HEROES

Project Title	Hybrid Eco Responsible Optimized European Solution
Project Acronym	HEROES
Grant Agreement No.	956874
Start Date of Project	01.03.2021
Duration of Project	24 Months
Project Website	heroes-project.eu

D3.4 – Best Practice Library

Work Package	**WP 3, Infrastructure as a Service**
Lead Author (Org)	**Davide Pastorino (Do IT Systems s.r.l.)**
Contributing Author(s) (Org)	**Benjamin Depardon (UCit), Anaelle Dambreville (UCit), Julita Corbalan (BSC), Elisabeth Ortega (HPCNow!)**
Reviewed by	**Elisabeth Ortega (HPCNow!), Julita Corbalan (BSC)**
Approved by	**Name (organization)**
Due Date	**01.09.2022**
Date	**12.10.2022**
Version	**V1.1**

Dissemination Level

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

HEROES

Versioning and contribution history

Version	Date	Author	Notes
0.1	25.07.2022	Davide Pastorino (Do IT Systems s.r.l.)	TOC and V0.1
0.2	30.08.2022	Davide Pastorino (Do IT Systems s.r.l.)	BP for WP3: 3-1 to 3-6.
0.3	12.09.2022	Benjamin Depardon (UCit)	BP WP4
0.4	14.09.2022	Davide Pastorino (Do IT Systems s.r.l.)	Added References section. Review of BPs for WP3.
0.5	15.09.2022	Anaelle Dambreville (UCit)	BP WP4
0.6	16.09.2022	Davide Pastorino (Do IT Systems s.r.l.)	Updated summary, added more details to WP3 BP.
0.7	19.09.2022	Julita Corbalan (BSC), Davide Pastorino (Do IT Systems s.r.l.)	Added BPs to WP4 (merged)
0.8	19.09.2022	Elisabeth Ortega (HPCNow!)	BP WP2
0.9	30/09/2022	Davide Pastorino (Do IT Systems s.r.l.)	Final pass & corrections before review
0.10	4/10/2022	Elisabeth Ortega (HPCNow!), Julita Corbalan (BSC)	Review
1.0	5/10/2022	Davide Pastorino (Do IT Systems s.r.l.)	Final pass after review.
1.1	12/10/2022	Corentin Lefevre (Neovia)	Version approved by the MB

Table of Contents

HEROES

References and Applicable Documents

[1] Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nature Biotechnology, 35(4), 316–319. doi:10.1038/nbt.3820

[2] Kurtzer GM, Sochat V, Bauer MW (2017) Singularity: Scientific containers for mobility of compute. PLoS ONE 12(5): e0177459. https://doi.org/10.1371/journal.pone.0177459

[3] Min.io security documentation and encryption capabilities, https://docs.min.io/minio/k8s/security/security.html

[4] Keycloak security guidelines for mitigation, https://www.keycloak.org/docs/latest/server_admin/#mitigating-security-threats

[5] SSH Man in the Middle scenario, https://www.ssh.com/academy/attack/man-in-the-middle

[6] Random UUID probability of duplicates, https://en.wikipedia.org/w/index.php?title=Universally_unique_identifier&oldid=755882275#Random_UUID_probability_of_duplicates

TERMINOLOGY

Terminology/Acronym	Description
KPI	Key Performance Indicator
HPC	High-Performance Computing
BP	Best Practice
ACL	Access Control List
SSH	Secure Shell
SFTP	Secure File Transfer Protocol
EAR	Energy Aware Runtime

Executive Summary

The HEROES Project is aiming at developing an innovative European software solution allowing industrial and scientific user communities to easily submit complex Simulation and ML (Machine Learning) workflows to HPC (High Performance Computing) Data Centers and Cloud Infrastructures. It will allow them to take informed decisions and select the best platform to achieve their goals on time, within budget and with the best energy efficiency.

This document contains the descriptions and the principles of the technical "best practices" that have been identified during the development of the project, covering the following topics:

- *Integration with existing infrastructures*
- *Integration with cloud platforms*
- *Data management*
- *Security and Identity management*
- *System auditability*
- *Decision module*

The "best practices" in this document are to be intended more as policies and guidelines rather than practical implementations and ready-to-use solutions.

The development of this document is closely linked to the ongoing developments of WP2 and WP3 of the HEROES Project. As such, new best practices are expected to be defined and adopted throughout the life of the Project.

1 Introduction

Best Practices will be grouped in the document by Work Package, with the Work Package originating them working as the main topic. However, the BPs should be intended as applicable to the HEROES project in its entirety.

