploits could be stored into two databases:

- Vulnerability Database: the list of vulnerabilities (targets) that are in the scenario. This list can be taken from an open vulnerability database as for example cve-details. A CVE from cve-details database has details about taxonomy, severity, references, products affected and exploit modules. A prototype database could have the following attributes: cve_id, product_type, vendor, product, version
- Exploit Database: the list of Metasploit attack modules that is attack techniques, malicious scripts or hacking tools

These two databases give a general scenario for any cyber exploitation.

The information gained from the opensource vulnerability databases can be used to enrich the scenarios of the exercises held in the E-FCR environment. By combining and escalating the vulnerabilities and the exploitations, the scenarios generated can have different characteristics; with different duration, persistency

and difficulty to provide the trainees to exercise their skills in attacks ranging from simple attacks to advanced persistent threats.

Depending on the level of detail, multiplicity and multi-nationality of the exercises, the scenarios can be based on red/blue teaming approach as well as other teams can also be integrated. This fact is also in alignment with the requirements of each client and the heterogeneity of the participants / trainees.

Regarding the acclamation of the open vulnerability databases, the E-FCR could offer the following capabilities for red and blue teams.

- The E-FCR has prepared attack scenarios, based on known vulnerabilities and exploits, that can be offered to the client in order to increase the skillset of the trainees in terms of intrusion prevention, incident detection and incident response.
  - A faster and more effective simulation is achieved
  - The client gains from utilizing scenarios created by experts with experience in attack scenarios
- The E-FCR can provide selected vulnerabilities and exploits from the databases so that the client can train users.
  - Trainees increase their capabilities in writing and running red/blue team scenarios.
  - The scenarios can have a better adaptation to the clients' needs and requirements.

The red team has the role of the attacker in those scenarios and is responsible for the generation of malicious traffic in order to exploit identified vulnerabilities. For the successful deployment of the red team the requirements are the following:

- Vulnerability searching from the open source databases
- Definition of the utilization of the vulnerabilities by the red team
- Identification of the vulnerable hosts
- Selection of matching exploits
- Scenario implementation, based on the aforementioned prerequisites and the desired difficulty level
- Decision on the individual characteristics of the scenarios (e.g. duration, number of participants, etc.)

Practical application using Metasploit:

Metasploit is used as the primary automatic attack method for training scenarios. Metasploit is the most commonly-used framework for exploiting vulnerabilities worldwide. It allows attackers to exploit security vulnerabilities in operating systems and software through custom code or Metasploit modules. The three main features of Metasploit that can be used are:

- Modules: Metasploit integrates many modules, updated periodically, for cyber-attacks which could be used for reconnaissance, exploiting vulnerabilities or privilege escalation. Therefore, using Metasploit is a good solution for automatic cyber-attacks;
- Modules generalization: Metasploit attack modules almost always have the same format; they therefore share the same commands and the same options. It is therefore possible to generalize the attack tasks;
- Automatic running: Metasploit supports script running so that execution can take place programmatically and the results of the program can be used for further actions

In an ideal workflow it should be possible to generate a script file as an input for Metasploit. The script should contain the name of vulnerability, the IP address of the victim and the log file location. The workflow should follow the following steps:

- The trainer chooses a vulnerability which want to study on from the Vulnerability Database
- The vulnerability is searched in the Exploit Database;
- The module found is executed in order to attack the victim;
- The output of Metasploit (if the attack was successful or not) is stored in the log file;
- The log file is parsed and analysed to give the attack result to the trainee in order to inform him/her whether the attack was successful or failed

The trainee, after checking the success of the attack, could connect to the victim, adopt procedures to counteract that type of attack and restart the automatic attack in order to see if this time it fails.

Defense scenarios are as of significant value as red-teaming scenarios for a CR. Training of Computer Emergency Response Teams (CERT) and members of Security Operations Centre (SOC) must include blue team exercises. These exercises must be repeated with the recent developments for predefined periods to keep the knowledge of the teams up to date. For this reason, the defense scenarios provided in FCR must include open source vulnerability databases that are being constantly updated and kept well-maintained.

For a successful defense scenario, the actions of the scenarios must include risk analysis, detection, mitigation and counterattacks. By utilizing the aforementioned open source databases, an participant can analyse its own vulnerabilities for its software and hardware arsenal and look for IOC provided by these databases for detection. With the community-maintained, crowdsourced knowledge bases about vulnerabilities, mitigation strategies can be developed and if necessary, counterattacks can be directed.

Another significant benefit that are originated by the open source vulnerability databases is the ability to calibrate security measures and detection/prevention mechanisms for attacks with diverse parameters such as duration, scale and technicality.

An attack is also very tightly coupled with the network configuration, testing different attacks on different configurations will make the systems ready for a wide range of possible attacks. The network configuration requirements might come from scenario dependencies or from user-based requirements. In this case, not only the required instances from scenarios are taken into accounts, but also the vulnerabilities that comes within the network configuration is also considered.

A direct requirement for defense scenario development is the inclusion of the open source vulnerability databases integrated with exploitation tools such as Metasploit as mentioned in the previous part of this section. To fully comprehend the weaknesses of the infrastructure provided by the E-FCR and the COTS applications that are being used on E-FCR, it is also vital to include risk analysis scenarios with open sources vulnerability databases. Tools like OWASP, OpenVAS and Nexpose can provide these integrations with the vulnerability databases. These databases are also revealing the indicators of the exploitations therefore providing useful information for detection.

Designing of the E-FCR should also consider the following:

- Most vulnerable COTS – These products generally contain the weak points of the software inventory in the enterprises and prioritizing these COTS will increase the actuality of the scenarios.

- COTS common among partners and enterprises – This will increase the relevance of all scenarios relative to partners and it will help training the personnel of the partners for the most relative scenarios.
- COTS common in the contemporary attack reports – This consideration will help training the personnel for the most up to date information about attacks and again will increase the timeliness of the developed scenarios.

E-FCR should include the necessary links between the Early Warning Systems. One of the most important consideration of cyber security training is the dissemination of the attack information to build stronger federated defense mechanisms leveraging the CTI.

3.3.2 CR Interconnection

When federated scenario is requested there are number of tasks to be executed before the scenario becomes available for use:

- Connectivity between participating CRs should be established
- Appropriate VMs/services should be spun and configured
- Monitoring should be established to track progress of the exercise(optional)

In the initial versions most of those tasks are to be performed by administrators of the CR, however this approach is resource consuming and prone to errors. Automation of those tasks would be very beneficial for the scalability and speed of operation of the E-FCR, however the major obstacle is that the E-FCR is connecting to many different cyber ranges, each with its own architecture and specifics. Thus, is very important when introducing automation to the CRs to keep the ability to fail gracefully and delegate the tasks that cannot be automated in that particular CR to human administrator.

In order to approach the automation let's look in more details the architecture of typical CR. Some of the major components are:

- Virtualisation engine
- Orchestration engine
- Participant Access/VPN
- CR-to-CR VPN
- (optional) monitoring engine

Out of those CR-to-CR VPN is almost always the same technology, as it's supplied by E-FCR in current version and is the ideal candidate for automation. The technical realization is as VPN VM, from which external interface should be made available for management. As COTS technology and OS is used for the VM the most effective approach would be to implement one of the publicly available infrastructure management systems.

Participant Access/VPN is the next candidate for automation, as depending on the technology, the function of the module is to provide access of the participants that are located outside of the CR to the CR available resources. As such usually the combination of username, password and participant role are enough for the participant access to be configured. As usually OpenVPN and RDP or VNC is used for participant access, again COTS technology is used and setup of users with their respective credentials and connectivity parameters (I.e. Red Team have access only to Red Team Workstations) can be implemented with standard infdrastructure management system

Orchestration Engine is used to simplify management of the number of VM systems that have to be spun and configured during setup of exercise. As such it overlays the Virtualisation engine and simplfies the complex tasks of creating virtual environments containing multiple servers, networks and communications devices. One major issue that can be tricky is provisioning of the VMs up to the point where they are capable to be managed by configuration management systems. This hurdle can be overcomed by working with the CR orchestration engine and not with the VM engine directly.

Infrastructure management systems

Comercially available Chef, Puppet, Ansible, and SaltStack are industry-wide used DevOps tools, included in some DevOps Certification paths. They are all configuration management tools, which means they are designed to deploy, configure and manage servers, which makes them ideal candidates for automation of the CR federation setup tasks

Ansible

As a latest entrant in the market compared with Puppet, Chef and Salt, Ansible was developed to simplify complex orchestration and configuration management tasks. The platform is written in Python. Architecture of Ansible deployment is single active node, called the Primary instance. If primary goes down, there is a Secondary instance to take its place. The Primary can reside in the E-FCR infrastructure with Secondary instances distributed among participating CRs.

The general approach is push configuration, i.e. all the configurations present in the central server will be pushed to the nodes. Ansible offers multiple push models to send command modules to nodes via SSH that are executed sequentially. Ansible doesn't require agents on every system, and modules can reside on any server. A centralized Ansible workstation is commonly used to tunnel commands through multiple Bastion host servers and access machines in a private network.

The configurations are stored in YaML files - The YAML acronym was shorthand for Yet Another Markup Language. But the maintainers renamed it to YAML Ain't Markup Language to place more emphasis on its data-oriented features. The language is a human-readable data-serialization language, commonly used for configuration files and in applications where data is being stored or transmitted. YAML has a minimal syntax and uses both Python-style indentation to indicate nesting, and a more compact format that uses [] for lists and {} for maps[1] making YAML 1.2 a superset of JSON.

From interoperability standpoint Ansible supports windows machines as well but the Ansible server has to be on Linux/Unix machine..

Saltstack

Salt was designed to enable low-latency and high-speed communication for data collection and remote execution in sysadmin environments. The platform is written in Python and uses the push model for executing commands via SSH protocol. Salt allows parallel execution of multiple commands encrypted via AES and offers both vertical and horizontal scaling. A single master can manage multiple masters, and the peer interface allows users to control multiple agents (minions) directly from an agent. In addition to the usual queries from minions, downstream events can also trigger actions from the master. The platform supports both master-agent and de-centralized, non-master models. Like Ansible, users can script using YAML templates based on imperative paradigm programming. The built-in remote execution system executes tasks sequentially.

From architectural standpoint It can have multiple masters configured. If one master is down, agents connect with the other master in the list. Therefore, it has multiple masters to configure salt minions.

From interoperability standpoint Salt Master works only on Linux/Unix but Salt minions can work on windows as well.

Puppet

Puppet is built with Ruby and offers custom Domain Specific Language (DSL) called Pupped DSL and Embedded Ruby (ERB) templates to create custom Puppet language files, and offers a declarative paradigm programming approach. Puppet uses an agent/master architecture—Agents manage nodes and request relevant info from masters that control configuration info. The agent polls status reports and queries regarding its associated server machine from the master Puppet server, which then communicates its response and required commands using the XML-RPC protocol over HTTPS.

Puppet server runs on the master machine and Puppet clients runs as an agent on each client machine. It also has multi-master architecture, if the active master goes down, the other master takes the active master place.

The general approach is pull configuration, so the slave nodes will automatically pull all the configurations from the central server without any commands.

From interoperability standpoint Puppet Master works only on Linux/Unix but Puppet Agent also works on windows.

Chef

Chef has a master-agent architecture. Chef server runs on the master machine and Chef client runs as an agent on each client machine. Also, there is an extra component called workstation, which contains all the configurations which are tested and then pushed to central chef server.

The general approach is pull configuration, so the slave nodes will automatically pull all the configurations from the central server without any commands.

Configurations are stored in RubyDSL language.

From interoperability standpoint Chef Server works only on Linux/Unix but Chef Client and Workstation can be on windows as well.

As conclusion out of the listed configuration management tools Ansible seems the most logical choice to be implemented in the context of E-FCR.

Request protocol

The requests created by the E-FCR Service Broker should be processed by the centralized configuration management system dependent on the particular CR capabilities.

- In case the selected CR do not have automation capabilities, the request should fall back gracefully to human readable format and send to person, capable of processing it further
- In case partial automation capabilities are present (I.e. just CR-to-CR VPN), then those are configured automatically, notifications are sent and remaining requirements are sent in human readable format.
- In case of full automation capabilities, all the configuration tasks are performed, and notifications are sent to appropriate services and persons.

3.4 E-FCR Platform

3.4.1 Virtualisation evolution

Introduction

Over the last years, there has been many developments regarding how computing services reach the end-user and how computational resources are made available. Beginning with mainframes, moving on to personal computers and finally reaching the point of another disruptive trend; virtualisation.

Virtualisation usually refers to an abstract layer that lies between a physical and a logical component. Through the virtualisation of an object, someone is able to achieve a greater degree of modularity, and gain control to only a certain set of resources, something impossible in the context of traditional computing and networking [25]. For example, virtual local area networks (VLANs), provide network performance and manageability through physical layer separation. Likewise, storage area networks (SANs) provide greater flexibility, improved availability, and more efficient use of storage resources by abstracting the physical devices into logical objects that can be quickly and easily manipulated.

Virtualisation and Hypervisors

By definition, a Virtual Machine (VM) can virtualise all hardware resources, including Central Processing Units (CPUs), memory, storage, and network interfaces. A Virtual Machine Monitor (VMM), commonly called a hypervisor, is the software that provides the environment in which the VMs operate. A simple overview of a VMM is illustrated in Figure 7.

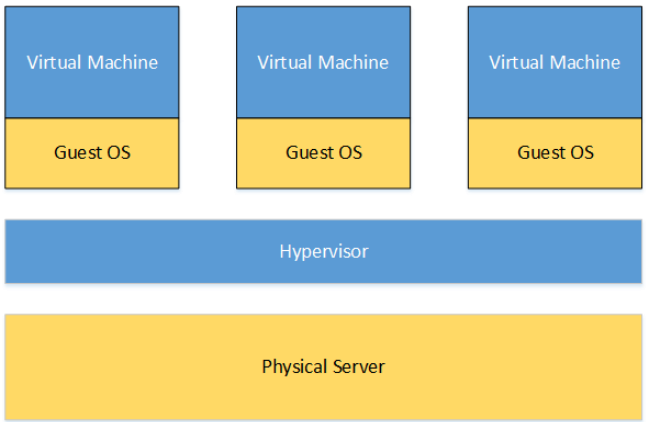

Figure 7: Simple illustration of a VMM

According to [26], a VMM must exhibit three basic properties in order to satisfy the definition. These properties are commonly referred to as the Popek and Goldberg virtualisation requirements:

Fidelity: The created environment should be identical to the original physical machine.

Isolation or Safety: The VMM should have total control of the resources.

Performance: There should be no difference between a physical machine and VM performance-wise.

Since most of the VMMs are able to achieve compliance with the first two properties, the third criterion is what differentiates an efficient VM from their inefficient counterparts.

Even though the idea and execution of virtualisation dates back to 1960, since IBM used it on their mainframes, the first commercially available solution targeted to x86 platforms came from VMware in 2001 [27]. Two years later, Xen was made available as an open-source solution, bringing the ability to condense numerous physical servers into fewer servers running many virtual machines. This condensing ratio is commonly referred to as consolidation and in its most basic form is calculated by dividing the number of VMs with the number of physical machines.

A second development besides consolidation, is containment. As companies realised the benefit of virtualisation, once their older servers reached their end-of-life (EOL), they staged their application workloads on their existing virtual infrastructure instead of purchasing new hardware.

There are two classes of hypervisors, namely Type 1 and Type 2, and their main difference is the way they are deployed.

Type 1 Hypervisors: A Type 1 hypervisor runs directly on top of the server hardware without having to rely on an underlying operating system. Since there is no intervening layer between the hardware and the hypervisor, they are also referred to as bare-metal implementation. Due to the absence of an intermediary layer, Type 1 hypervisors can directly communicate with the hardware resources beneath them, which offers greater efficiency than Type 2 hypervisors. A simple architecture of this type of hypervisors is illustrated in Figure 8.

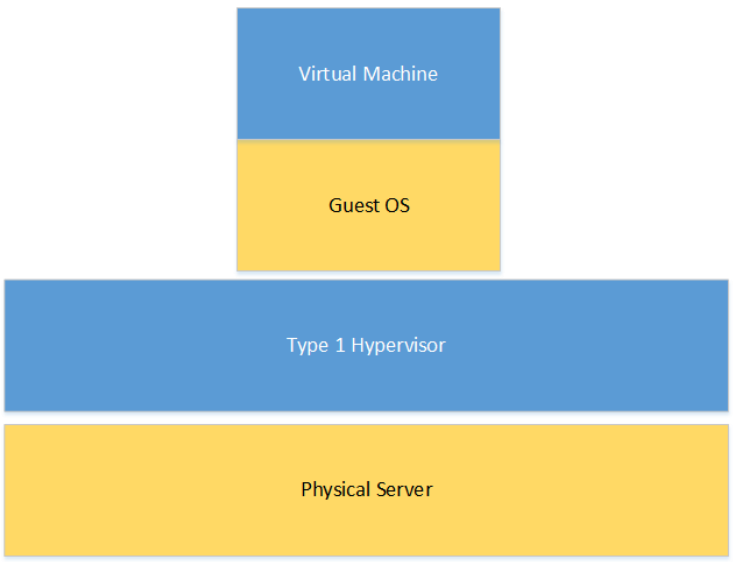

Figure 8: A simple illustration of a Type 1 Hypervisor

Type 2 Hypervisors: Type 2 hypervisors are applications that run in a traditional operating system. They are easier to deploy since they do not differ from a typical application, and most of the configuration has already been set up by the operating system itself. They also support a large variety of hardware, as long as the operating system it uses supports it. This, however, is also the main drawback of type 2 hypervisors since the extra layer between the physical server and the hypervisor creates an additional performance overhead to almost every hardware interaction. Each request has to be forwarded to the operating system which handles the I/O requests and then passes the information to the hardware adding two additional steps in the process. As a general rule, type 2 hypervisors are less reliable because they contain an additional point of failure, since

anything that affects the host operating system also affects the guest operating systems above it. A simple architecture of a type 2 hypervisor is illustrated in Figure 9.

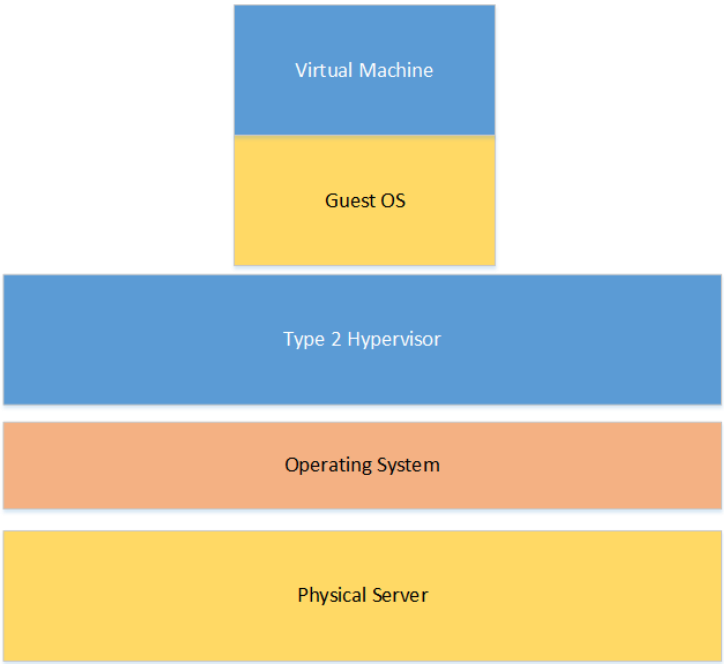

Figure 9: A simple illustration of a Type 2 Hypervisor

Virtualisation in E-FCR

A cyber range is defined within ECHO as a multipurpose virtualisation environment and has the potential to help strengthen the stability, security and performance of cyberinfrastructures and information technology (IT), operations technology (OT), and industrial control systems (ICS) by facilitating high-fidelity simulations of operational conditions in this virtual environment. These virtual environments can be used as practice ground for hands-on education and training purposes or for advanced prototype development and cybersecurity certification testing.

The use of virtualisation in E-FCR can be found in two occasions:

VPN VM: Cyber Range inter-connectivity is realised by VPN VM technology, developed by RHEA. VPN VM allows to inter-connect two cyber ranges with *client-to-server Layer 3* virtual private network (VPN). The VPN software and its configuration are contained within a virtual machine on both client and server side of this point-to-point connection. According to deliverable D6.5, VPN VM technology is agnostic to the virtualisation technology, used by a cyber range itself. However, it requires certain technical readiness from the cyber range's networking point of view. It has also some limitations, mainly stemming from the fact it is an L3 segmentation technology.

VMs: Each scenario requires a number of VMs to be deployed in every single one of the participating cyber ranges. These scenarios must contain a router/firewall (virtual appliance or VM with routing functions) acting as a transit hop between the VPN VMs and the rest of the VMs/networks inside the "local" part of the scenario, but also VMs that will contain the actual nodes of the aforementioned networking.

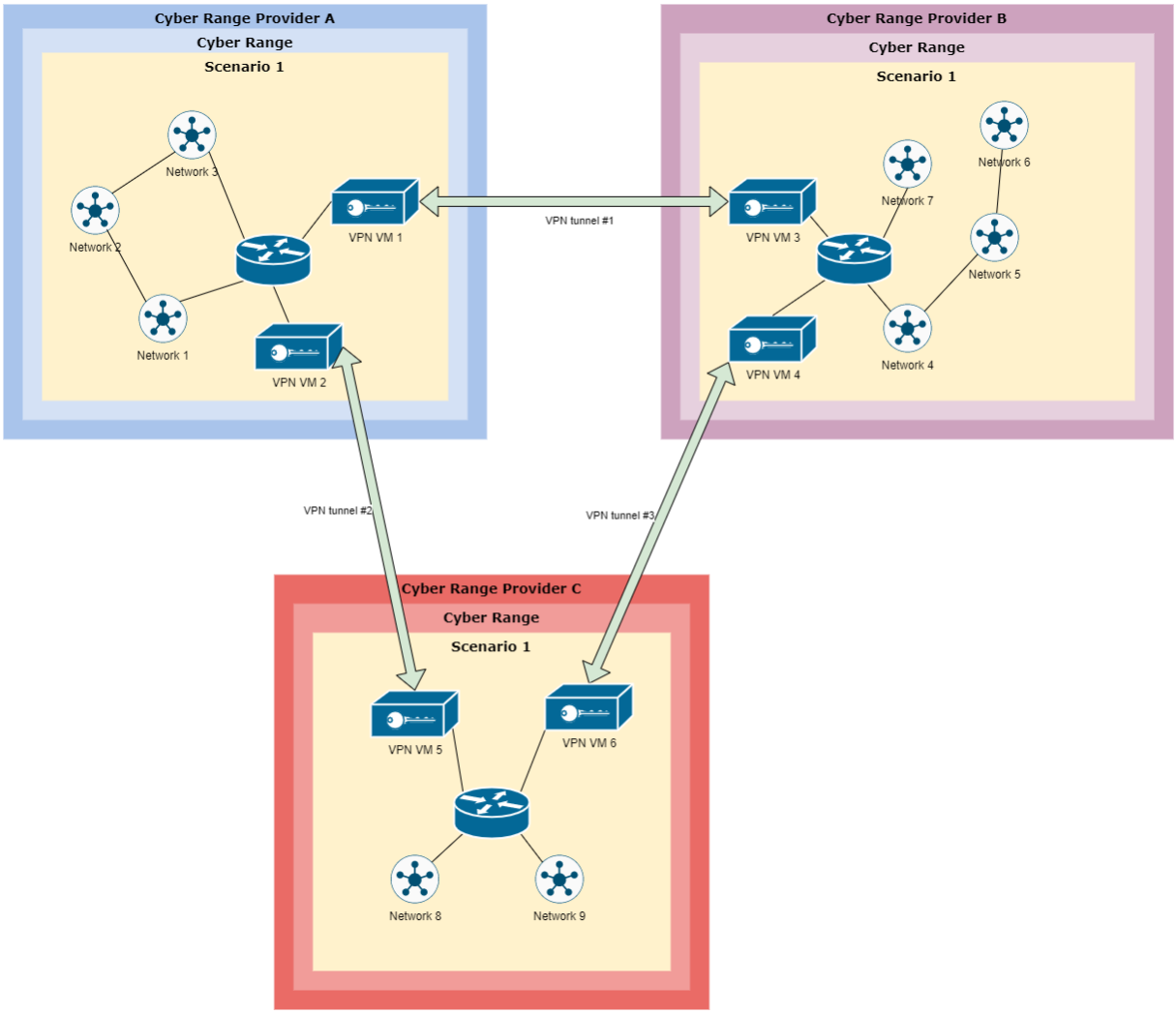

Figure: 10 Virtualisation in E-FCR

Evolution and future opportunities

Among emerging virtualisation technologies, the most common one is storage virtualisation (also called software-defined storage), which according to [28] has a 40% adoption rate. The Storage Virtualisation provides a combination of storage resources for Storage Area Network with multiple servers as well as the storage devices. The main aim for storage virtualisation is to be inexpensive without any direct impact to the performance. Currently, the storage virtualisation can be divided into three different architectural levels: (1) the storage device, (2) the host and (3) the SAN fabric hardware as the central management unit. Some of the benefits of adopting software-defined-storage technology are the following:

- Enhanced flexibility with storage infrastructure
- Improved fault tolerance
- Improved storage performance
- Decreased hardware investment compared to alternatives
- Improved manageability of storage
- Increased storage scalability

Another emerging virtualisation technology is intent based networking, which is defined as a network control in form of policies, explaining what to achieve, rather than mechanism, or how to achieve a result. Intent based network requires the capability to provide the intent in case of network failures or changes in the state of the network. Taking the example of communication between hosts, if a given device in the chosen path is removed from the network making it incapable of forwarding traffic anymore, the intent has to be recomputed to provide again a feasible end to end path with a new configuration that will be sent to devices in the network. Such capability assumes awareness of the system to network state at any given point in time [29].

To keep up with the needs of the market, virtualisation technologies, like hypervisors, have been making major development steps. However, VMs are not as efficient in specific use cases where smooth scaling and density of applications becomes a major factor. These shortcomings are mostly related to the large size of the guest OS which waste a huge amount of system resources. To overcome this, new virtualisation techniques such as container-based virtualisation and Unikernels [30] have emerged as efficient and lightweight alternatives to traditional VMs for specific cases of small or medium size application deployments. Containers, contrary to VMs, avoid deploying a full-blown copy of an OS for each instance and instead depend on sharing the same base OS among themselves. This hybrid container/hypervisor approach, offers unprecedented agility in developing and running applications and enables the design of multiple tenancy applications with less overhead, compared with the traditional hypervisor solutions [31].

Comparative analysis of commercial Cyber Ranges roadmaps

As Cyber Ranges are relatively new technology one can expect that the development of new platforms has not matured yet. New platforms utilizing different architecture approaches, novel services and addressed sectors as the market is not yet at a satisfying level. As evolution and changes of the underlying technologies of those platforms is inevitable, it is instrumental for the E-FCR platform to be able to accommodate and follow those changes in order to maximize the potential of participating CRs and thus creating opportunity of growth of the marketplace

In order to explore what are the trends in the CR market a survey was produced and 53 CRs, identified as part of Annex 3 Cyber Range Providers Market analysis in [D6.5] Update - E-FCR High-Level Design and documented in Annex 3 – Cyber range Providers Market Analysis were invited to participate to share details about their directions of development. The survey was deliberately conceived in a way not to request excessive information, as it can be perceived as an attempt to obtain business plans or other trade secrets. Moreover, all questions are marked as optional in order to produce maximum amount of data, in case a participant would prefer to avoid answering any questions. Survey questions were as follows:

Environment

- In what geography you are located: Europe, South America, Asia, North America, Middle East, Africa, Australia
- Is the Cyber range dedicated to specific sector with options: Not dedicated, Yes – Defence sector only, Yes – Law enforcement only, Yes – Academic sector only
- What services is your Cyber range providing currently and do you plan expansion in the future with possible answer for each service – “current services”, “not right now, but plan to have in 2 years” and “not provided with no plans to be introduced’. Services chosen were “Personnel training”, “Development/testing of cyber-security products”, “academic research”, “Benchmarking of cyber security products”
- “Is there any other service that you are currently providing or plan to provide in next two years, that is not mentioned before?” with open text answer

Resources and Cyber Range Platform

- "Do you use commercial Cyber range platform, or you are developing your Cyber range platform in-house?" with options "Commercial platform" and "in-house developed platform"
- How many cores your Cyber range can currently use: 0-50, 50-250, 250-1000, other (open text)
- In next two years how many cores you plan for your Cyber Range to use: 0-50, 50-250, 250-1000, other (open text)
- "How much RAM can your Cyber Range currently use" : Up to 1 Tb, 1-5 Tb, More than 5 Tb
- "In next two years how much RAM you plan for your Cyber Range": Up to 1 Tb, 1-5 Tb, More than 5 Tb
- Do you currently use Virtualisation or containerization technology: Virtualisation – unspecified, Virtualisation -VMWare, Virtualisation -Microsoft, Containerization – Docker, Containerization – OpenShift, Containerization – unspecified
- "Do you plan changes in your platform in next couple of years?":
    - We will switch from Virtualisation to containerization
    - We will switch from containerization to Virtualisation
    - We will introduce containerization, but will not remove Virtualisation
    - No, our platform Is adequate as it is
    - (open text answer)
- Can you currently utilize Public Cloud resources in your Cyber Range:
    - Yes – unspecified Public Cloud
    - Yes – Amazon
    - Yes – Google
    - No and we don't plan to
    - No, but we plan to start using
- "What technologies are currently supported and what you plan to introduce in the future" with possible answer for each technology "Can be used as virtual devices in platform", "Can be used as physical devices, connected to platform", "Not right now, plan in next 2 years", "Not right now, plan in next 5 years" and "No and no plans""
    - ICT/SCADA/OT in Manufacturing
    - ICT/SCADA/OT in Utility
    - Telecom level devices and services
    - DDoS attacks
    - IoT technology
    - Mobile devices as part of attack/defence scenario
- "Is there specific technology that is supported in your Cyber range that doesn't fall in above categories?" with open text answer
- "What other devices or technologies you plan to introduce in next 2 years" with open text answer

Out of all invited participants only six of them chose to provide answers to the survey and the analysis of the provided answers can be found in following sections.

