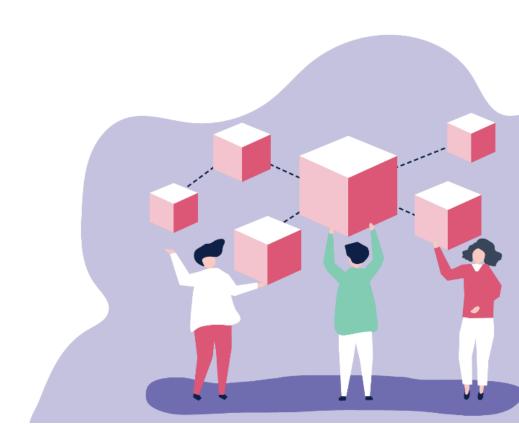




## D4.7 Microservices Containerization & Deployment

Deliverable No.	D4.7	Due Date	31/03/2022
Description	Microservices Containerization & Deployment		
Туре	Other	Dissemination Level	PU
Work Package No.	WP4	Work Package Title	GATEKEEPER Things Management Infrastructure & Development
Version	1.0	Status	Final





## **Authors**

Name and surname	Partner name	e-mail
Claudio Caimi	HPE	claudio.caimi@hpe.com
Mirko Manea	HPE	mirko.manea@hpe.com
Giovanni Saponara	HPE	giovanni.saponara@hpe.com
Mario Bartolucci	HPE	mario.bartolucci@hpe.com
Fausto Peverelli	HPE	fausto.peverelli@hpe.com
Patrizia Ciampoli	HPE	patrizia.ciampoli@hpe.com

## **History**

Date	Version	Change
01/20/2022	0.1	Table of content and initial content
02/02/2022	0.2	Revision of ToC and contributions
02/21/2022	0.3	Integration of content up to the date
02/28/2022	0.4	Integration of 1 <sup>st</sup> round of contributions
03/10/2022	0.5	Integration of contributions
18/03/2022	0.6	Incorporated comments from internal reviewers
21/03/2022	1.0	Revision of content after quality review

## **Key data**

Keywords	Data Centre Infrastructure	
Lead Editor	See Authors (HPE)	
Internal Reviewer(s)	ternal Reviewer(s) Carlo Allocca (SAM), Alessio Antonini (OU)	



## **Abstract**

This deliverable is the second version of D4.1 and contains a description of the GATEKEEPER Infrastructure that is currently hosting the GATEKEEPER Platform and the execution of the Pilots projects.

It describes the status of the setup of the Cloud Services of the GATEKEEPER Data Centre, their technical and organizational security measures, the Continuous Integration and Continuous Deployment setup, the Containerization platform, and the Ticketing Support System.

The GATEKEEPER Infrastructure is continuously updated and user manuals are provided to the users to allow the optimal exploit of the provided resources. A new updated and final version of this document is expected at M40, describing the final version of the infrastructure.

This deliverable accompanies the release of the actual hardware and software release of cloud service infrastructure for the GATEKEEPER platform.

## Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.



## **Acronyms**

Table 1 – List of acronyms

Acronym	Description
2FA	Two-Factor-Authentication
CA	Certification Authority
CE	Community Edition
CIFS	Common Internet File System
CPE	Common Platform Enumeration
CPU	Central Processing Unit
CVE	Common Vulnerabilities and Exposures
CWE	Common Weakness Enumeration
DoW	Description of Work, i.e. the GATEKEEPER proposal document
DC	Data Centre
DDI	DNS, DHCP, and IPAM
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name Service
GK	GATEKEEPER
iSCSI	Internet Small Computer Systems Interface
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
IPAM	IP Address Management
1/0	Input/Output
IT	Information Technology
IRF	Intelligent Resilient Framework
KVM	Kernel-based Virtual Machine
LDAP	Lightweight Directory Access Protocol
NIC	Network Interface Card
OS	Operating System
OSS	Open Source Software
OWASP	Open Web Application Security Project



PKI	Public Key Infrastructure
RAID	Redundant Array of Independent Disks
RAM	Random Access Memory
RHEL	Red Hat Enterprise Linux
SCA	Software Composition Analysis
SMTP	Simple Mail Transfer Protocol
ToC	Table of Content
VLAN	Virtual Local Area Network
VM	Virtual Machine
VPN	Virtual Private Network
VTL	Virtual Tape Library



## **Table of contents**

TABLE OF CONTENTS	7
LIST OF TABLES	8
LIST OF FIGURES	9
1 INTRODUCTION	10
2 GATEKEEPER DATA CENTRE INFRASTRUCTURE	11
2.1 PROVISION AND SETUP OF CLOUD SERVICES	11
2.1.1 Data Centre Layout	
2.1.2 Security Organizational Measures	13
2.1.3 Security Technical Measures	19
2.2 CONTINUOUS INTEGRATION AND CONTINUOUS DELIVERY	36
2.2.1 CI/CD Implementation	
2.3 CONTAINERIZATION IMPLEMENTATION	
2.4 SUPPORT AND TICKETING SYSTEM	
3 GATEKEEPER USER MANUALS	45
3.1 GATEKEEPER DATA CENTRE INFRASTRUCTURE	45
3.1.1 Data Centre Access User Manual	45
3.2 CONTINUOUS INTEGRATION AND CONTINUOUS DELIVERY	45
4 CONCLUSIONS	47
5 REFERENCES	48
APPENDIX A ANNEXES	
APPENDIX B USER REGISTRATION FORM TEXT FOR S2S VPNS	
APPENDIX C USER REGISTRATION FORM FIELDS FOR S2S VPNS	~
	•
APPENDIX D INTERNAL HPE PORTAL HOME PAGE	53



## **List of tables**

Table 1 – List of acronyms	5
Table 2 – Authorization list (as on March 2022)	17
TABLE 3 – DPA SECURITY REQUIREMENTS DESCRIPTION	
TABLE 4 – INTEGRATED SOURCES ON SIDECARS	22
Table 5 – Backup Structure	25



## **List of figures**

Figure 1 – DC Physical Layout (Ro=Rack 0, R1=Rack 1)	12
FIGURE 2 – DC LOGICAL LAYOUT	13
FIGURE 3 – PREFACE OF THE S2S USER REGISTRATION FORMFORM	14
FIGURE 4 – FIELDS OF THE S2S USER REGISTRATION FORMFORM	15
FIGURE 5 – EXTENDED FIELD FOR OPEN CALLERS USER REGISTRATION FORM	16
FIGURE 6 – INTERNAL HPE PORTAL	19
FIGURE 7 – LOG MANAGEMENT ARCHITECTURE	21
FIGURE 8 – GRAYLOG CUSTOM DASHBOARDS	22
FIGURE 9 – AUTHENTICATION DASHBOARD	23
FIGURE 10 – CI/CD MANAGEMENT DASHBOARD	23
FIGURE 11 – BRUTE FORCE ATTACK DETECTION	24
Figure 12 – Bareos Dashboard	25
FIGURE 13 – BAREOS CONFIGURED CLIENTS	26
Figure 14 – Bareos Storage VTL	26
Figure 15 – Bareos Restore	27
FIGURE 16 – HPE STOREONCE DASHBOARD	29
FIGURE 17 – BACKUP PERFORMANCE EVALUATION	29
FIGURE 18 – MAIN DASHBOARD	30
Figure 19 – All-hosts Dashboard	31
Figure 20 - Nexus Repository	38
Figure 11 - Jenkins Infrastructure	38
FIGURE 22 - GATEKEEPER CI/CD PIPELINES DIAGRAM	38
FIGURE 23 - STAGES OF THE BUILD PIPELINE	39
FIGURE 24 – STAGES OF THE DEPLOY PIPELINE	39
FIGURE 25 – STAGES OF THE SECURITY PIPELINE	40
FIGURE 26 – SEGREGATION FOR GATEKEEPER PILOTS	40
FIGURE 27 – LIST OF GATEKEEPER PILOTS PROJECTS (TENANTS)	41
FIGURE 28 – SAMPLE PROJECT WITH AVAILABLE COMPONENTS	42
FIGURE 29 – SUPPORT PROJECTS' LABELS	43
Figure 30 – Issue boards.	44
FIGURE 31 - ISSUE TEMPLATE PROTOTYPE	
FIGURE 32 - GATEKEEPER-WP4-GK_DATA_CENTRE_ACCESS_SITE_TO_SITE_HP TOC	
FIGURE 22 - AGENDA HDE CL/CD WEBINAD	16



## 1 Introduction

This deliverable is the second version of D4.1 and contains the updates, improvements and novelties that have been implemented until M30 of the GATEKEEPER project execution, under the WP4 activities umbrella (in particular T4.1).

It defines the current GATEKEEPER Infrastructure architecture with the components and services that build the HPE GK Data Centre for GATEKEEPER.

The GATEKEEPER Infrastructure at the time of this deliverable (M30) has greatly matured with respect to the previous reporting period (M18) and it is currently satisfying the needs of the GATEKEEPER Pilot's projects at production level as its main target objective.

With respect to T4.1 objective of "provision and setup of high-performance cloud services", the GATEKEEPER Infrastructure is now hosting 15 Pilot's projects (as independent tenants), as well as a development, a testing and a general production environment.

Both the Pilot's projects tenants and the developers' tenants are hosted on the OpenShift/OKD Kubernetes environment ("the containers clustering techniques and orchestration mechanisms") and are deployed thanks to the "definition of the continuous integration and continuous delivery pipelines" which has been fully implemented and is daily used by the GATEKEEPER developers' team. Each tenant hosts part or all of the GATEKEEPER core components defined in T4.2 (Thing Management System), T4.4 (Data Federation Framework), T4.5 (GATEKEEPER Trust Authority), depending on the specific Pilot needs and requirements.

Specific activities are also in progress with respect to the integration of T4.1 and T4.3 (about Big Data Services), where T4.1 data from tenants will be integrated with the Big Data platform for AI modelling. These will be reported on the next version of D4.3 which is expected at M36.

The T4.1 task results have been organized following the structure described below:

Section 1 - this introduction.

**Section 2** – focuses on describing the Data Centre setup, with the provided tools and services.

**Section 3** - reports the user manuals useful for the GATEKEEPER partners to access and work with the Data Centre tools and services.

**Section 4** - shows the conclusions and future work.

**Section 5** - is the bibliography.

Finally, **Appendices** show the collateral material (annexes) and additional information.



## 2 GATEKEEPER Data Centre Infrastructure

This document describes the GATEKEEPER Infrastructure setup at HPE Data Centre for GATEKEEPER. The infrastructure runs the GATEKEEPER Platform as a show case of the project-developed functionalities and provides the environment for the execution of the GATEKEEPER Pilots.

The Data Centre and the associated equipment are dedicated to the GATEKEEPER project and are configured following the project requirements as well as the applicable regulations and HPE policies, as you can see from the sections related to the technical and organizational security measures.

All hardware equipment is composed by devices manufactured by Hewlett Packard Enterprise (HPE) group. We are in the progress of extending the hardware devices to introduce enhanced computing capabilities for Artificial Intelligence/Machine Learning, as demanded by the AI/ML requirements of GATEKEEPER partners. This will result in the addition of new servers and Graphical Processing Units (GPUs) specifically optimized to address such need.