Best Practices (BP) are identified with the following structure:

> BP<WP#>-<BP#> - <Name of the BP>

Example: *BP3-1 - Deployment of HEROES Runtimes through Ansible* is the first BP of WP3.

For each Best Practice we identified the following list of items as sufficient to explain the motivation, the purpose, the scope, the impact and the limitations of each BP:

- Name and brief description of the BP
- Scope of application of the BP
- HEROES module(s) impacted by the BP
- Relevance & expected benefits
- Intended Outcome
- Evidence of implementation
- Limitations

The BPs in this document are in no particular order within each work package, as they have been collected during the organic development of the project.

2 Best Practices for Work Package 2

2.1 BP2-1 – Isolation of workflow steps according to their hardware requirements

CAE and AI simulation processes can be fragmented into different steps according to their individual requirements and be later described in a pipeline where the execution order should be defined. Those steps and their specific configuration related to the cluster where it will be executed , the scheduler options (partition name, number of CPUs, required memory, etc.) and the data storage should be properly described as Nextflow[1] inputs and configuration files.

2.1.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

2.1.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 3 – Energy Management & Optimisation
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 7 – Decision Module & Performance Metadata

2.1.3 Relevance & expected benefits

This BP was adopted to ensure that the computational resources demanded for each step are the most suitable ones in terms of computational power and energy efficiency.

2.1.4 Intended Outcome

Workflow executions will use all the potential of the HPC resources accessible by the user, minimizing its computational cost and carbon footprint and maximizing its performance.

2.1.5 Evidence of implementation

Separation of the pipeline steps involving the user interactivity from the ones involving more computational power, memory or disk, enhanced the speed of a CAE simulation. In this mentioned case the more demanding steps will be executed in an HPC cluster involving several cores, leaving the user interaction to a less powerful (hence, less power consuming) node.

2.1.6 Limitations

More than one HPC resource should be available for the execution of the workflow. If only one HPC resource is available, there is no reason to split the execution of the workflow into different steps.

2.2 BP2-2 – Workflow containerization

The software stack needed to execute the workflow should be containerized to maintain the immutability of the environment. In the case that the workflow execution involves more than one container, the container platform should be the same and the software stack should be as similar as possible (operating system and compilers). According to the actual design of the HEROES platform, the use of Singularity [2] container platform will be strongly recommended.

2.2.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

2.2.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management

2.2.3 Relevance & expected benefits

As the HEROES platform peculiarity is the execution of the different steps of a pipeline using the most suitable HPC resources for each one, the use of different software origins is prone to errors. Using the same software stack within a container and sharing it among the HPC facilities where the pipeline will be executed ensures reproducibility.

> **Commenté [A1]:** ugly expression. Find a new one
>
> **Commenté [A2R1]:** Slightly modified

2.2.4 Intended Outcome

This BP ensures the standardization of the software stack used to execute the different steps of a pipeline, avoiding the errors related to the usage of different software and library versions.

2.2.5 Evidence of implementation

The first mockup of a multi-step workflow execution was performed using the same container built in a Singularity and later defined to be used by Nextflow using the *-with-singularity* flag.

2.2.6 Limitations

The main limitation of the application of this BP is that the Singularity software is required to be installed on the host HPC facility.

2.3 BP2-3 – Access and usability of the user interface

The credentials to access the HEROES portal should be unique for each user, ensuring that the data buckets, cluster information and other configuration parameters are not modified without the consent of the owner of the account.

2.3.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

- System auditability
- Security and Identity management

2.3.2 HEROES module(s) impacted by the BP

- Module 1 – Identity Management
- Module 2 – Deployment & Integration
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 10 – User Interface & API

2.3.3 Relevance & expected benefits

One of the concerns of HEROES platform is data protection and privacy. The administrator of an organization will generate different users with different policies for data buckets or HPC resources to be used, which will be later displayed in the user interface and used to prepare new workflow executions.

2.3.4 Intended Outcome

This BP ensures that the user will use only the resources which they have access to.

2.3.5 Evidence of implementation

Please refer to Deliverable D2.3 Workflow Interfaces, currently under development, for details about this BP implementation. This BP has also been implemented throughout the architecture of the HEROES platform, for example in the Data Management module, the Deployment module and HEROES pseudo-FileSystem.

Commenté [A3]: @Davide can I say this?

Commenté [A4R3]: I think so, let's put this in the references table

2.3.6 Limitations

This BP does not apply if the owner of the account shares the login information with anyone under their own responsibility.