Respondents background and provided services

The six responded CRs are split equally with three in Europe and three in North America. All participants provide services to the general market and are not focused in a particular domain. All participants develop their own CR platforms and do not rely on any commercial platform that is developed by an external organization.

All participants provide personnel training services like Red Team or Blue Team training with various level of maturity, courses, etc. All except one also provide services for academic research. Out of the six, four are currently providing services related to development/testing/benchmarking of cybersecurity products with one participant planning to introduce those services in the coming years.

Platform

In terms of capacity and resources available, only a part of the respondents chose to provide answers to currently available resources and plans for expansion; one CR is using less than 50 CPU cores and two others are using between 250 and 1000 CPU cores. One of the participants is cloud based platform that routinely handles more than 1000 cores as well. In terms of available RAM for simulations; one participant is working with less than 1TB of RAM and two are using between 1 and 5 TBs.

All respondents claim that they will expand their hardware resources available to their CR platforms in the following years which will allow more complex and advanced simulations to be supported in real time. One respondent is planning to dedicate more than 1000 CPU cores and more than 5 TB of RAM to this task in two years and the respondent that is based in public cloud is aiming for a range between 10,000 – 100,000 cores worldwide.

Another way to ensure that CRs have enough resources available is to enable the technology platform to use public cloud services and their elastically reserved resources. This provides an economically efficient way to support massive simulations without bearing prohibitive cost of acquiring, maintaining and operating large computational facilities. Three of the respondents are already using a relevant approach, one of those is fully based in the cloud with the rest planning to introduce such capability in the coming two years by providing either a CR, completely functioning on a public cloud or a hybrid scenario with part of the range working on on-premise servers and a part on the virtual cloud servers.

In the core of CR platforms is the capability to simulate multiple diverse systems in real time. Currently, there are several approaches on how this can be achieved:

- **Physical servers connected to separate network:** This allows the seamless addition of different technologies and devices to the simulation; a setup which is very common across the industry.
- **Virtualized servers:** Using either commercial (such as VMWare and Hyper-V) or open source (such as KVM) hypervisor.
- **Containerized services**: Using platforms like Docker which allow increased performance and better utilization of the available resources at the expense of less separation between the virtualized services.
- **Hybrid solutions:** Combining several or all of the approaches above in order to accommodate broader set of simulations.

Four out of six respondents are currently using Virtualisation platform with three of them supporting the hybrid capability of connecting of physical devices. Two respondents are currently using containerized services. Out of those three, one respondent is planning to switch its CR main platform from Virtualisation to containerization in the next two years and two others will introduce containerization as an addition to already built virtualized architecture. CR platforms that currently use Virtualisation dominantly utilize open source Virtualisation and orchestration engines.

Supported technologies in simulation

The third domain to analyze development trends of CRs is the technologies and devices that are available to be used during performed simulations – either for training of personnel or in cases of testing and certification

of devices. In this realm, hybrid platforms gain advantage as they allow to connect any type of device easily and present it in virtually any point of the simulated network. Such devices should be configured and managed separately, outside of the centralized CR monitoring and management; which is relatively low price to pay for such flexibility.

Another approach is to virtualize all devices that are part of the simulation. This is advantageous for the vendors that chose commercial Virtualisation software due to the bigger number of supported virtual architectures and better support from vendors of the hypervisor and the device manufactures as well.

Our respondents were asked about their support on six specific fields: SCADA systems in manufacturing, SCADA systems in utility sector, telecom level devices, DDoS attack simulation, IoT and mobile devices. All of them supported SCADA systems in their simulations, either as a virtualized system or as a physical device connected to the virtual CR environment. Three respondents supported telecom level devices and IoT in the same fashion, either virtualized or physical. One respondent is planning to introduce the telco devices in next two years and another will support IoT devices in the next five-year interval. DDoS attack simulation seems to be the least popular feature with two CRs supporting it and one is planning to implement in the next two years.

3.4.2 Use of Cyber Threat Intelligence

Adapting to changes

It has become evident in recent years that the traditional, reactive approach to cyber threats is no longer adequate and even anti-virus vendors, software manufactures and security experts are finding it hard to keep up with the pace of threat actors and their tactics, techniques and growing expertise, which makes the timely collaboration and cyber threat intelligence (CTI) sharing between organizations, vendors and security professionals all the more important in this battle.

Gartner defines threat intelligence as the evidence-based knowledge, including context, mechanisms, indicators, implications, and action-oriented advice about an existing or emerging menace or hazard to assets [33]. It is usually thought to come in three forms, as follows:

- Strategic – includes broader trends and adversarial motives typically leveraged in strategic security and business decision making.
- Tactical - outlines some technical aspects, including Indicators of compromise such as IP addresses, file names, or hashes, useful to a more technical audience.
- Operational – provides insight about the specifics surrounding particular attacks, such as intent, timing, and sophistication of tactics, techniques and procedures.

Many CTI solutions, driven by machine learning, are designed for integration with an array of security technologies such as vulnerability scanners, IDS/IPS and SIEMs to name a few. Such integration allows cyber security professionals to be proactive and take advantage of tactical and/or operational data to protect their environments before an incident occurs.

Threat Intelligence in the E-FCR

The availability of actionable threat intelligence data will be a substantial addition to the E-FCR features, especially when it comes to:

- Enabling cyber-security professionals to rapidly adapt to an evolving threat landscape.

- Enhance threat data with real-time context for determining the most suitable response.
- Reducing the time spent on processing threat data which is not relevant to an organization, network infrastructure, or industry.
- Accurate estimations of the impact of cyber-risks.

CTI data could be supplied by a diverse set of security-oriented tools and applications, for example: there are numerous open source and commercial feeds/platforms with different levels of automation and distinct threat reports, considered external sources, which could be utilized for this purpose. There is also a variety of internal ones, in the form of monitoring and prevention systems such as firewalls, intrusion detection systems(IDS), SIEMs and endpoint protection solutions that could pass attack information to a centralized CTI platform, which will turn it to actionable data.

Steps for introducing CTI to the E-FCR

Defining intelligence requirements:

With the constantly increasing amount of threat intelligence data, which could be derived from any of the aforementioned methods, it is practically impossible and unnecessary to analyze, act or even collect each and any alert available from a CTI source. Therefore, identifying threat data of interest, prior to gathering, is highly recommended, and prioritization could be based on:

- Industry Sector, including: Financial Services, Healthcare, eCommerce, Oil & Gas, etc.
- Attack type
- Attack vectors
- Impact
- Targeted assets

Dashboards

Data visualization in common E-FCR dashboard panels is important for providing a seamless access to threat activity by matching intelligence source content to events observed either within the cyber ranges, or externally in the public domain. The following visualization filters would be considered a plus:

Filter by	Description
Data Source	Origin of the data, e.g. SIEM, IDS, public feeds, etc
Threat Group	A named group or entity representing a known threat, such as a malware domain or known malicious actor.
Threat Category	A category of threat, such as advanced persistent threat, financial threat, or backdoor.
Search	Used for searching on a value related to fields: destination, source, threat group, key indicators and artifacts, threat match, etc.
Threat Artifact	A collection of objects, such as network, domain, file, user, process, registry, and service.
Time Range	Select the time range to represent.

Table 4: CTI Filters

SIEM

A bidirectional data flow between the two systems will ensure that indicators from the CTI platform are automatically sent to the SIEM for alerting, during exercises and/or tests, conducted within the E-FCR. Furthermore, the telemetry and specific events, logged by the SIEM during any activities, could be in turn sent back to the CTI platform for aggregation, correlation, analysis and visualization. Processing intelligence data allows analysts to verify the interrelation among different security events, complementing each other so that the complete relevant information, necessary to describe attacks could be obtained. Automating the processes of threat data import/export and rules configuration/updates to a reasonable extent is a prime requirement.

Intrusion Prevention/Detection Systems

Enhanced security monitoring capabilities through renewing the information about threats configured in the IDSs, are an extended application of CTI data within the E-FCR, which should be pursued. By inserting novel rules and blacklists built automatically from intelligence data, the complete cycle of IDS knowledge update - from collecting threat intelligence, renewing signatures, generating rules, and installing them is closed, allowing prevention techniques to be tested and confirmed without much effort. The figure below illustrates the process, where the OSINT feeds could be replaced by any other CTI source, offering integration.

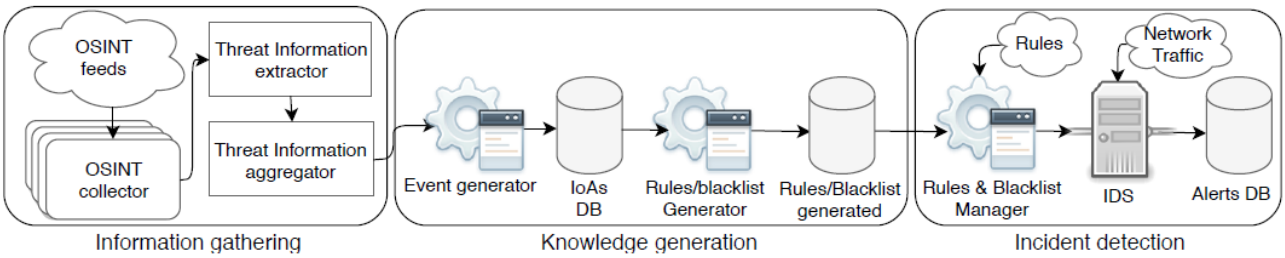

Figure 11: CTI + IDS flowchart [16]

Honeypots

A step further in the proactive approach, honeypots provide excellent means for surveilling attack behaviour in the wild. Moreover, depending on their strategic distribution around the world, observations about the malicious activities in particular region/s of interest could allow analysts to recognize local trends far more effectively. Cloud or on-prem deployments exist, and it seems the former would fit better the E-FCR needs for threat intelligence, as they are more scalable, come with out-of-the-box integration capabilities, reside outside any internal networks, and could be provisioned in various geographic locations with ease. Whatever the deployment may be, the figure below illustrates the process of turning data collected from honeypots into actionable intelligence.

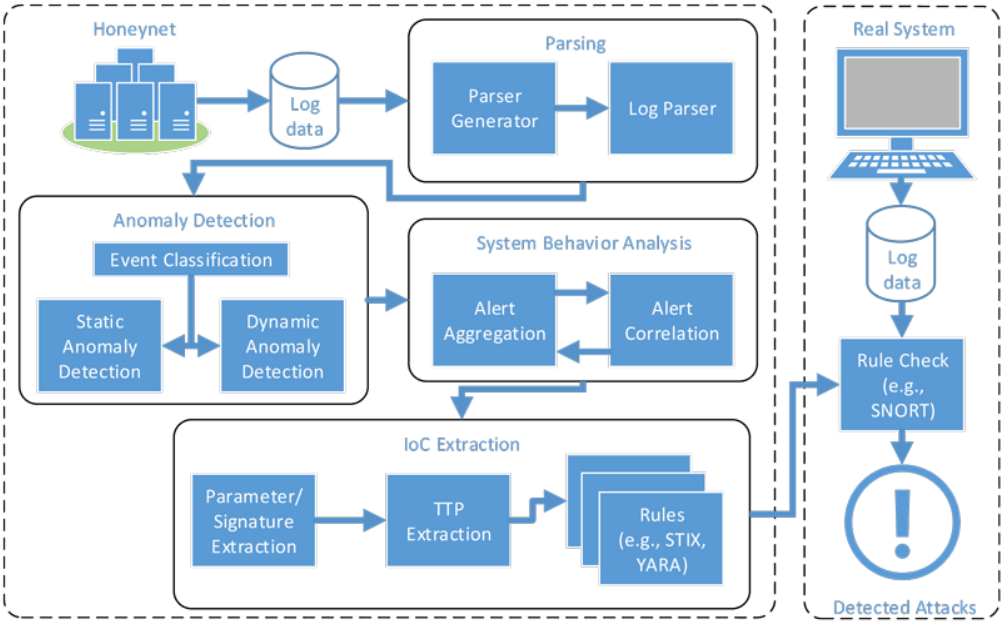

Figure 12 CTI Extraction from Honeynets Example[34]

The need for CTI data consistency, resulting from the layered extraction process above, makes a strong case. This is mainly due to the fact that event log, or alert data, all by itself, does not constitute viable threat intelligence, capable of exposing the sometimes-wide range of adversary tactics, employed in multi-staged attacks. To address this concern, the observed deviations in log data are checked against (and tied to) several data sources, which allows for a comprehensive analysis of even short-lived detections. This examination is likely to paint a complete picture of the intrusion phases, indicating entry points and lateral movement.

It should be noted that E-EWS can also be used as CTI source and the two options of implementing this opportunity are discussed in detail in UserStory 3.4.3 Integration with E-EWS of this document.

Selecting CTI sources

External sources

For sources of this type to be considered credible in the E-FCR context, the following requirements should be met:

- Support for STIX/TAXII[35], IODEF[36] and OPENIOC[37] standards – the availability of specifications for describing and sharing cyber threat information in a common language that can be easily understood by humans and a wide range of security technologies is an important factor.
- Automation – manual user interaction for sharing and acquiring valuable intelligence should be kept at minimum.
- Flexibility – support for feeds, direct, or API integration and HTTP GET scripts to pull data from a remote web source, allowing the continuous update of short-lived indicators of compromise lists.
- Presentation - proper visualization in dashboards, is mandatory to avoid the accumulation of outdated and/or inaccurate information on one hand and provide operators with a centralized view on the other.

With the characteristics and high-level integration prerequisites mentioned above, a specific evaluation checklist, against which CTI platforms and feed providers are evaluated, could improve the selection of providers.

Internal sources

Specific threat data could originate from various points within the E-FCR itself too, as illustrated by the examples in Table 2 below. Although it is not always possible to fully automate the export/import process from all internal sources, the attack behaviour insight they bring should be considered valuable and is worth including into the CTI Platform.

Source	Examples
Network Data Sources	
Router, firewall, Wi-Fi, remote services (such as remote login or remote command execution), and Dynamic Host Configuration Protocol (DHCP) server log	Timestamp, source and destination IP, domain name, TCP/UDP port number, MAC address, hostname, action (deny/allow), status code, other protocol information
Diagnostic and monitoring tools (network ID/PS, packet capture & protocol analysis)	Timestamp, IP, port, and other protocol information, network flow data, packet payload, application-specific information, type of attack (e.g., SQL injection, buffer overflow), targeted vulnerability, attack status (success/fail/blocked)
Host Data Sources	
Operating system and application configuration settings, states, and logs	Bound and established network connection and port, process and thread, registry setting, configuration file entry, software version and patch level information, Hardware information, user and group, file attribute (e.g., name, hash value, permissions, timestamp, size), file access, system event (e.g., startup, shutdown, failures), command history
Endpoint protection products	Hostname, IP, MAC address, malware name and type (e.g., virus, hacking tool, spyware, remote access), file name, file location (i.e., path), file hash, action taken (e.g., quarantine, clean, rename, delete)
Other	
Security Information and Event Management (SIEM)	Log data of interest from all devices reporting to it.
Forensic toolkits, sandboxes, and honeypots	Malware samples, system artifacts (network, file system, memory, etc.)

Table 5: Internal Sources

Aggregating data

The potential of all the input gathered through the various security tools cited so far is realized only when there is a method for centralized storage and interaction with the information of interest, from a single screen within the E-FCR. In certain scenarios a SIEM solution could serve that purpose to some extent, but for the need of applied threat intelligence throughout the partner community, discussed here, a separate CTI Platform seems to be more appropriate. It completes the intelligence cycle by facilitating the data aggregation, analysis, and dissemination processes. Scalable ingestion and automation to normalize, correlate, enrich and qualify intelligence at large, are all challenging tasks and considerable computing resources would be required, but ultimately achievable with the right approach, including:

- Select data sources
- Identify, develop, and utilize vendor-specific APIs to suite the various data flow needs
- Create a scheduling plan for data injection to avoid system overload and exhaustion of API polling limits
- Enforce input verification to ensure well-structured IOCs are supplied for analysis and correlation
- Confirm deduplication to guarantee data uniqueness and limit overlaps
- Specify retention

Utilization of CTI in E-FCR

Consider the following exercise scenario, in which a Red Team (RT) conducts a staged attack against a simulated enterprise environment from the financial sector. The attacked infrastructure consists of an AD server, a couple of Web and file servers, an email server and some workstations running modern operating systems. For monitoring purposes within the used cyber range there is a SIEM, collecting security and access logs from servers and workstations, along with application logs from a business-critical web application; Web proxy, monitoring clients' internet activity; IDS, inspecting traffic for potential attack patterns. It starts with a highly targeted and very convincingly crafted phishing e-mail, sent to an identified accountant in the organization. The message states that a new version of a company-wide procedure has been added and there is a masked hyperlink in the message body, which leads to a malicious replica of the cloud document sharing enterprise portal. Upon submitting his domain credentials in the authentication form, the user is redirected to an attacker-controlled page, from where a drive-by download is initiated and a malicious backdoor executable is dropped and executed on the victim's machine. As a result, an internet connection to the C&C server is established and the attacker is now presented with various system details, including the running OS version, and is also able to access the victim's system with non-admin rights remotely. The OS details information allows the attacker to determine that the system is susceptible to a recently published privilege escalation vulnerability. After Exploiting the same, he now has system-level privileges on the machine and ensures persistence, scans the network for other hosts and services, compromises them, installs additional tools (e.g. credential harvesting software) and eventually obtains domain admin credentials, granting him administrative access to most domain-joined servers and workstations.

In this example the attack phases could be examined by correlating data from mail, web proxy, IDS and security/system logs from servers and workstations to uncover the specific actions and resources utilized, provide a timeline and context, and produce relevant threat intelligence data to help partners recognize similar techniques. Extracted indicators would include domains, addresses, urls, file hashes, process names and vulnerabilities. The newly acquired threat data is aggregated, normalized and shared with the community (using standards like STIX/TAXII, IODEF or OPENIOC), it may be presented via a dedicated portal/dashboard and/or feed service, as well as made directly available for import to security monitoring solutions as a pre-defined rule, based on the standards mentioned above.

3.4.3 Integration with E-EWS

The E-EWS and E-FCR are part of the vital technologies developed within the ECHO project. Both can exploit each other in order to maximize their capabilities and offerings to the users. The following are the three most relevant use cases the team analysed with regards to the integration of the two systems.

As a premise, it is important to remember that E-EWS is able to provide support at multiple levels: single organization (as incident management and ticketing management tool), multi organizations and national/EU level (tickets, early warnings, threat intelligence and knowledge base sharing platform, with advanced sectorial filtering and search capabilities). The Consortium envisions a single, EU-wide, E-EWS instance able to connect all nodes of the Network of Competence Centres on a multi-layered federation where each tenant can benefit from the notifications of early warnings and the gathered threat intelligence data. These possible deployment models can be leveraged by E-FCR: E-EWS could be installed on a single cyber range scenario (providing incident management capabilities, for example, for a training service), or it could be installed in multiple scenarios and in multiple cyber ranges and provide other types of services.

E-EWS used within a cyber range scenario

The E-FCR allows Providers to publish their services via its Marketplace. Some services will encompass more than one Providers, some others will be offered by a single Provider. E-EWS could be offered as a commodity provided by E-FCR and installed within the Providers' cyber ranges, to be part of their services (Cyber Coalition exercises in NATO often use MISP to share incident information within the exercises, for example). Since E-EWS provides incident management and sharing capabilities, a training service could use E-EWS as an incident management tool and knowledge base data source. Utilizing the E-EWS within a training service, for example, allows the trainer to follow what the trainees are doing, how they respond to the threats, step by step. E-EWS acts like a recorder of actions and helps members to collaborate and the trainer to evaluate the performance of the trainees. The data and incident reports that are produced from the service can be fed into the local E-EWS which will then make an analysis of this. This analysis can be used by the organizers of the exercise or training to maximize the impact of the exercise/training on the participants.

It is possible to envision two main deployment strategies:

1. E-EWS as a totally independent instance, local to the cyber range service.
2. E-EWS deployed locally in the service and acting as incident management tool, but connected to the main (EU-wide) E-EWS in order to download recent threat intelligence data and early warnings (this setup would probably be in read mode, with the EU-wide E-EWS instance operating as the data source)

In the following Figure 11, case 1 is modelled within a multi-range scenario:

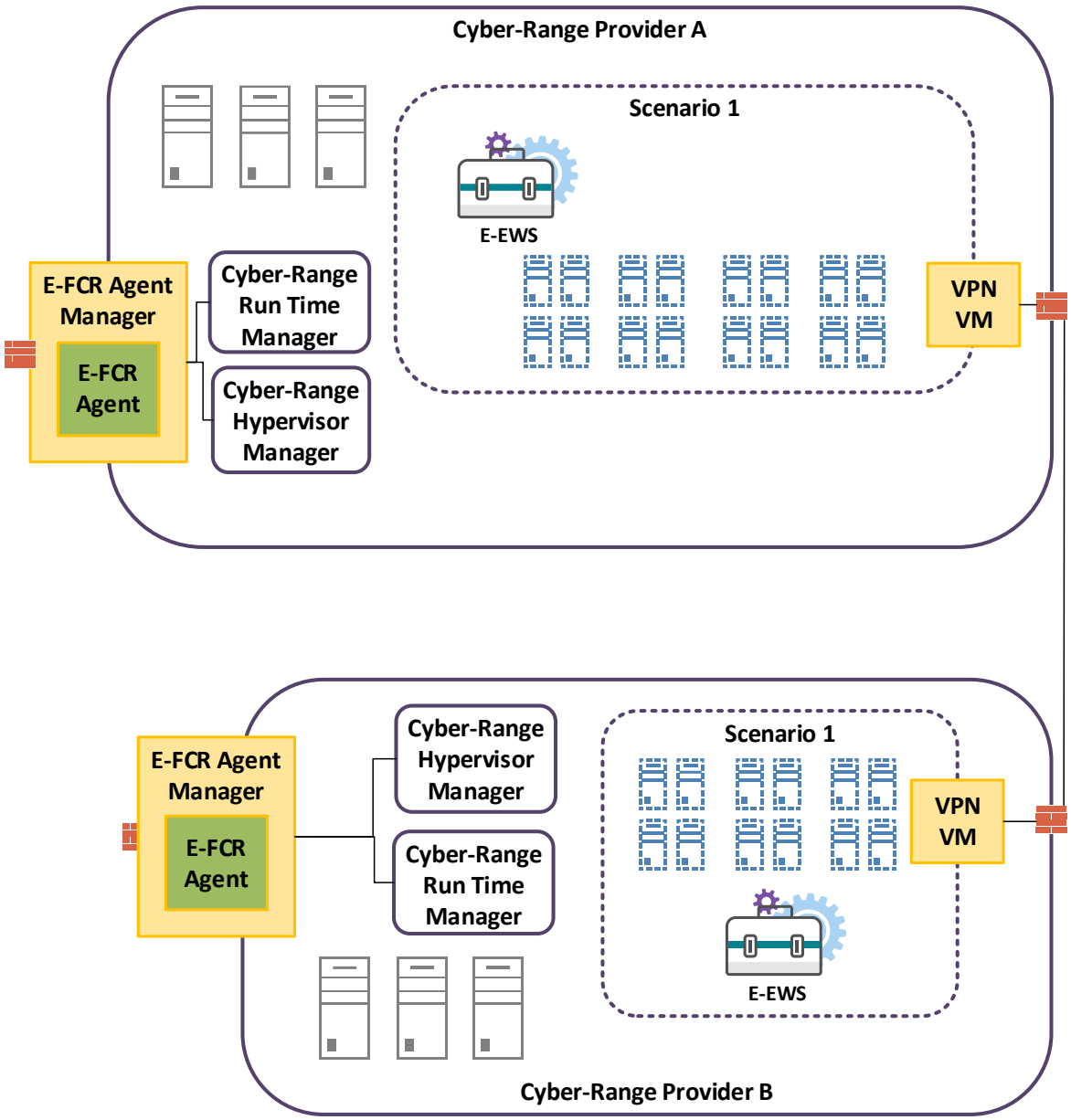

Figure 13: E-EWS used within a cyber range scenario encompassing two Providers

In the following Figure 12, the second deployment possibility is shown:

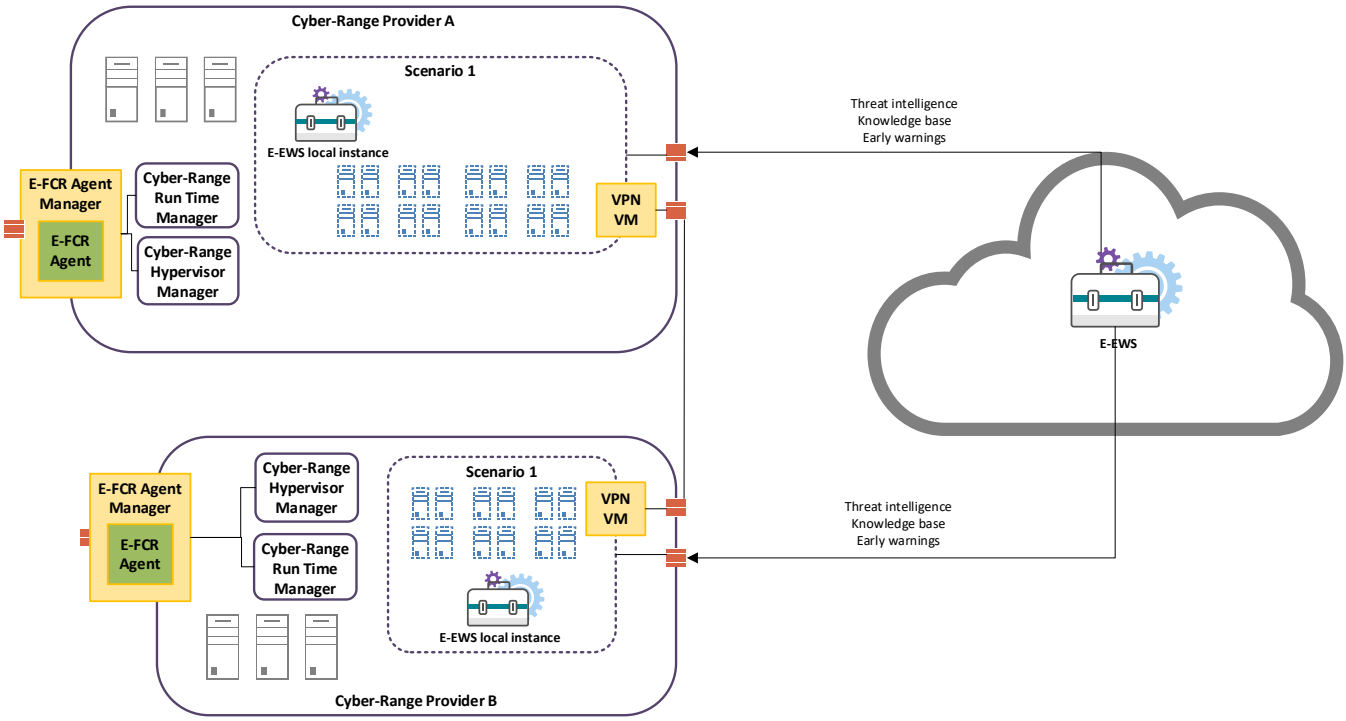

Figure 14: E-EWS used within a cyber range scenario encompassing two Providers (connected to main E-EWS)

In conclusion, this integration option allows cyber range providers to benefit from a modern incident management and threat intelligence tool. This possible integration of E-FCR and E-EWS is already technically possible as we write this text. E-EWS instances are fully virtualizable and a KVM-based VM with E-EWS is already complete and it has been tested on some ECHO cyber ranges. However, there is actually no direct support of E-EWS from a E-FCR perspective (as discussed earlier, E-EWS VMs already preconfigured could be offered as a downloadable commodity from the E-FCR Portal). As part of the future E-FCR roadmap, this integration with E-EWS should be taken into consideration.

E-EWS used to collect threat intelligence data from cyber range scenario

As mentioned in previous section, E-EWS can be source of CTI for E-FCR. One opportunity is to utilize the emulation environments as an possible source of threat and attack information. Within a controlled environment it is possible to train and experiment on a simpler way than using an operational environment or a physical testbed. While some E-EWS plugins act at the level of the main E-EWS instance (e.g. threat intelligence collection from multiple sources, knowledge base management and collection), others can be used to collect threat information and cyber exposure data from single organization's networks and share them with the EU-wide E-EWS instance as early warnings and knowledge base data. In general, E-EWS can be used with a local instance to register the progression of attack scenarios and provide intelligence data to the wider community. This may have a value for zero-days attacks research or for sector-specific threats.

Information related to the progression of attack scenario in cases of emulated context could be used by E-EWS for further analysis of attack scenarios progression and evolution, providing benefits to the whole E-EWS community.

Within this context, E-EWS could be offered from the E-FCR Portal as a commodity for cyber range providers: since they would provide a service to the E-EWS community with their collected data, it is possible to imagine that this would trigger a discount on the E-EWS subscription fees, for example (assuming there will be subscription fees – it is not clear yet at the time of writing this deliverable how the pricing model of E-EWS will be).

In the following Figure 13, a modelled representation of this possible integration.