We follow the same road described in D4.1, where mostly Open Source software is used along with HPE-branded software for *infrastructure management*, *network* and *storage devices*, and *big data services*. As of now, the experience with Open Source software has demonstrated to be very positive and we consider very high the level of maturity of the selected tools (e.g., OKD for Kubernetes, oVirt for virtualization, etc. - more in the next sections). Both Open Source tools and HPE hardware/software are operated by a team of highly-skilled HPE professionals that has an extensive experience from the field in the different required sectors.

The choice of both hardware equipment and software tools have been a trade-off between security and reliability of advanced hardware technologies (commercial off-the-shelf HPE products), the associated costs, and the openness of the software services.

## 2.1 Provision and Setup of Cloud Services

This section describes updates to the provision and the setup of the Cloud Services offered to the GATEKEEPER project by HPE partner hosted at its Italian premises.

#### 2.1.1 Data Centre Layout

In this section we describe the updated configuration of the Infrastructure, where the GATEKEEPER resources reside. Please refer to D4.1 for all the other details.

Mainly we are going to add 2 new Synergy servers to the Compute Block with GPU support to enhance the computation of AI/ML algorithms for both training and inference needs. These design for these new servers plans to extend the oVirt virtualization cluster to both add new computing capabilities (CPU and RAM), exploit the GPU processing power and reserve it for HPE Ezmeral Big Data and Machine Learning service (part of T4.3).



#### 2.1.1.1 Physical and Logical layout

The updated Physical Layout of the DC is represented in the Figure 1 diagram:

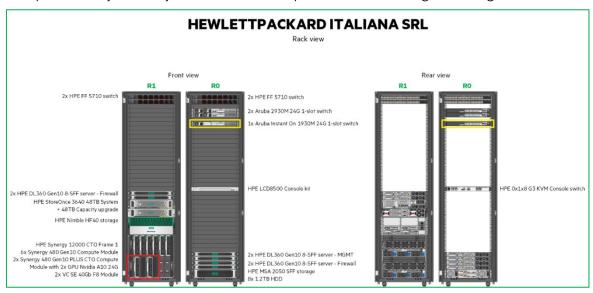


Figure 1 – DC Physical Layout (R0=Rack 0, R1=Rack 1)

As you can see in Figure 1, there are two Racks shown both in front and rear view. With respect to D4.1, we added a new low end switch to connect to Uninterruptible Power Supply (UPS) monitoring card (boxed in yellow). In this way we are able to assess the UPS health and when UPS goes to battery mode from the GATEKEEPR Monitoring System (see 2.1.3.4). Also the schema has been already updated to host two new computing node with GPUs (boxed in red, which will be available starting from April/May 2022).

Figure 2 shows the Logical Layout where you can see on the right side Service Block components, that host general Data Centre services, and on the left side the GATEKEEPER Block components (Compute and Storage Blocks). With respect to D4.1, we added the UPS and UPS switch connection, and prepared the design for Bay #7-#10 that will host new computing nodes with Graphical Processing Units (GPUs), that will be used to speed up ML training and inference algorithms.

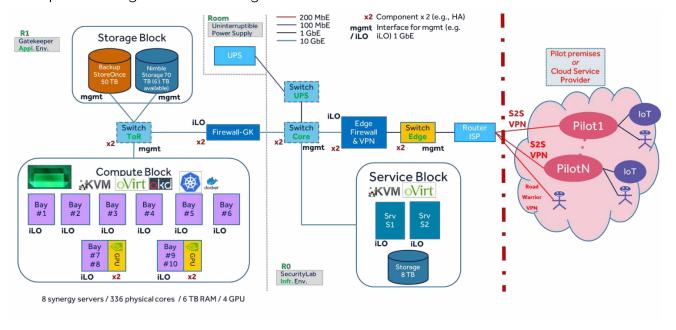




Figure 2 – DC Logical Layout

#### 2.1.2 Security Organizational Measures

In D4.1 we discussed the security and organization measures for controlling the Data Centre, in particular related to physical access and roles definition. Here we update to describe how we manage the new scenarios coming from Project Pilots' requirements related to **site-to-site (S2S) connections** and the addition to the **open callers** as new actors in the GATEKEEPER project. Also we describe the impacts of the Data Processing Agreement (DPA) that is currently being established as a legal basis for the proper Pilot execution.

#### 2.1.2.1 User Registration Process for S2S VPN

In addition to the VPN services the Data Centre already provides for users (called **road warrior VPN**), the GATEKEEPER Infrastructure has been extended to provide also the functionality to establish a secure encrypted channel for machine-to-machine interaction between a Pilot service and the Data Centre. We call this channel a **Site-to-Site VPN** and is used by those Pilots that have their own hardware/software infrastructure to connect to the HPE-operated Data Centre.

For those Pilots requiring to obtain an access to setup a direct VPN connection from the Pilot's site to the HPE GK Data Centre, it has been defined a process of request, validation and provision called user registration for S2S VPNs. Process includes de-provision in case of revocation of user access.

Users' accounts are of these kinds:

- GATEKEEPER Pilots for S2S VPN;
- Open Callers.

**Note**: for the specificities regarding the declination of this process for HPE Administrators and GATEKEEPER partner users, see D4.1, Section 1.1.2.2.

Request for user registration to S2S VPN is made by partners filing a registration form. HPE validates the form and then proceeds with the creation of the S2S VPN on HPE GK Data Centre, including credentials and S2S VPN configuration.

Figure 3 shows the preface of the S2S registration form, where the person is informed about the rules of HPE Data Centre access and the Privacy statements (form and text has been approved by HPE Legal Office):





# GATEKEEPER HPE Data Centre Site-To-Site Remote Access Request

Please complete the following form. Your data will be collected by HPE as GATEKEEPER partner for the sole purposes of account management and user access provisioning to HPE Data Centre infrastructure hosting the GATEKEEPER project services. You must give your consent to HPE for processing your data according to the HPE Privacy policy available using the link located at the bottom of the form. The HPE responsibilities are described in the link.

This computer network is provided for authorised use only. Users have no explicit or implicit expectation of privacy. Any or all uses of systems of this computer network and all data in this Data Centre may be intercepted, monitored, recorded, copied, audited, inspected, and disclosed to authorised sites and law enforcement personnel, as well as authorised officials of other agencies. By accessing to this computer network, the user consent to such disclosure at the discretion of authorised site personnel. Unauthorised or improper use of this network may result in administrative disciplinary action, civil and criminal penalties. By accessing to this computer network, you indicate your awareness of and consent to these terms and conditions of use. Users are responsible for ensuring that they act in accordance with this policy and other policies and legislation applicable to the HPE network.

Please note that by submitting this form you also agree to have your activities on HPE Data Centre monitored for compliance and security purposes. You must not disable (or attempt to) or change any security control, as well as respect the assigned privileges and not escape the perimeter of your role.

Should you have any inquiry please contact HPE project partner for further information.

Technical details about how to access HPE Data Centre services are published on Alfresco project document management system. You will receive an email as soon as your access will be granted with more information.

Notice: it is strictly forbidden to share this registration link outside involved GATEKEEPER project partners.

HPE GATEKEEPER Team

-

https://www.gatekeeper-project.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857223

Figure 3 – Preface of the S2S User Registration Form

For the sake of readability and accessibility, the full text is reported in Appendix B.



#### Figure 4 shows the data collection form fields:

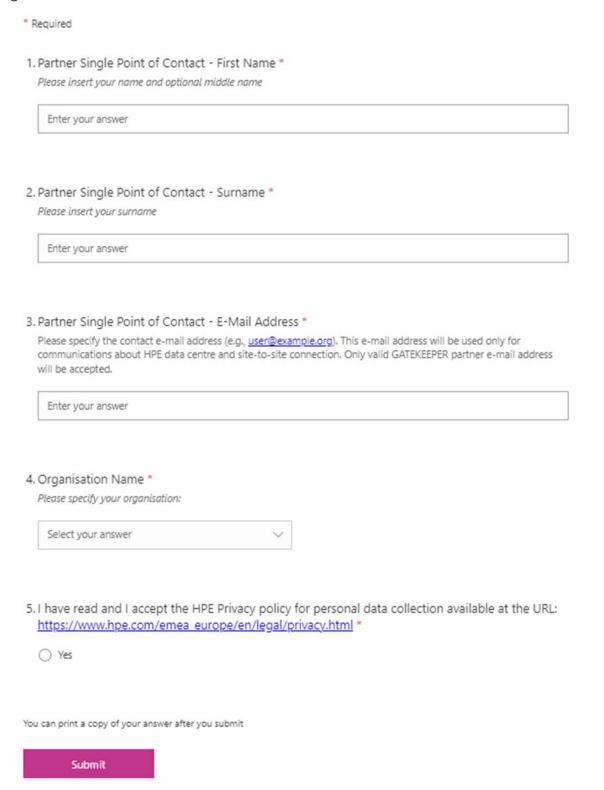


Figure 4 - Fields of the S2S User Registration Form

For the sake of readability and accessibility, the full text of the form is reported in Appendix C.



#### 2.1.2.2 User Registration Process for Open Callers

In order to give Open Callers (OCs) access to HPE infrastructure a process of request, validation and provision of a user registration has been defined. Each OC has been assigned to a specific GATEKEEPER partner, taking the role of **Mentor** of the Open Caller. A request for user registration to GATEKEEPER Data Centre services is made by OCs filing a registration form. HPE representative validates the form and asks, with an email, confirmation by the responsible OC Mentor. After approval by the Mentor, HPE proceeds with the creation of the user account for the specific user's role.

The registration form for Open Callers is the same as for GATEKEEPER partner users that has been extended to show the choices of the Open Caller participating organizations. The form fields and screenshots are available in D4.1, Section 1.1.2.2. In particular, field #4 (Organisation Name) now lists also the Open Callers are available organizations.

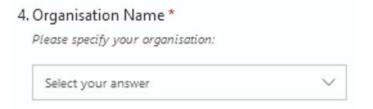


Figure 5 – Extended field for Open Callers User Registration form

#### 2.1.2.3 Pilots authorisation for partners

Pilots have to authorise the partners that will need to access their tenant (including their data) for performing administrative tasks, such as GATEKEEPER platform installation and management, or are the technological system integrators for a Pilot business partner (i.e., an hospital that leverages the services of a partners).

The activities in the GATEKEEPER LSP Cluster working group have designed a document related to the **Criteria for pilot qualification** (Toolkit) in which several information has been gathered and shared from and among all partners. In consequence a table has been created to collect authorizations decisions where Pilots' partners grant access to one or more of their available GATEEKEEPER Pilot-dedicated services. In particular, the access can be granted to:

- **GATEKEEPER Platform Pilot Tenant**: this grants a partner the access rights to the dedicated tenant of the Pilot on the OKD platform. It is where the GATEKEEPER Platform instance for the Pilot is running (includes the collected data);
- **Big Data Analytics/AI/ML Pilot Tenant**: this grants a partner the access rights to the Pilot Tenant on the Big Data platform (HPE Ezmeral). It is where the AI/ML programs and algorithms are developed and models are trained (includes the collected data).

The process flow is that every Pilot elects a responsible person who authorizes partners willing to grant access to Pilot data, then send the information to HPE which updates the table and perform the technical configurations.



Table 2 illustrates the information about which GK partner is supporting which Pilot(s).