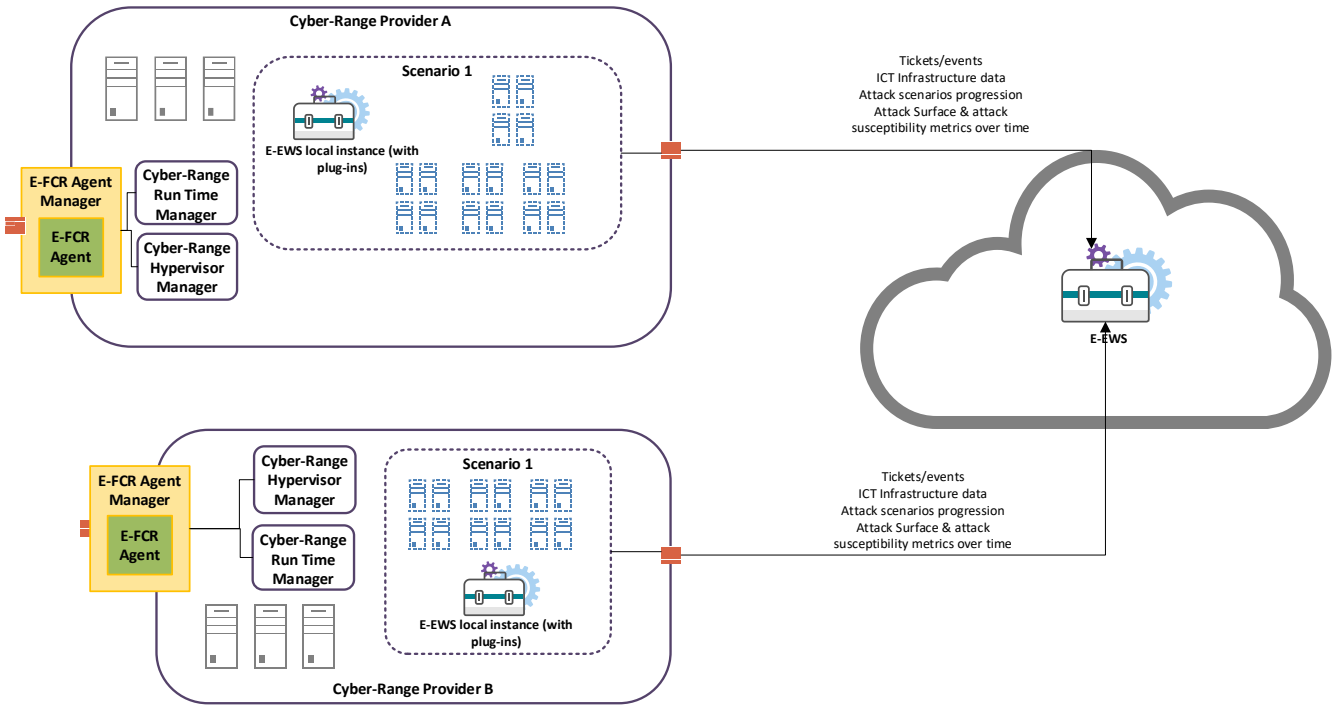

Figure 15: E-FCR cyber range services used as source of data for E-EWS

In conclusion, this integration would allow E-EWS to benefit from multiple scenarios hosted in cyber ranges part of the E-FCR. At the moment of writing this deliverable, E-EWS and its plugins could be easily installed in some ranges of the E-FCR. However, there is actually no direct support of E-EWS from an E-FCR perspective (as discussed earlier, E-EWS VMs already preconfigured could be offered as a downloadable commodity from the E-FCR Portal). In addition, E-EWS should be configured in order to use data from the Scenarios in the proper way (since these data are not 'operational' or related to real attack scenarios, they should be somehow highlighted and not merged with operational data). Such activities should not be extremely complex and could be considered as part of the future features of E-EWS and E-FCR.

E-EWS to share trends and threat intelligence data with Providers

E-EWS can also collect valuable data from many nodes of the federation and will constitute a main source of trends and threat intelligence data (e.g. top threats for a specific period or critical sector, zero-days attacks). These collections of information could be key on providing inputs to cyber range and content Providers while designing new cyber range services (e.g. trainings) to be published in the E-FCR. If the E-FCR core system will be connected to the E-EWS, Providers could leverage specific subscriptions on E-FCR to retrieve valuable, tailored suggestions for their content development activities, based on real world needs. This could be an additional service provided by E-FCR on regular bases or via periodic reporting.

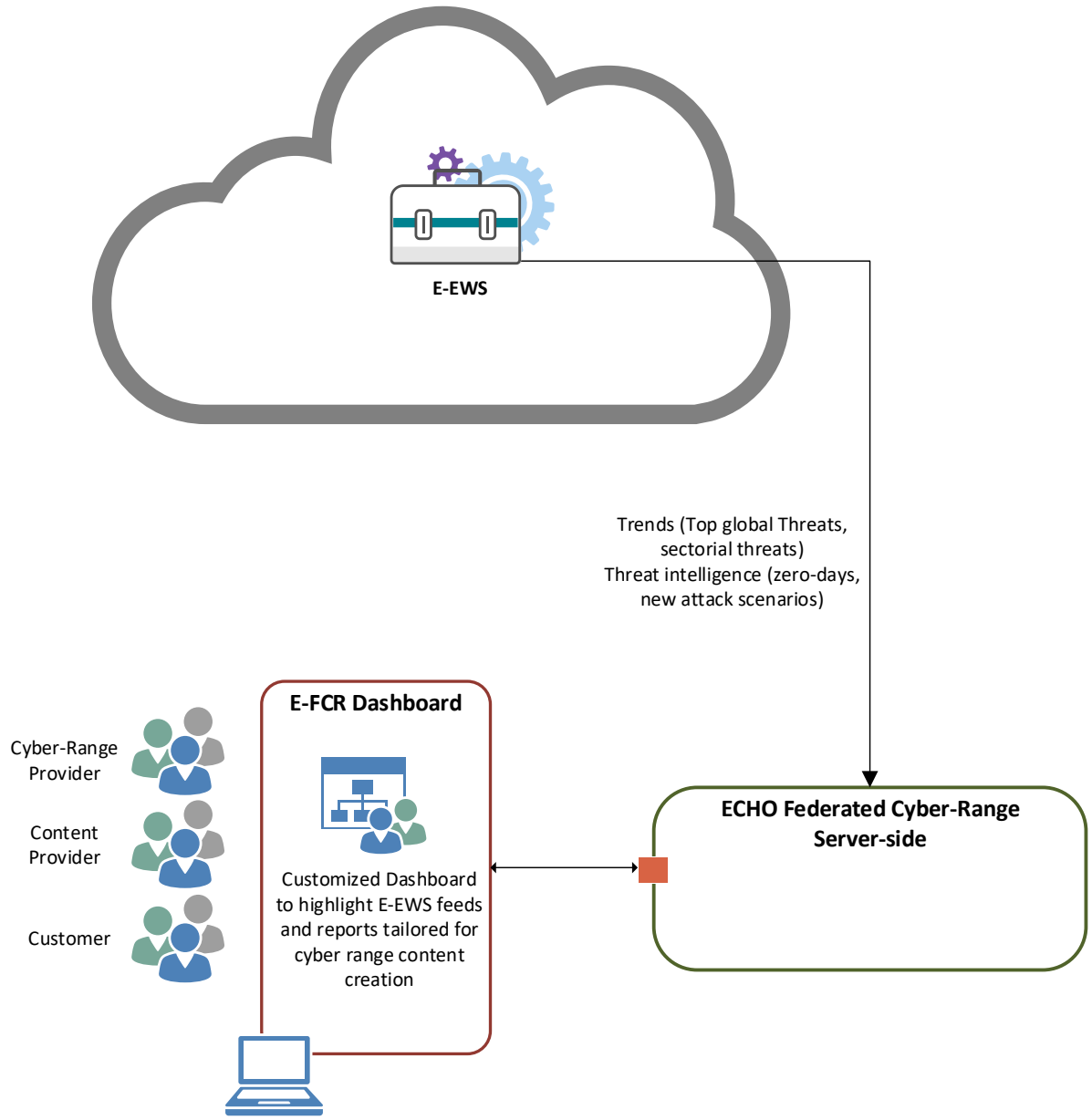

Figure 16: E-EWS as a source of input for E-FCR services creation and design

In this configuration, E-EWS would constitute a great source of input for the creation of new content (Service) to be published on the E-FCR by Providers: this would allow a mean for Providers to keep their services up to date with respect to the market needs. As part of the roadmap, ad-hoc queries and data analysis methods could be added in order to tailor the threat intelligence information into specific services (e.g. trainings), in order to help the Provider design new content based on the E-EWS feeds. This service could be supported by specific modules in the Service Designer of the E-FCR.

At the time of writing this deliverable, this feature is not fully supported and will require additional effort to be completely implemented this is highlighted as opportunity for future development that should be seriously taken into consideration once further resources are available to be invested in E-EWS and E-FCR.

Conclusions

As from the previous Sections, there are several use cases where a tighter interconnection between E-FCR and E-EWS would provide valuable additional functionality: other cases should be studied for the next iteration of this deliverable. Of course, these three activities could technically be combined as they are not mutually exclusive. Some of these alternatives are already partially covered, some other will require some further work to be implemented, but generally they represent features which the team could potentially add to E-EFR and its federated Services.

3.4.4 ECHO – Service Description Language - ESDL

The case for a Service Description Language

By design, the E-FCR centralizes the services offered by various providers in the Marketplace. Customers connected to the access portal have the possibility to explore the existing services and place a request selecting one of them. It may however occur that a customer does not find a specific service in the Marketplace and wishes to request a service of its own design. In such a case, the E-FCR should offer the customer the possibility to describe this particular service in the most precise way, before forwarding the request to the appropriate providers.

The complexity of this apparently simple task comes from the need to interface with both humans and machines. Faced with the task to translate the desired service in free text, customers might express essentially the same request in a very different way, with diverse words, depending on their experience, the level of their technical knowledge and their mastery of the English language. The extend and complexity of natural human language makes it challenging for machines to parse it: extracting precise information out of unstructured text, given the riches of the natural language, is notoriously difficult for computers.

A solution to that problem is to channel the expression of the request by laying constraints on the language. Restricting the syntax and vocabulary considerably eases the automatic processing of a text. The ECHO Service Description Language (ESDL) aims at finding a balance between the need to fix a textual framework that is manageable for machine, but still feels natural for humans.

ESDL must fulfil two main goals. The first one is its *ease of access*. It must function as a description language that is intuitive and effortless to read, even for customers with relatively limited technical knowledge. The second aim is to be provide the possibility to describe all aspects of a service, from a high-level description to technical constraints on machines. These two objectives set ESDL apart from other existing description languages that are usually either human- or machine-oriented, and typically focus on a subset of the scenario.

As such, ESDL will provide a *Lingua Franca* for the different actors and operations taking place around a service request.

How the ESDL will fit in the E-FCR

Customers logged into the E-FCR access portal will have the opportunity to either select an existing service from the Marketplace, or to design a specific service. The Marketplace will present the existing services in ESDL format.

On the other hand, the customers might prefer to design their own services. Given that one cannot be expected to master ESDL to script their service request, they will be presented with a wizard that will guide them through a series of questions, the answers to which will be constrained to be ESDL compliant. This wizard will invite the customers to provide a maximum of relevant information. The data collected will be processed by the Service Broker to select appropriate cyber range- and/or content-providers based on a matching exercise between the requirements of the service request and the Capacity/Capability Map.

An ESDL file will serve as common ground to discuss and negotiate the service between the parties and can evolve in the course of the negotiations. Given the versatility of ESDL, this file can be a central point to articulate different subsets of a service, as shown on the figure below. In the hypothetical use case presented, a service necessitates the collaboration of different partners. Once the customer has expressed its

requirements, the Service Broker selects the relevant providers: the request needs be fulfilled by the collaboration between a cyber range provider, a content provider and an executing agent. In this instance, a cyber range provider can supply an infrastructure, and describes the topology in the appropriate section of the ESDL file. Based on the topology description, a content provider can design a scenario, such as a penetration testing exercise for instance, to carry out on the cyber range, and outputs the formatted pentesting scenario to the file. Finally, an agent could consume the ESDL description of the scenario to automate its execution.

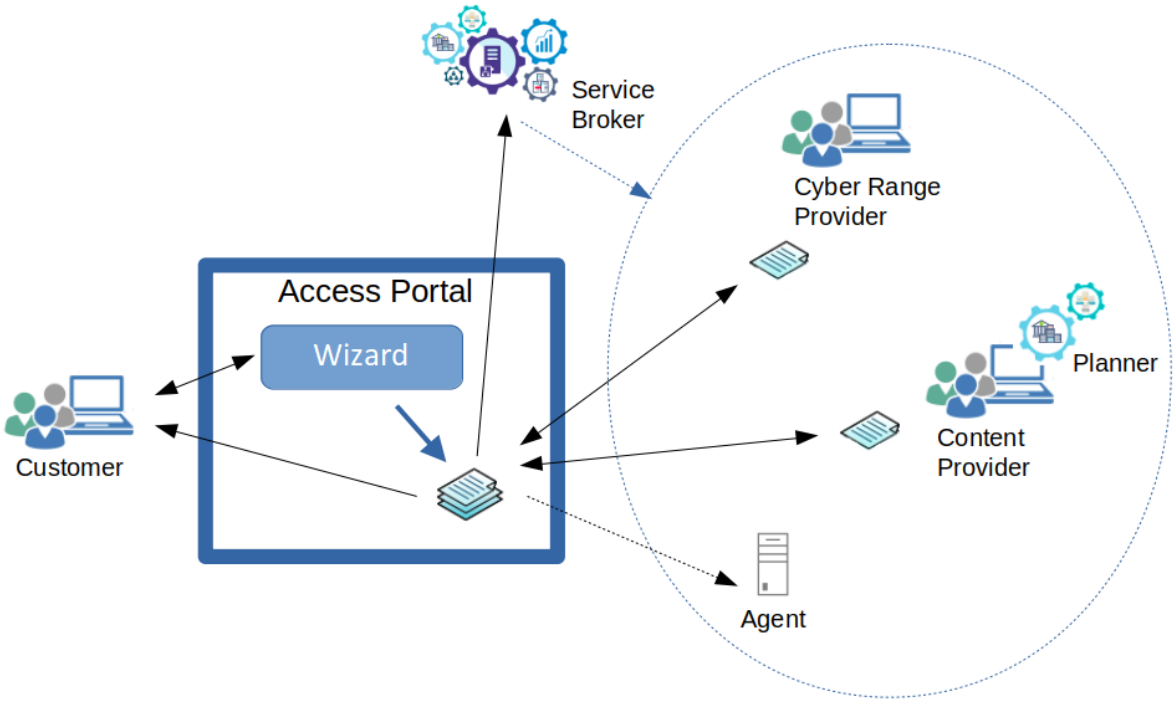

Figure 17: An instance of collaboration articulated around a ESDL service description file.

In conclusion, the ambition of ESDL is to function as a language that can be leveraged by the many actors around the service request, while remaining easily readable by machines and humans.

Roadmap for ESDL

Domain-specific languages are defined by their grammar. It has been decided that the ESDL grammar will be written in xtext, a prevalent software framework for developing programming languages and domain-specific languages, available at https://www.eclipse.org/Xtext/.

To meet its objectives, ESDL will evolve in time. While a first version of ESDL will be designed to answer the immediate needs of the present E-FCR functionalities, new grammatical rules might be developed to sustain additional emerging services, or possibly requirement shifts. It should hence be understood that ESDL will keep improving and expanding over time.

ESDL will be designed to remain as general as possible, to be useful outside the E-FCR implementation. Eventually, ESDL will the provided to the public, open-source, under the license CC BY-SA[42].

3.4.5 Adversary Simulation Capabilities

Introduction

Adversary simulation does not have yet a concrete definition. The term can be related to the cybersecurity sector, where it is defined as threat simulation or purple teaming[81]. One basic goal of adversary simulation is to prepare the IT staff of an organization for highly sophisticated zero-day attacks that may occur in the future. Specifically, adversary simulation focuses on the customer's ability to deal with an attack, post-compromise[80].

Exercise Simulations and more advanced Adversarial Attack Simulation Exercises (AASE) is a key feature the E-FCR provides to its users and customers for training purposes. However, due to the volatility of the Cyber Security landscape and the general technological advancement, threats, adversaries and their capabilities (knowledge & tools) evolve respectively in a rapid manner. Therefore, in conjunction with threat intelligence - on emerging real-world cyber threats - from the E-EWS, the E-FCR should be able to provide its users with effective adversary capability modeling.

More specifically, the E-FCR can be used by content providers for cyber threat exercises and scenarios according to the needs of a customer's organization. First of all, the content provider will investigate the customer's organization systems about vulnerabilities, zero-day exploits, unpatched services etc. This can be done by using a team of penetration testers.

Secondly, after the investigation of the customer's systems, the content provider can use E-FCR for the deployment of cyber exercise scenarios dedicated to the systems and vulnerabilities of the current customer. The scenarios can be created by using the reports of the penetration testing team. The content provider will create Virtual Machines with the identified vulnerabilities, then by using the E-FCR functionalities the provider will create the network of the customer's organization systems.

Lastly, the content provider will perform the cyber exercise created on E-FCR in order to teach the customer's IT staff how to handle the identified vulnerabilities, how the penetration testing team found these vulnerabilities and exploits contributing in this way to the education of the IT staff.

Alternatively, the content provider could create scenarios hosted in E-FCR platform to test the information sharing processes and the incident response of a customer organization.

Furthermore, in Red Team Exercises, adversary profiles and their capabilities should be able to change, reflect the threat model and needs of each user or customer of the E-FCR, and the most likely attack vectors for the simulated infrastructure. Adversary profiles according to their technique sophistication and their capabilities may include:

- **Individual hackers – simple penetration testing**
  Individual hackers not associated with any organization can launch attacks independently. Even if they are equipped with very basic expertise and technological means, they may participate in operations through the use of automated attack tools available on the web.

- **Insider threats**
  An insider threat is a security risk from within the targeted organisation. This does not necessarily mean that the actor has to be an existing employee or officer in the organization. They could be consultants, former staff members, business partners or board members.

- **Organized groups - Industrial espionage**
  The most common type of industrial espionage is the active gathering of intelligence about a company or organization. Such a practice is intended to create a competitive advantage for a party that has the information.

- **Hacktivists - Terrorist groups**
  There are numerous hacktivist groups, with different social and political agendas. The hacktivists commit various types of cybercrimes including website defacements, website redirects, DoS attacks, malware distribution, data theft and disclosure, and sabotage.

- **Nation state sponsored groups**
  There are no global standards to determine “adversary nations”, but APT hacking groups are widely considered to be sponsored by nation states. Looking at a list of these APTs, we see Russia and North Korea making headlines and Iran expanding their capabilities over the last few years. Vietnam seems to be focusing on local geopolitical objectives while China has the most APT groups, all focusing on industrial spying.

That being said, as new exploits, public CVEs, tools and frameworks emerge and as resources become more readily available (i.e. computer hardware, AI technologies), the aforementioned adversary profiles will have to shift in order to mirror the ongoing real-world attackers and the evolution of tactics, techniques and procedures (TTPs) they use.

Additionally, to better outline the adversary capabilities and stay up to date with ongoing trends, globally-accessible knowledge bases of adversary TTPs, like MITRE ATT&CK, can be used in a supporting role to the already available threat intelligence provided by the E-EWS.

Cyber Defence Exercise using E-FCR

An adversary simulation exercise is in general a cooperative exercise. The adversary coordinator executes the attack scenario with the blue team present.

Specifically, the adversary coordinator could be the content provider and the blue team could be the IT staff of a customer company. The adversary coordinator can use the E-FCR platform to host cyber exercises according to the needs of the customer.

During the exercise the coordinator discusses with the blue team the executed actions, presents their attack vectors and techniques. In this point of view the blue team gain experience in attacks and learn ways to compromise serious zero day attacks. Furthermore, the coordinator usually executes the same exercise with a few changes and new indicators to measure the improvement of the blue team. Adversary simulation in a few words is discussion and cooperation between a red team and a blue team, thus it can be seen like the middle discussion layer of these two teams.

The main difference of an adversary coordinator and a penetration tester is that the adversary coordinator will simulate the exercise and help the blue team customers to understand and improve their skills against a specific attack vector. Contrariwise, the penetration tester will perform penetration testing on the systems of the customer and will give to the coordinator-service provider an analytical report with the identified risks, exploits and vulnerabilities. The content provider will then create cyber incident scenarios based on the reports identified by the penetration testers. Furthermore, the content provider will use the E-FCR platform to host the created cybersecurity scenarios.

Therefore, the E-FCR platform will host cyber exercises or CTF-like procedures created by the coordinator-service provider based on the needs of the blue team-IT staff. The E-FCR platform will give the IT staff the opportunity to understand, handle and mitigate risks on their company's systems in an interacting, gamified way. Moreover, the IT staff could even perform attacks on their virtual system hosted in E-FCR with the help of the coordinator-service provider in order to deeply understand the problems.

Adversary simulation capabilities can go beyond the limitations of traditional penetration testing. Below an example scenario is presented of a phishing attack accomplished by the supervision of the coordinator hosted in E-FCR platform.

Phishing attack scenario – Stages of an attack scenario

The penetration testing team of the service provider found unusual spamming emails on the systems of the customer's organization. The mission of the service provider is to inform and teach the IT staff of the customer organization on how to mitigate and handle these suspicious emails. The service provider will use the E-FCR platform to create a virtual test environment based on the systems of the organization. Furthermore, the service will create Virtual machines based on technical characteristics of the customer's computers, then the service provider will create a cyber scenario based on the vulnerability of the unusual spamming emails. The scenario will be deployed on the E-FCR platform.

1. (Introduction) The coordinator will provide to the staff the high-level description of the phishing mechanism along with the social engineering techniques that could be used from a potential attacker. Also, some very basic (or famous) examples of past phishing attacks can be used here. Furthermore, the coordinator will present to the IT staff cyber exercise scenario hosted in E-FCR and will explain its network characteristics.

(Reconnaissance) The coordinator will then inform the organization about the technical characteristics of some famous phishing attacks, addressed in the previous stage. Also, the coordinator then will explain the reconnaissance tools and techniques used by the penetration testing team. The E-FCR platform and the cyber exercise example will provide to the IT staff an actual gamified environment for testing. This will help the IT staff to deeply understand the problem of the phishing email and how it bypassed their systems.

2. (Techniques) The coordinator at this stage presents to the staff the most common phishing techniques:
    - domain typosquatting
    - domain shadowing
    - maliciously registered domains
    - domain spoofing
    - subdomain services

    Some of these techniques can be hosted in Virtual machines inside the cyber exercise hosted in E-FCR. The coordinator will try to give a global view of the phishing attack problem, by introducing the phishing techniques on the IT staff. Further, the IT staff will learn how to mitigate and handle the risks of the phishing attacks in the secure testing environment of E-FCR platform.

3. (Mitigation) The coordinator will provide to the staff ways to prevent or mitigate a phishing attack. The education of the staff can be achieved by actually attacking the systems of their organization hosted in the virtual environment of the E-FCR platform. Specifically, the coordinator will present to them some mitigation actions inside the exercise like:
    - The use of the three basic security standards of spamming (SPF, DMARC, DKIM)

- Implementation of an anomaly detection system
- Disabling automatic execution of code
- Use of security email anti-spam, anti-malware scripts

In the same way as the previous example a coordinator can create similar adversary simulation scenarios in various topics based on the needs of each organization, hosted in the E-FCR platform. The other key topics for the security of an organization are:

- Threat modelling for key assets
- Evaluation and coordination in their defensive technologies, antivirus, anti-malware protection, network monitoring protection (IDS)
- Review and evaluate security procedures and manual of the organization
- IT staff security awareness

Conclusion

In that frame, this user story shows possible developments for the adversary simulation capabilities with the functionalities of training in advanced realistic scenarios, discovering indicators of compromises(IOCs), and preparing effective defenses. More specifically, a service provider, using the E-FCR platform, can demonstrate a sequence of training sessions according to the needs of a customer organization.

3.5 E-FCR Exploitation

3.5.1 Certification of products

Introduction to the problem

Cyber capabilities can be defined as the resources and assets available to a state to resist or project influence through cyberspace. Nation states are investing in technologies, methods and processes to develop those cyber capabilities and cyber ranges are being looked at as the equivalent of a traditional firing range. On March 12, 2019 members of the European Parliament approved the **Cybersecurity Act**. It establishes an EU-wide certification scheme for products, processes and services to guarantee they meet **common minimal EU cybersecurity requirements**.

The Cybersecurity Act is a fundamental component of the EU Cyber Strategy and complements the European Directive on Network and Information Security (NIS). The introduction of the Cybersecurity Act will permit the development of European Certification Standards and accreditation schemes for *evaluators* and *certification labs*. A wide adoption of certifications will also reduce the cost of tests, also enabling an economy of scale. For this reason, it is important to start developing product and service certification as soon as possible, making them mandatory for selected critical devices/products.

ENISA has released the EUCC candidate Cybersecurity Certification Scheme. The EUCC scheme is based on the CC and the CEM, with an additional set of supporting elements. Certification under EUCC scheme shall cover the assurance levels ‘substantial’ and ‘high’ of the CSA Group (formerly the Canadian Standards Association) for Information Communication Technology (ICT) Products. Under this scheme, both certification bodies (CBs) and testing laboratories (ITSEFs) shall be assessed for authorisations to perform certification and evaluation at the assurance level ‘high’ of the CSA.

According Art. 2 of CSA, an ICT product is an element or a group of elements of a network or information system.

How to use Cyber Ranges to certificate products

Cyber ranges can be used by a wide range of target users: Corporates (private and government), Strategic decision makers (private and government), Security professionals, Military agencies and CNOs, Security Operations Centres (SOCs), Educators, Students, Researchers, Event organisers.

With respect to certification, the following application area are envisaged.

Conformity Assessment

A customer would certify ICT products. Along with security research, security testing is the most traditional use case of cyber ranges where system and application simulations are tested and security attacks are carried out against them, in a controlled way, to identify potential vulnerabilities before deployment and use. A cyber range can set up security testing according the Certification Scheme to assess the ICT product conformity. Certified Protection Profiles may be defined as applicable or reference standards for specific stakeholders’ communities.

Certificates issued shall indicate which version/release of the CC and CEM have been used for the evaluation and certification.

Article 54 1. A European cybersecurity certification scheme shall include at least the following elements: “g) the specific evaluation criteria and methods to be used, including types of evaluation, in order to demonstrate that the security objectives referred to in Article 51 are achieved.”

In addition to some specific provisions of this scheme, the CC strongly contribute to meet the security objectives defined by Article 51 of the CSA. This shall occur through the selection of relevant components within the following classes/families of the catalogue of Security Functional Requirements (SFRs) and Security Assurance Requirements (SARs).

The assignment to the assurance levels of the CSA shall be based on the use of the assurance components for vulnerability assessment defined in CC Part 3 as follows:

- AVA_VAN.1 and AVA_VAN.2 map to the assurance level ‘substantial’ of the CSA;
- AVA_VAN.3 to AVA_VAN.5 map to the assurance level ‘high’ of the CSA.

ENISA may provide guidance as how to select the proper assurance level based on risk assessment. The assurance level shall be commensurate with the level of the risk associated with the intended use of the ICT product, ICT service or ICT process, in terms of the probability and impact of an incident.

Building competence and security education

Cyber Range can contribute to the increase of the number of cybersecurity experts and the cybersecurity skills: without this valuable additional element, the risk is that the Cybersecurity Act will become a good piece of legislation with limited or no impact. Cyber Range simulation-based exercises are aligned with the main roles of cybersecurity experts in critical sectors:

- Cyber Defense analyst
- Cyber Defense forensic analyst
- Cyber Defense responder
- OT Operator
- Exploitation analyst
- Infrastructure support specialist
- Vulnerability assessment analyst

Cyber Range can provide a convenient and more cost-effective way of delivering hands-on training within an organization, as well as the associated training assessment and people certification. One of the recurring complaints from industry is the lack of hands-on experience by young graduates: universities around the world have begun looking at cyber ranges as a means of improving teaching and learning. Cyber ranges can be used to organise large scale cyber exercises involving hundreds to thousands of people.

Meanwhile the second application can be exploited without bonds, the Conformity Assessment is possible only if the Cyber Ranges are accredited environment labs. Indeed, a testing laboratory (ITSEF) including its staff performing evaluations for a certification body, be it the internal testing laboratory of a conformity assessment body or an external testing laboratory in the case where testing is performed by a subcontractor, shall be technically competent for the related tasks.

For the assurance level ‘substantial’, this technical competence shall be assessed through the accreditation of the testing laboratory according to ISO/IEC 17025.

These Information Technology Security Evaluation Facilities (ITSEFs) and their concerned personnel shall be required to meet the following requirements:

a) to have the necessary expertise and experience in performing the specific testing activities to determine the product's resistance against specific attacks (penetration testing) assuming an attack potential of 'basic-enhanced' as described in the CC (AVA_VAN.3 Focused Vulnerability Analysis).
b) For the technical domains defined in Chapter 8, SPECIFIC EVALUATION CRITERIA AND METHODS:
    - to have the necessary expertise and experience in performing the specific testing activities necessary to methodically determine the resistance of the product against attacks exercised in the product's operational environment assuming the corresponding attack potential of either 'moderate' or 'high' as described in the CC (AVA_VAN.4 Methodical vulnerability analysis, AVA_VAN.5 Advanced methodical vulnerability analysis);
    - to be able to demonstrate the specific following technical competences
        - ✓ for the 'Smart Cards and similar devices' Technical Domain, required capabilities from Annex 8 EUCC Scheme;
        - ✓ CAB, including their testing laboratories, are subject to specific requirements in addition to their accreditation for the 'high' assurance level of certification.
        - ✓ for the 'Hardware Devices with Security Boxes' Technical Domain, required capabilities from Annex 10 EUCC Scheme.

Further guidance for accredited certification laboratories is provided in EUCC Scheme.