Pilot name	Partner Name	Partner access to: GK Platform Pilot Tenant (OKD/Openshift)	Partner access to: Big Data Analytics/Al/ML Pilot Tenant (Ezmeral)
Puglia	ENG	TRUE	TRUE
Puglia	UPM	TRUE	FALSE
Puglia	UOI	FALSE	TRUE
Puglia	SAMSUNG	TRUE	TRUE
Puglia	CSS	TRUE	TRUE
BANGOR	SAMSUNG	TRUE	TRUE
BANGOR	UoW	TRUE	TRUE
BANGOR	UOI	TRUE	TRUE
BANGOR	OU	TRUE	TRUE
Milton Keynes (UK)	SAMSUNG	TRUE	TRUE
Milton Keynes (UK)	SPIROCCO	TRUE	FALSE
Milton Keynes (UK)	OU	TRUE	TRUE
Greece UC #a	CERTH	TRUE	TRUE
Greece UC #a	UPAT	FALSE	TRUE
Greece UC #a	UOI	FALSE	TRUE
Greece UC #a	BIO	TRUE	TRUE
Greece UC #b	CERTH	TRUE	TRUE
Greece UC #b	UOI	FALSE	TRUE
Greece UC #b	UPAT	FALSE	TRUE
Greece UC #b	BIO	TRUE	TRUE
Cyprus UC #a	CERTH	TRUE	TRUE
Cyprus UC #a	UOI	FALSE	TRUE
Cyprus UC #a	UoW	FALSE	TRUE
Cyprus UC #b	CERTH	TRUE	TRUE
Cyprus UC #b	UoW	FALSE	TRUE
Cyprus UC #b	UOI	FALSE	TRUE
Saxony	TUD	TRUE	TRUE
Saxony	CCS	FALSE	FALSE
Saxony	SAMSUNG	TRUE	TRUE
Basque Country	SAMSUNG	TRUE	TRUE
Basque Country	MyS	TRUE	TRUE
Basque Country	TECNALIA	TRUE	TRUE
Basque Country	OSA	TRUE	TRUE
Basque Country	S4C	TRUE	TRUE
Basque Country	IBERMATICA	TRUE	TRUE
Basque Country	BB UPM	TRUE TRUE	TRUE TRUE
Basque Country	KRONIKGUNE	TRUE	TRUE
Basque Country Covid-19	UPAT	TRUE	TRUE
	UPM	TRUE	FALSE
Aragon	ENG	TRUE	TRUE
Aragon	UOI	TRUE	TRUE
Aragon	001	IKUE	IKUE

Table 2 – Authorization list (as on March 2022)

#### 2.1.2.4 Data Processing Agreement (DPA)

Under the terms of GDPR [34], a data processing agreement is a legally binding contract that states the rights and obligations of each party concerning the protection of personal data. It is described in Article 28 of the GDPR. The GATEKEEPER project, in particular the Pilots projects, collect personal data as described as reported in the DoW. With respect to WP4 and T4.1 in particular, HPE owns the Data Centre where (part of) the data collected from the Pilots are sent for being used by the GATEKEEPER Platform services. It is worth mentioning here that the DoW explains that personal data will be (pseudo-)anonymised before being sent to the Data Centre services.

Nevertheless, HPE takes the role of data processor, while the Pilots are the data controller under the terms of GDPR.



The reader will find more legal details in WP6 deliverables, in this section we concentrate on the HPE security measures (both organizational and technical) for the GATEKEEPER Infrastructure with respect to the DPA document that is currently being finalised.

Table 3 – DPA security requirements description

Section5 - Security	Description
5.1	D4.1 and D4.7 contain the list of physical (D4.1 Section 2.1.2.2), technical and organizational measures (respectively Sections 2.1.2 and 2.1.3) put in place to safeguard collected personal data. Also Pilot isolation measures have been defined in Section 2.3.
5.2	D4.7 contains updated and enhanced security measures with respect to D4.1, and the same will be true for the final version of these documents (D4.8 at M40). In this way, we plan to follow the security evolution of threats and adopt new or updated measures as needed.
5.3	We follow a patch management process that apply as soon as possible the security patches to operating systems, devices and other infrastructure software components. HPE also collects logging information for security and compliance purposes as explained in Section 2.1.3.2.
5.4	People in the HPE group follow periodic and mandatory cybersecurity trainings to improve their skills and react to new threats, as well as to understand how to deal with data privacy aspects.
5.5	The HPE Project Team set up a notification process to follow whenever it is recognised a disruption of the service, loss of data, data breach, etc. Notification is sent to the dedicated HPE Data Centre registered users mailing list or Pilot owner (if it is a disruption related to Pilot only). Registered users and Pilot owners are available as part of processes described in Sections 2.1.2.1, 2.1.2.2, 2.1.2.3 of this deliverable, and also Section 2.1.2.2 of D4.1.

#### 2.1.2.5 User Data Centre Landing Page

The internal HPE Portal has been updated respect to the previous version in D4.1 by adding services for HPE Ezmeral Console for Big Data, Artificial Intelligence and Machine Learning. Furthermore it has been provided the access for the configured CI/CD tools such as GitLab, Nexus and Jenkins.



(Note: some content has been removed due to its security sensitiveness and links are available only internally).



Figure 6 - Internal HPE Portal

For the sake of readability and accessibility, the full text of Figure 6 is reported in Appendix D.

## 2.1.3 Security Technical Measures

Several technical security measures have been improved with respect to D4.1 to protect the Data Centre infrastructure where the GATEKEEPER Platform runs. This process is still ongoing, because these measures are frequently assessed, refined and improved over time, as the global security landscape evolves.

**Note**: several technical details are not included in this document, because of the sensitivity of such security information. This is mainly due to the Public scope of this document (i.e., Dissemination Level = PU), as specified in the DoW.



#### In particular:

- **Secure Access**: in addition to user VPN connections handling (i.e., road warrior VPN) reported in D4.1, we extended the VPN service to support site-to-site (S2S) connections for pilots and open callers (for more see Section 2.1.3.1);
- **Identity Management**: we are using successfully the tools implemented in D4.1 and we are managing over 150 user accounts at the time of this deliverable. Users have been enrolled by following the organizational process described in Section 2.1.2, as well as Section 1.1.2 of D4.1;
- **Log Management**: several new log sources have been integrated to support security and compliance needs (for more see Section 2.1.3.2);
- **Backup**: keeping safe copy of installed services, components, and data are crucial for the resiliency of the GATEKEEPER Infrastructure and safeguard against human errors and malicious cyberattacks (see Section 2.1.3.3);
- Monitoring: having the thermometer of functioning hardware equipment is important to plan support activities, discover broken components, monitoring the most critical software services. This is part of the effort to provide a high degree of resiliency to the delivered Data Centre services.

#### 2.1.3.1 Secure Access for Pilots and Open Callers

Access to HPE Data Centre for GATEKEEPER can happen only via secured an encrypted VPN connections. As described in D4.1, users (e.g., GATEKEEPER developers, GATEKEEPER data scientists, etc.) need to use a VPN client and authenticate via a Multi-Factor Authentication (user/password + one-time passcode).

In addition to this methods, we had the need to enable machine-to-machine VPN connection (as described in Section 2.1.2.1). This allows a server (e.g., at Pilot premises) to connect via a VPN connection to the HPE Data Centre and then use the GATEKEEPER Platform services. It still uses a Multi-Factor Authentication, now based on user/password and digital certificates.

A detailed how-to guide to setup S2S VPN has been delivered to Consortium partners (see Appendix A), describing step-by-step the procedure to follow under several operating systems for servers (Ubuntu Linux, CentOS, AWS Linux).

At firewall level, the S2S VPN connections can connect only to the services where the GATEKEEPER Platform can be installed and operated, guaranteeing a level of segregation with the rest of the Data Centre services.

#### 2.1.3.2 Log Management

The term Log Management identifies the processes of generation, transformation, storage and analysis of logs. In particular, our attention is focused more on security logs. The benefits of implementing such a solution are manifold. First you make sure you keep security logs for an appropriate period of time. Moreover, thanks to Log Management it is possible to review and analyse events in order to identify security incidents, violation of internal policies, fraudulent activity and operational problems.



The previous deliverable D4.1 illustrated the Open Source Graylog as a tool for Log Management services. Figure 7 shows the high-level architecture based on Graylog components that has been implemented:

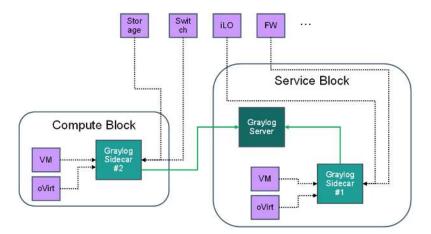


Figure 7 - Log Management Architecture

The implemented solution is made up of three main components (in green), one master node named "Graylog Server" (dark green) and two worker nodes named "Graylog Sidecars" (light green).

The master node represents the core of the architecture from which the orchestration of the other components is based. It is mainly used for operations such as creating or deleting an index on the log database, managing log services, performing queries, and creating dashboards. On the other hand, *sidecars* are a lightweight systems that allows for centralized and stackable configuration, utilizing a log collection agent.

In our case we implemented two *sidecars* following a "proximity" concept:

- A Service Block sidecar (Graylog Sidecar #1): to collect logs from all the devices present on the Service Block zone of the infrastructure;
- A Compute Block sidecar (Graylog Sidecar #2): to collect logs from devices running on the Compute Block zone of the infrastructure.

The Service and Compute Block are two logical zone (see D4.1) that group general services and for Access/Security/Network/Services management and specific Project block that includes also mass storage services. These blocks are physically grouped in terms of servers and racks. Having a sidecar in each zone implements the "proximity" concept, which allows log collection nearest to log generation and advanced mechanisms like bandwidth optimization, data channel encryption, and queuing mechanism to improve resiliency in case of "Graylog Server" downtime.

The following table shows in detail the split of all sources belonging to the GATEKEEPER project scope. It also give the feeling of the number of difference source types (both physical hardware devices and applications) that have been integrated and from which logs are collected:

Sidecar	Source Device		
	OPNsense VPN logs	HPE Flex Switch	
Sidecar 1 – Service Block	Linux OS	HPE iLO	
<b>Dissi</b>	oVirt	PWDMGMT service	



	Nexus YUM/APT repo HPE MSA	HPE iLO Amplifier FreeIPA	
	Bareos	Checkmk	
Sidecar 2 – Compute Block	OPNsense Access logs Linux OS oVirt GitLab Jenkins	HPE Flex Switch HPE iLO HPE Nimble Nexus GitLab user provisioning tool	

Table 4 – Integrated sources on sidecars

The process of integrating a source into Graylog is highly structured and involves the iteration of different configurations.

First of all it is necessary to create an index for storing the received logs from the device. Then create a stream to send events to the index and on the stream create a rule to direct the received device events. The next step is create an input of type beats for the *filebeat* (i.e., the agent collecting events) in order to accept log messages. Then it is necessary to configure the sidecar to make it listen on a chosen receiver port (all collected logs are pushed from the source to the filebeat agents on the sidecars). Last but not least create a pipeline that implements rules to parse the required events. Among all of them, the parsing step is the most important since it allows extracting structured fields from a raw event, that can be later searched and insights can be generated. All these operations can be done easily by accessing the Graylog GUI.

Once the resources have been integrated following the procedure described above, it is important to be able to actively manage, filter and monitor the log events. To do this, in Graylog it is possible to create Dashboards to have clear visualization of your search query results for discovering potential issues and intervening guickly.

In our environment we created two Dashboards, one to keep trace of Authentication events for security purposes and another to monitor CI/CD components:



Figure 8 – Graylog custom dashboards

Each dashboard can contain several tabs within it, useful for monitoring specific use cases.

Currently on the "Authentication" Dashboard there are three tabs (see Figure 9), but new ones can be added to meet with the continuous project evolutions:

 The first tab "Users Authentication" is used to trace all authentications on every source device in scope;



- The second tab "VPN Users" is focused on users accessing the project VPNs. As shown in the figure it has been added a custom widget with the localization on the map from which the access is done using the IP obtained from the log generated by the VPN;
- The third tab "HPE Portal Password Change" traces the users that have performed the change of the domain password.

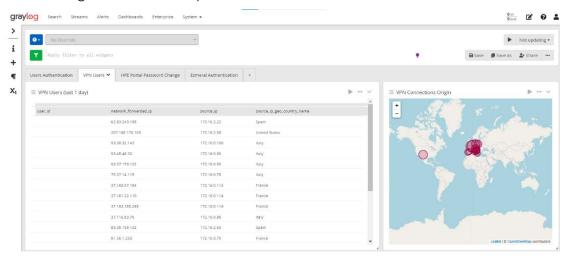


Figure 9 - Authentication Dashboard

On the other Dashboard, "CI/CD Management", at this moment there are five tabs to monitor specific authentications on CI/CD components. We added tabs for GitLab, Nexus and Jenkins authentication. Moreover, for the GitLab component it has been realized a script that synchronizes LDAP users with GitLab users. The result of this script is shown with more attention on the failed user addition on the first two tabs of the figure below. Also in this case widgets are created to help to understand and read the data by giving a more high-level overview.

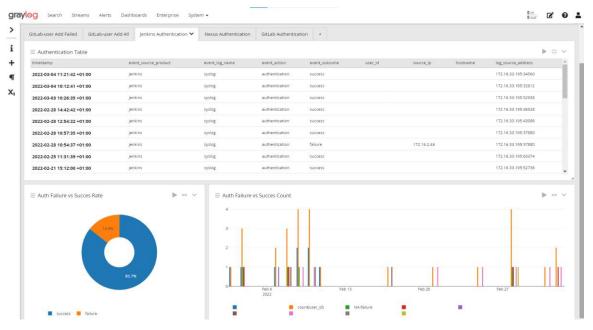


Figure 10 - CI/CD Management Dashboard.



Graylog also gives the possibility to create aggregation rules that are called "Event Definitions" for cyber-attack detections. We created a **brute force rule¹** based on events with fields that have identical values in a given time range. Going deeper, a filter is applied to the events restricting the analysis only to authentication events with failure outcome. For security reasons, we do not report here further details about how this use case is implemented and configured. When an alert event for this rule is generated, it notifies us of this attack attempt for further investigation.

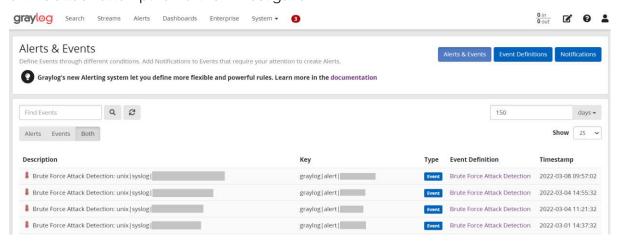


Figure 11 – Brute Force Attack detection

The Log Management service needs to be extended in the next months to be able to recognise others attack patterns from the collected log data.

#### 2.1.3.3 Backup

The Backup service is crucial for the resiliency of the whole Data Centre, its services and data. It provides the ability to recover after hardware failures, security incidents, or human mistakes. We started implementing the backup strategy and solution based on the Open Source tool called **Bareos** [16]. This enables to save system files, protect data, assets configuration and others artefacts such in a way it is possible to quickly restore them in case of need. Bareos has currently been integrated with **HPE StoreOnce** solution, to store data files and virtual machines snapshots. Integration with all the Data Centre services is mostly completed and is updated whenever a new service is requested.

Bareos itself is a program tool, Server/Client based, that permits the system administrator to manage backups and restores of virtual machines, disks and files and it supports a lots of Operative Systems and host virtualization Solutions. Bareos can backup to various types of media, including tape and disks.

In this scenario the HPE StoreOnce has been integrated and the same for the oVirt hosts for which a plugin has been implemented to save VMs' snapshots.

HPE StoreOnce offers a set of **CIFS** secure shares (encrypted) by which Bareos creates its volumes to store backups.

Version 1.0 | 2022-03-31 | GATEKEEPER ©

<sup>&</sup>lt;sup>1</sup> A Brute Force Attack is an attempt to gain illegal access to a system by trying to guess the password of a specified user.



In the same time, in order to follow the **3-2-1** implementing backup rule<sup>2</sup>, a **VTL** (Virtual tape Library) has been implemented to store bigger data.

For most of the Virtual Machines a backup policy has been implemented in which a file system backup and Snapshot (via **oVirt Plugin** installed in Bareos) policy has been defined to guarantee a hypothetical restore as granular as possible.

The following Table 5 shows the Backup Structure for the integration of Bareos with the HPE StoreOnce to accommodate both the CIFS (shares) and VTL types of backups:

Туре	Bareos		StoreOnce		
	Path	Use	NAS Shares	VT Libraries	
Share (CIFS)	/SHARE <b>xx</b>	Production	SHARExx	-	
Library Tape	VTL via iSCSI	Production	-	MSL Library	

Table 5 - Backup Structure

#### **Bareos**

#### https://www.bareos.com/

Here is reported how Bareos **Dashboard** looks like. It allows inspecting the status of the backups with all the details such as start and end time, size, etc.

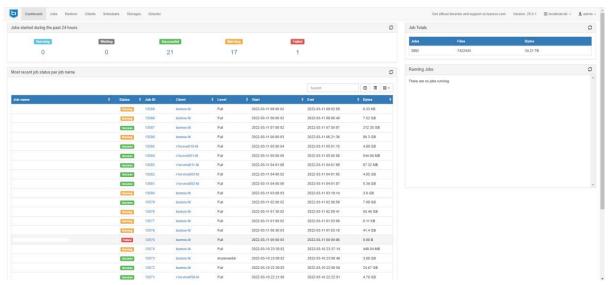


Figure 12 - Bareos Dashboard

Version 1.0 | 2022-03-31 | GATEKEEPER ©

<sup>&</sup>lt;sup>2</sup> https://en.wikipedia.org/wiki/Backup#3-2-1\_rule



The next figure shows the configured clients (e.g., servers) that are currently under backup. Names have been hidden for security reasons.

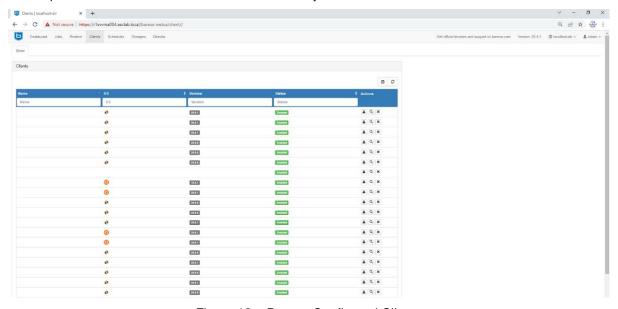


Figure 13 - Bareos Configured Clients

Bareos GUI allows to inspect also the status of the Storage Tape VTL, with all the backups that are in the virtual tapes:

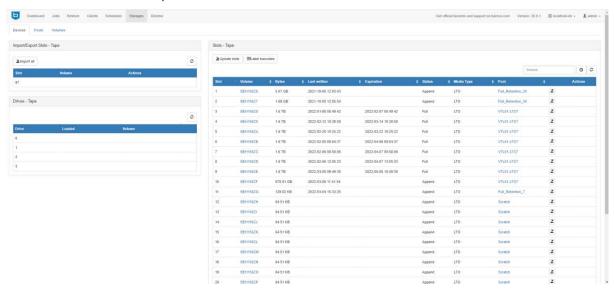


Figure 14 – Bareos Storage VTL



Bareos finally allows to quickly **restore** an individual file via the GUI, as briefly shown in the following figure:

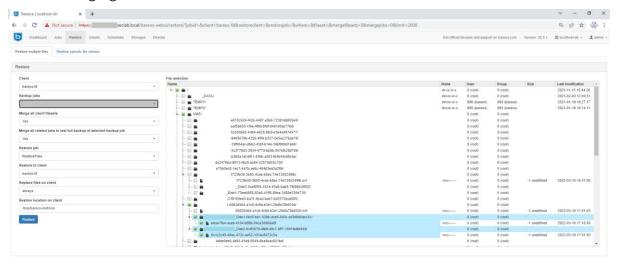


Figure 15 - Bareos Restore

From the GUI you can select the restore of an entire VM or volume or file simply exploring the navigation bar.

#### **HPE StoreOnce Gen 4**

The New StoreOnce GUI (**Grommet**) lets the user have an easy access to all of the library functions as reported below:



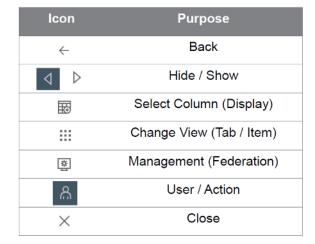
https://www.hpe.com/us/en/insights/articles/getting-to-know-grommet-an-open-source-ui-dev-tool-1808.html

https://v2.grommet.io/components

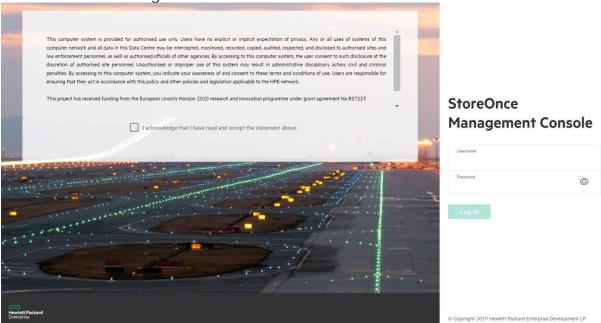


#### Here is a list of the general icons used in HPE StoreOnce Gen4 GUI:

Icon	Purpose	
묘	System Selector	
Q	Search	
	Menu/ Action	
+	Add	
#	Option / View	
0	Edit	
7	Filter	



#### The Grommet GUI - login:





On the HPE StoreOnce **Dashboard** we can inspect the usage of the storage resources in terms of VTL (**Virtual Tape Library**) and shares (**Catalyst**, **SMB** or **NFS**):

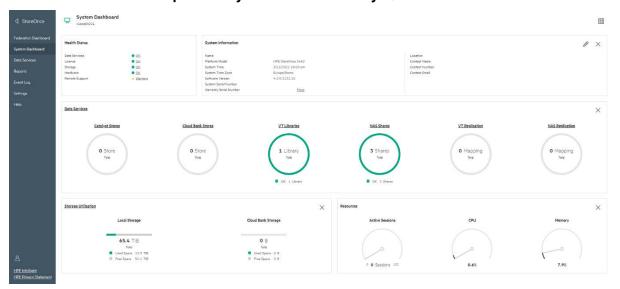


Figure 16 - HPE StoreOnce Dashboard

From here (Figure 16) the *Data Service* status, **Reports**, **Event log** and **Settings** Environments are accessible.