Added value by a Federation of Cyber Ranges

The Federation brings together multiple cyber ranges in a way to increase or improve simulation capabilities as well as other capabilities beyond what can be offered by a single cyber range. The concept of federation is based on the assumption that it is highly complex and costly for a specific cyber range to be able to provide all the required capabilities and functionalities and it is therefore conceivable that multiple cyber ranges federate, each with its area of specialisation. It is important to note here that **federation does not imply integration**, which instead requires that two or more cyber ranges must be able to communicate with one another in order to deliver a scenario.

A federation of cyber ranges can offer users a one-stop shop for all their needs and requirements: it acts as a broker between the customer and the cyber range providers by giving the following added value:

- Overview of available cyber ranges suitable for formal certification to be used as certifications labs;
- Regularly check and update of new certification labs within the Federation or past accredited cyber ranges not renewing the accreditation;
- Hybrid environment (multi sector or physical components simulation) in order to work together to offer end users the ability to achieve multiple use cases and different types of scenarios;

- Complex scenarios of *critical sector*[1] customers, for which a traditional vulnerability assessment and penetration testing are not enough;

In this sense, the FCR can provide a customer interface in which they can fill put different forms for:

- Security Training;
- Environment Simulation for certification;

According to the information provided, an overview of possible services is presented by specifying the type of ranges, the number of ranges, the accreditation or not for certification tests, the envisaged time to accomplish the service, the envisaged time of availability to carry out the service and the proposed pricing an discounts.

A FCR cannot be a Certification Body (CB) but the ranges belonging to the federation can be used by CB to perform the conformity assessment if they are accredited certification labs.

3.5.2 Evolution of learning methodologies

This section considers the methodological perspectives of the Knowledge and Skills Development function of the ECHO-Federated Cyber Range. They align the customers' expectations and the service providers' capabilities to meet those expectations. The consideration of the evolution of learning methodologies in terms of selection of the most appropriate traditional and innovative techniques, tools and models is fundamental for the decision and definition of the learning design (the technologies that should be used for the deployment of the cyber range scenario) and assessment methods (the solution of the customer needs).

Evolution of learning methodologies theoretical overview

This chapter aims to briefly overview the evolution of learning methodologies leading to the necessity of a structured approach to bridge the gap between theory and practice in the cybersecurity domain. Furthermore, within this summary we attempt at establishing a common understanding of the profound impact technology has had on learning methodologies and the importance of developing novel learning practices and instruments where simulations, playgrounds, polygons, and hands-on cybersecurity theoretical instruction meet, to respond to the cybersecurity industry needs for knowledge, competences, skills, and abilities to perform various tasks, including to respond to sector-specific cybersecurity challenges.

The traditional meaning behind the term ***learning methodology*** considers the combination of tools, principles, and methods used by educators to enable learners to acquire, retain, and recall knowledge and skills, or in a more contemporary outtake on this, to enable learners to improve or maintain qualification and proficiency. The technology and theory behind the topic of instructional design, however, have undergone somewhat of a paradigm shift, especially with the increasingly digitized society and the implications this rapid technological development has in terms of education, and knowledge, skills, and abilities demand. Yet, three fundamental

[1] Critical sector customer are Operators of Essential Services (OES) and Digital Service Providers (DSP).

DSP are: Online marketplace, Online search engine, Cloud computing service. OES are: Energy, Electricity, Oil, Gas, Transport, Air, Rail, Water, Road, Banking, Financial market, Health sector, Drinking water supply

schools of instructional design, often labeled generically **behaviourism, cognitivism, and constructionism**, and their underlying approaches, provide the contemporary outlook on educational psychology.

For the sake of simplicity, we could generalize these learning theories as follows.

- Behaviourism reflects on learning as a function related to conditioning, thus advocate a system of rewards and targets in education[60]. The process of acquiring, retaining, and recollecting new knowledge, skills, and abilities are regarded as a behavioural change, thus a process, requiring an extensive environmental orchestration to achieve the desired learning outcomes.
- Cognitive theory undertakes the approach of interrelation between existing knowledge and new knowledge acquisition, through the lens of believing that the learner, rather than the environment, is the enabler of the learning process. Particular attention is paid to memory and how we retain and recollect knowledge.
- Constructivism steps on the previously explored learning theories, combining them and further expanding them with methods to actively involve learners in the process of constructing meaning for themselves, which could be achieved through social, environmental and situational orchestration, linked with pre-existing knowledge, interests, emotional and experience-based fragments. Constructivists stand behind the idea that the acquisition of knowledge is a personal and intimate process, requiring exploration of meaning through an individually tailored methodology.

Without exploring them in detail, we could generalize that the differences between these three pillars of learning theory lay mainly within the specifics of the interaction between the participants in the process and their experience, along with the cognitive, emotional, and environmental stimuli through the orchestration of the learning dynamic [57], [58].

Against this backdrop, there is a recognized need within the cybersecurity domain of knowledge and practice for a more hands-on approach to training the cybersecurity professionals and in terms of maintaining and expanding upon their proficiency levels in their field of practice. Slightly more than 60% of companies, consider their cybersecurity applicants to lack the practical qualifications related to their responsibilities [59]. By the same report, survey respondents by far look to specific training opportunities (76 percent) and professional development organizations (71 percent) to build knowledge, skills, and abilities (KSAs), rather than security certifications, where “organizations can also employ more sophisticated and continuous training, such as “just-in-time” online training, and focus on the practical development of specific skills and map these into training plans for overall career path development”.

Furthermore, within the context of the ever-changing landscape of IT and cybersecurity respectively, there is a recognized need for a flexible, problem-based learning environment, which could support the constantly evolving cybersecurity curriculum for different domains of practice. This need is likewise recognized by the ECHO consortium with the structured approach towards creating the ECHO Federated Cyber Range.

Within the context of the roadmap for the development of the federation of cyber ranges, the most natural choice of learning methodology, through the lens of which further development and improvement will occur, the consortium considers being **simulation-based learning**. Simulation-based learning is a learning model categorized under the constructivist theory. It is a well-known approach in medical, engineering, and military studies to the instruction of domain-specific skills, knowledge, and abilities, along with 21st-century skills, such as problem-solving skills, creativity, and critical thinking.

Through simulation-based learning, the participants are offered a controlled, fault-tolerant, and scaffolding-offering environment, where learners could practice pre-existing knowledge and concepts along with learning the implementation of its principles and the consequences of their application.

As the focus of the present deliverable is to highlight the roadmap for the development of the federation of cyber ranges, the authors employ here only the constructivist methodologies related to computer-supported learning of adaptive learning, and how the technology can support the interaction between the participants or so-called collaborative learning[2]. The Computer-Supported Collaborative Learning (CSCL) is a learning model focused on collaboration among learners through the use of connected computers and the software built on them. CSCL understands learning under a constructivist foundation, where one learns by doing and social engagement. It doesn't include cognitive symmetry between learner and trainer but the trainer or learning provider should design the system and tools thus facilitating the interaction between the learners and the co-creation of meaning. The different stages and models for the design of the learning syllabus will be described in the next section. The main objective of this section is to describe the evolution of the specific methods that correspond with the cyber range concept.

Features of the cyber ranges

The simulations and exercises on cyber events are the methodological aggregations of interconnected technologies (applications, devices, systems) and orchestration of the human-human and human-machine interactions. The cyber ranges are considered as one of the most effective instruments for improving the learning experience, increasing the cognitive processes, and acquiring new skills, mindset, and from the perspective of learners. The cyber ranges, as a tangible representation of this methodological aggregation behaviour, contribute significantly for filling the gaps in cyber professionalism allowing to young graduates and junior professionals faster and effective gaining hands-on experience who, otherwise, following the traditional path for acquiring practical skills, should spend years working in different settings and facilities.

The defined by the National Initiative for Cybersecurity Education (NICE) of the United States National Institute of Standards and Technologies reveals that together with research and test functions of the cyber range as an instrument, most of the features are dedicated to education and training purposes. According to NICE, a cyber range should provide the stakeholders with:

- An environment where new ideas can be tested safely and teams and work to solve complex cyber problems
- Performance-based learning and assessment
- A simulated environment where teams can work together to improve teamwork and team capabilities
- Real-time feedback
- Simulate on-the-job experience

The development and deployment of simulations of real-world cyber events for the training and professional development of cybersecurity professionals are relatively new phenomena, as well as the adolescent age of the cybersecurity domain [52]. As was mentioned before, in the evolution from the instructional to adaptive learning methodologies the main problem is how to create a suitable environment that encourages effectively the communication and interaction between the participants in the exercise in the "classroom" (physical or virtual) with the ultimate goal to apply a pre-existing knowledge in real-world scenarios. Likewise, the provision

[2] One of the 3 instances of digital learning (e-learning, Computer-Supported Collaborative

Learning (CSCL) and online learning) defined by [61]

of a fault-tolerant system, where mistakes and curiosity, supports the learner's free exploration within a given framework or structure.

Many different tools and methods can be explored and tested. We can classify them provisionally as methods or techniques taken by other disciplines and tools in terms of models used only in the creation of cyber ranges. Below, only the methodological tools exploited in other disciplines will be concerned. The technological models will be considered in the design section.

Evolution of the methods for knowledge creation

If the paradigm of Jonh Moravec about the knowledge production in higher education could be adapted to the context of cyber ranges, its evolution would follow three overall stages:

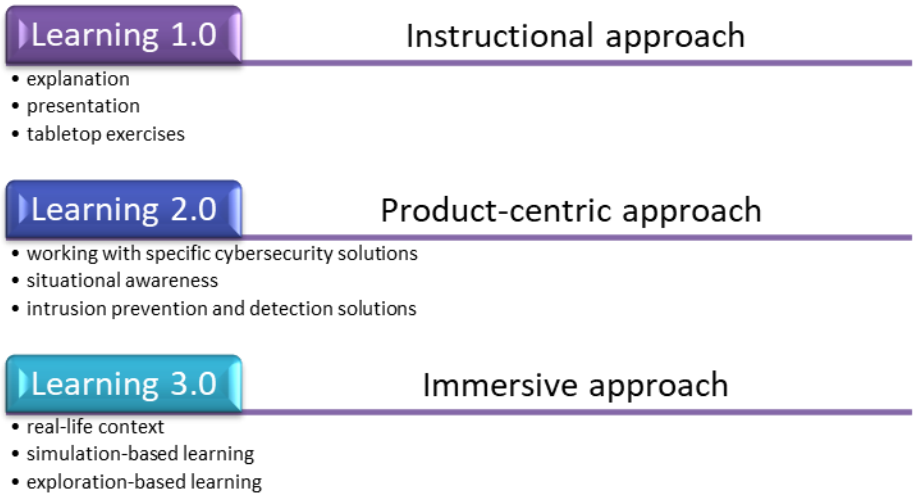

Figure 18: Jonh Moravec's knowledge production in higher education paradigm, image adapted by the ECHO Consortium, developed by ESI CEE

The concept for Learning 3.0 involves the use of a variety of tools (interconnected technology assets) and methods (performance measurement, learning orchestration, teamwork measurement, alignment of scenarios with competences and job roles) that implement many different tasks and reflect the nature of the cyber range: reusable, repeatable, scalable, adaptable (configurable) to different needs and objectives. When those tools, techniques, and functionalities are consolidated under the concept of one-stop-shop or the federation of many interconnected and interoperable services the cyber range providers are enabled to act in one common environment with shared abilities to analyze, distribute and sell services. Thus, the price of the development processes decreases and the designers are allowed to focus on the content aspect rather than on distribution and marketing.

Future development within ECHO-Federated Cyber Range

According to the general and specific goals of each cyber range scenario, the educator is empowered to select the most appropriate method for its deployment. Educators are also able to justify the exploitation of the methods with the additional benefits that may be brought by any specific technique or a combination of methods. For example, in an exercise, related to cybersecurity process management or strategy design, the educators can choose the project-based learning method or combination between project-based and design thinking techniques based on the assumption that those methods support best the demonstration of teamwork

skills and enable to better identify gaps in the team constitution. Therefore, as part of the preparation process of the service catalogue within the Federated Cyber Range, it is necessary to include the description of the most commonly employed methodologies, the benefits, advantages and disadvantages related to their employment, as well as the types of tasks that present a relevant combination in order to reach the desired learning outcomes.

The most exploited learning method envisaged in the design of the cyber range is ***gamification***. Used in many business, educational, academic and government settings, this term refers to "the use of game design elements in non-game contexts". The method of gamification involves principles from design thinking, strategic thinking, and creative thinking – thinking out-of-the-box, dealing with uncertainty, predictive thinking, holistic approaches, heuristic techniques, etc. Machine learning and artificial intelligence technologies can be applied as part of the gamification toolbox to provide a better adaptiveness to the learner's needs, skills and abilities, to assist in the processing of big data sets for analytical tasks, and ensue the gradual complexity of the tasks in alignment to the proficiency and qualification level of the participants. The gamification method further improves the collaboration between the participants by encouraging their involvement in the process and their commitment to the team's performance and success and it case of E-FCR is explored in UserStory 3.1.1 Gamification of Cyber Ranges

3.5.3 Evolution of assessment methodologies

The Assessment Methodology is a tool to support the customer and learner to obtain a better understanding of whether the expectations and needs of an educational activity are met and whether the learning objectives are achieved. This section concerns mainly the evaluation of the cyber range exercises within a training pathway and their implementation. It does not cover an assessment of the skills and knowledge gained during the entire training or education program.

The evolution of assessment methodologies involves customized approaches and connects the specific evaluation criteria with the abstract learning objectives aimed at receiving thorough insights about the progress of learners rather than simple pass-fail results. Thus, having an accurate assessment will empower customers, end-users, and training service providers alike, to better plan and obtain a relevant prognosis regarding requirements and needs for future training and professional development initiatives. Likewise, a robust assessment methodology serves the E-FCR content providers in the instructional design process to create better and more fit-for-purpose training activities and practical tasks. The assessment methodology is an integral part of the learning methodology and training service design processes, consequently posing a necessity for those processes to be considered in parallel, in alignment with the learning objectives, and continuously communicated with the end-users of the training activities.

Due to the collective and collaborative nature of the cyber range exercises, the learners' achievements can be measured individually and collectively, or provide assessment strategies for considering both individual and collaborative performance. Designing assessment methodologies for individual performance and satisfaction is less challenging based on the premise that the designer has to consider and compare a smaller amount of data and information about the profile of the learner and the feasible goals. Within the assessment process of collaborative performance, more complex assumptions about the possible synergies and implementation of the deployed scenario are involved. Correlations between experimental data can be made but they would usually require a larger amount of data obtained from the exercises, simulations, or instructional activities, as well as the deployment of rather more complex scenarios.

Undoubtedly, structuring a more robust, comprehensive, and multi-faceted assessment paradigm is imperative for the continuous improvement and sustainability of the educational content, as well as of paramount value for the content providers and customers alike. However, to achieve this robustness of assessment is easier said than done. Therefore, a vastly adopted approach is adopting standardized assessment methodologies

with easily customizable parameters, and continuously build and integrate assessment items based on them. Those assessment methodologies present an amalgam of customer feedback, self-assessment tools, activity evaluation paradigms, as well as input from the implementation and maintenance teams, to ensure a common understanding of the strengths, weaknesses, and potential areas of improvement.

Following the establishment of a flexible assessment methodology and based on the collected data and experiences, the evaluation methodologies should be further elaborated to achieve the necessary level of granulation of the impact, activity, and performance analysis, and to enable the further betterment of the professional development initiatives and milestones.

Individual assessment – domain-specific competency-based approach

Whether following a standard or customized assessment process, a training provider can better design learning objectives, as well as specific skills, competencies, and abilities, that learners should be able to cultivate and perform, following the activity completion and thus to determine the performance and proficiency level or gained skill(s) in the context of the entire learning pathway. Therefore, both parties (provider – customer/end-user) will be objectively aware of the preparedness of the professional to take their place in an organization, specific tasks, or processes.

A generic assessment methodology should involve the following steps, which are generic and will require

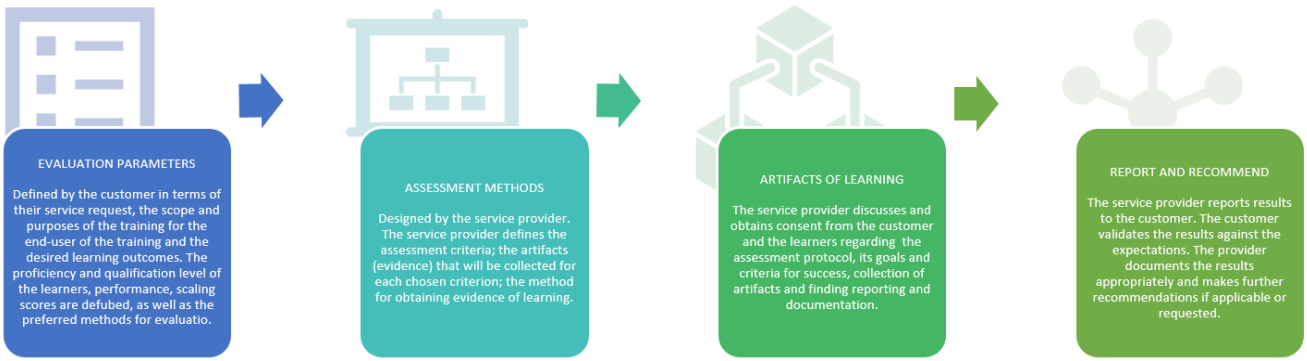

Figure 19: Generic assessment methodology. Image by the ECHO Consortium, developed by ESI CEE.

further customization and alignment to the evaluation criteria and evidence of learning from the activity, through the chosen learning method:

The first two steps are key for outlining the progress that learners could expect after the completion of the exercise. Those steps support the definition of the assessment methodology (in terms of tools, methods, and criteria) as well as the selection of tasks that are appropriate for the learner's skill profile prior to the training. To achieve any proper automation of this process and assert whether the assessment methods are optimal, the ECHO consortium recommends the implementation of pilot assessment, experimentation and evidence of learning collection, to evaluate whether the process fits the purposes for which it is established.

Steps III. and IV. ensure the objectivity of the assessment methodology and in-depth analysis of the learners' achievements in terms of mastering a set of skills. It wouldn't be enough to conclude whether the learner can "*apply relevant standards, best practices and legal requirements for information security*", but to specify which standards and best practices the learner knows and at what level – to apply, to implement, to observe, to

analyze, to adapt, etc., as well as to assess what else could support the reaching of higher goals and integrity of the profile.

The analysis of specific levels of acquired skills and knowledge can be done using different models such as Bloom's, SOLO, etc. [51] on higher and lower (more detailed) level of analysis. The higher level is what are the exercise goals or what the exercise will measure, and the lower level represents how the achievements are proved. Behind the lower level of assessment are different scenarios for measuring the different levels of competence. For instance, the verb "identify" … the type of attack reflects basic, awareness skill level and the verb "uses external, incl. developed by himself, instruments for handling/ remedying/ ensuring the availability … of the system" reflects mastery level of proficiency. The authors propose a mapping of the designed within the scenario tasks with "a range of cognitive learning and skills layers (based on Bloom's taxonomy) as Remembers (1), Understands risks/ attacks (2), Applies attacks (3), Applies defenses (4) and Masters defenses (5)" but those actionable verbs can be different, an object of a consensus between the cyber range providers within the Federation. Thus, the service and training designers have evidence-based information about the competence/ skill/ knowledge achieved, and not only about the passing rate which is not so meaningful in terms of the constitution of the entire training/ professional pathway. This approach ("competency-driven") provides learners with feedback regarding their skills and enables them to identify the current strengths and weaknesses modifying her future training activities accordingly. The authors conclude that the assessment designer should follow iterative steps searching for a balance between the feasibility of the technical deployment of the scenario and tasks considering the initial skill set of the end-user, the measurement criteria, and the realism of the scenario to avoid any misconfiguration of the cycle scenario-tasks-technical infrastructure-targeted skills.

The evaluation and measurement criteria for the assessment and further analysis could combine factors as implementation time, strict following the pre-defined procedures (rules, steps), scores and weights of the different parts (steps) of the exercise, ability to use and apply specific tools/ applications/ methods, putting clear boundaries of a technology domain through knowing when to ask support, etc.

In conclusion, the assessment design should follow a standard process iteratively passing from the abstract layers of collecting data for better definition of the learning progress going deeper to knowing if the achieved proficiency level satisfies the goals and enables further planning of the learning/ professional pathway in correspondence with the appropriate learning methodology.

Team performance assessment – involving transversal competency-based approach

When the goal of the customer and provider is to assess how a team handles with the exercise or mission, the parties need to assume, compare, and address accordingly:

- the different proficiency and qualification (initial) levels of the participants
- the combination of the necessary technical infrastructures to deploy the scenario
- the combination of types, methods, and criteria for assessment and measurement of the individual and team performance applicable to each task of the scenario
- involving tasks related to the assessment of teamwork skills

In this context, to receive meaningful information and data for accurate assessment and analysis, the competence-based approach is also applicable with the inclusion of a supporting method, proposed by Sten Mäses in a framework [50] for evaluating cybersecurity-related skills based on computer simulations. In connecting the exercise results with information about the achievements of the participants in terms of gained

skills, Mäses applies a “Cybersec-Tech Window” that divides the 4 different types of skills. This division reflects the holistic nature of the cybersecurity domain and corresponds with the exercise objectives to provide insights for the needed technical infrastructure, measurement criteria, and appropriate tasks for more than one professional profile. The 4 skill types placed in two two-dimensional axes are:

1. Technical vs non-technical
2. Cybersecurity specific vs Not cybersecurity-specific

Although the distribution of abstract tasks that may correspond simultaneously to technical and non-technical type requires further elaboration and is challenging for the definition of target skills, the method supports and facilitates the addressing the infrastructure and measurement criteria. This challenge could be handled by the design of the latter steps and going back for a more detailed specification of the skills and competences and more efficient distribution of the tasks between the participants in the cyber range exercise.

Regardless of the requirement for human intervention in the design of the assessment methodology, some of the steps can be automated in support of selecting the right measurement method. The Cybersec-Tech Window as a definition tool can be provided within the Service Designer to the cyber range and content providers for a more accurate definition of the tasks and for supporting the subsequent selection of the assessment criteria, technical environment, and additional, transversal skills.

3.5.4 Evolution of design methodologies

Overview

This section of the current report focuses on establishing a common vision for the possibilities for the design of cybersecurity educational services through the federation of cyber ranges in ECHO.

The E-FCR supports training and simulation services for a multitude of proficiency levels, ranging from fundamental information security knowledge to sustainable models for competence level maintenance and domain-specific improvement of skills and abilities with a focus on cybersecurity. Thus, the topic of how effective instructional design is implemented against the backdrop of a system, versatile in educational opportunities and methods, could contain only suggestions of a somewhat generic nature and well-proven strategies and recommendations for specific use-cases of proximity to the characteristics of the E-FCR.

Therefore, in this section, we investigate several methods and frameworks for instructional design throughout the evolution of instructional design methodologies, which could be further employed by the content providers to design activities, relevant to the needs, context and proficiency levels of the relevant stakeholders, as well as to adapt to the technological developments, challenges, and trends, throughout the E-FCR roadmap.

Following the discussion from the previous chapter, the three main pillars of the contemporary learning theories, namely behaviourism, cognitivism, and constructionism, have greatly influenced the evolution of instructional design. Nevertheless, other fields of science have also been found by scholars as foundations of instructional design, apart from psychology. Such fields are cybernetics and systems engineering and design [47], and the intersection between these disciplines in the field of instructional design, have historically greatly influenced the training-by-simulations models, developed by, for instance, the US Army and the medical instruction worldwide.

The field of instructional design has had a common development and history with the information and communication technologies domain, which has led prominent scholars, such as [48][49] whose principles of instruction we will be discussing later on within the same chapter of the current document, to consider the field

of instructional design as a discipline, tightly interwoven with software engineering and computer sciences in a broader perspective.

Consequently, contemporary instructional design has evolved to follow a somewhat conventional software development process, especially, when considering instructional design for digitally empowered learning or simulation activities. Curriculum design, especially considering simulation-based learning, gamification, and other learning methods, mostly native to the constructionist theory of instruction, which the ECHO consortium generally recognizes as the most effective in the context of the E-FCR, have been considerably following a general heuristic consisting of several generic phases, namely:

- Analysis
- Design
- Development
- Implementation
- Evaluation

These steps comprise the general backbone of the widely recognized ADDIE Model for instructional systems design, a variation of which, we would like to put forth as one of the propositions for design methodologies for the E-FCR content and simulations. Many of the contemporary models and frameworks for curriculum design and/or instructional systems design, follow variations of the ADDIE model to best tailor the process to the needs of the content under design.

A common alternative to the ADDIE model, although not as widely-adopted, is the Rapid Prototyping Model for instructional design, which follows a little more of an iterative approach to instructional design, which could be more suitable for other use-cases of content creation for the E-FCR, and especially for cases which envisage a more intensive end-user involvement in the design of the simulations or content. The Rapid Prototyping Model for instructional design mainly follows and builds upon the James Martin Rapid Application Development method for software engineering, and is commonly outlined in the following four phases:

- Requirements planning
- User design
- Construction
- Cutover

In the following section of this chapter, we will present those two models in more detail, provide several modifications to better fit the context of the E-FCR, and complement them with several design principles for other contemporary instructional design methods, to ensure the simulation or content creators within the E-FCR have a variety of tools to choose from when it comes to the activity and/or curriculum design, as well as a supporting mechanism for each of the methods proposed.

ADDIE Model

Although quite innovative and complex, the E-FCR supports inherently constructionist learning activities and content. Therefore, the classical ADDIE Model could be applicable, especially for the more “conventional” training activities and in the cases where less end-beneficiary engagement is envisaged in the initial design and development of the activity. The ECHO Consortium would like, however, to propose to the E-FCR Content developers the PADDIE+M modification of the standard ADDIE Model.

The PADDIE+M modification is an instructional design methodology developed by the United States Navy [74] for designing a scenario for their simulation program and to incorporate input from the entire content and

activity delivery team. The entire PADDIE+M instructional design framework, through the prism of the E-FCR, would look as follows:

- ***P****lanning*
  During the planning phase, the goals, learning objectives, specific needs, schedules, and budget are clarified. Based on the output of those, the implementation team is selected.
- **A**nalysis
  The analysis phase goes into detail about the learners' needs, their characteristics, the domain in which they would apply the learning objectives, and the learning constraints. During the analysis phase, interviews with potential beneficiaries might be conducted, to obtain a structured view on their proficiency levels, and based on that – the best learning methods and delivery options that will be implemented through the E-FCR
- **D**esign
  During the design phase, the learning objectives are refined, based on the analysis performed, and structured into exercises of gradual complexity, simulations, educational content, and lesson plans. The design phase also includes the elaboration of appropriate assessment methods, which will be further discussed within the next chapter of this document. More on the design phase will be discussed further below within the same segment.
- **D**evelopment
  Within the development phase, the implementation team is expected to develop everything that is outlined within the design phase. This includes programmers and instructional designers alike.
  The development phase further includes a testing phase, where testers and/or potential users of the content and/or activity go through the activity and provide structured feedback regarding strengths, weaknesses, and areas of improvement. Based on the feedback provided, the implementation team debugs and revises the activity and contents and conduct a pilot of the materials with the implemented improvements.
- **I**mplementation
  The actual activity delivery takes place.
- **E**valuation
  The developed assessment mechanisms and tools are deployed immediately after the completion of the exercise.
- **M**aintenance
  During the maintenance phase, the content and activities are improved, updated, elaborated upon, and continuously refined.

Visually, the PADDIE+M Model we propose as an instructional design framework looks as follows:

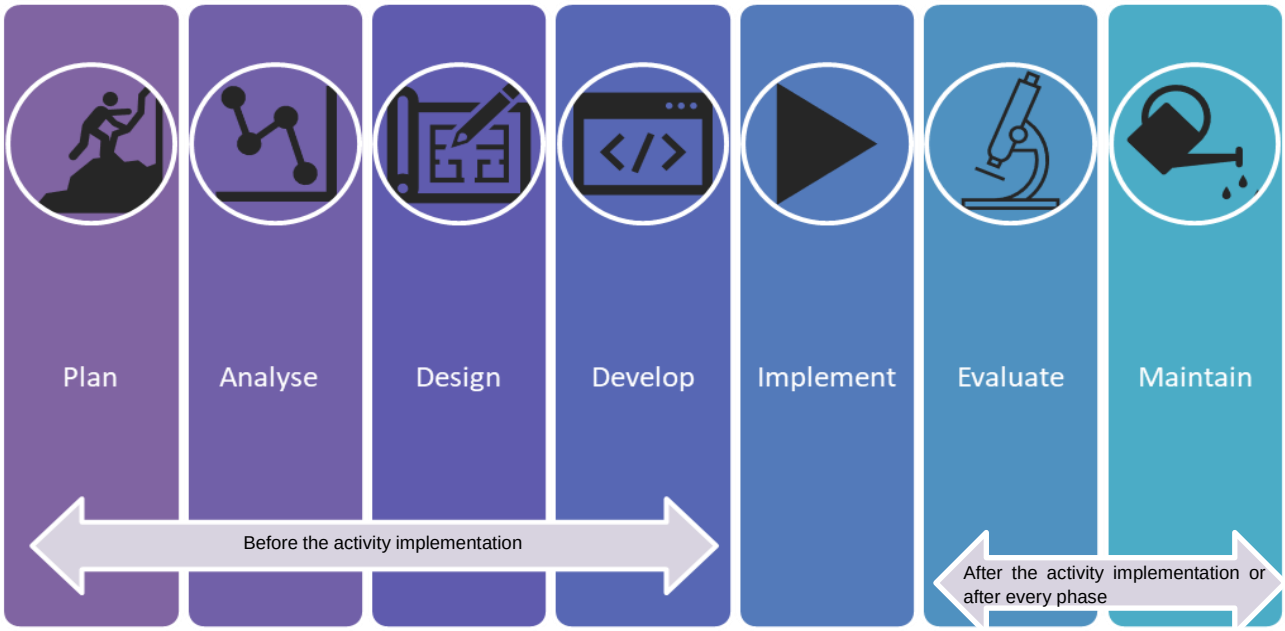

Figure 20: PADDIE+M Model E-FCR Adaptation proposed. Image by the ECHO Consortium, developed by ESI CEE.