We also inspected the performance of the backup solution by looking at the HPE StoreOnce indicators during the backup operations. An example is reported in the next figure:

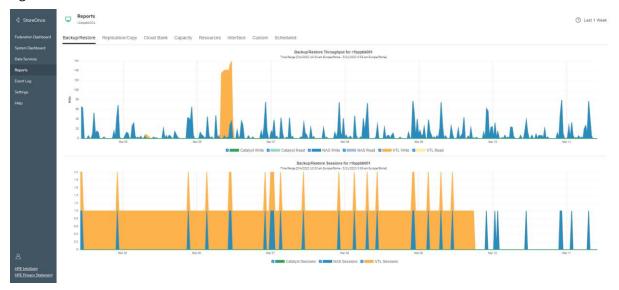


Figure 17 – Backup performance evaluation

In the previous figure (Figure 17) you can see when backups are going and how they are distributed in the backup window.

Besides, you can generate reports with **year/month/week/day/hour** granularity, but you can also customize your backup window for past backups or see them in real time (**live**).



Reads and writes are divided but also **VTL/Catalyst & NAS** are distinguished by colour in order to be able to read all the backup activities more easily.

As next step, we do plan to add native a backup solution for **OKD** (**Kubernetes**), in addition to **Bareos** general purpose solution. Also we do plan to use **HPE RMC** (**Recovery Manager Central**) to have an additional layer of protection.

#### 2.1.3.4 Monitoring

The monitoring is the process to gather information about the operations of an IT environment's hardware and software to ensure everything functions as expected to support applications and services.

Basic monitoring is performed through device operation checks, while more advanced monitoring gives granular views on operational states, including average response times, number of application instances, error and request rates, CPU usage and application availability.

The software chosen as tool to monitor the entire environment of the GATEKEEPER project is **Checkmk Raw Edition** because it is open source and free to use and it meets all the requirements.

Checkmk is a centralized monitoring system that lets us monitor the entire network in real time from a single machine. If an error occurs, we will be notified immediately and can fix the problem at once.

The central displays for your monitoring in Checkmk are Dashboards. They provide you with both overviews and detailed insights into specific areas. For example, you can visualise the general status of entire network segments, but also simply list which services are generating a load or overload of certain system resources. Checkmk comes with some standard dashboards, such as for problems, Checkmk server statistics and of course a general overview.

Main Dashboard is displayed in Figure 18 and gives the overview of the entire environment.

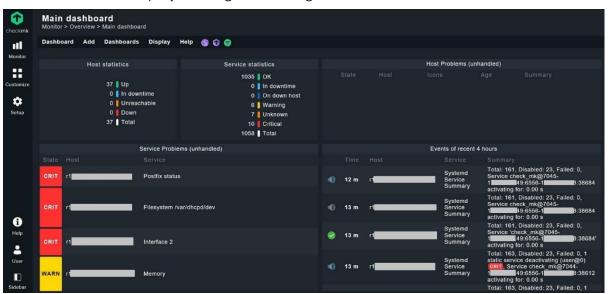


Figure 18 - Main Dashboard



Another useful Dashboard is the All-hosts that gives in a single view the details of all monitored hosts with the number of events divided by type.

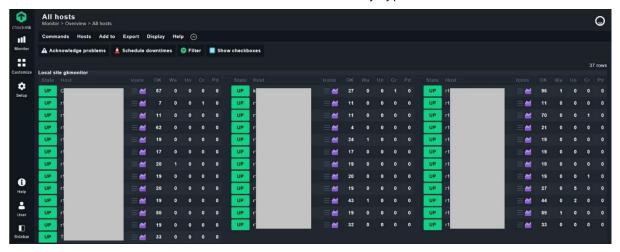


Figure 19 - All-hosts Dashboard

The advantage of incorporating the monitoring process into the project is that it serves a dual purpose, that of identifying real-time problems with systems in production and that of helping to discover bottlenecks or slowdowns on the network that would otherwise be difficult to identify.

Since it was integrated into the infrastructure, the checkmk tool has already helped to solve problems, for example:

• Slowdowns due to errors generated by faulty network cables:



• Prevent filling the filesystems of some machines:



The guideline to implement monitoring into GATEKEEPER project is to integrate at least the lowest layers of the infrastructure (hardware and systems). We are now gradually approaching the integration of other more advanced components (like containers and applications).

The various layers of the infrastructure that are on our integration radar for the monitoring service are:

- Hardware components of Servers/Storage/Switches;
- Hypervisors, e.g., Virtual Machines (VMs);
- Containers orchestration engines, e.g., OKD (containers);
- Other Applications (including databases).



We have setup Checkmk to monitor the first bullet (see next Section 2.1.3.4.1) and we are proceeding with the others in the next period.

#### 2.1.3.4.1 Hardware components of Servers and Storage

The role of a hardware monitor is to help to check hardware health by providing instant visibility of the status (up, critical, or warning) and monitor power supply unit (PSU), battery, fans, motherboard, CPU, memory, disks, etc. with the added ability to set baselines and use these values as future points of reference to compare configurations.

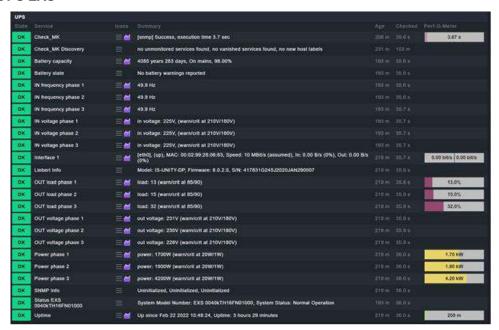
Each type of hardware device needs a specific way of integration into Checkmk that depends on the method by which the device makes information available to the outside world. The most common way to query a device is through SNMP, but being an old protocol, some modern systems allow integration only through special agent that use specific APIs.

The systems that make up the GATEKEEPER infrastructure and that have been integrated so far within Checkmk are:

- Uninterruptible Power Supply (UPS): Vertiv UPS EXS;
- Servers: HPE DL360 Gen10 Proliant Server (SNMP agent on ILO);
- Storage: HPE StoreOnce 3640 (SNMP agent on ILO + Special agent via REST API 4.x);
- Storage: HPE Nimble (SNMP agent);
- Storage: HPE MSA 2050 (Special agent via Web Interface);
- Servers: HPE Synergy 12000 (SNMP agent on ILO);
- Switches: HPE 5710 24XGT (SNMP agent).

The next screenshots provide a glance on such integrations:

#### **Vertiv UPS EXS**





## HPE DL360 Gen10 Proliant Server (SNMP agent on ILO)

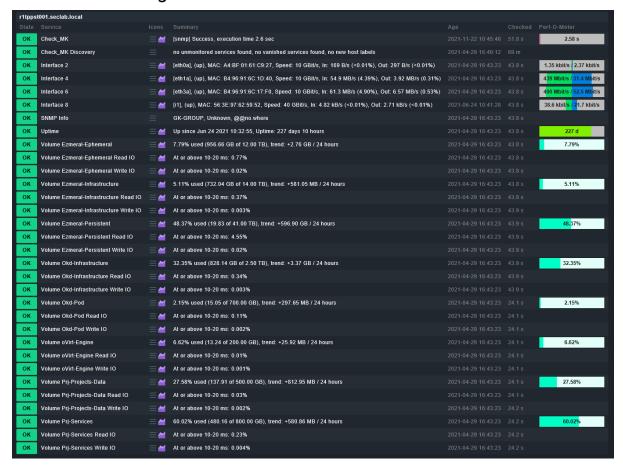
r1spsb	s002-i.seclab.local				
State					
ок	Check_MK	$\equiv$	[snmp] Success, execution time 5.4 sec		5.38 s
ок	Check_MK Discovery		no unmonitored services found, no vanished services found, no new host labels		
ОК	HW Controller 0		Condition: ok, Board-Condition: ok, Board-Status: ok, (Role: other, Model: 92, Slot: 0, Serial: PEYHB0ERHDP47U)		
ок	HW CPU 0		CPU0 "Intel(R) Xeon(R) Gold 5215 CPU @ 2.50GHz" in slot 0 is in state "ok"		
ОК	HW CPU 1		CPU1 "Intel(R) Xeon(R) Gold 5215 CPU @ 2.50GHz" in slot 0 is in state "ok"		
ОК	HW FAN1 system		FAN Sensor 1 "system", Speed is normal, State is ok		
ок	HW FAN2 system		FAN Sensor 2 "system", Speed is normal, State is ok		
ОК	HW FAN3 system		FAN Sensor 3 "system", Speed is normal, State is ok		
ок	HW FAN4 system		FAN Sensor 4 "system", Speed is normal, State is ok		
ок	HW FAN5 system		FAN Sensor 5 "system", Speed is normal, State is ok		
ок	HW FAN6 system		FAN Sensor 6 "system", Speed is normal, State is ok		
ок	HW FAN7 system		FAN Sensor 7 "system", Speed is normal, State is ok		
ок	HW Mem 7		Board: 0, Number: 7, Type: unknown (19), Size: 64.0 GiB, Status: good, Condition: ok		
ок	HW Mem 9		Board: 0, Number: 9, Type: unknown (19), Size: 64.0 GiB, Status: good, Condition: ok		
ок	HW Mem 19		Board: 0, Number: 19, Type: unknown (19), Size: 64.0 GiB, Status: good, Condition: ok		
ок	HW Mem 21		Board: 0, Number: 21, Type: unknown (19), Size: 64.0 GiB, Status: good, Condition: ok		
ок	HW Phydrv 0/0		Bay: 1, Bus number: 0, Status: ok, Smart status: ok, Ref hours: 11237, Size: 286102MB, Condition: ok		
ок	HW Phydry 0/1		Bay: 2, Bus number: 0, Status: ok, Smart status: ok, Ref hours: 11237, Size: 286102MB, Condition: ok		
ок	HW Phydrv 0/2		Bay: 3, Bus number: 0, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ок	HW Phydrv 0/3		Bay: 4, Bus number: 0, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ОК	HW Phydrv 0/4		Bay: 5, Bus number: 1, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ок	HW Phydry 0/5		Bay: 6, Bus number: 1, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ок	HW Phydry 0/6		Bay: 7, Bus number: 1, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ок	HW Phydry 0/7		Bay: 8. Bus number: 1, Status: ok, Smart status: ok, Ref hours: 11237, Size: 1716957MB, Condition: ok		
ок	HW Power Meter	=	Current reading: 212.00 Watts	2021-05-06 17:56:13	
ок	HW PSU 0/1	=	Chassis 0/Bay 1, State: "ok", Usage: 97 Watts		
ок	HW PSU 0/2		Chassis 0/Bay 2, State: "ok", Usage: 115 Watts	2021-05-06 17:56:13	
ок	HW PSU Total		Usage: 212 Watts		
ОК	Logical Device	=	Status: OK, Legical volume size: 279.37 GB		
ок	Logical Device 2		Status: OK, Logical volume size: 8.19 TB		
ок	SNMP Info		r1spsbs002-i seciab local, Rome, Achille Campanile,		
ок	Temperature 1 ambient	=	17.0 °C		17 °C
ок			40.0°C		40 °C
ок	Temperature 2 cpu				40 °C
	Temperature 3 cpu		40.0 °C		
ОК	Temperature 6 memory		25.0 °C		25 °C
ок	Temperature 10 memory		31.0 °C		31 °C
ок	Temperature 12 system		35.0 °C		35 °C
ок	Temperature 15 ambient		18.0 °C		18 °C
ОК	Temperature 16 system	=	25.0 °C		25 °C
ок	Temperature 17 system		30.0 °C		30 °C
ок	Temperature 18 system		23.0 °C		23 °C
ок	Temperature 19 system		22.0 °C		22 °C
ок	Temperature 20 system		25.0 °C		25 °C
ок	Temperature 21 system		24.0 °C		24 °C
ОК	Temperature 22 system		28.0 °C		28 °C
ОК	Temperature 23 system		61.0 °C		61 °C
ОК	Temperature 24 system	==	33.0 °C		33 °C
ок	Temperature 25 system		40.0 °C		40 °C
ок	Temperature 26 system		22.0 °C		22 °C
ок	Temperature 27 system	=	48.0 °C		48 °C
ок	Temperature 28 ioBoard		85.0 °C		85 °C
ок	Temperature 29 system		24.0 °C		24 °C
ОК	Temperature 31 ioBoard		24.0 °C		24 °C
ок	Temperature 33 ioBoard		24.0 °C		24 °C
ок	Temperature 37 system		28.0 °C		28 °C
ок	Temperature 38 powerSupply	≡≝	20.0 °C		20 °C
ок	Temperature 39 powerSupply		28.0 °C		28 °C
ок	Temperature 40 powerSupply	$\equiv$	40.0 °C		40 °C
ок	Temperature 41 powerSupply	$\equiv$	40.0 °C		40 °C
ок	Temperature 42 powerSupply	$\equiv$	21.0 °C		21 °C
ок	Temperature 43 powerSupply	$\equiv$	28.0 °C		28 °C
ок	Uptime	$\equiv$	Up since Sep 19 2021 23:40:45, Uptime: 139 days 21 hours		140 d