For those of the implementation teams, that are used to working through agile development paradigms, we recommend that evaluation and maintenance are included after each phase of the instructional design lifecycle.

About the actual design phase within the context of the E-FCR, we recommend that once the learning method is selected and the cyber range and content providers start designing the service as an entire training event or part of a curriculum or training program they use the Scenario Description Language taxonomy, developed in ECHO WP5, to define the parameters of the activity.

Within the ECHO Federated Cyber Range, the design phase should include parameters, such as session(s) description (duration, access, schedule), involved teams (red, blue, forensic, analyst, scientist/ researcher, etc.), employed technological tools (that enable and facilitate the interaction between actors), and playtime (goals, tasks, steps/ actions, expected results).

The learning methodology, instructional design, and assessment methodology are strongly interdependent in a way, which is why we recommend that those are all outlined in parallel as much as possible, and preferably, during the design and maintenance phases of the activity design lifecycle.

Another point we would like to make, regarding the actual activity design phase is that the customer and learners' training needs vary from very specific (prediction and recognition of zero-day attack vulnerabilities in specific applications or interconnected cyber-physical systems) to more generic (threat awareness), thus all methodologies, frameworks, and tools for instructional design, need to be customized and modeled by educators.

In the design of a cyber range exercise the following interconnected elements are involved:

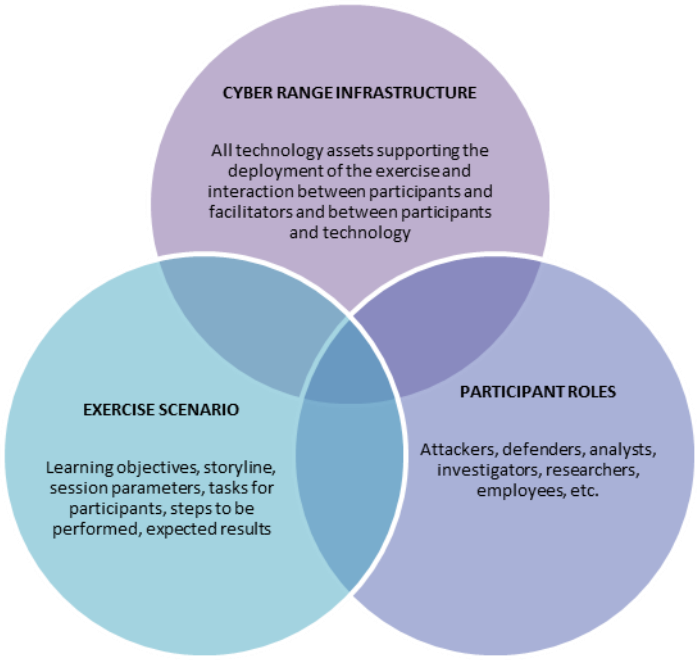

Figure 21: Interconnection visualization between CR exercises. Image by the ECHO Consortium, developed by ESI CEE.

In the related literature, those components are mostly considered as isolated parts of the design model, however in the context and due to the specifics of the E-FCR, those need to be parallelly considered.

To respond adequately to the market demands, experiments for connection and correlation of the existing instruments (SecGen), frameworks, standards and models (CyTrONE, Cyber Mission System Development Framework, Cyber Risk Bow-Tie, Stages of Process for Attack Simulation and Threat Analysis (PASTA) threat modeling methodology, Cyber Range Interoperability Standard (CRIS) (MITRE 2014), Common Vulnerability Specification System (CVSS) (NIST), OCTAVE method (Carnegie Mellon University Software Engineering Institute), EISTIX/TAXII-MITRE) should be implemented, as well as the research results need to be properly documented. Therefore, one of the tasks that the federated environment should implement is to analyze the approaches of the cyber range and content providers and based on the collected data to propose design phases that involve collecting of different elements from different (related or similar) cyber ranges. The processing of such kind of information should provide insights for adaptability, transferability, reusability, and ability for automated design based on all three elements. For example, what kind of tasks and level of difficulty are appropriate for a learner group with prior knowledge and skills on application development. Another question that such an algorithm should be able to answer is which specific technologies and instruments can be used for the deployment of cybersecurity analytical tasks.

A correlation and design guidance engines apply to the concept of the Federation, where many cyber range and content providers develop scenarios and deploy them to different infrastructures dedicated to different types of players. The engine can collect data from the design experience of the presented providers, correlate this data and provide guidance and hints for complementing a given service or scenario with tasks, methodologies, tools, or creation of new simulation services/ scenarios. This will guarantee a better quality of the requested service and will enrich the capabilities of the service and content providers.

A possible use case could be a request from CSIRT to assess the skills of its team to analyze data from the behaviour of a given digital service for a specific period. The correlation and design guidance engine, in this case, proposes the technical capabilities of several cyber range providers, appropriate learning methods, and further specification of the optional learning outcomes, content providers for the design and structuring of the storyline, tasks, and actors. Based on the selected combination of the components and deeper specification of the service design, and analytical report that consists of the main design components will be sent to selected providers who will identify development and collaboration opportunities for improving the quality of the service.

Rapid Prototyping Model for Instructional Design

Some E-FCR content providers might find the PADDIE+M model might not apply to their internal team dynamics, which is why here we propose an alternative instructional design methodology, namely a rapid prototyping model for instructional design.

The rapid prototyping model for instructional design is known since the early 1990's and developed by Tripp and Bichelmeyer [75].

The rapid prototyping model for instructional design focuses on two main qualities, mainly 1) speed and 2) prototypes. This means that the design phase of the activity would be significantly shorter compared to a standard ADDIE model or the proposed PADDIE+M model and that at the end of the design phase, a prototype of the activity will be ready for testing, ideally by an end-user.

This model envisages a heavier end-customer engagement in the activity design process and is more iterative than the PADDIE+M framework proposed, which might make it more suitable for activities and content requiring heavier software development or for implementation teams, which are more accustomed to the agile paradigm of software development.

The rapid prototyping instructional design paradigm envisages the quicker creation of a working prototype for end-user testing and preliminary validation, as well as verification against the requirements for learning outcomes. By combining the design, development, and evaluation phases in one generic design phase, this model might further suit better content providers that already reuse materials, previously developed throughout the E-FCR or fragments of previously created exercises, code, etc.

The generic phases of the model include:

- Requirements planning
  The needs and objectives are stated and clarified, much as the planning and analysis phase of the PADDIE+M model with a stronger focus on creating a short course of action for the next phase.
- User design
  This phase envisages the creation of models and prototypes that represent different aspects of the system. Users are heavily involved, testing small prototypes of working elements of the bigger prototype to combine them all.
- Construction
  The construction of the first version of a working prototype happens in this phase, which still envisages heavy user involvement and testing.
- Cutover
  The design and development are finalized, and a working model is in place, ready to be deployed and further include assessment frameworks.

Visually, the model could be represented like this:

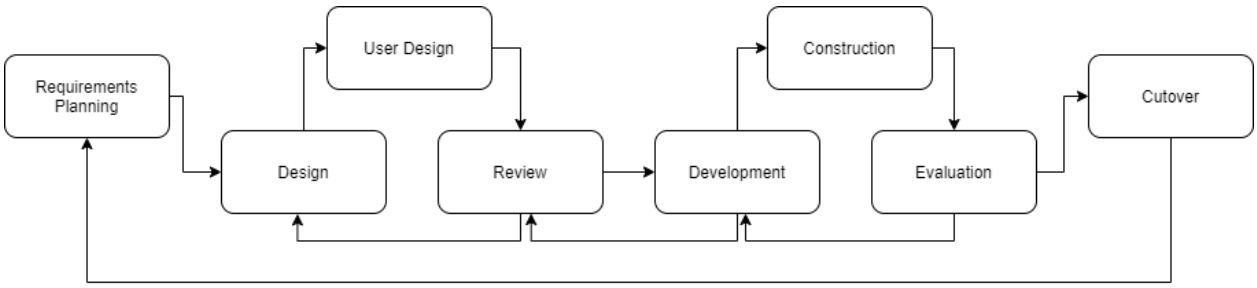

Figure 22: Rapid Prototyping Model for Instructional Design. Image and adaptation to the E-FCR context by the ECHO Consortium, developed by ESI CEE.

As it is evident from the model, this design framework provides more opportunities for iterations, as well as for user testing and validation of requirements.

If this model is adopted by E-FCR content developers, we recommend that special attention is paid to the requirements planning phase, where it would be highly appreciated to use the terminology for the Scenario Description Language taxonomy, developed in ECHO WP5, to define the parameters of the activity.

It is of utmost importance that the activity design language is aligned to a common baseline and should include, at the very least, parameters, such as session(s) description (duration, access, schedule), involved teams (red, blue, forensic, analyst, scientist/ researcher, etc.), employed technological tools (that enable and facilitate the interaction between actors), and playtime (goals, tasks, steps/ actions, expected results).

This effort will also facilitate the work that needs to be done during the cutover phase, including customer assessment, final feedback, activity documentation, etc.

Further instructional design considerations

A few further research-informed considerations could be proposed to training service providers, using the E-FCR, to create more robust educational experiences.

One of these considerations lies in the foundation of the five principles of instruction, identified by Merrill [49] at the beginning of this century. Those principles could be adopted within both of the proposed instructional design frameworks and introduced in the learning activities, developed by training service providers. The five principles could be visualized as follows:

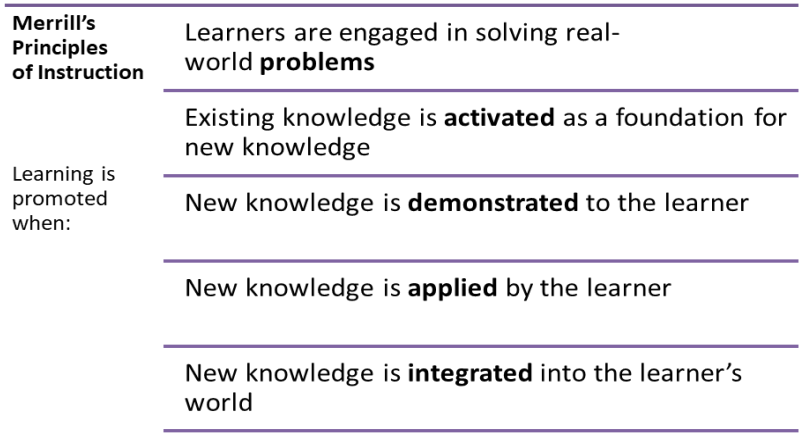

Figure 23: Merrill's Principle of Instruction [49]. Image adaptation by the ECHO Consortium, developed by ESI CEE.

Consequently, we recommend the following principles to be observed in the E-FCR, when integrating new educational content:

1. Content providers, who choose to include multiple exercises or educational topics, will benefit the learning experience when including in the design of the materials a progression of complexity. Through gradual complexity, learners will be able to more quickly solidify their understanding of the previously considered topics, view a practical application for the already learned principles, and retain knowledge quicker.
2. Aim at engaging prior knowledge into the solving of challenges, requiring new knowledge as well. If possible, and in case the end-users of the training services are not from the cybersecurity domain of practice, involve use cases, demonstration, and challenges, that include information from their context and domain of practice.
3. Provide demonstrations or examples for as many statements you make, as possible.
4. Try to integrate challenges in simulations or contexts, as near to real-life, as possible.
5. Make challenges as engaging as possible, add hints and help for those that cannot solve them, to not discourage them from the new topics they are learning, however always explain everything that you hint at.

Following these principles when engaging learners will optimize their experience with the learning environment and will help facilitate engagement and retention.

Another direction, that we recommend that E-FCR content providers follow as much as relevant when producing content and activities, is to include in the E-FCR content creation Gagne's nine events of instruction [45]

In his research, Robert Gagne concluded that there is not one single style or type of learning. What is more, he identified several levels, at which learning happens. To facilitate the learning experience, he recommended educators to follow and note several events, that would point us to the direction of where the learner is at.

What is more, in his book Conditions of Learning, he proposes that there are several types of learning, each of which has its conditions and styles of learning. Whether the learning outcomes consider the development of intellectual skills, motor skills, verbal skills, attitude change, or cognitive strategies, different practices are proposed to be adopted to best facilitate the desired learning outcome.

For the E-FCR, as defined within the ECHO consortium, the most appropriate strategies for the accumulation of new knowledge, skills, and abilities, we recommend that the content creators facilitate the nine events of instruction, as described by Gagne:

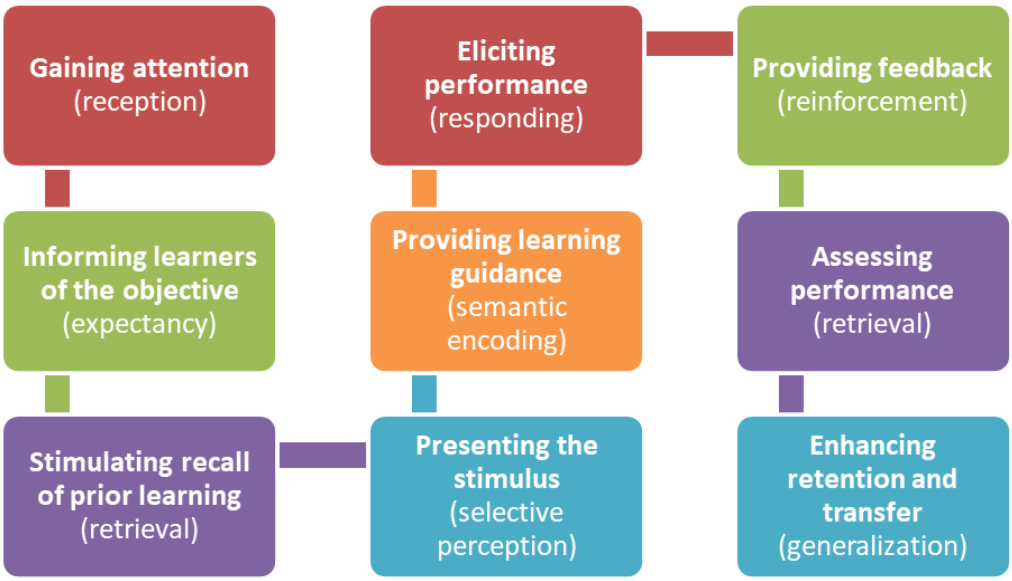

Figure 24: Gagne 's nine events of instruction [45]. Image adaptation by the ECHO Consortium, developed by ESI CEE.

These events should satisfy or provide the necessary conditions for learning and serve as the basis for designing instruction and selecting appropriate media [46]

Sample Use Case and Debrief

To illustrate the more theoretical approach, with which we outlined the previous pages, the following use case might be considered as an example for the use of the PADDIE+M instructional design methodology.

The IT department of a hospital wants the hospital administration and the general practitioners to learn more about ransomware, as they send and receive a lot of e-mails with attachments and there is a recognizable risk. The hospital administration and general practitioners have no prior experience or knowledge in information security or information technologies in general.

A training service provider wants to create an immersive simulation for them through the E-FCR, where they could practice recognizing and mitigating the risk of ransomware and spark a discussion between the healthcare professionals.

The training provider chooses to follow the PADDIE+M instructional design paradigm, so they start with planning. During the planning phase, they outline the project goals, scope, budget, and schedule for implementation. Following this discussion, the training provider decides on an implementation team, consisting of 5 people, namely, 3 software developers, 1 network specialist, and 1 researcher experienced in pedagogy and instructional design.

During the analysis phase, they analyze closely the pedagogical considerations, learning constraints, and the delivery options, choosing simulation with educational prompts and content on ransomware. Following the

analysis phase, they step onto designing and outlining the actual activity. They take into consideration the scenario description language developed under ECHO, to create a structured activity plan and explore other efforts in the E-FCR, to use as examples.

The development of the actual simulation takes place, during which process, they decide to involve representatives from the hospital management or the IT department to validate whether the activity developed is up to par with the expectations, requirements and learning objectives they have for their team members. Their feedback is obtained and based on that, the activity is improved, debugged, packaged, and deployed.

During the actual implementation of the activity, the team provides technical support and facilitates the user experience of the participants. The facilitators go through the learning objectives, to enable self-assessment of the participants on a later point, go over the method of delivery and the testing procedures.

Based on the implementation of the activity, the participants and the implementation team alike, evaluate the experience, following which, during the maintenance phase, improvements, updates, and upgrades are implemented.

`With these design principles, as well as recommendations for specific practices, E-FCR content creators have a practical framework to follow and contribute to a more practically-oriented and adequate to the needs of the end-users cybersecurity training.

3.5.5 E-FCR Legal and Intellectual Property considerations

This section will look into the E-FCR roadmap in terms of legal and intellectual property (IP) considerations. First it will identify what the current and future iterations of the E-FCR are and then it will give an overview of the legal and IP questions that can be identified at this stage.

The E-FCR is a digital cloud platform, offering cyber range providers and content providers from different countries the opportunity to showcase their cyber range offerings to customers. Basically, it is an international hub of cyber exercise and training resources and information.

In terms of the specific cyber exercising offering, the E-FCR will evolve over the years. The initial E-FCR model that is being developed during the lifespan of the ECHO-project is that of a digital platform, offering a catalogue of services, collecting the customer requirements with the scenario description language, collecting and transferring the payment to the selected cyber range and/or content provider and monitoring the quality of service.

In this initial model the customer is the cyber range user, which uses the E-FCR client interface to select a cyber range provider or content provider that provides what the customer needs.

The cyber range provider offers the virtual environment consisting of actual hardware and software, simulating organization's local network, system, tools, and applications that are connected to a simulated Internet level environment[43]. The content provider is for instance the scenario developer or trainer building sector specific scenarios/trainings and distribute them through a catalogue of scenarios/trainings.

The offering of the cyber range and content provider can be either a ready-to-use standard cyber range service or a custom designed service offered by one of multiple cyber range and content providers (this is the so-called pooling of offerings).

This offering will be more than just a simple address book. Basically, in this first iteration, the E-FCR is a one-stop-shop where:

- Cyber range providers and content providers can list their offerings.
- Customers can browse the different propositions and buy a proposition.
- The E-FCR provides the technical interconnection between the cyber range provider and the customer.
- Service fees are collected and transferred from the E-FCR to the cyber range provider(s)/content provider(s)
- The quality of the offered cyber services is monitored by the E-FCR.

Future envisioned iterations of the E-FCR could include infrastructural components. Currently, each cyber range provider has its own hardware and software environment (physical or virtual servers, cloud-based storage as well as applications) on which cyber exercises are running. From an economy of scale and European strategic autonomy point of view it could be interesting to explore the development of a European federated cyber range infrastructure, where cyber range providers pool their infrastructural components or buy these as a service from a single provider.

Another future iteration could envision linking and integrating the E-EWS and the E-EMAF with the customer requirements and cyber range content and capabilities. For instance, the attack vectors identified by the E-EWS will be incorporated in the cyber range providers' offerings and the customers' E-MAF risk assessment profile will be matched automatically with the cyber range offerings.

As for the legal framework of the E-FCR, current and future versions of the E-FCR digital cloud platform do operate in a legal context, with applicable rules and regulations. In addition, the digital platform is an offering of cloud computing, per definition operates across different jurisdictions as the service provider and customer operate in different countries, each with their own legal requirements (civil, criminal, administrative and tax legislation).

The E-FCR digital cloud platform operations will require that a number of legal aspects have been addressed, such as:

- Data privacy
- Data security
- Data property
- Data portability
- Contract documents
- Fees and VAT treatment
- Record keeping and audit rights
- Intellectual property
- Contractual Liability
- Variation of terms, suspension and termination, governing law and jurisdiction

Below these points will be outlined and briefly analysed in terms of the E-FCR roadmaps and development work ahead for future iterations of the federated cyber ranges.

Data privacy

In terms of **data privacy,** the key legislation in place is the European General Data Protection Regulation (GDPR). It defines in article 4 personal data as:

'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

In the design phase of the E-FCR the following personal data has been identified:

- Cyber range user accounts data (name, surname, email address, organization name, role, address, login credentials including passwords).
- Cyber range exercise results (user proficiency, user tracking).
- Cyber range users survey results.
- Information generated in the E-FCR, such as the aggregated customers search queries.
- Data used in simulation of virtual systems during exercise, if real environment is simulated.

Out of those cases preliminary risk analysis can conclude that the processed data is not high risk and mainly consist of contact information and information about cyber-security skills of users with the exception of test data, where if used without careful assessment/anonymization can contain virtually anything, not only limited to PII, but also company confidential information.

This means that on both the E-FCR as the providers need to be compliant with the GDPR requirements. However, what needs to be further investigated is 'who is the E-FCR'? What entity is the data controller for the E-FCR? In the initial phase during the lifespan of the ECHO-project, it seems logical that one partner will take on this responsibility. However, in the post-project exploitation phase this needs to be further analyzed, ideally per offering level (i.e. the E-FCR model of a services catalogue probably has a different legal set-up than the E-FCR model of a cyber range infrastructure as a service offering).

Another relevant issue is the geographical scope of the E-FCR. According to article 45 of the GDPR no data can be transferred outside the European Union without an adequate level of protection. This means that the E-FCR data can only be stored in European data centres or transfer should happen after careful analysis. Special attention should be paid to Cloud services as well, following the ruling of Court of Justice of the European Union on data transfers that invalidated the Privacy Shield agreement.

It is important to notice that compliance with GDPR is not optional and possible fines for non-compliance can go up to €20 million or 4% of the worldwide annual revenue (whichever is higher), but also for the reputation of E-FCR and cyber range and content providers.

Data security

In terms of **data security,** the E-FCR also needs to consider the consequences of data breaches, loss of data, service interruption and illegal content on the digital platform. According to the GDPR data breaches have to be declared to the supervisory authority and the data subject. It must be clear from the outset of the operation of the E-FCR what organization will do this and what procedures are in place (e.g. appointing data controller for the E-FCR dealing with all E-FCR related data protection requests). Also as E-FCR like ECHO spans multiple countries in EU, there might be some freedom in selecting under witch supervisory authority the platform operates.

Furthermore, some arrangements have to be made of a legal entity that runs the E-FCR and for instance can be held liable/asked for financial damages due to loss of training data or interruption of a cyber exercise.

As for potential illegal content on the E-FCR digital platform, the EU Directive 2000/31/EC (E-Commerce directive) protects online service providers with an exception for liability (articles 12 and 13) in specific cases of not modifying of having actual knowledge and control over the information. However, it is unclear if this is the case. This point has to be investigated in terms of architectural design of the E-FCR.

Data property

In terms of **data property,** the E-FCR also provides some interesting challenges. On the one hand there is the customer who provides digital information such as search queries for the cyber range and content, inputs specific cyber range needs and receives cyber range training and evaluation results. On the other hand, there are the cyber range provider and the content provider, which provide the virtual environment, training scenario's and injects and exercise feedback. On top of this there is information generated in the E-FCR, such as the aggregated customers search queries and the survey feedback that generate new content.

A number of ownership status challenges can arise as there are legal grey areas between what data is generated inside the digital platform by whom. Ideally, these issues should be tackled by contract law, the terms and services of the E-FCR and the end user license agreements.

Data portability

The concept of **data portability** in the context of the E-FCR relates to the right of users to receive personal data they provided to a data controller in structured, commonly used and machine-readable format and to transmit those data to another controller (cyber range provider). This right stems from article 20 of the GDPR.

As stated before, the E-FCR handles personal data which is more than user names and addresses that are needed to create an account. There is also the personal data gathered while observing the individual cyber range users' activities, such E-FCR customer search queries and survey results. All these aspects fall also under the right to data portability.

Contract documents

Digital platforms do have **contracts or agreements** in place in place, governing the relationship between the cloud customer and the cloud service provider. These are sets of documents regulating key points such as the relationships with the customers (Customer Agreement), the acceptable use of the service (Acceptable Use Policy), the service level description and service levels (Service Level Agreement).

For the E-FCR these contract documents need to be developed, looking into detail into questions such as:

- Contractual relationship between E-FCR, customer and cyber range and content provider.
- Ways in which the E-FCR customer may use the service and what actions the provider may take in the event of a breach.
- What is excluded and what constitutes illegal use within the framework of the E-FCR.
- What business guarantees the E-FCR provides (e.g. 99.999% of service availability, time-to-repair, escalation procedures, etc.).
- Support and planned maintenance conditions, provisions and described services.
- When/how payment is to be made, late payment penalties, service credits payments for outages and VAT regimes.
- Sector specific standards for regulated industries like healthcare.
- Data policies (see previous headings on data privacy, security, property and portability).

Fees and VAT treatment

Ultimately, after the end of the ECHO-project, the E-FCR ultimately provides a service that at exploitation phase will be paid for. Such sales are subject to taxes. The role of digital platforms in the collection of VAT/GST on online sales is a complex matter. The different VAT/GST liability regimes are outlined in a 2019 OECD report, describing the digital platform liability and tax authority needs in terms of administering, policing and collecting VAT/GST[44]

Further exploitation phases of the E-FCR need to investigate matters such as between whom the transaction takes place, what tax law is applicable, and how the collection of the taxes is implemented.

Record keeping and audit rights

The E-FCR is a digital platform that a customer and cyber range and content service provider is using to engage and exchange data. For legal, administrative and fiscal compliance purposes a set of policies and implementing system of record keeping and audit rights will have to be developed.

This is a complex area, that touches for instance the GDRP (and for instance the right to be forgotten), data retention, evidence law in matters of litigation and tax authorities record keeping requirements. It is also partly governed by national laws and for certain sectors such as the health sector there are sector specific requirements for record keeping and audit rights.

Intellectual property

In terms of the E-FCR and intellectual property, there are globally four aspects that have been identified at this stage and that should be further analysed at exploitation stage:

- IP of the customer
- IP of the cyber range provider
- IP of the content provider
- IP of the E-FCR and data aggregation

The IP of the customer relates to for example the customer software installed on the platforms of the cyber range providers in order to simulate exactly the customers virtual cyber environment. The IP of the cyber range provider relates for example to the virtual simulation environment set up by the cyber range provider, the scenario's and injects provided by the cyber range provider, the UI components hereof, the assessment tools, etc. The IP of the content provider relates to the scenario's and injects, the training programs, the evaluation methodologies, the support tools linking the cyber range outcomes to the customers risk assessment profiles, etc. Finally, there is the IP of the E-FCR, which relates to for instance the UI, the search engine connecting the customer queries with the providers offerings as well as aggregated data such as the aggregated customers search queries.

Each one of these aspects needs to be further detailed and analysed in terms of IP-protection (e.g. patents, trademarks, copyrights, trade secrets and utility models).

Contractual liability

Another legal question to investigate is what the contractual liability regimes will be applied to the E-FCR. Basically, what liability can arise and how this is dealt with in legal terms (circumstances, exclusions, monetary caps, etc.).

The E-FCR will have to develop standard terms, which should address all these (and more) legal issues. These terms should take into account the role division between the digital platform as a broker and the legal relationship between the customer and the provider. Also, it is advisable that the E-FCR liability terms should continually be monitored in light of the evolution of the E-FCR and its offered services.

Variation of terms, suspension and termination and governing law and jurisdiction

The E-FCR and all involved providers will deliver a service, which will be governed by a contract. Such a contract is never set in stone and evolves over time as for instance the service offering and marketplace develops and its regulatory environment mutates. This means that there will be clauses in the contract, detailing how to deal with the evolution of its terms. Usually, such variation needs to be accepted by both parties in writing (or online, by providing the updated text for review and allowing the user to click on 'I agree'). Such mechanisms need to be built into the E-FCR online platform (read the terms, review of the amended terms).

Another aspect dealt with in contracts are the suspension and termination clauses. This is basically about the right of the vendor to suspend the service or to terminate it altogether upon certain events or conditions. It is unclear in terms of the E-FCR who is the vendor in this regard – is it the E-FCR or the cyber range and/or content provider? Also, what are these events that will suspend or terminate the contract? Legally speaking, such a clause should be limited in scope in order to be defendable in litigation cases.

A further aspect dealt with in contracts are the clauses regarding the governing law and jurisdiction. In the case of the E-FCR this would mean that the law of the consortium partner exploiting the E-FCR would govern the E-FCR and the courts of that country will have exclusive jurisdiction over any disputes arising out of the contract. The question is whether this is acceptable for a public institution customer base, located in a different member state. Such institutions often have legal restrictions on what governing law and jurisdictions clauses they are allowed to accept. One way out is to not have such a provision in the contract and leave the question open for when litigation arises. Other solutions are to refer to alternative dispute resolution (ADR) such as international arbitration, negotiation and mediation.

These reviewed legal aspects of the E-FCR digital cloud platform described are not exhaustive and more will arise in future iterations of the E-FCR. The analysis shows that the legal aspect of the E-FCR is complex and needs a detailed investigation at exploitation phase as it is a key component of the business offering. An important selling point for larger companies and public institutions is to have legal compliance above the law (for data privacy) and a system of terms that is acceptable to the legal departments of these companies and institutions. An option to explore could be a ‘Legal terms generator’ that would dynamically generate customized terms (and is connected to pricing as terms and financial risks are closely connected).

It will not be easy to develop such as flexible system but it can be seen as a differentiation strategy from other providers. At the end of the day such an investment is closely connected to what type of customer the E-FCR wants to target.

4. ECHO Early Warning System Roadmap

ECHO Early Warning System (E-EWS) is an platform to be used to establish network of partners, sharing cyber-security information and one of the most important instruments in achieving the ECHO objectives – to strengthen the proactive cybersecurity defence of the European Union through effective and efficient collaborative information sharing. As the cyber-threat actors are changing their tactics constantly and the attacks are growing in numbers and evolving more and more sophisticated and complex, there is a need to identify future development options for E-EWS to sustain the leading edge and adapt to the new challenges. The roadmap also aims to identify innovative and horizontal capabilities to be implemented in the future version of E-EWS platform, after development under ECHO project is complete.

The goals of the ECHO Early Warning System within ECHO are:

- Deliver a secure sharing support tool enabling personnel to coordinate and share cyber-sensitive information in near real-time.
- Support information sharing across organizational boundaries and between disparate information repositories as may be used by partner organisations, including granular control of data and functionality access.
- Provide sharing capability of both general cyber information and specific incident management data.
- Secure connection management from clients accessing the E-EWS, to ensure only personnel with a valid certificate can access E-EWS functions and data.
- The secure information sharing model will account for sector-specific needs including GDPR compliance and others related to health care, banking, insurance and other sectors dealing with personal data.

Using above goals as base level, roadmap description efforts focus on the next opportunities for expansion, growth and increased functionality of the platform, following the methodology described in Chapter 2. This section outlines the consolidated major technological areas, called Epics, that were identified during the brainstorming session of extended ECHO experts panel and their respective supporting “user stories” – technology drivers, targets, alternatives that affect future development of E-EWS after initial period of ECHO Project.. Three major domains were chosen to capture different aspects of technology, environment, and requirements evolution:

- **User experience** domain deals with how users of the EWS interact with the platform
- **Platform domain** deals with how the EWS can be grown and integration of new tools
- **Exploitation domain** focuses on uses and adoption of the E-EWS platform

4.1 E-EWS User Experience

The below chapter describes one of the development opportunities considered with regards to the enhancement of the user experience in the E-EWS platform. Taking into consideration development of related technologies in other work packages of the ECHO project, practical exercises and feedback collection sessions planned for upcoming years, it is foreseen that EPIC E-EWS User Experience will be further developed and described in the next update of this document due in January 2023.

4.1.1 Response aware AI suggestions

AI in Cyber-security

As [79] describes Artificial Intelligence (AI) as "an enabling technology being widely applicable across many different domains to achieve cybersecurity-related goals. In the context of cybersecurity, it can be considered as 'technological solutions: Integrating machine learning approaches and capabilities to process large amounts of information and derive insights that can inform a course of action relevant for cyber-related purposes.
Following concurrent advances in computer power and the ability to treat large amount of data in the last years many new applications of Machine Learning have been developed by researchers in order to address many cybersecurity problems."
One application of AI in Cybersecurity is utilizing the potential to detect Malware where the traditional blacklists methods are not able to detect new malware or lack exhaustivity. The ability of AI components to produce cyber intelligence can be used for the development of specific defensive applications. These applications cover the tactical/ technical and operational level of cybersecurity. From operational perspective AI could be used to retrieve and process data gathered from network security analysis programs. From a tactical point of view, AI can increase the capacity of cyber threat detection, analysis, and, possibly, prevention by supporting the upgrading of Intrusion Detection Systems (IDS) aimed at discovering illicit activities within a computer or a network, the same goes for spam and phishing detection systems as well as malware detection and analysis tools. Moreover, AI components have the capabilities to integrate multi-factor authentication or verification systems. These aspects are helpful for the detection of a pattern of behaviour for a user into the identification of changes in those patterns.
Another key defensive application of AI is automated vulnerability testing, also known as fuzzing, which could be used both to improve cybersecurity services and as to train staff through very interesting tools, able to fight "on the field".

Risks of use of AI

By analysing large amounts of data and identifying links among them, AI may also be used to retrace and de-anonymise data about persons, creating new personal data protection risks even in respect to datasets that per se do not include personal data. AI is also used by online intermediaries to prioritise information for their users and to perform content moderation. The processed data, the way applications are designed and the scope for human intervention can affect the rights to free expression, personal data protection, privacy, and political freedoms (https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf).

Storyline AI

Employees of one of the companies decided to connect more users to the corporate network. The network includes a router (gateway), routers with wired and wireless access, switches, programmable logic controllers, a server, workstations. Microphones, video cameras, smartphones are also connected to the network. The main data transmission technology for connecting devices to the network is Ethernet, and Wi-Fi is additionally used. The gateway provides access to the Internet. For quick access to terminal devices, and interaction with

terminal processes, the Telnet protocol (port 23) is used. The server stores personnel databases, financial statements, and other important documents. File transfer is carried out using the FTP protocol, documents are sent via regular e-mail, without data encryption. The network administrator did not use virtual circuits for this network, as he believed that the manufacturers of network equipment had already configured security policies, including the use of tunnels for data transmission. The company's routers were purchased from the same model and from the same manufacturer. Server operating systems were installed on the server. The network administrator knew that they have more cyber defences than conventional operating systems. A firewall built into the operating system was also used.

On July 10, a user working at a workstation saw that the cursor was moving across the monitor screen, and the files were flipping through on their own. He turned to the administrator, who promised to investigate the situation.

The network administrator has urgently downloaded updates for the server operating system. And checked other workstations for similar incidents. It turned out that only one workstation had an incomprehensible situation. Just in case, the administrator gave the user an access password, advised him to remember it. The user remembered it, and, just in case, wrote it down in a notebook on the table.

On July 11, users began to contact the administrator with a question why they did not have access to the data stored on the server. The administrator ran a check and it turned out that the database was corrupted. Some data has disappeared.

On July 12, the director summoned one of his subordinates to his office and showed the video with audio recordings, which came by email to the director's reception today. On them, a subordinate talked to someone on a mobile phone and provided important confidential information about the company in which he works. The conversation took place in one of the company rooms, where no video surveillance cameras were installed and there were no microphones.

On July 13, a cyberattack was carried out on the company's routers, most of the information on the server was erased, and the latest versions of documents disappeared from workstations. After that, the network stopped functioning.

On July 14, a cybersecurity specialist was invited to help in this situation. He found that port 23 was not closed, and through it, the company workstation was remotely controlled. The access, authentication, authorization, and accounting policies were not configured correctly on the server operating system. Additionally, there are no firewalls installed in the form of hardware, software and software on workstations. No tunnels are installed on the network equipment (routers, softswitches), passwords are used by default from the manufacturer, two-step authentication is not configured when accessing the e-mails of company employees. Additional anti-virus programs were not installed and periodic testing of all network components was not started, network analyzers were not used. Access passwords were written down in a prominent place by many. The suspicion arose that there was a man engaged in industrial espionage in the company. also recommended using two-factor authentication, and encrypting important documents before sending them by email, and using a digital signature. Passwords must be changed frequently, and the password must contain at least 8 characters, have letters of different registers, in different languages, with numbers and symbols. It is better to generate it automatically (for example, artificial intelligence will help with this). You also need to set up Radius and AAA servers, set up group security policies, do not write passwords on paper that other employees of the company can see. You also need to have protection against keyloggers, and spyware, software and hardware bugs. Smartphone cameras and their microphones can be subject to cyber attacks. Some applications that are downloaded from the Internet and installed on a smartphone have additional software through which you can invisibly turn on the cameras and microphones of the smartphone. For data transfer, it is better to use the TCP protocol instead of FTP, since the largest number of cyber attacks occur on files transferred using this protocol.

It is also advisable to update the software of all devices on the network. Patching can be done using artificial intelligence algorithms. It was also proposed to use artificial intelligence algorithms so that similar situations do not arise in the future.

USEREXPERIENCE-AI-1: When setting up the procedure for transmitting information on network equipment, the user did not programmatically close ports 23, through which Telnet protocol transmission works. By default, all ports are potentially at risk of attack, no port is natively secure.

Each port and base service has its own risks: depending on the version of the service, the correct configuration settings, whether strong passwords for the service are created, whether they are reliable.

Also, some vulnerable services have a permanent utility, outdated services, such as Telnet on TCP port 23, and were initially insecure, in particular, Telnet sends data completely without a mask in clear text. While some network ports are good entry points for cybercriminals, others create good escape routes, so TCP / UDP port 53 (for DNS) can be used to organize exit.

USEREXPERIENCE-AI-2: When organizing and operating network equipment on the network, access policies for network resources and equipment are not configured.

AI can be used for:

- Spam filter applications;
- Network intrusion detection and prevention;
- Botnet detection;
- Secure User Authentication;
- Cybersecurity Ratings;
- Hacking incident forecasting, etc.

AI can help in the analysis and exploitation of vulnerabilities by extracting information from services running on target systems. AI helps generate metrics, discover network infrastructure, and report results faster.

USEREXPERIENCE-AI-3: The user did not update the software of the softwwitch, router, program-logic controller, server in order to patch the vulnerabilities of their software.

To eliminate vulnerabilities in the software of servers, program logic controllers, routers, software switches, updates with patches for vulnerabilities are used. They are offered to download by the equipment manufacturers themselves. AI can periodically search for updates on the corresponding sites (this is prescribed in the training program for the neural network), download them, thus eliminating vulnerabilities.

USEREXPERIENCE-AI-4: No tunnels, virtual circuits, and virtual networks have been configured by user for secure data transmission.

Machine learning behaviour analytics technology may indicate an unusual virtual networks activity, including:

- Abnormal VPN connections from the user.
- Abnormal VPN session duration.
- First VPN connection from an unknown device.
- VPN connection from an anonymous proxy.

- Abnormal amount of data uploaded during a VPN session.
- Increase of company-related data files access.
- MFA from a new device for a user.
- Physical badge access after VPN access.
- Too many failed VPN logins.
- VPN access from a disabled account.
- Source IP from unauthorized location.
- Malicious VPN source IP.

USEREXPERIENCE-AI-5: The user did not installed antivirus programs and software and hardware-software firewalls.

AI and ML can be used to threaten the security but also can be used to improve it with the new means.

There are several tasks that can improve security systems and prevent attacks:

- Detection of anomalies (potential threats);
- Detection abuse based on training with tagged data;
- Data research, often using visual research, which helps security analysts increase the readability of incoming requests.
- Assessing the risk of a particular user's behaviour, either by establishing an absolute risk indicator, or by classifying users based on the likelihood that they are bad actors.
- Mitigation solutions based on AI and ML can help identify and correct DDoS accounts quickly and without human intervention. Human intervention can cause a delay in detection and response, resulting in a loss of time. Automatically identifying a suspected DDoS attack can help quickly deploy appropriate countermeasures and risk reduction filters for superior downscaling.

USEREXPERIENCE-AI-6: When connecting to the Internet and working with an account, the user was not configured with two-factor authentication.

Machine learning techniques can provide an analysis of the user's environment via the monitoring of hundreds of parameters. Automatic identification can help quickly deploy appropriate countermeasures and risk reduction filters for superior downscaling, allows one to implement multi-factor (in particular, two-factor) authentication when a user logs into his account while working on the network.

USEREXPERIENCE-AI-7: When sending electronic documents, the user did not use cryptographic protection methods, in particular, a digital signature.

Machine learning techniques can be used to:

- indicate the relationship between the input and output data created by cryptosystems;
- create the private cryptographic key over the public and insecure channel (Boosting and Mutual Learning);
- classify the encrypted traffic and objects into steganograms used in steganography (Naive Bayesian, support vector machine, AdaBoost).

USEREXPERIENCE-AI-8: When creating databases, the user did not take measures to protect information.

To protect information we can use a machine learning system capable of extracting valuable information from the log files regarding the access patterns of queries in the same transaction. To utilize this system for databases with a large user population, the roles can be employed for training a classifier.

USEREXPERIENCE-AI-9: Data transmission in the network was carried out using insecure FTP protocol.

FTP protocol has many vulnerabilities, such as anonymous authentication, directory traversal, and crossite scripting, which makes port 21 vulnerable. The AI can be used to automatically mine potential features when the sample set is large enough, breaking the restriction of artificially designed detection features. In this way, hidden cyberattacks can be identified.

Compared with traditional traffic detection solution, the AI-powered traffic detection solution has following highlights: by combining multi-dimensional semantic modelling with AI, the solution can discover deeper hidden features and different granularity of data can be used to find abnormal behaviours. User behaviours can be modelled, and unusual access to service resources can be detected.

USEREXPERIENCE-AI-10: When using network equipment, the user did not conduct periodic testing for vulnerabilities and software trojans.

To detect software Trojans using AI we can use the following approaches:

- Traditional machine learning approaches.
  - Static-based approaches extracting features derived from a piece of program without involving its execution.
  - Dynamic-based approaches including those approaches that extract features from the execution of malware during runtime.
  - Hybrid approaches combining static and dynamic analysis to extract features
- Deep learning approaches
  - Methods performing feature engineering to extract a feature vector representing the executable;
  - Methods taking the gray scale representation of an executable as input;
  - Methods that are fed with the sequence of API function invocations;
  - Methods modeling a program as a sequence of instructions.
  - Methods representing a computer program as a sequence of bytes.
  - Methods aiming to classify a program from its network traffic.
- Multimodal approaches.
  - Early fusion methods creating a joint representation of the unimodal feature vectors.
  - Late-fusion methods training one model per modality and fusing the output decision values.
  - Intermediate-fusion methods constructing a shared representation by merging the intermediate features obtained by separate models.

USEREXPERIENCE-AI-11: When working with network equipment, servers, workstations, the user did not conduct a visual scan for hardware trojans.

To detect hardware Trojans using AI the following approaches can be used:

- Reverse engineering improvement analyzing the circuit aiming to determine its structure and operation and comprising five stages: decapsulation, delayering, imaging, annotation and schematic creation, organization and analysis.
- Real-time detection used to disable the infected circuits or to bypass it in order to enable the smooth operation of the circuit.
- Golden model-free approaches using golden models of the circuits as a guide for the detection of anomalies produced from hardware Trojans.
- Gate-level netlists aiming to detect infected circuits and abnormal behaviours based on features of gate-level netlists by analyzing the gates and their interconnections, inputs-outputs for the detection of hardware Trojans.
- Classification approaches aiming to detect and classify different types of hardware Trojans and trojan-free circuits.

USEREXPERIENCE-AI-12: The user did not configure the AAA and RADIUS servers on the server, router, softswitch, work stations.

Neural network (AI) can be applied to an AAA (or RADIUS) server that covers all the seven layers of the OSI 7 model and the physical user credentials, user authentication, data encryption.

USEREXPERIENCE-AI-13: When working with video cameras, laptops containing cameras and microphones, the user did not test for the presence of spyware, in particular keyloggers.

A machine learning technique to model normal behaviour of application and detect cyber attacks, spy-attacks with use of penetration tests technique. Penetration testing generates large massive datasets. AI can be used to filter this data and eliminate noise.

USEREXPERIENCE-AI-14: The user did not perform a periodic password change, and the password itself was not complex and insecure.

Artificial Intelligence may be used for periodic change of passwords, generation of a strong password, which you can send to the user as sms, e-mail, or in the special application.

USEREXPERIENCE-AI-15: The user utilized uncontrolled machine learning to detect rare or abnormal patterns to increase the detection of new attacks. However, false positives and warnings triggered periodically, requiring a significant amount of analysis effort first to investigate the accuracy of these false positives.

Solutions that can overcome these difficulties should help the early stages of new and evolving attacks that will help analysts to use their time effectively, and to reduce the reaction times between attack detection and attack prevention, with extremely low false positive rates. An Artificial Intelligence based cyber security platform can provide these solutions. The purpose of using AI in early warning and intrusion detection is to develop an intelligent assist system to detect attacks from the Internet as early as possible on both local area networks and wide area networks.

USEREXPERIENCE-AI-16: To subvert the prestige of the organization, the user took actions which were potentially harmful to the organization; e.g., unsanctioned data transfer or sabotage of resources.

To cope with the problem, we can use a system employing an online deep learning architecture that produces interpretable assessments of anomaly for the task of insider threat detection in streaming system user logs. The system models “normal” behaviour and uses anomaly as an indicator of potential malicious behaviour.

The system supports the streaming scenario, allowing high volume streams to be filtered down to a manageable number of events for analysts to review.

USEREXPERIENCE-AI-17: The user utilized the traditional antivirus that depended on signatures of known malware variants and did not help him to cope with zero-day malware – new viruses that are not yet known to researchers or antivirus providers and malware "mutations" where a virus is deliberately modified to evade detection.

AI-based approaches can be applied to analyze malware source code or activity to understand if the software is legitimate or may be doing something malicious.

USEREXPERIENCE-AI-18: When organizing a cyber defense system at his enterprise, the user (administrator) decided to use old, well-proven software tools for cybersecurity.

AI algorithms can be implemented in applications to identify and respond to cyber-attacks before they take effect. The assessment of a real-time system from a security point of view to determine the cyber vulnerabilities proves to be challenging due to the fact that the current conventional technology is computationally expensive and slow for systems such as the power grid, transport, IoT, IIoT, maritime, healthcare (where the time is much expensive). The machine learning techniques, having pattern recognition capabilities and high-speed learning, to be most suitable for security assessment of these systems.

4.2 EWS Platform

4.2.1 Vulnerability and threat management

As the main focus of the E-EWS is to provide an early feedback to involved partner about new and most used vulnerability in-the-wild, the most recent and active threat and TTP (Tactics, Techniques and Procedures) of cyber criminals whit the aim of prevent or detected cyber-attacks before that happen. For this reason, the process of collecting information and data is critical for give an effective early feedback to the users.

For gathering as much information as possible, the E-EWS need to interact and accept data from different sources (OSINT and CLOSINT). It is also clear that not all the information can be trusted only by the source, but they need also a human validation in order to avoid false positive or incorrect data. For this reason, the E-EWS must have the chance to upload information both in an automatic and manually way. Information collected via manually way have to be consider more reliable than the other collected via automatism and the users of E-EWS need to know how much an IOC (Index of Compromise) or a security repot about a new vulnerability are reliable.

The two use cases (OSINT ad CLOSINT sources) can be distinguished and analysed separately

OSINT

OSINT (Open Source INTelligence) sources as a public available source were the information is collected, exploited and disseminated in a proper way and with determinate time in order to respond of an intelligence requirement.

In the case of E-EWS, and most in general for every intelligence service dedicated for cyber-security, OSINTare all the resource in all the public documents and reports about cyber-attacks, new vulnerability and threat actors profile and TTP.

E-EWS must have the capability to gather data from the public available sources both in automatic and manual way. For the definition of automatism, E-EWS need to natively integrate the various sources via system like API or web crawling method. Otherwise E-EWS need to implement also a manual method to add data from those sources, in this way a human need to perform some action in order to add the data. In this case would be useful to add some kind of information validation process in order to give more value to the uploaded information. For example checking the age of the information, as if the data are too old maybe are not yet useful and can generate incorrect feedback.

E-EWS need also to provide a score of reliability for each source (in the case of information acquired via automatism) or for each data import (in the case of manually upload), in this way every user of E-EWS can have a comparison parameter of the information and could help in the process of decision making. For example if a user receive a feedback of a malicious email address (IoC) and he have to decide if block the address, block the entire domain or put in place some monitoring on it the reliable parameter will help the user in this process.

CLOSINT

In addition of public source there also close source (CLOSINT) that are information available form close source like internal analysis performed by cyber-security expert on suspicions file, monitoring dark and deep web communities, responding cyber-incident or conduct research on a particular group of cyber-criminal.

Otherwise, this information can also be gathered from a partnership with other cyber threat platforms that collect this kind of data.

It is clear that CLOSINT require more effort for collecting data and need dedicate resources that have to work on it but it also increases the value of the stored information:

- CLOSINT data are usually validated by a human.
- CLOSINT source are usually more reliable than the public.
- CLOSINT source are usually more up to date and take more care on what about is most used by attackers in the specific period.

Tailored early warning

Most of the modern attack exploit vulnerability that are already reported and fixed by the vendor but the owner of the system did not patch yet. This is usually related to the missing of a patch/vulnerability management but also because the vendor did not notify directly about the availability of the patch or the presence of a specific bug.

For this reason, the E-EWS have to give the possibility to the user to receive a dedicate feedback about the technologies that they use. In this way, a E-EWS user, can define the software and the systems adopted by their organization and receive a specific early warning when a vulnerability is discovered, a patch is released or there is an ongoing attack that are impact a specific system.

Users that receive this type of alert will take more care about the report and will apply quickly the patch, arrange dedicate monitoring and most in general react to a threat that can affect him directly.

Threat invalidation

In addition, of give early feedback about vulnerability and threat an Early Warning System has also to keep track about the invalidation of a threat. Every user can handle and early warning in the way which he believes is better (block IoC on his perimeter, put in place some dedicate monitoring, etc.) but he have to take in care that what is related to a threat today it may not be tomorrow. For this reason the E-EWS must implement a decay method for report and advise when an early warning have to be considered as deprecate.

There are various method for to this, one for example is to calculate a TTL (time to live) of the warning, in this way an user will know at moment of early warning reception until when he have to consider the threat. Another way to perform this review is to send or make available a revocation list in which every user can check if a specific threat can be consider deprecate or not.

4.2.2 Data Governance Model

In essence, the ECHO Early Warning System (E-EWS) is a system that collects, processes, stores, generates and distributes data and information. This means that proper data management is extremely important for its successful functioning. Some of the main characteristics of the data for which data management is responsible are the following:

- Quality – the data quality must ensure that they are correct, consistent and suitable for use and analysis;
- Availability – the data availability must ensure to the system that the data is easily accessible at any time by any business function that has the right to use it;
- Usability – the data usability must ensure that the data is structured, labeled, documented and easy to search and use by system-approved plug-ins and tools;
- Integrity – the data integrity must ensure that data retains its quality when stored, transferred, viewed and / or converted.
- Security – the data security must ensure that the data are classified according to their sensitivity. It also defines the processes related to their protection against loss or leakage (theft).
- Ownership – the owner of the data has full rights over them. It may delegate full or partial rights over them to other users, programs, processes and services.
- Accessibility – the data access defines the rules according to which certain users, programs, processes and services have the right to use this data.
- Transfer of ownership – this is a process of changing the owner of a given data.

The main elements that influence the Data Governance Model are Cyber Data Gathering / Sharing and User Data Management.

Cyber Data Gathering / Sharing Governance Model

One of the main functions of E-EWS is the collection and sharing of cyber data and information. For this process to be beneficial, the data must be accurate, up-to-date, relevant and addressed to the right users. In order to achieve this, it is necessary to ensure the following processes:

- Keeping structured / unstructured data up to date;
- Adding new source of structured / unstructured information;
- Updating of structured information;
- Removal of a source of structured / unstructured data;
- Adding a ticket;
- Automatic update;
- Automatic generation of tickets (based on predefined rules) by external / internal sensors;
- Automatic sharing and notification.

Keeping structured / unstructured data up to date

A single download of data from a source is not enough to ensure the reliable operation of the E-EWS. Over time, the data becomes obsolete and needs to be updated. This means that new data from the source, whether structured or unstructured, must be added to existing data. There are two approaches to when this happens:

one is to update at a certain time, and the other is to start at a certain event, requiring an update. It is best to combine the two approaches – to have periodic scheduled updates, but to be able to initiate an additional update when necessary (for example, in the presence of a new vulnerability that is highly threatening, and at the same time the scheduled update is after an unacceptably long time, which puts the users of the system at risk). Scheduled updates should have a well-chosen execution period, because if they are too infrequent, this will lead to work with outdated data, and if they are too frequent – this will lead to unnecessary load on system resources.

Adding new source of structured / unstructured information

Another important approach in keeping data up to date and reliable is the ability to add new data sources (structured and unstructured) that have not been used so far. In order to retrieve this data, as well as to convert it to a suitable format for E-EWS, it may be necessary to process existing plug-ins and / or system tools or even to create new ones. Another important thing in this case is whether these sources are external or internal to the ECHO network. When the sources are internal, the data from them will be easier to bring to the required format and the sources themselves will have a higher degree of trust (reliability). On the other hand, external sources, although they will require more effort to retrieve and process data, provide much more information due to the limited amount of internal sources. For both types of sources (external and internal) when they are added as E-EWS data sources, they should be periodically updated, as already described above.

Updating of structured information

Another type of update related to structured data is when you need to change the data structure itself (add a new data field or modify an existing one). When doing this, it is very important that it does not lead to loss of information. Appropriate adjustments should also be made to plug-ins and tools related to periodic updates of data from these structured sources of information so that the update is in line with the new structure.

Removal of a source of structured / unstructured data

In certain circumstances, it may be necessary to remove a data source, be it structured / unstructured, internal / external to the ECHO network. The reasons for this can be various, such as: the data source no longer exists (its support /update by its creators has been suspended); the data source has a new version or has been merged with another; the data source is already unreliable or has been compromised in some way, etc. There are several ways to act in these situations:

- data recorded so far from this source are left in the database and used by E-EWS, but no new updates are made from this source;
- data recorded so far from this source are marked as low reliable or unreliable, i.e. if they are used in the future, they must be subject to additional checks or comparisons with data from other (reliable) data sources;
- data recorded so far from this source should be completely deleted from the E-EWS database.

Adding a ticket

Tickets are the other main source of structured data for E-EWS. In order to keep up with changing cyber security needs, they must be easily modified as needed. This means that it must be possible to add new fields and / or modify existing ones. The same should be done with ticket facets, including the creation of new types of facets. If necessary, it should be possible to make changes in the work process related to ticket processing.

Automatic update

In the future, the system must have the functionality to check itself for new information that contains data and descriptions of attacks (e.g. CVEs or specialized WIKIs) in the context of the sectors in which E-EWS operates and with the help of appropriate tools, plugins and scripts automatically (without human intervention) to select and add new data to existing data.

Automatic generation of tickets (based on predefined rules) by external / internal sensors

An excellent opportunity to significantly improve the performance and capabilities of E-EWS is the development of capabilities for automated ticket generation (by program, process, service, script, etc.) through the use of external or internal sensors (based on IDS, SIEM, Firewall, Nessus, Suricata, Snort, OpenVAS, etc.). This would have a positive effect on various factors such as: speed in submitting the ticket itself; objectivity in assessing the circumstances surrounding the cyber event; elimination of the probability of human error in filling in the ticket, etc.

False-positive indications of cyber events would be a major problem in introducing the ability to automatically generate tickets.

Automatic sharing and notification

Another direction for the development of E-EWS is the integration and development of mechanisms for fully automated information sharing with the partners of the ECHO network, including the possibility for automated application of protection measures. For example, if a dangerous cyber vulnerability exists and the system has data on how to deal with it (for example, by adding certain rules to the firewall), then the system should be able to provide these rules directly to at-risk users, and if prior consent on their part, even if they put these measures into practice fully automatically.

User Management Governance Model

Users of data from the system can be organizations or people, and people in turn can be in the capacity of employees of an organization or be in the capacity of independent users (not under the umbrella of an organization).

Organizations, users of E-EWS

With regard to organizations, the cases when access to certain information is defined or changed are the following:

- Adding a new organization as an E-EWS user. When adding a new organization, the following things must be very clearly defined for it:
    - what access to the E-EWS services the organization will have – full or partial;
    - user of how many and which sectors the organization will be;
    - what data will be available – all, even those obtained before the accession of the organization, to data acquired only after its accession or some combined approach;

  - what data the E-EWS organization itself will provide and whether this data will be available to other users of the system.
- Modifying the conditions of an organization that is already a user of E-EWS. All the above characteristics must be able to be changed (renegotiated) at any time:
  - the organization's access to E-EWS services must be able to change from full to partial, from partial to full or to change the parameters of partial access;
  - the organization must be able to add new ones and subtract from the declared sectors in which it is interested;
  - the organization must be able, if desired or necessary, to change the scope of the data to which it wishes to have access;
  - the organization must be able to change the conditions under which it provides its data for use by E-EWS.
- Disable and enable an organization as a user of E-EWS data. It must be clear under what conditions, when and how an organization can be temporarily excluded as a user of the system and under what conditions, when, how and to what extent it can be restored as a user.
- Delete an organization as an E-EWS user. The mechanisms by which an organization will leave E-EWS and what happens to data received from it must be clearly regulated. In addition, when the leaving organization is not just a user of the system, but is part of its structure and work, it must be known in advance which other organizations will take over its functions and how the services related to these functions will be provided after leaving the partner.

People who use E-EWS as employees of an E-EWS-related organization

It is necessary to know who sets the rights of such users. The organization selects a user who will have the maximum rights of the organization as a user and he will create user accounts of other people in the organization and will determine what rights they will have or this will be done by people serving E-EWS on applications submitted by the organization. Also there must be a clear and fast procedure for changing the rights or deleting a user when necessary.

People using E-EWS who do not belong to an E-EWS affiliate

These are freelancers. They may be researchers, scientists, etc. As they are not part of an organization that has concluded a contract for the use of the services provided by E-EWS, their access to data and rights in the system should be regulated in a contract with their direct participation. It is also important to know whether non-organizational people will only have access to certain sectors (according to predefined needs) or will be able to use data from all sectors for which the system maintains information. You need to know the timeframe in which this will happen and possibly the extent to which they can use this data and to what extent they can share it with third parties and organizations.

4.2.3 E-EWS Scalability

Scalability refers to capabilities of every infrastructures or application to expand their capacity in order to handle increased load. In the specific case of E-EWS this need, it is bound to number of users and the integrations with third party systems. Having a scalable infrastructure and web application ensure that it, can scale up to handle the load and not crash or lower the user experience of E-EWS. Another factor is that various plugins can interact with different layers of the organizational infrastructure, generating significant load on the network. Thus EWS instance location can be factor in the platform performance.

The scalability problem can be addressed in two different ways.

Vertical scaling

The vertical scaling is the process of adding resource to an existing instance in order to increase the performance. Usually this is considered the easier and simpler approach for handle the scalability problem. In this case the problem is addressed by increasing the power of the used server instance by adding more memory (RAM), increase disk capacity and performance (using SSD disks) or increase the computing performance (CPU). The reason of thought that vertical scale is an “easy approach” is due to the fact that upgrade this component is easy, especially in virtual environment where you do not need to physically upgrade server components.