#### HPE StoreOnce 3640 (Special agent via REST API 4.x)

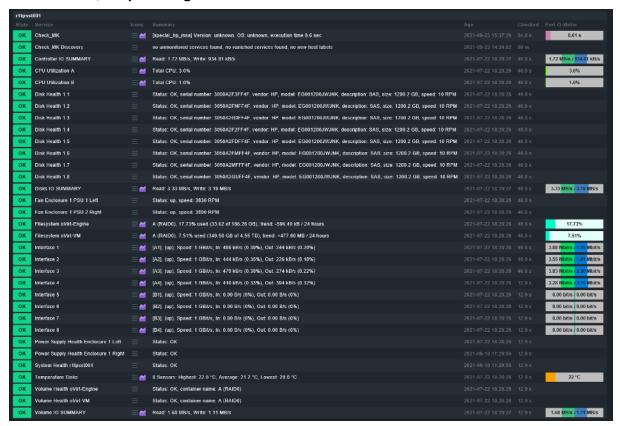


#### **HPE Nimble (SNMP agent)**

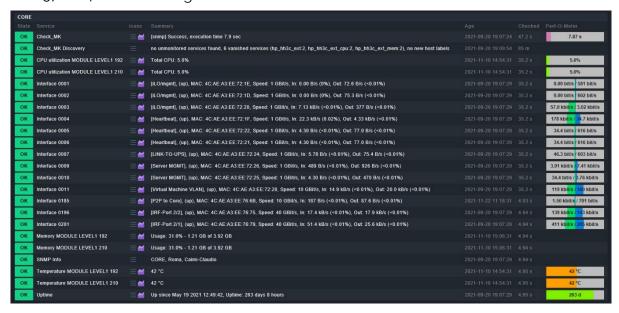




#### **HPE MSA 2050 (Special agent via Web Interface)**



#### HPE 5710 24XGT (SNMP agent)





#### 2.1.3.5 IDS & Anti-malware

We are internally prototyping the deployment of an Intrusion Detection System (IDS) at network level, based on open source software. The IDS will detect suspicious patterns and alert the Log Management solution (see Section 2.1.3.2) for further actions. We will start deploying first the IDS system and later we will evaluate to turn into Intrusion Prevention System (IPS), which will block malicious software payloads automatically. The prototyping solution is based on **Suricata** [35], a well-known IDS and IPS threat detection engine. It is a signature-based solution that works by using regularly updates patterns to spot security threats. The result of this activity will feed the Log Management system, where it will be analysed and reported.

Another technical security measure we are prototyping is the deployment of an anti-malware/anti-virus solution, which will run on all the HPE GATEKEEPER hosts. We are using **ClamAV** [36], a well-known "open-source antivirus engine for detecting trojans, viruses, malware & other malicious threats". It works by employing always updated signature to detect viruses and other malicious content. It can be periodically run (e.g., daily or weekly) to scan the whole server filesystem for detection. The scan result will be sent to the Log Management, where it will be analysed and reported.

## 2.2 Continuous Integration and Continuous Delivery

In D4.1 we outlined the CI/CD strategy and architecture, including the concepts of SecDevOps and CI/CD tools identified in the previous phase of the project.

The aim of the following section is to describe the implementation of the CI/CD architecture focusing on the interaction between components, how the solution guarantees security requirements and what is the output for the final user.

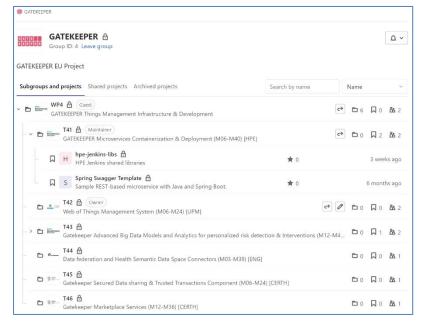
#### 2.2.1 CI/CD Implementation

In this section we describe the tools used to implement the CI/CD pipeline starting from the code versioning system, through the software repository, the ci/cd server until the tools for security testing and reporting. The idea of creating the pipeline is to make automatic the software development process instead of doing each step manually. Also, our solution introduces security at every level of the process.

As said in D4.1, the tool selected for managing code versioning is **GitLab**.



All partners have to push code in GitLab and for convention all projects are organized



following the division of work packages (WP) and tasks (T).

To automate the access to GitLab to authorized users only, a custom user provisioning tool creates on GitLab only those users and assign them the guest role.

Every WP has a Task Leader Maintainer (TLM) which is a developer that will get the maintainer role of the corresponding GitLab Task Group. TLM will appoint users to work in projects under the Task Group with the developer or maintainer role. GitLab roles are

#### explained in [33].

**Nexus** has been chosen as software repository for components such as packages, libraries, binary and Docker images available to end users such as the Figure 20 below shows. For uploading the images we followed the same structure of GitLab, with the division into WPs and Tasks. This convention is also followed in the **Jenkins** pipeline to call the respective images.

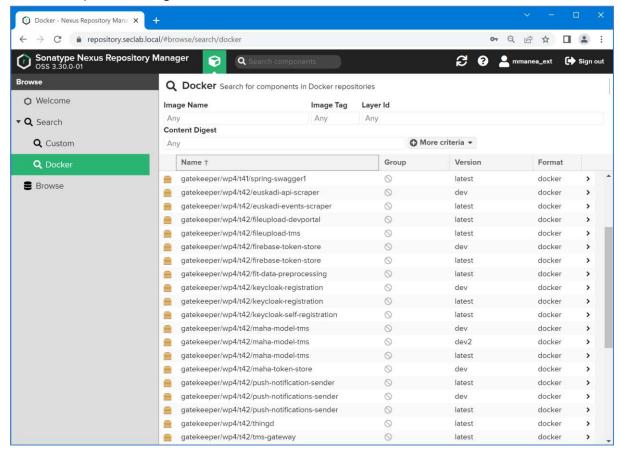




Figure 20 - Nexus Repository.

As a CI/CD orchestrator tool to monitor source repositories, build software, run tests and deploy software it has been chosen the open source tool **Jenkins**.

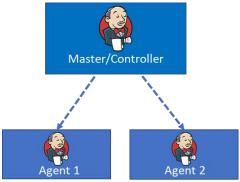


Figure 21 - Jenkins Infrastructure.

For Jenkins we realized a multi node infrastructure made up of one Master node and two Agent nodes: the Master node is designed to coordinate and provide the Web Access and the API endpoints while the Agent nodes are designed to perform and balance the work.

In particular, Agent 1 builds the software artefacts in a container, such in a way that the tools needed to build are part of the container instantiation and not directly installed on the Agent.

Agent 2 is identified as the security node that analyses the code of the software artefacts and it performs different levels of security analyses.

According to the CI/CD pipeline concept described in the D4.1, in order to automate the deployment process three pipelines are implemented.

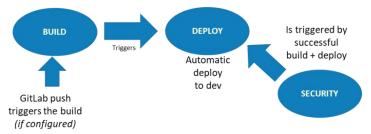


Figure 22 - GATEKEEPER CI/CD pipelines diagram

As shown in the Figure 22 the Build pipeline is triggered automatically at every push of the project in GitLab and it automatizes the build of the project, the creation of the Docker image and its push on the Nexus Docker Registry. Then, if the previous pipeline succeeded, the second Deploy pipeline is triggered and automatically deploys the component to the development environment. Finally, the Security pipeline is triggered if the Build and the Deploy pipelines are succeeded.

For automating the creation of the three pipelines we use the Jenkins plugin *Seed Jobs*. This utility has been configured to permit to fill out a form by entering parameters such as:

- Work Packages/Task folder where the Jenkins Jobs will be created;
- Job basename, that typically is the component name;
- GitLab URL, retrieved from the GitLab web interface:
- Build template, chosen from a preconfigured template or customize it manually;
- Dockerfile, the name of the dockerfile to build the container image:
- Image, the name of the container image pushed to the private Nexus Docker registry;
- Yaml files, used for deployment onto OKD (one or more).



Once these details are provided, the *Seed Job* automatically creates the three standardized pipelines for build, deploy and security.

### **Build pipeline**

In the Build pipeline there are different phases where first the code is checked out from GitLab and the container used to build is setup. It follows compile, testing and package stages where it is possible to customize the build process depending of the used tools. The next three stages are referred to the container images managing, building and pushing to the images to the Nexus Docker repository. There is an optional phase to tag the image before the push if it is not used the latest tag. When no errors occur the Deploy Job is automatically called.



Figure 23 - Stages of the build pipeline

### Deploy pipeline

In the deploy pipeline the attention is focused on the release of the components in the development environment. To make it possible, we need credentials to access to the OKD container orchestrator. In the next steps are applied the YAML deployment files to OKD where the configuration of the multiple Kubernetes resources are defined. The security pipeline is automatically triggered upon a successful build and deploy.



Figure 24 – Stages of the deploy pipeline.

#### Security pipeline

The security pipeline performs different types of security analysis of the software artefacts.

The first security control is performed by Semgrep [1] which is used to scan static source code to spot code vulnerabilities. The second security control is the container scanning by the Grype [2] tool to spot well-known vulnerabilities in container packages. The last control is Software Composition Analysis (SCA) done by OWASP Dependency Check [3] to spot security issues in third party libraries.

Every security tool creates a specific report. We added in the pipeline a step that make use of DefectDojo [4], an open source vulnerability report aggregator, where all security analysis results obtained by the previous scans are merged. These results are displayed in dashboards and diagrams that give an overall view on vulnerabilities to the developer.



Figure 25 – Stages of the security pipeline.

## 2.3 Containerization Implementation

As described in the containerization techniques overview presented in D4.1 (Section 1.3), we finally select OKD [4] as the container orchestration engine to host and run the GATEKEEPER Platform services developed by the partners as part of the other WP4 tasks.