Horizontal scaling

Once the simple vertical scaling is limited due to hardware limitations or geographical requirements horizonal scaling approach should be taken. Scaling an system horizontally mean add more server instances or node to the infrastructure in order to spread the load across multiple machines. This process is more complex than vertical scaling because add more node involves extend the perimeter that result in a more complex management of the servers and the related security. Another issue is if different modules of E-EWS are spread across different servers (this is the typical scenario if there is requirement for E-EWS collecting massive amounts of data from organizations infrastructure). In such cases special attention needs to be paid to the communications between the modules, as factor such as communication protocols, bandwidth, delay or even jitter of the link between the different servers can hinder or break the functionality of the system,. Moreover even simple task like backup or software update can now became more complex and requiring more effort to be executed. However horizontal scaling is considering a long-term advantage instead of vertical scaling that is often used when you need to upgrade the performance quickly.

One of the specific cases of horizontal scaling is utilizing cloud resources by either completely implementing E-EWS in the cloud or adopting hybrid scenario where platform partially operates on premise (i.e. due to legal requirements) and partially – on cloud server (due to financial/efficiency considerations) – i.e. in case of organization that deals with classified information, typically EWS core have to be on premise. This is not true for all plugins, especially the Knowledge providers, as they are mainly data collectors. Such hybrid setup would be one of the most technologically demanding and special effort should be directed to enable such capability by extensively developing communication resilience between the E-EWS modules and also testing the platform for various communication degradation scenarios between the modules.

4.2.4 Certificate Authority Manager:

E-EWS in its core is information sharing platform. As the sharing is done between legally separate organizations, that would be in different geographies, environment and may not even established formal or informal contact between them with the exception of E-EWS, such communication can become prime target of an attacker. Such attacker can inject false information in the data stream of E-EWS and in such way to sabotage operations, perform attack of the “Snowblind” type or even seek to infect workstations of SOC team with malicious software.
One approach to address this concert is to establish central trust authority, that extends the chain of trust to participating organizations in managed and verifiable way. Technologically this can be implemented by creating centralized certificate authority manager under the control of ECHO governance organization.
This authority will create the certificates for each E-EWS Node. Those certificates can be utilized in digital signature to assure identity and enable other nodes in the network to trust the newly installed one. Alternatively

the certificates can be used in establishing encryption tunnel that is protected from eavesdropping even during the initial setup (Perfect Forward Secrecy)
The Central Certificate Manager should also provide other related services, for example:

- Certificate Revocation List to ensure that in event of compromised certificate, the network can be protected
- Public Key Directory to ease the distribution and crypto

Further development of this centralized authority manager can be controlling the instances of E-EWS and ensuring better maintenance of the E-EWS network, increased fee collection (if E-EWS is implemented as paid service) etc.

4.3 E-EWS Exploitation

4.3.1 Sustainability

When creating an E-EWS system, servers that store databases of ongoing attacks and failures to the system are used. Access to the cloud infrastructure of different companies is assumed, client applications must be installed. In order for the information in the databases on the servers to be relevant and useful, it is necessary that users of the system perform timely notification of ongoing events (failures, accidents, cyber-attacks) in their company. The sustainability of the system depends on whether the administrators of the companies' networks that are connected to the EWS correctly enforce security policies, whether they troubleshoot network equipment failures, keep logs, patching, backup data, etc. Also on whether the software of the network devices and the server are updated, as they may contain vulnerabilities.

Consider the situation of connecting to the EWS network of a new company. The employees of this company are only planning to improve their qualifications and undergo training and certification in the settings of network equipment, access policies, and work with databases.

The router has failed and no virtual circuits (VPN) have been configured. The stability of the system is under the question. In this case, the company is recommended to set up virtual channels of information transfer and have backup equipment in case of its failure.

A situation may also arise if a user from any organization that connects to the E-EWS has not configured security policies and has not enabled firewalls on the network equipment. Now the information that is transmitted between the nodes over the network is not crypto-protected, and the network equipment can be subject to cyber-attacks.

A user from the organization did not make sure that server data were being completely backed up on a regular basis or regular backups were being created but the recovery images have not been tested to work as expected. In this case, some server data can be not recovered and the organization can have the problem with ensuring fault tolerance and contingency against accidents and unexpected damage of critical data.

An organization was dealt with an unauthorized access and an organization’s network was hacked but hacker was not identified because of low level of log management. It blocks access to the local database or modifies data in it. Some steps must be taken to access the shared database. A user used system log storage only and did not use infrastructure log storage. During an incident, the EWS’s logs were changed by attackers. Incident response staff can’t use the data from the infrastructure logs to compare the infrastructure and system logs to determine what data was changed or removed. Also, a user from the organization who was responsible for log management did not use log compression and did not use event filtering, or did not use log clearing, and it decreased the amount of storage space.

A user from the organization tampered with start-up processes, such as antivirus software and certain server scripts and these processes are not configured correctly now.

Also, there may be a violation of the sustainability of the system due to malicious actions on the part of company employees or intruders who entered the company office. A user from the organization who was responsible for the server room did not make sure keys to the server room were kept secure. An attacker got copies of these keys.

The user should take into account that the source of noise and interference from the rooms where the network equipment and the server are located must be removed.

The sustain operation of the system being created depends on many factors. The resistance of various subsystems to attacks and malicious influences leads to a more sustainability of the entire system. To ensure the sustain of the network, it is necessary to ensure the sustainability of its individual hardware components (servers, workstations), software, network equipment.

For sustain data transfer between network nodes, it is necessary to eliminate sources of noise and interference, ensure the sustain operation of network devices and a connecting cable (wireless network), configure security policies, firewalls, virtual channels on each node and server, and use encrypted data transmission.

Each organization that will access the replicated database and cloud storage must provide for the sustain operation of network components and their cyber protection.

Also, it is necessary to control user access, automatically download updates and perform patches, and provides for the network sustainability. It is necessary to provide measures to protect from cyber-attacks databases, network equipment, server from malicious influences, human errors in settings, hardware and software failures. Considering that hardware and software trojans and spyware can be embedded in software and hardware, it is necessary to provide testing for their absence.

The software of the network equipment (router, software-controlled switch) allows you to configure it so that it functions sustain, with configured protection against cyber-attacks. It is also necessary to patch the firewall software (software and hardware / software).

Some steps user (administrator) have to do to make E-EWS system sustain:

1. The user and the network administrator must provide for the organization of VPN networks to access the main server.
2. The user must configure the maintenance and storage of information about events in a special logbook.
3. It is necessary to provide for data backup in the databases and backup of the databases themselves to ensure the integrity of information, the implementation of snapshots and the use of backup media.
4. Provision must be made for the correct configuration of security policies (certificate issuance, authorization, authentication and accounting) in Server Manager.
5. It is necessary to provide redundant hardware components for scalability and convergence of networks when connecting new organizations, such as: servers, network devices, RAM and hard drives of servers in the cloud.
6. It is necessary to take into account that the system may have vulnerabilities that must be eliminated through periodic software updates, patching. This procedure can be performed automatically or by a user (administrator).

7. It is necessary to exclude all sources of interference, noise, which can affect the quality of information transmission in the network.
8. It is necessary that the organization that connects to the system has an installed client application, access to the Internet (regardless of the information transfer technology - cellular, satellite, using fiber-optic, copper cables, wireless) to share information about the attacks that have occurred, and access to databases of attacks.
9. At each level of the system, security policies must be configured correctly.
10. The staff of company (users, system administrator have to make higher their qualify by certification and trainings).

USEREXPERIENCE-S-1:

A user from an organization that connects to the E-EWS has not configured security policies on the network equipment, has not enabled firewalls. Now the information that is transmitted between the nodes over the network is not crypto-protected, and the network equipment can be subject to cyber-attacks. The stability of the organization's network is now in question.

It is necessary to install firewalls, apply anti-virus programs on the nodes, configure security policies on switches and routers, and programmatically close the ports most frequently attacked. Also, for the stable operation of the network, it is necessary to provide backup power supplies, uninterruptible power supplies. It is necessary to provide for the absence of sources of noise and interference, it is necessary to update the software installed on the local server, network equipment and nodes. It is also necessary that only trained qualified personnel are allowed to work with the system, the rest must set the access time and limit the resources available for access.

USEREXPERIENCE-S-2:

User from the organization connecting to the system did not ensure the stability of the server in the organization that connects the E-EWS system. The user did not update the software in time, which could contain vulnerabilities, and did not patch it. Attackers can successfully attack a server through these vulnerabilities.

To solve this problem, it is necessary to update the software in a timely manner and patch its vulnerabilities. Otherwise, an attacker can enter the server settings with administrator rights, download databases, and even programmatically disable the main server components, or disable them, as well as make requests to the main server and try to attack it.

USEREXPERIENCE-S-3:

A user from the organization did not make sure that server data were being completely backed up on a regular basis. In this case, the organization can have the problem with ensuring fault tolerance and contingency against accidents and unexpected damage of critical data.

To solve this problem, all critical data should be included in the backup process and backed up in at least three separate locations. The process for securely backing up the data can be as: check backup logs to ensure regular backups are being created, backup server data onto local drive, backup server data onto secondary local drive, ensure secondary drive is stored at a secure separate location, and backup server data to a cloud-based storage.

USEREXPERIENCE-S-4:

A user from the organization ensured regular backups were being created but did not test that the recovery images was working as expected. In this case, some server data can be not recovered.

To solve this problem, the user should test that the recovery images is working as expected. The test process can be as follows: take three of the most recent backup images, and attempt to access data from all three backup images. If there are problems with the test images, then the user should perform extensive testing to get to the route of the problem. This may include re-making and re-testing system-wide backup images or switching the backup process that currently in use to a new one. For example, the user should: troubleshoot the whole backup process, evaluate what might have caused the problem, test alternative solutions to the problem at hand, and update the backup process accordingly.

USEREXPERIENCE-S-5:

An organization was dealt with an unauthorized access but hacker was not identified because of low level of log management.

If your organization dealing with an unauthorized access, collect all app server logs, application logs, web server logs, database logs, firewall logs, switch or router logs, and any other logs where an authorization was present.

USEREXPERIENCE-S-6:

An organization's network was hacked but hacker was not identified because of low level of log management.

In case of a network hack, collect logs of all the other devices found in the route of the hacked one. ISP router logs are also useful.

USEREXPERIENCE-S-7:

The attacker successfully implemented a cyber-attack that blocks access to the local database or modifies data in it. Some steps must be taken to access the shared database.

Creation of logical channels of information transmission (virtual private networks – VPN) allows protecting network users from a huge number of different types of attacks. Therefore, most of the well-known router manufacturers provide for the possibility of Virtualisation and the use of virtual links. To ensure system sustainability, you must properly configure virtual links.

4.3.2 User data collection management

The General Data Protection Regulation (GDPR) is a EU law that requires organizations to safeguard personal data and uphold the privacy rights of anyone in EU territory.

This safeguarding requires organizations to meet both technical and organizational standards regarding personal data. For the moment, this document will simply discuss some questions about where difficulties in meeting these standards may arise, where plans to meet them and for future developments should be considered and put in place and a general overview of how to meet some of these challenges. The questions, plans, possible future developments, etc. are expected to be detailed and expanded.

From the technical side, the first thing to note is that the GDPR requires that systems dealing with the personal data of EU persons must be developed with security and privacy in mind. This means that best industry

practices in privacy and security should be implemented and documented in the earliest stages of product development. A technical privacy and security assessment should be conducted, to include such items as:

- methods of communication used,
- methods of login used,
- database systems used,
- methods of encryption used,
- allowable passwords,
- location of servers and databases,
- methods of pseudonymization or aggregation used,
- implementation of roles / data access in system,
- backup procedures, etc.

These should be well documented and comply with industry standards. It is, however, quite likely that the regulations on privacy (and technology) will also evolve in the future. And these evolutions should be predicted and planned for to the extent possible. A very simple example of this is not using the very lowest level of encryption possible. Another might be maintaining a flexible system of roles in the event of changes in the future requiring a reorganization or differing kinds or levels of information access.

As far as organizational safeguards, clearly each organization is responsible for its own internal compliance with GDPR, however, there are questions which the EWS raises.

One of the most significant is how data is shared with parties or partners outside the EU. What information is shared? How is it aggregated or pseudonymized? What plans are there for deletion of the data should that be necessary? What data on EU persons can be shared outside the EU and how? What data on persons outside the EU (or perhaps non-EU persons) is collected or shared? How is such data stored, handled, processed, etc.? Also the matter is made more complicated after ECJ rule to invalidate Privacy Shield as of July 2020.

Again, these questions should also be answered with a view to the future – where member states of the EU may change, the meaning of "EU person" may change or the law may simply apply to all personal information stored by an EU entity.

Company Data

Besides compliance with GDPR for personal data and privacy, there is the question of how to manage and process the data of the various organizations taking part in the EWS. Even assuming that the GDPR gives no protection to this kind of data (not being of private individuals), many organizations will desire to know how the information they provide will or could be used.

Clearly, at one end of the spectrum, the data they share would give complete information about the organization practices and structure – which most organizations would not consider ideal. While at the other end of the spectrum, organizations might obtain information, but provide none – giving almost no information about their internal structure or practices, but not benefiting the EWS except monetarily.

While it may be possible to build in various types of memberships for the EWS, what information is shared at each of those levels and an honest assessment of how it might be used to profile an organization (taking into account organization size) may go a long way in allaying potential concerns and helping to build the network for the EWS.

As with personal data, how, where and by whom the data is stored, processed, communicated should be documented and made very clear – as well as any backup procedures and requests for deletion.

4.3.3 Training monitoring tool

As every new technology on the market, striving for wide adoption, E-EWS can benefit from training. The exact approach, curriculum, cost etc of the training is to be determined by WP3 ECHO Governance work group with support of T2.6 “Derivation of ECHO Cyberskills Framework and related trainings” and possibly T2.7 “Certification schemes” should E-EWS people certification be defined. In any of the above approaches however technological tool to monitor trainee performance can be extremely beneficial.

By implementing as new plugin, utilizing existing API and programing infrastructure, such tool should be able to visualize, monitor and measure the activities performed by trainees. At least three modes of monitoring/assessment should be supported:

- Table-top exercise
- Training exercise using one E-EWS instance
- Training exercise using multiple E-EWS instances, including such located in different organizations

Optionally the training monitoring tool can provide dashboard or interface to all participants in ECHO network and some sort of gamification to be implemented, following the model of Gamification in Cyber Ranges.

5. Conclusions

ECHO Consortium has identified several major technology domains for the potential development of E-EWS and E-FCR platforms – each discussing specific technology sector or aspect of the performance, user experience, reliability and expansion. Attention was also paid to new functionalities that can expand the reach of the platforms and in such way, create an opportunity for growing the ECHO network.

Several chapters of this document focused on the advancement of one of the main E-FCR goals – providing training and education, by exploring practical training framework, possible future development for adversary simulation capabilities, so that the training environment can be built upon advanced realistic scenarios and AI could be used for training scheduling. Other user stories explored future development of the concept of the federation of cyber ranges in the aspects of tighter integration between the platform and the different CRs using innovative state-of-the-art network technologies, automation and also simulation description language that can become the future standard for CR simulations and be used by multiple actors in the ecosystem.

Tighter and better integration between the two ECHO platforms was also studied with several use cases that explored the synergies within ECHO project and network of companies. Other user stories concentrated on the expansion of the platform functionality by introducing vulnerability and threat management and response-aware AI suggestions to the operators. Platform scalability and better, more secure integration of E-EWS network of instances was detailed, as well as the study on sustainability, training and data collection management and governance.

Apart from technical aspects, the two innovative platforms also bring net legal and privacy issues that need to be addressed in order to create sustainable operation and provision of services in essentially global marketplace.

Once the new development resources are available for E-EWS or E-FCR platforms, either by addition of new ECHO partners willing to invest or by securing funding, the derived opportunities for development should be discussed. Based on specific needs and available resources, appropriate functionalities should be selected, and the software development plan should be created.

The next version of the document shall also include roadmaps for technology prototypes, developed by T4.3 “Early prototypes selection, research and development”.

6. Annexes

6.1 Annex 1 – Customer Feedback form for E-FCR

Proposed customer feedback form structure:

Proposed Structure

Are you an expert in cyber range tool?

- ○ <1 year
- ○ >1
- ○ >5
- ○ >10

How easy was it to configure and use E-FCR?

	1	2	3	4	5	
Very Difficult	○	○	○	○	○	Very Easy

How quick was the inscription and configuration process?

	1	2	3	4	5	
Slightly quick	○	○	○	○	○	Very quick

How user-friendly is the E-FCR's interface?

	1	2	3	4	5	
Slightly user-friendly	○	○	○	○	○	Very user-friendly

What or which functions have you liked most of the prototype? Why?

How successful is our software in performing its intended task?

	1	2	3	4	5	
Slightly successful	○	○	○	○	○	Very successful

How useful is the included documentation for our software?

	1	2	3	4	5	
Slightly successful	○	○	○	○	○	Very successful

How can we improve our software?

How likely are you to recommend our software to others?

	1	2	3	4	5	
Slightly likely	○	○	○	○	○	Very likely

6.2 Annex 2 - 5G technology concepts

Fifth generation networks are in the final stages of standardisation, but we can already find commercial deployments with an almost complete set of functionalities. Compared to the current 4G Long Term Evolution (LTE) technology, 5G technology will offer high speeds close to 1 Gbps; low power which will mean much lower energy consumption, allowing the deployment of sensors that will not require power supply and whose lifetime can be several years; and above all low latency (1ms or less) which will allow the deployment of services that require real-time information such as voice communications or autonomous car guidance.

In our immediate environment, the European Commission's analyses forecast that the estimated benefits of introducing 5G in four production sectors (automotive, health, transport and "utilities") would progressively increase to reach 62.5 billion euros of direct annual impact within the European Union in 2025, which would amount to 113 billion euros adding up the indirect impacts.

5G networks will facilitate:

- Very high speed and capacity mobile broadband, with mobility speeds above 100 Mbit/s and peaks of 1 Gbit/s.
- Ultra-reliable and low latency communications, around 1 millisecond (ms) compared to 20-30 ms for 4G networks. This condition could make them suitable for applications that have specific requirements in this area, such as the connected or autonomous vehicle, telemedicine services, security systems and others such as intelligent manufacturing.
- Massive machine-to-machine (M2M) type communications, including the Internet of Things (IoT). The capacity to manage simultaneous connections will be increased, allowing, among other things, the massive deployment of sensors.

With the arrival of 5G, another set of paradigms are enabled at the same time, such as the extension of cloud computing towards edge computing and the Virtualisation of network infrastructures that will no longer be associated with a given Hardware(HW)/Software(SW) but will become a functional software element running on cloud servers.

The International Telecommunication Union(ITU) has classified the services of future 5G networks into three categories: Enhanced Mobile Broadband(eMBB), Ultra Reliable Low-Latency Communications(uRLLC) and Machine Type Mass Communications(mMTC).

- eMBB aims to meet the demands of an increasingly digital lifestyle and focuses on services that demand extremely high bandwidths such as high definition (4K) video, virtual reality (VR) and augmented reality (AR).
- uRLLC aims to meet the expectations of demand from the digital industry and focuses on services that are sensitive to latency and reliability, such as assisted and automated driving, and remote management or real-time robot control or coordination
- mMTC aims to meet the demands of a more developed digital society and focuses on services that include high connection density requirements, such as smart city, smart metering and smart agriculture.

Added to all of the above, the 5G network will, if necessary, promote more efficient intelligent production. According to the vision of the transformation of industry 4.0, the factories of the future will be based on cybernetic systems. These systems will integrate computing, networks and physical processes to improve the ways in which manufacturing companies operate. The entire manufacturing supply chain will be largely interconnected via wireless networks capable of meeting demanding bandwidth and latency requirements.

Data will be shared between different locations on key aspects of the business, such as design, manufacturing and distribution. The factories will be populated with robots with high manufacturing capacities and densely equipped with sensors and automated systems. Manufacturing on demand will increase, and flexibility and efficiency will improve.

The 5G network will also allow the simultaneous increase of the number of home broadband users in the future by means of the Wireless Broadband Access(FWA) solution combined with the use of the millimetre wave band (mmWave), offering home users broadband services similar to those offered today by optical fibre and will increase up to 38 times the number of concurrent users with 4K video services (compared to the service of a current LTE network). These services, however, can start to be provided with 4G technology (LTE-Advanced and Advanced Pro) with the application of massive MIMO, in bands below 6GHz, although the performance is not as good as in mmWaves due to the lower bandwidth available.

The scenarios considered by the ITU for 5G respond to different needs and therefore also have different performance requirements for the technological solution that supports them:

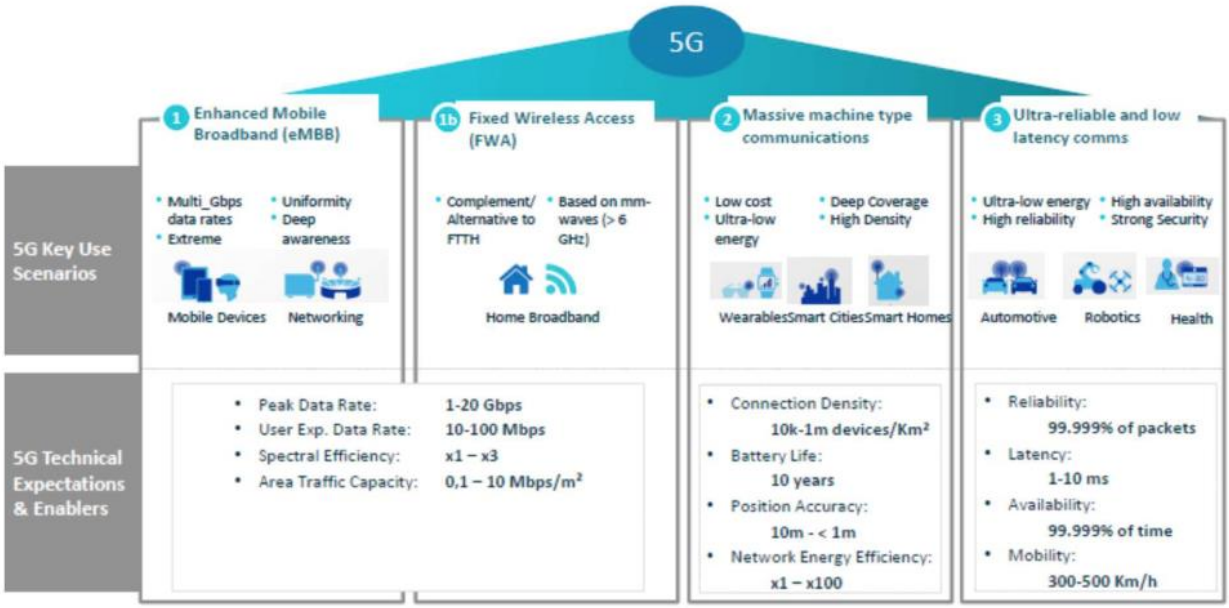

Figure 25: ITU 5G key use scenarios. Image taken from: [1]

The extension of 5G will be articulated, to some extent, on a Small Cells architecture. The deployment of small outdoor cells (canopies, street lights, facades, etc.) is not a new concept. It already existed in 2G and 3G networks.

This solution is appropriate in environments with high traffic concentration as a means of increasing network capacity. The very evolution of traffic in the coming years does not suggest a significant demand for small cells, except in very localized environments such as large cities or tourist areas.

On the other hand, the coverage that a small cell provides indoors is small because of the power with which they transmit and because they use high bands. It is more effective to provide a higher density of cells in the macro network as a first option.

In addition, proper location selection is necessary, because placing the small cell on one corner of a street or another can have very different results in capturing traffic due to the effects of propagation and obstacles.

Between 2021 and 2025 a very important growth in traffic is expected, which, depending on the context, could force an increase in the deployment of small cells (4G and 5G).

The greatest capillarity would be reached indoors (80% of mobile network traffic is indoors), both in public or office environments and at home. This type of solution is already being deployed in commercial areas, offices, airports, etc. as a need for coverage and capacity. Similarly, these environments will be the most suitable for an initial deployment of the 5G small cell. In the residential environment, there are femto cells (small cells with less power) that will become 5G solutions in the future, extending the technology to the interior of homes.

Software-Defined Networking and Network Function Virtualisation

Network softwarization seeks to transform the networks using software- based solutions. Through technologies like Software-Defined Networking (SDN) and Network Function Virtualisation (NFV) which will be absolutely crucial and essential elements in 5G. SDN and NFV can provide the programmability, flexibility, and modularity that is required to create multiple virtual networks, each tailored for a given use case, on top of a common network. These logical networks are referred to as network slices [2] which will be explained in the following section.

NFV and SDN will be absolutely crucial and essential elements in 5G. NFV and SDN will also enable the reduction of network equipment costs and automate their management improving efficiency, as well as allowing the entry of new actors in the ecosystem capable of providing purely virtualized network elements.

NFV/SDN are operational management levers for both the core (Virtual Core) and the access (Cloud RAN) or edge of the network (Cloud Edge) and are key not only to improve current performance but also to enable functionalities that will only exist in future 5G networks, allowing different capacities to be delivered to different verticals in an agile and dynamic way, supporting traffic growth and connected devices that are difficult to handle with existing architectures. In short, it guarantees sufficient versatility and smooth evolution to support new cloud services such as edge computing, the Internet of Things (IoT) or the concept of Networks as a Service (NaaS), all of which are supported by 4.5G and 5G networks.

The next figure shows the core of the SDN architecture based on ONF TR-521, “SDN Architecture”

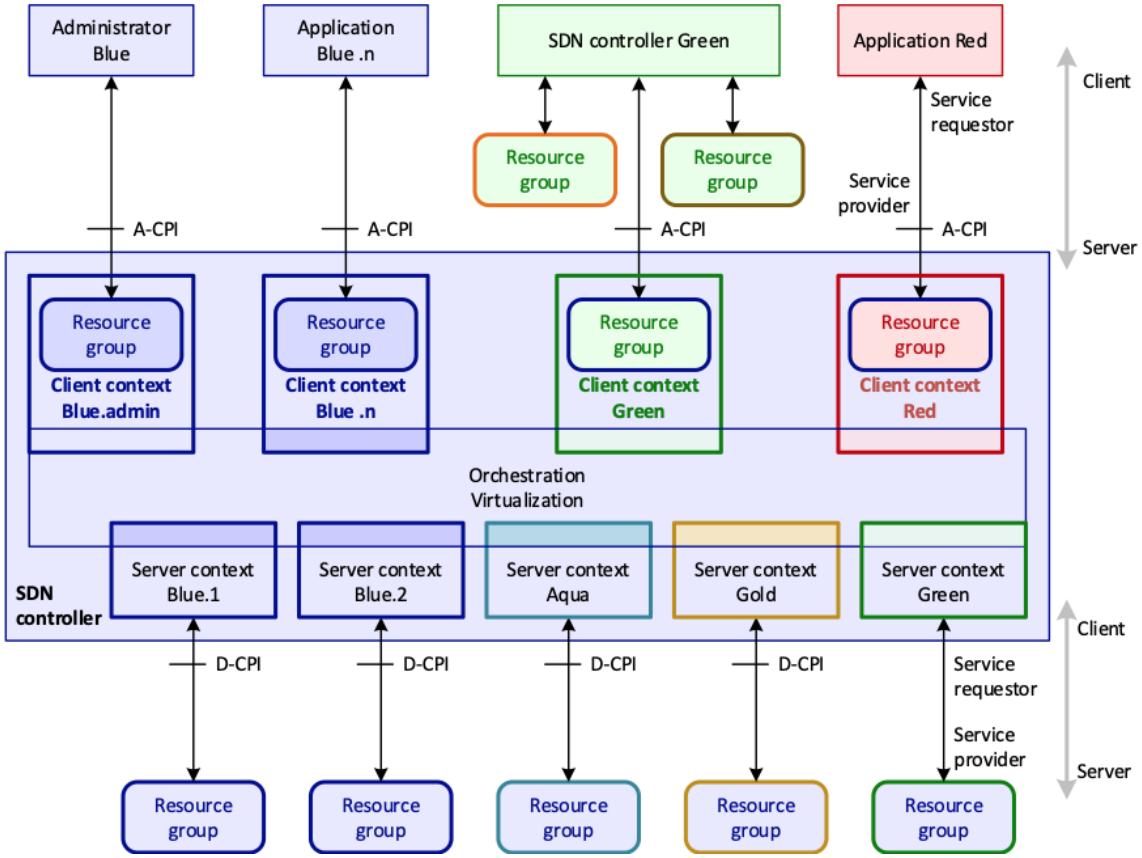

Figure 26: SDN Architecture. Image taken from [3]

The next figure shows the Telefonica's end-to-end virtualized network vision to illustrate the concepts:

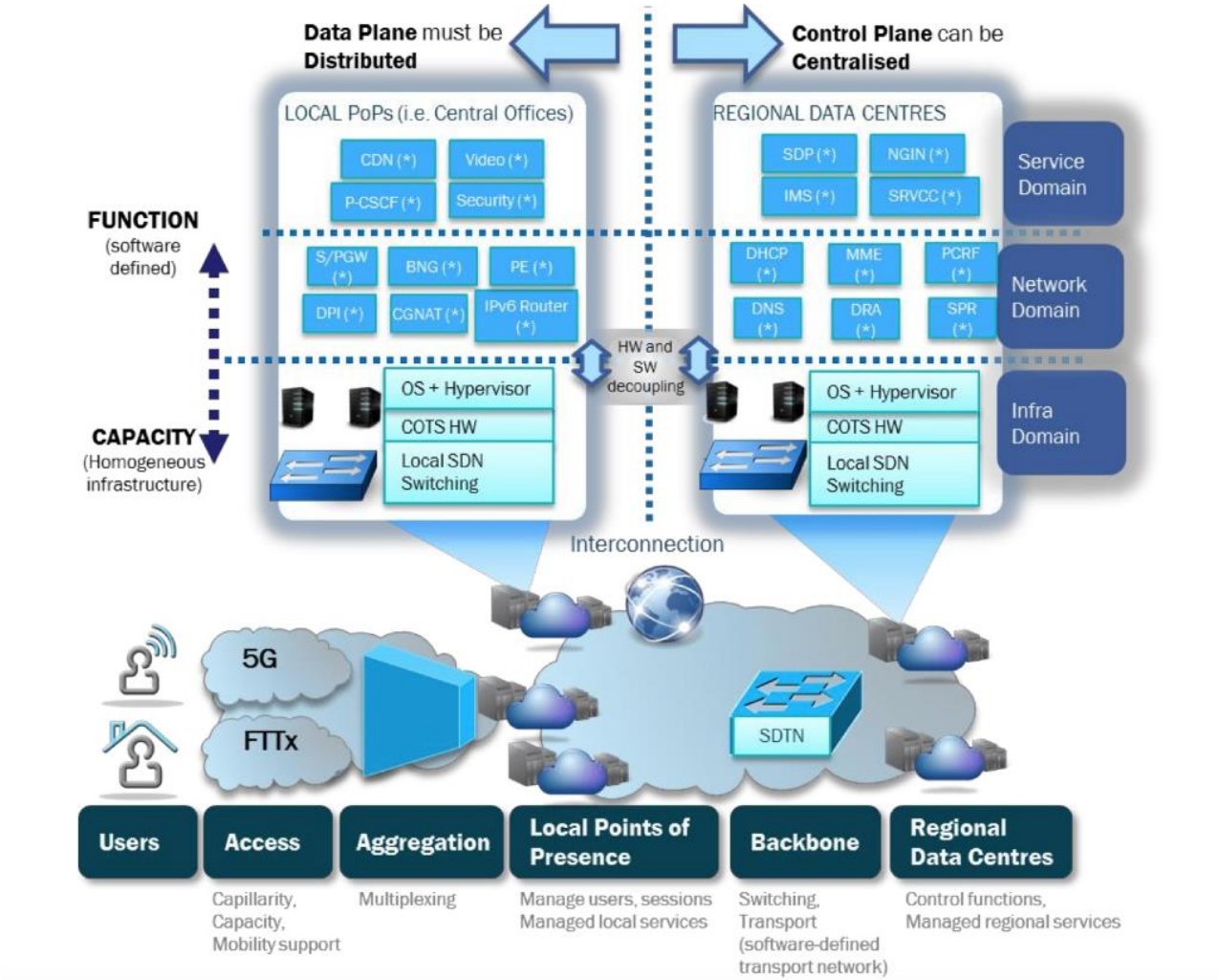

Figure 27: Telefonica's end-to-end virtualized network vision. Image taken from [4]

Network slices

SDN and NFV will be facilitating elements for the implementation of Network Slicing, which represents a milestone in network architecture and will make possible the attention of multiple verticals and use cases on a common platform and with a centralized network management based on software.