OKD leverages on the de-facto standard Kubernetes engine [7], but adds many features including a stricter security environments (by default containers cannot run with administrative privileges) and a sophisticated but user-friendly graphical interface that let users interact with it without the need to go through the command line.

OKD has been configured to segregate with separated tenants all the GATEKEEPER Pilots. This makes use of the project and namespace concepts that allow to isolate the deployed components. Components are packages as containers and run in so called *Pods* (the Kubernetes resource to host one or more containers). Pods in different projects cannot talk to each other because we defined specific virtual network level policies for isolation.

The following Figure 26 illustrates the isolation between Project Pilots and their running components in the Pods<sup>3</sup>:

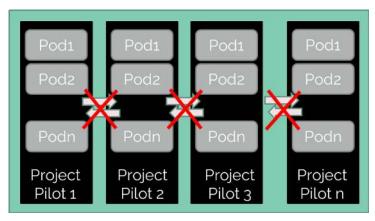


Figure 26 - Segregation for GATEKEEPER Pilots

Version 1.0 | 2022-03-31 | GATEKEEPER ©

<sup>&</sup>lt;sup>3</sup> The running components of the GATEKEEPER Platform services can be considered a set of Pods.

**Projects** 



Project Pilots can be accessed only to authorized Pilot users and administrators. Accounts are defined on the centralized Data Centre Identity Management service and follows the authorization process defined in Section 2.1.2.3. Also access to Pilots tenants is granted via Multi-Factor Authentication to authorized people only.

Similar to the Pods, also the Pilots' data is segregated in different volumes (Kubernetes calls them Persistent Volume Claims - PVC). These volumes are maintained in the HPE Nimble storage solution, which guarantees a high-level of availability and also has been configured to use data encryption at rest as a technical security measure.

Figure 27 shows the list of OKD Projects defined to support the project Pilots execution:

#### gatekeeper Clear all filters Name gatekeeper X Name 1 Display name Status PR gatekeeper-dev Active Dev PR gatekeeper-pilot1 PUGLIA Active PR gatekeeper-pilot2 ARAGON Active PR gatekeeper-pilot3 SAXONY Active PR gatekeeper-pilot4 BASQUE\_COUNTRY Active PR gatekeeper-pilot5a Greece\_UC\_#a Active PR gatekeeper-pilot5b Greece\_UC\_#b Active PR gatekeeper-pilot6a Cyprus\_UC\_#a Active PR gatekeeper-pilot6b Cyprus\_UC\_#b Active PR gatekeeper-pilot7 Milton\_Keynes Active PR gatekeeper-pilot8 Poland Active PR gatekeeper-pilot9 Covid-19 Active PR gatekeeper-pilot10 Bangor Active PR gatekeeper-pilot11 SINGAPORE Active PR gatekeeper-pilot12 TAIWAN Active PR gatekeeper-pilot13 HONG KONG Active PR gatekeeper-production Prod Active PR gatekeeper-test Active Test

Figure 27 – List of GATEKEEPER Pilots projects (tenants)



In addition to the Pilots, we have three special projects:

- **gatekeeper-dev**: it contains always the latest version of the GATEKEEPER Platform services that is under active development (this is also automatically populated when the CI/CD executes, see 2.2);
- **gatekeeper-test**: it is used to perform integration testing for the Pilots, before moving to their specific project;
- **gatekeeper-prod**: it is used to host some common components that are shared among all the pilots, like the GATEKEEPER Marketplace (T4.6).

The following Figure 28 shows a specific Pilot project (tenant) with all its required running components of the GATEKEEPER Platform:

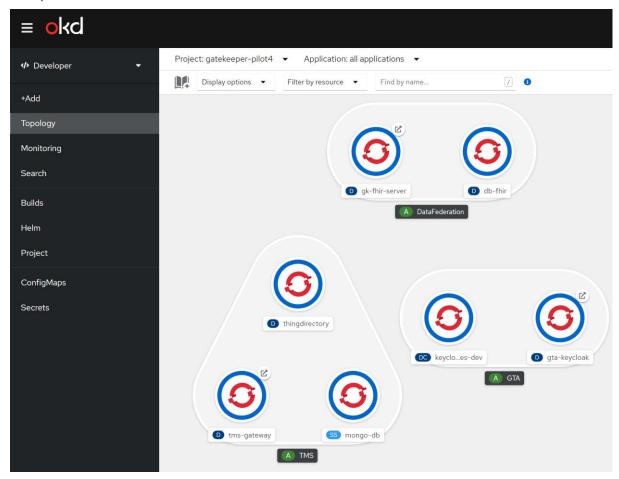


Figure 28 – Sample project with available components

The figure shows the Pods grouped by component, like TMS (Thing Management Systems, part of T4.2), Data Federation (part of T4.4), and the GTA (GATEKEEPER Trusted Authority, part of T4.5).

## 2.4 Support and Ticketing System

The Support and Ticketing System in GATEKEEPER responds to the need to implement a system to handle infrastructure problems across the entire Data Centre services and to collect new feature requests in a structured and reportable way.



To address this purpose we utilized the GitLab Issue functionality, by creating a specific project under a GitLab which is accessible to all GATEKEEPER users that want to report a problem or a feature request.

This is useful both for the users that keep trace in a single place of their issues and both for the person assigned to resolve it that can check easily the status of the issue.

It has been created a Wiki landing page for the GATEKEEPR users with instructions about how to report issues, the provided functionalities and some guidelines to follow to standardize the process.

The issue reporting makes use of custom labels (shown in the Figure) to permit a better identification of the issue lifecycle status and to classified them by type.

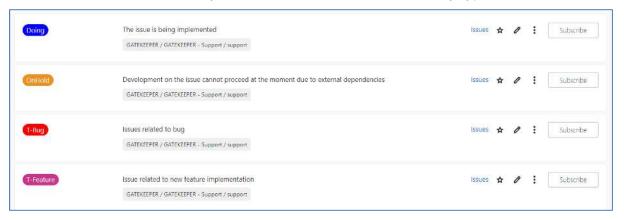


Figure 29 – Support projects' labels.

For example, we identify two type of issues, depending it concerns a **Bug** or a new **Feature** and we created the respective labels:

- **T-Bug**: this label is assigned to an issue that reports an incorrect behaviours in a Data Centre component.
- **T-Feature**: this label is assigned to an issue that is a new or change request to develop new functionalities or enhance existing ones.

Regarding the status, in addition to the Open/Closed label given by default, we decided to create other labels for handling the issue lifecycle by adding two new states:

- Doing: a label for an issue that is in progress;
- **OnHold**: a label for an issue that is temporary stopped for external dependencies, which can be an action asked to the user.



Based on these labels it is possible to create multiple boards on the Issue tab with a different custom view that follows our criteria to process the issue.

For example, we created two boards: one by Type and one by Status.

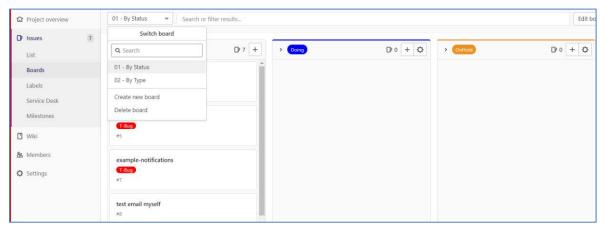


Figure 30 – Issue boards.

When a user creates an issue, following the instructions on the Wiki page, s/he has to select in the picklist the proper template that is created for each type of issue (bug or feature, as mentioned before). To simplify and best fitting the problem, it has been created **BugTemplate** and **FeatureTemplate** with a suggested form to fill.

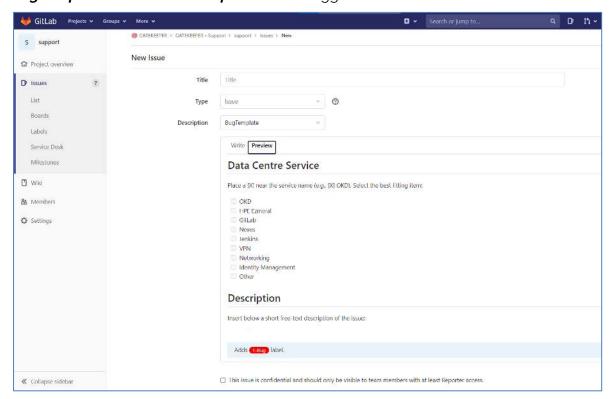


Figure 31 - Issue Template prototype.

There is also the possibility to keep an issue confidential (visible only to the issue creator and to HPE staff).

Currently the Support and Ticketing System has just been released and it is expected it will be enhance in the future period with the feedbacks coming from the users.



## 3 GATEKEEPER User Manuals

## 3.1 GATEKEEPER Data Centre Infrastructure

## 3.1.1 Data Centre Access User Manual

In D4.1 we delivered the Access Manual for partners' access via VPN (road warrior VPN). We now also release the manual for site-to-site VPN connections used for project pilots and open callers.

As part of this deliverable, we release as an Annex the **confidential document** that describes in detail the operating procedure GATEKEEPER partners must follow to use the S2S VPN. We report in Figure 32 the index of the document.

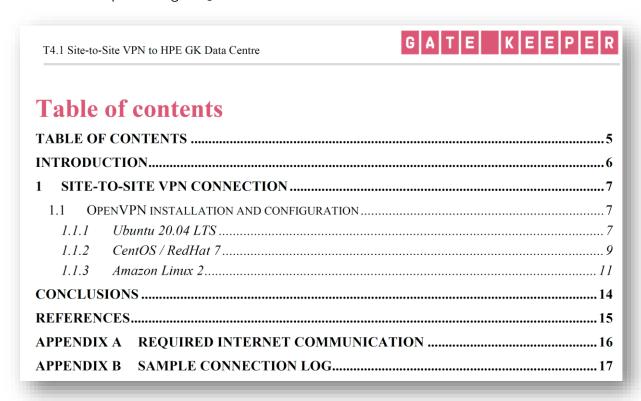


Figure 32 - GATEKEEPER-WP4-GK Data Centre Access Site To Site HPE document ToC

The document guides the GATEKEEPER partner users through the steps of getting the necessary software to setup the S2S VPN connection to the Data Centre for various operating systems. Part of this information has already been discussed in Section 2.1.3.1.

This document is marked as confidential because it contains Data Centre details that must not be made public.

## 3.2 Continuous Integration and Continuous Delivery

Since the implementation activities related to the CI/CD task have been delivered we released a manual for CI/CD attached in Appendix A.



In the manual we describe a high-level workflow for developers and present a demo to get started with a sample micro-service. We go deep in the use of the versioning tool, the CI/CD tool and the Seed Job to create automatically the build, deploy and security pipelines. Below we report the agenda of the webinar that HPE delivered to the GATEKEEPR partners (developers):

### SECURE CI/CD PLATFORM

DEVOPS FOR GATEKEEPER PLATFORM - WEBINAR AGENDA



- Pre-requisites
- Quick hands-on
  - CI/CD for GATEKEEPER
  - High-level workflow for developers
  - Demo: how to get started with your own micro-service
- All the details
  - Versioning Tool: GitLab
  - CI/CD Tool: Jenkins
  - Security analysis
- Tuning the seed job output
  - Build Pipeline
    - Deploy Pipeline
    - Security Pipeline

Figure 33 - Agenda HPE CI/CD webinar

To use this platform for CI/CD this webinar explains the whole steps that a user have to follow:

- Log in to the Web applications;
- Create a project into GitLab;
- How initialize pipelines with Seed Job in Jenkins;
- Run the generated CI/CD pipelines;
- Generate a security report.