The concept of separate virtual networks deployed over a centralized network, although not new (e.g. VPN), has specific aspects that make network slices a novel concept. We define network slices as end-to-end logical networks running on a common underlying (physical or virtual) network, mutually isolated, with independent control and management, and which can be created on demand to accommodate different business-driven use cases from multiple players on a common network infrastructure. In the next figure, it is shown 5G network slices running on a common underlying multi-vendor and multi-access network.

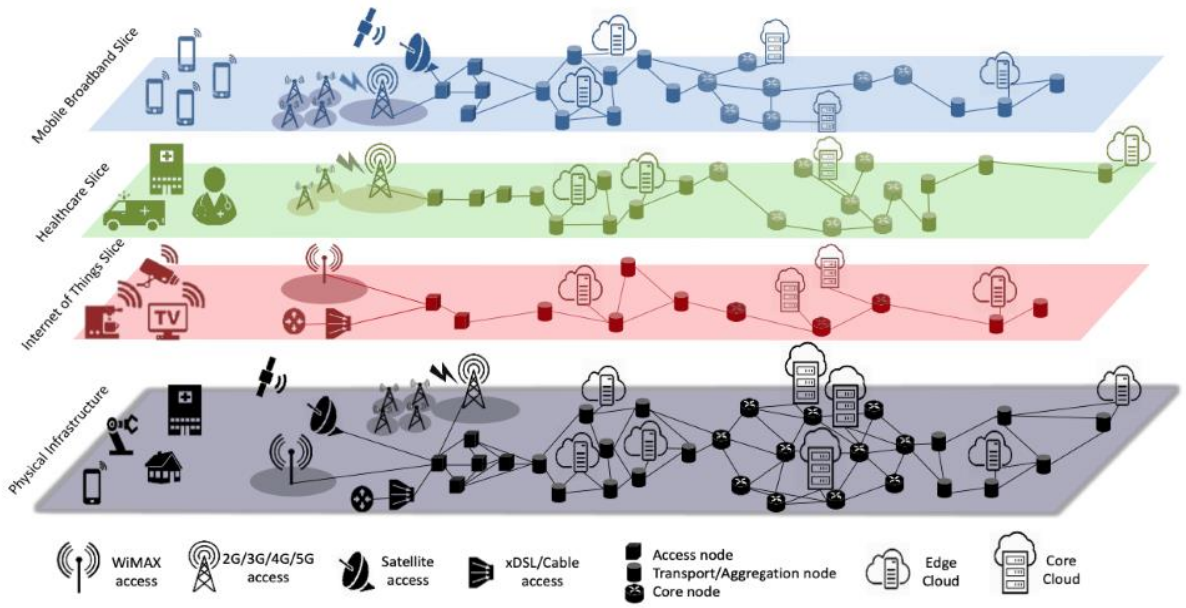

Figure 28 5G Network slices. Image taken from [4]

In network slices, there are 4 issues considered: resources, Virtualisation, orchestration and isolation.

A network slice is composed of a collection of resources (a resource is manageable unit, defined by a set of capabilities that can be used to deliver a service) that, appropriately combined together, provides the service requirements demanded by the use case. In network slicing, there are two types of resources:

- Network functions (NF): functional blocks that provide specific network capabilities to support and execute the particular service(s) that each use case requires. NFs are typically implemented as software instances running on computer resources. They can be physical (hardware and software sold on a physical device) and/or virtualized, which is a network function software, using general purpose hardware on which it runs.
- Infrastructure resources: term referring to the general-purpose hardware and software needed to host and connect NFs. These resources include both computer resources, network resources (switches, routers, etc.) and radio physical resources (antennas, etc.).

Virtualisation is a key process for network slicing as it enables effective resources sharing among slices. Virtualisation is the abstraction of resources; this is the representation of a resource in terms of capabilities that match predefined selection criteria with the aim to simplify the use and management of that resource in some useful way. The resources to be virtualized can be physical or already virtualized (recursive pattern).

The orchestration function is key for network slicing According to the Open Network Foundation (ONF) [6], "orchestration is defined as the continuous process of selecting resources to meet customer service demands in an optimal manner". Centralized management of orchestration is unfeasible because of its complexity and scope as a key function, but also because each network slicing requires independent management.

Isolation is a key requirement to operate parallel slices on a common shared underlying network. Isolation must be considered in terms of management (each slice must be independently managed as a separate network), security and privacy (each slice must have independent security functions that prevent attacks or unauthorized accesses) and performance (each slice is defined to meet particular service requirements- KPIs that the client is paying for). Isolation is achieved based on consistent policies and mechanisms to be tackled by each Virtualisation level.

The next figure shows the SDN-based slice abstraction based on [5]. As it can be seen, the SDN control plane can involve multiple hierarchically arranged controllers enabling recursion to support 5G slicing.

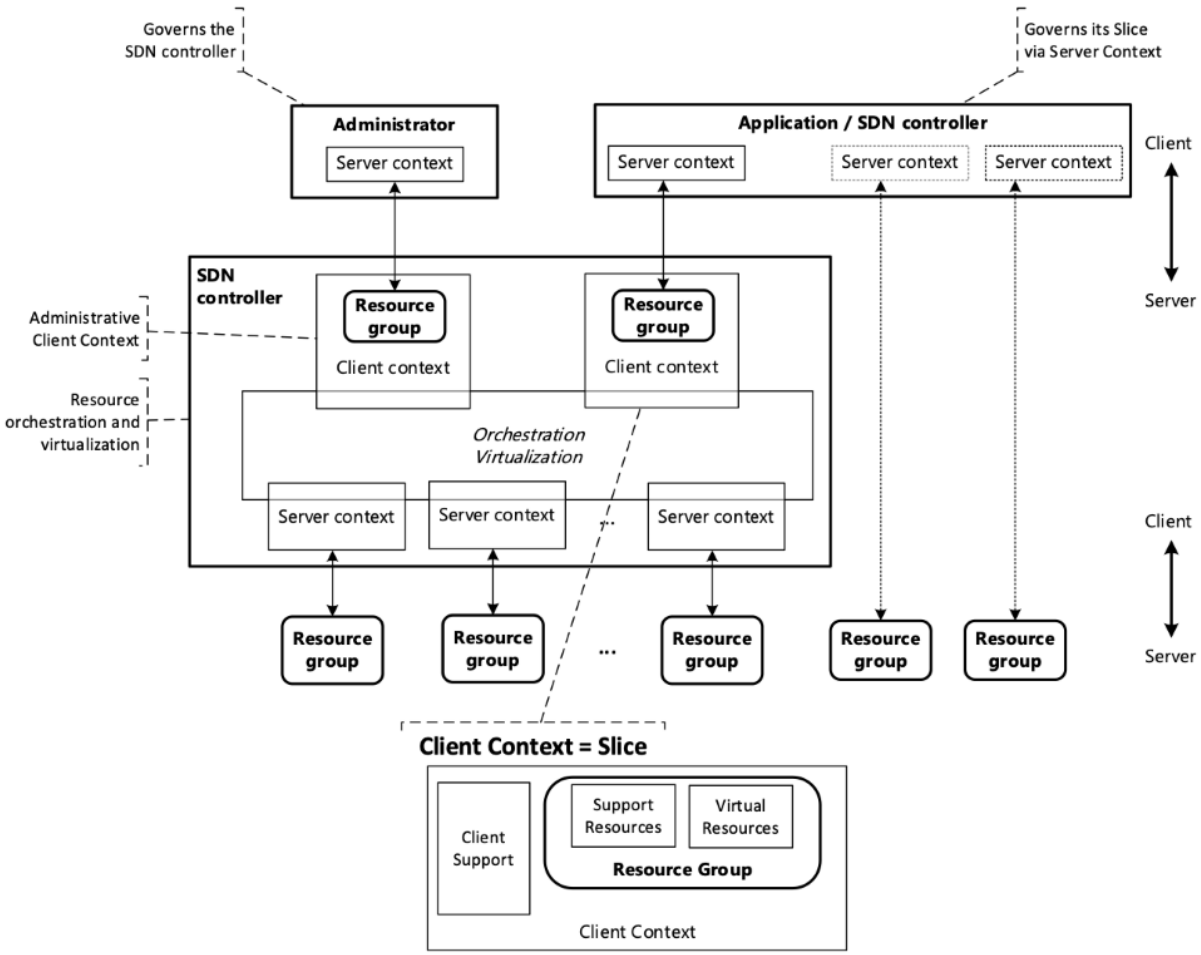

Figure 29: SDN-based slice abstraction. Image taken from [5]

In order to manage network slices efficiently, NFV architecture [6] plays an important role to manage the infrastructure resources and orchestrate the allocation of such resources needed to provide Virtual Network Functions(VNFs) and network services. The following figure shows the NFV reference architectural framework based on ETSI definition.

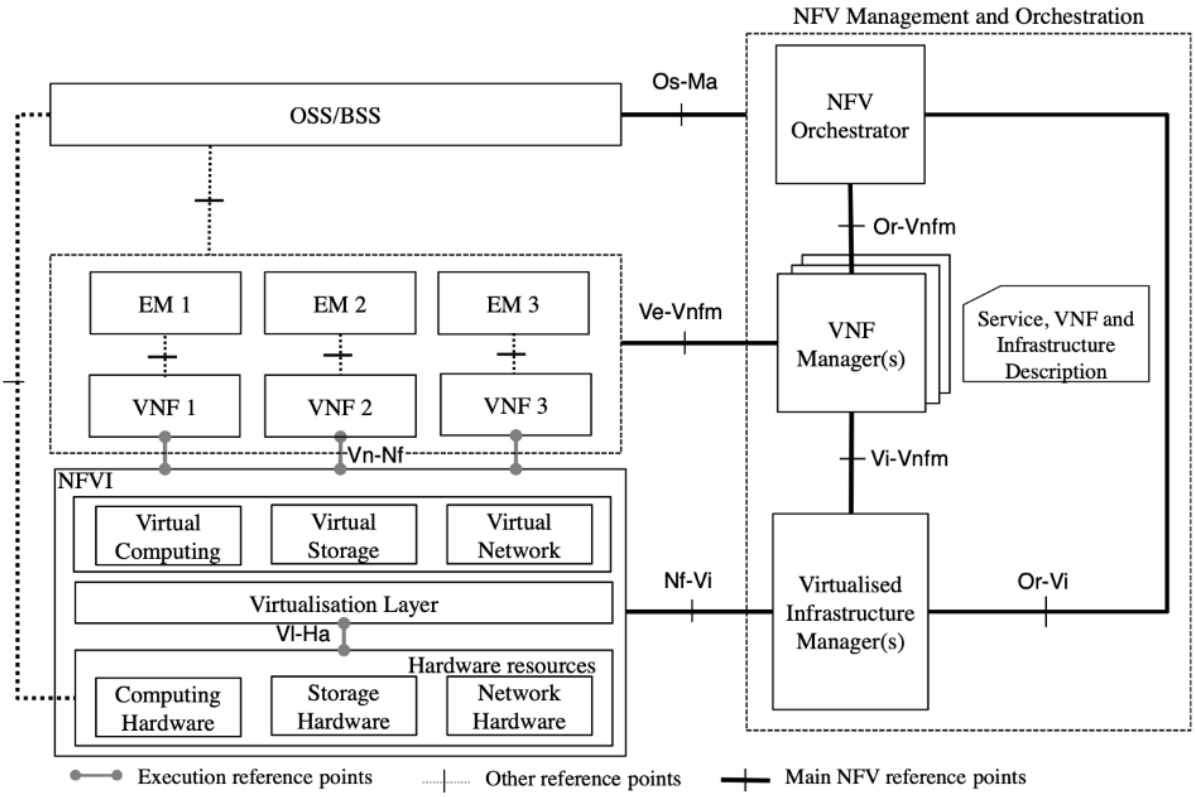

Figure 30: NFV architecture. Image taken from [6]

The NFV architecture comprises the following entities:

- Network Functions Virtualisation Infrastructure (NFVI): a collection of infrastructure resources used to host and connect the VNFs.
- VNFs: virtual implementation of network functions (NF) running over the NFVI.
- Management and Orchestration (MANO): performs all the Virtualisation-specific management, coordination and automation tasks in the NFV architecture. It comprises three functional blocks: Virtualized Infrastructure Manager (VIM) is responsible for controlling and managing the NFVI resources; VNF Manager (VNFM) is responsible for lifecycle management of the VNF(s) on its domain; NFV Orchestrator is in charge of the orchestration and management of NFV infrastructure and software resources, and realizing network services on NFVI.
- Element Management (EM): performs the management functionalities (Fault, Configuration, Accounting, Performance, and Security) of a VNF.
- Operation/Business Support System (OSS/BSS): Operator's systems and management applications.

IoT

Telecommunications operators, aware of the expectations and interest that 5G is generating, are already anticipating some of the benefits that 5G networks will bring to the currently deployed 4G network. This network is based on standards and, as aspects intrinsic to its nature, considers quality of service, safety and interoperability as priorities. The objective is to develop an ecosystem from this moment on the current LTE network that will allow us to explore, conceptualize and launch pre-5G use cases that will make the market more dynamic, generate viable and sustainable business models and that will cement the arrival of 5G, so that this technology will fit naturally among our clients, who will already enjoy services that 5G will later enrich with new use cases and improved performance thanks to its low latency and high speed features.

A good example is technologies such as narrowband communications for the Internet of Things(NB-IoT) and Machine Type Communications (Cat-M), which are in the process of being deployed in the network. These technologies allow millions of sensors to be efficiently connected to the LTE network.

The main features of IoT world are:

- Very high device volume, which requires inventory and self-management facilities.
- Unattended devices. This implies the need to optimize operation and maintenance costs while providing monitoring and diagnostic facilities of the devices and their communications; and facilities for remote updating of devices.
- Secure communications mainly based on mobile connectivity by cost and simplicity and uniformity of deployment.
- Information processing: collection, storage of a very high volume of data which implies to adapt the customer's business processes and decision-making support, using Big Data to leverage in data to increase revenues
- Criticality of applications (payments, logistics, supplies...) which demands robust and redundant solutions.

Currently billions of devices are connected to Internet, generating, collecting and sharing data. Any kind of sensors are being deployed in practically any sector. Factories are connected creating Connected Industry 4.0 concept which is extended with the use of robots to increase performance and improve production lines.

The application of IoT in day-to-day tasks is infinite, but always complemented with the use of AI techniques to process data and get relevant information to be shared. Let’s imagine our smart watch waking us up at the best time to have a good rest and it automatically connects with your coffee maker to turn on and start making coffee. Now imagine that your fridge knows the food inside is running out and automatically orders more. Another big area of work is what is so-called Smart cities. Cities are becoming smarter due to the use of sensors to manage waste services, intelligent irrigation, quality air control, etc. Actually, IoT is present in our lives through home automation systems, smart home alarms, GPS trackers, intelligent locks, etc.

When talking about IoT services, we understand there are the following functional levels (layers):

- IoT final solutions which are specific M2M solutions adapted to each type of industry
- Information layer (Data Collection)
    - Collection, conversion and aggregation of information.
    - Application Development Tools (APIs)
    - Business Intelligence and Big Data.
- Device Management Layer
    - Configuration, update and remote maintenance of the IoT devices (configuration of data connections, reboot, battery and coverage measurements, firmware update...)
- Managed Connectivity Layer
    - Inventory and SIM Life Cycle Management
    - Consumption Control
    - Management, Supervision and Diagnosis of Communications
- Connectivity Layer
    - Mobile communications of the IoT devices
    - Secure connection between IoT devices and customer applications.

Edge Computing

In the 1990s, when telecommunications companies, which previously offered point-to-point data circuits, began to offer other Virtual Private Network (VPN) services at much lower cost. This is how the term cloud computing came about. In August 2006 Amazon Web Services (AWS) introduced its Elastic Compute Cloud and it was the first public cloud computing service available.

With the strong emergence of the Internet of Things, data volumes grew exponentially and in parallel with the cloud computing services needed to process and obtain value from the data generated by the millions of sensors deployed worldwide. While these computing capabilities, available at sites remote from where the sensor is located, in many cases take a reasonable time for service. The problem comes in certain use cases where every millisecond that passes is crucial and we need the latency, the response time of the server, to be as low as possible. This is where the concept of bringing computing closer to the place where data is generated comes from and edge computing.

The term 'edge' depends on the environment, the time and the device we are talking about. For example, in mobile networks the edge can be a smartphone or a telephone antenna or also a connected car. In a fibre connection, the edge can be a router, an internet-connected sensor or the nearest processing centre of your operator.

What matters when we talk about the edge of the network is that all these elements have one feature in common: processors. They all have the ability to process and handle data. Thanks to that, with Edge Computing we can virtualize the server capabilities and enable the processing power to occur in those edge devices. The combination of edge computing and 5G is an alliance with many advantages since 5G offers a latency of very few milliseconds, a key aspect when it comes to sending data, performing processing and giving a real-time response as demanded by services such as the autonomous car.

Multi-Access Edge Computing (MEC), a.k.a. Edge Computing, is a telecommunications networks architecture that enables the placement of cloud and IT resources, methods and technologies in data centres within a telco operator network. These data centres can vary in size, location and capacity and can be deployed within mobile, fixed, TV and/or enterprise networks.

MEC is originally a standard from ETSI that was designed for mobile networks and was evolved to fixed and convergent networks. Deploying computing and storage resources closer to the customer, enables real time processing, guaranteed bandwidth and increased privacy and security while reducing latency, diminishing devices' computational needs, and lowering ineffective use of communication capacity vs. centralized Cloud. It plays an important part in delivering some of the promises of ultra-low-latency and ultra-reliability in 5G standards. MEC promises to deliver the best of both worlds: cloud affordability and scalability, with on-premise performance and convenience.

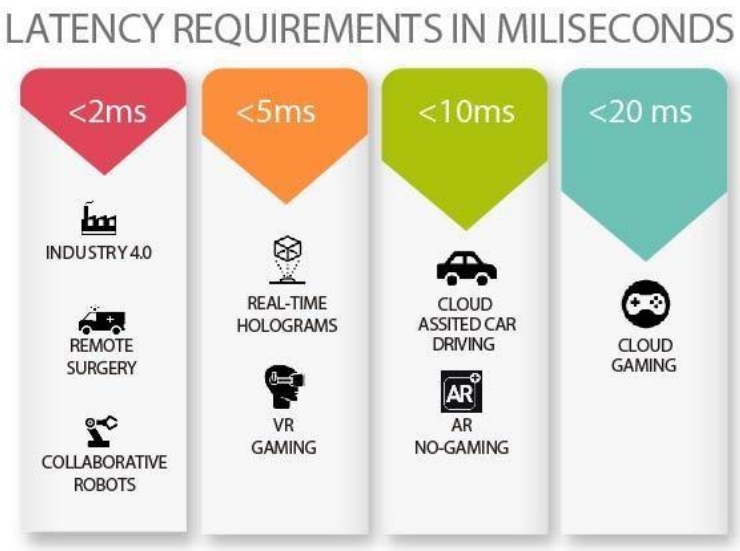

Figure 31: Estimation of latency requirements. Image taken from [8]

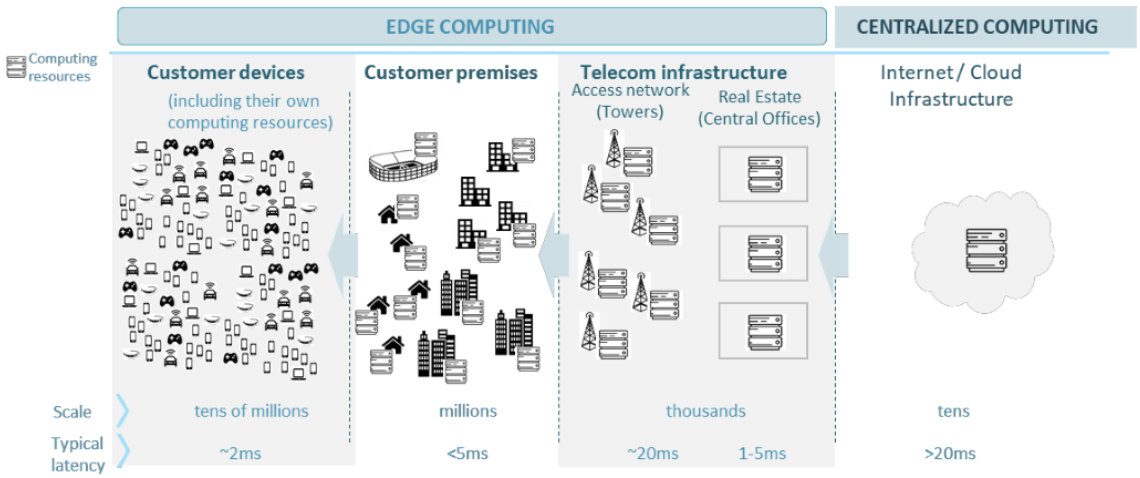

Figure 32: Edge computing vs Cloud Computing. Image taken from [8]

With multi-access edge computing, telcos can make compute and storage capabilities available to customers at the edge of communications networks so workloads and applications are closer to customers, enhancing experiences and enabling new services and offers.
MEC implies improvement for existing applications such as content/application delivery or caching by bringing these closer to the user in a geographically distributed way. It is the key enabler for emerging scenarios such as Industry 4.0 (control and monitoring of industrial machinery), connected and autonomous vehicles, Augmented Reality (AR) and Virtual Reality (VR), federated artificial intelligence and consumer blockchain.

These boundaries are starting to blur, as customers demand better performance from the continuum between internet, public, private, hybrid clouds and telco networks. One aspect of providing better performance is to reduce latency. MEC is specifically designed to reduce transport latency by reducing the distance between the device and the compute and storage capabilities. This implies deploying compute and storage capabilities closer to where content and services are created or consumed.

Some of the possible locations for these data centres can be:

- The telco cloud: centralized data centres within the operator's network, serving a country or a region. Typically, one per region. Latency 20 – 100ms
- The telco core network: centralized data centres serving mobile traffic. Typically, 2-3 per country. Latency 20 – 50ms
- Central offices, transport aggregation points: decentralized traffic commutation centres, typically designed for copper or fibre aggregation, between the core network and the enterprise / home FTTH. Typically, hundreds per country. Latency 5 – 20ms
- Base stations, telco towers: Compute and storage capabilities on telco towers. Typically, 10's to 100's of thousand per country. Latency 1ms (5G) – 20ms (4G)
- On-premise, on-device: Computing solutions ranging from mini data centres in enterprise premises, to storage or compute appliances, to apps and software programs in devices. Typically, millions to 10's of millions per country. Latency >2ms

Autonomous vehicles

The connected car of the future will include a series of cameras and sensors that will capture environmental information in real time. That information can be used in a variety of ways. It will be able to be connected to the traffic network of an intelligent city, for example, to anticipate a red light. It can also identify vehicles or adverse situations in real time or even know the relative position of other cars around it at all times.

This approach will transform how we travel by car and improve road safety, but the road to it is not without its pitfalls. One of the most important is that all the information collected by the different cameras and sensors ends up being of considerable dimensions. It is estimated that a connected car will generate about 300 TB of data per year (about 25 GB per hour). That information needs to be processed but moving that much data quickly between the servers and the car is unmanageable, we need the processing to occur much closer to where the data is being generated, at the edge of the network.

As an example, let us imagine a road of the future on which 50 cars are connected and which are also completely autonomous. This involves sensors that measure the speed of the surrounding cars, cameras that identify road signs or obstacles on the road, and a whole range of other data. The speed at which communication must take place between them and the server controlling that information must be kept to a

minimum. It is a scenario where we simply cannot afford to have the information travel to a remote server in the cloud, be processed, and return.

What happens may have been an accident, a sudden change in traffic conditions (an animal crossing the road, for example) or some other unforeseen event. We need the processor that operates with the information produced by the car sensors to be as close to the car as possible. With the cloud, this should go to the antenna (the operator) from there it travels over the Internet to the server and then back, triggering the latency. With Edge Computing, as part of the server's capabilities are at the edge of the network, everything happens right there.

In each of the stages of the last few decades, we have seen the evolution of connected car technology, as well as advances in the ecosystems where this technology is developed. At each stage new features and services have been added to the growing catalogue of connected car products. The following figure shows the evolution of a connected vehicle from its origins to the present, as well as new developments in the future.

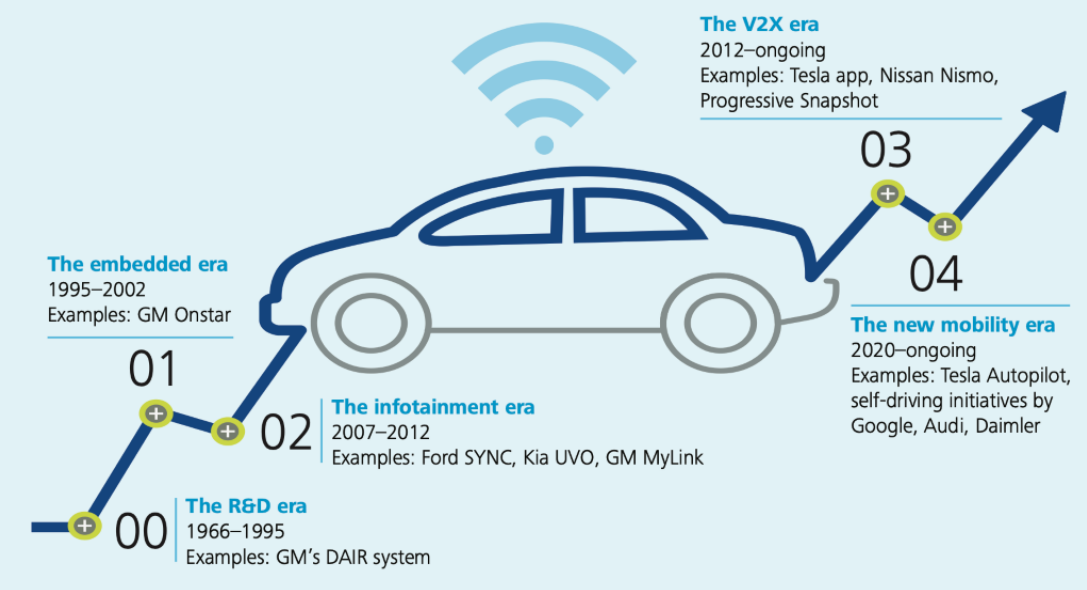

Figure 33: Evolution of the connected vehicle. Image taken from [7]

The five big stages according to Deloitte are:

- Research and development era. It is the longest, about 30 years, during which great ideas were proposed but not implemented due to lack of technology.
- The embedded era. It was the age of integration: modules integrated into cars, for example, mobile phones, were used to communicate wireless information to a telematics service provider.
- The infotaiment era. It was about information and entertainment. This is where in-vehicle information and entertainment applications are introduced and where a remarkable growth begins in the automotive industry, where new members are added to the ecosystem, software providers, content providers and third-party applications.

- The Vehicle-to-Infrastructure (V2X) era. Here integrated technology and services are combined, where the key elements are multiple sensors in the vehicle, in intelligent devices and home infrastructure, which communicate and share data within an integration called: V2X integration.
- The new mobility era. This era of autonomous vehicles is expected to begin to take off in 2020. Some autonomous driving prototypes are already on the road. This is also an era in which automakers and software providers competing for total dominance of the industry can trigger a shift away from vehicles as a commodity.

However, for the development of autonomous cars, it is necessary to have a 5G network deployed along with edge computing capabilities that allow lower response times than currently available technologies for autonomous driving to be possible.