## 4 Conclusions

This deliverable provided the description of the result of the activities spent for the setup and operation of the GATEKEEPER Data Centre by HPE to support the project consortium. It described the current data centre layout, the organizational and technical security measures in place, and the supporting services, including CI/CD tools, container orchestration engine for running the GATEKEEPER Platform, and the Support and Ticketing System.

For the upcoming period, we plan to continue operating the Data Centre and provide support to partners. We will also put effort in maintaining and improving the overall security posture of the data centre, by patching and updating already in place tools and evaluating new countermeasures to lower the changes of successful cyberattacks.



## **5** References

- [1] Flex Fabric Virtual Technology Configuration, <a href="https://support.hpe.com/hpesc/public/docDisplay?docId=a00050312en\_us">https://support.hpe.com/hpesc/public/docDisplay?docId=a00050312en\_us</a>
- [2] HPE Synergy composable bladed infrastructure, https://www.hpe.com/emea\_europe/en/integrated-systems/synergy.html
- [3] oVirt free open-source virtualization solution for enterprise, <a href="https://www.ovirt.org">https://www.ovirt.org</a>
- [4] OKD community distribution of Kubernetes, <a href="https://www.okd.io">https://www.okd.io</a>
- [5] Red Hat OpenShift hybrid cloud platform, <a href="https://www.openshift.com">https://www.openshift.com</a>
- [6] CentOS, Community Enterprise Operating System, <a href="https://www.centos.org">https://www.centos.org</a>
- [7] Kubernetes production-grade container orchestration, <a href="https://kubernetes.io">https://kubernetes.io</a>
- [8] HPE Ezmeral, Run, manage, control and secure the apps, data and IT that run your business—from edge to cloud, <a href="https://www.hpe.com/ezmeral">https://www.hpe.com/ezmeral</a>
- [9] Docker, Open Source containers, <a href="https://www.docker.com/community/open-source">https://www.docker.com/community/open-source</a>
- [10] HPE Nimble, storage solution, <a href="https://www.hpe.com/us/en/storage/nimble.html">https://www.hpe.com/us/en/storage/nimble.html</a>
- [11] HPE StoreOnce, backup solution, https://www.hpe.com/us/en/storage/storeonce.html
- [12] HPE Networking Switches, network solution, https://www.hpe.com/emea\_europe/en/networking/switches.html
- [13] IEEE 802.1Q, Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks, <a href="https://standards.ieee.org/standard/802\_1Q-2018.html">https://standards.ieee.org/standard/802\_1Q-2018.html</a>
- [14] IEEE 802.3, Standard for Information Technology Local and Metropolitan Area Networks - Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications-Aggregation of Multiple Link Segments, <a href="https://standards.ieee.org/standard/802\_3ad-2000.html">https://standards.ieee.org/standard/802\_3ad-2000.html</a>
- [15] HPE Proliant DL Servers, rack-optimized secure industry-standard servers, <a href="https://www.hpe.com/us/en/servers/proliant-dl-servers.html">https://www.hpe.com/us/en/servers/proliant-dl-servers.html</a>
- [16] Bareos, Open Source Data Protection, <a href="https://www.bareos.org">https://www.bareos.org</a>
- [17] Synopsis Black Duck Open Hub, <a href="https://www.openhub.net">https://www.openhub.net</a>
- [18] Git, distributed version control, <a href="https://git-scm.com">https://git-scm.com</a>
- [19] Linux KVM, Kernel-based Virtual Machine, <a href="https://www.linux-kvm.org">https://www.linux-kvm.org</a>
- [20] VMware ESXi, Bare Metal Hypervisor, <a href="https://www.vmware.com">https://www.vmware.com</a>
- [21] Xen, Type-1 Virtual Machine, <a href="https://xenproject.org">https://xenproject.org</a>
- [22] FreeBSD Chroot Jails, <a href="https://docs.freebsd.org/en\_US.ISO8859-1/books/handbook/jails.html">https://docs.freebsd.org/en\_US.ISO8859-1/books/handbook/jails.html</a>
- [23] Solaris Zones, Oracle Solaris Zones Introduction, https://docs.oracle.com/cd/E36784\_01/html/E36848/zones.intro-1.html
- [24] Virtuozzo Containers, product retired, <a href="https://www.virtuozzo.com/support/all-products/virtuozzo-containers.html">https://www.virtuozzo.com/support/all-products/virtuozzo-containers.html</a>
- [25] OpenVZ, Open source container-based virtualization for Linux, <a href="https://openvz.org/">https://openvz.org/</a>
- [26] Cgroups (previously process containers), <a href="https://lwn.net/Articles/236038/">https://lwn.net/Articles/236038/</a>
- [27] Linux Containers (LXC) project, Infrastructure for container projects, <a href="https://linuxcontainers.org">https://linuxcontainers.org</a>
- [28] Docker Swarm, Swarm mode key concepts, https://docs.docker.com/engine/swarm/key-concepts
- [29] Etcd, a distributed, reliable key-value store for the most critical data of a distributed system, <a href="https://etcd.io/">https://etcd.io/</a>
- [30] Jenkins, free and open source automation server, <a href="https://www.jenkins.io/">https://www.jenkins.io/</a>
- [31] Source-To-Image (S2I), toolkit and workflow for building reproducible container images from source code, <a href="https://github.com/openshift/source-to-image">https://github.com/openshift/source-to-image</a>
- [32] Checkmk Raw Edition, <a href="https://checkmk.com/product/raw-edition">https://checkmk.com/product/raw-edition</a>



- [33] GitLab roles, <a href="https://docs.gitlab.com/ee/user/permissions.html">https://docs.gitlab.com/ee/user/permissions.html</a>
- [34] GDPR, regulation 2016/679, https://eur-lex.europa.eu/eli/reg/2016/679/oj
- [35] Suricata, opensource threat detection engine, <a href="https://suricata.io/">https://suricata.io/</a>
- [36] ClamAV, open-source antivirus engine for detecting trojans, viruses, malware & other malicious threats, <a href="https://www.clamav.net/">https://www.clamav.net/</a>





# **Appendix A** Annexes

The following documents are released as confidential documents restricted only for members of the consortium (including the Commission Services):

- Slideset of the CI/CD Webinar held: GATEKEEPER-WP4-GK\_CI-CD\_Webinar\_HPE.pptx
- Document about how to setup a S2S VPN: GATEKEEPER-WP4-GK\_Data\_Centre\_Access\_Site\_To\_Site\_HPE\_v1.05.pdf





# **Appendix B** User Registration Form Text for S2S VPNs

For the sake of readability and accessibility, we report the full text of the User Registration Form shown in Figure 3 – Preface of the S2S User Registration Form.

#### *GATEKEEPER*

HPF Data Centre Site-To-Site

Remote Access Request

Please complete the following form. Your data will be collected by HPE as GATEKEEPER partner for the sole purposes of account management and user access provisioning to HPE Data Centre infrastructure hosting the GATEKEEPER project services. You must give your consent to HPE for processing your data according to the HPE Privacy policy available using the link located at the bottom of the form. The HPE responsibilities are described in the link.

This computer network is provided for authorised use only. Users have no explicit or implicit expectation of privacy. Any or all uses of systems of this computer network and all data in this Data Centre may be intercepted, monitored, recorded, copied, audited, inspected, and disclosed to authorised sites and law enforcement personnel, as well as authorised officials of other agencies. By accessing to this computer network, the user consent to such disclosure at the discretion of authorised site personnel. Unauthorised or improper use of this network may result in administrative disciplinary action, civil and criminal penalties. By accessing to this computer network, you indicate your awareness of and consent to these terms and conditions of use. Users are responsible for ensuring that they act in accordance with this policy and other policies and legislation applicable to the HPE network.

Please note that by submitting this form you also agree to have your activities on HPE Data Centre monitored for compliance and security purposes. You must not disable (or attempt to) or change any security control, as well as respect the assigned privileges and not escape the perimeter of your role.

Should you have any inquiry please contact HPE project partner for further information.

Technical details about how to access HPE Data Centre services are published on Alfresco project document management system. You will receive an email as soon as your access will be granted with more information.

Notice: it is strictly forbidden to share this registration link outside involved GATEKEEPER project partners.

HPE GATEKEEPER Team

---

https://www.gatekeeper-project.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 857223





# **Appendix C** User Registration Form Fields for S2S VPNs

For the sake of readability and accessibility, we report the fields of the Partner Registration Form for S2S VPNs shown in Figure 4 – Fields of the S2S User Registration Form.

- \* Required
- 1. Partner Single Point of Contact First Name

Please insert your name and optional middle name

2. Partner Single Point of Contact - Surname \*

Please insert your surname

3. Partner Single Point of Contact - E-Mail Address

Please specify the contact e-mail address (e.g., <u>user@example.org</u>). This e-mail address will be used only for communications about HPE data centre and site-to-site connection. Only valid GATEKEEPER partner e-mail address will be accepted.

4. Organisation Name \*

Please specify your organisation: **[here we show a drop down list of consortium partners' acronyms]** 

5. I have read and I accept the HPE Privacy policy for personal data collection available at the URL: https://www.hpe.com/emea\_europe/en/legal/privacy.html \*





# Appendix D Internal HPE Portal home page

For the sake of readability and accessibility, we report the text of the Internal HPE Portal home page shown in Figure 6 – Internal HPE Portal (in **green** the additions with respect to D4.1).

#### **HPE SECLAB**

This page provides helpful links for HPE SECLAB made for GATEKEEPER.

### **Identity Management**

Go here to manage your user account.

Access Password Change: [internal link]

You can change your password or reset it if it expires.

#### **SECLAB Internal Certification Authority**

Import the HPE SECLAB into your browser or other services. Download the SECLAB Internal CA to trust the provided services: *linternal linkl* 

## **For Developers**

Tools provided for authorised developers only.

Access OKD Console: linternal linkl

OKD is a distribution of Kubernetes optimized for continuous application development and multi-tenant deployment.

Access HPE Ezmeral Console for Big Data/AI/ML: [internal link]

HPE Software platform designed to run both cloud-native and non-cloud native applications in containers. More info.

Access CI/CD Tools

GitLab: [internal link] - source code repository

Nexus: [internal link] - container registry

Jenkins: [internal link] - build and deploy automation server

Tools to enable project software development.





### Important notice

This computer network is provided for authorised use only. Users have no explicit or implicit expectation of privacy. Any or all uses of systems of this computer network and all data in this Data Centre may be intercepted, monitored, recorded, copied, audited, inspected, and disclosed to authorised sites and law enforcement personnel, as well as authorised officials of other agencies. By accessing to this computer network, the user consent to such disclosure at the discretion of authorised site personnel. Unauthorised or improper use of this network may result in administrative disciplinary action, civil and criminal penalties. By accessing to this computer network, you indicate your awareness of and consent to these terms and conditions of use. Users are responsible for ensuring that they act in accordance with this policy and other policies and legislation applicable to the HPE network.

#### **About this Project**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857223

[GATEKEEPER] [Hewlett Packard Enterprise - HPE]

(c) Copyright 2021 Hewlett Packard Enterprise Development Company, L.P. Valid agreement required.