Commenté [A5]: The limitation is true in any case where the users share credentials. Generalizing.

3 Best Practices for Work Package 3

3.1 BP3-1 - Deployment of HEROES Runtimes through SSH

The deployment of HEROES runtimes on remote, existing infrastructures should be performed with the minimum impact on the existing infrastructure. One of the most common interfaces provided by existing HPC infrastructures with public access is through SSH channel with key-pair authentication. This same channel is also widely used by cloud providers to provide access to newly deployed instances.

This common channel is also the main interface used by Ansible, one of the most widely adopted tools for system automation and configuration management.

The HEROES team therefore identified Ansible as the main tool to be used to deploy and consistently manage the systems that are and will be part of the project's infrastructure.

3.1.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

3.1.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 8 – Logging & Accounting

3.1.3 Relevance & expected benefits

This BP was developed to answer the access and security requirements of remote HPC infrastructures. The adoption of this BP has the additional benefit of enabling the "Infrastructure as code" approach for the management of remote HEROES endpoints. By making use of existing configuration data within the HEROES main infrastructure this BP also enables a simple way to deploy different endpoints with a high degree of consistency and confidence: the code/operative procedure is static while the configuration data of each endpoint is dynamically fetched and used at run time from the Identity Management module.

This BP also helps with:

- Minimizing errors during deployments of new HEROES endpoints
- Providing a smaller surface for troubleshooting when errors do happen
- Enabling consistent configuration management and status auditing of endpoints

3.1.4 Intended Outcome

This BP enables the "infrastructure as code" approach for HEROES infrastructure endpoints. This approach has numerous advantages such as decreased risks, increased stability and consistency, faster iteration, cost optimization and self-documentation.

3.1.5 Evidence of implementation

This BP has been implemented in the tools used by the HEROES project for deployment and configuration of remote resources (ref. D3.2 – HEROES Deployment Suite).

3.1.6 Limitations

The BP cannot be used as-is in the following scenarios:
- Multi-factor authentication (MFA) is mandatory to access a remote resource
- SSH key-based authentication is not supported on a remote resource

However, the current requirements for integration of remote resource with the HEROES platforms include SSH key-based authentication as mandatory.

3.2 BP3-2 – Data Encryption in the main HEROES infrastructure

The project aims to deliver a multi-tenant capable infrastructure, so it is vital to have encryption for all data managed within the HEROES platform.
The main HEROES infrastructure is currently hosted, as a prototype, in the Do IT Systems laboratory environment. It is planned to move the infrastructure in the cloud once the platform development will have been completed.

Thus, the project aims to have all user data encrypted both "at rest" and "in transit". There are however a few exceptions due to intrinsic technical limitations of certain scenarios or software components.

3.2.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Data management
- Security and Identity management

3.2.2 HEROES module(s) impacted by the BP

- Module 1 – Identity Management
- Module 2 – Deployment & Integration
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 8 – Logging & Accounting

3.2.3 Relevance & expected benefits

This BP was developed to answer the multi-tenancy requirements of the HEROES platform. The adoption of this BP has the additional benefit of making the platform itself more secure and enabling better privacy of user data stored within the platform, especially in a hybrid cloud environment.

3.2.4 Intended Outcome

This BP's intent is to prevent data leaks from the HEROES platform.

Encryption at rest is designed to prevent the attacker from accessing the unencrypted data by ensuring the data is encrypted when on disk. If an attacker obtains a hard drive with encrypted data but not the encryption keys, the attacker must defeat the encryption to read the data.

Encryption in transit protects data content if communications are intercepted while data moves between two endpoints on a network.

3.2.5 Evidence of implementation

This BP has been implemented in the tools used by the HEROES project for data management and identity management. This includes the back-end components (min.io [3], keycloack [4]) and the APIs and interfaces at the core of the HEROES platform. Further details on both the back-end components encryption features and HEROES APIs capabilities are available in the project documentation. Wherever possible, the back end services and their communication endpoints make use of certificates and encryption in transit.

3.2.6 Limitations

There are some limited circumstances when data encryption is not possible.

For example, on a remote resource such as a HPC Center, data encryption is not directly managed by the HEROES project (if at all).

However, to be integrated with the HEROES platform, remote HPC Centers should provide:

- Means and to ensure data privacy for the project account on the remote HPC infrastructure (e.g., filesystems permissions and ACLs, strong passwords, key-based authentication, secure shell as main channel for multi-process communications, etc.)
- Secure remote access through VPN or SSH

3.3 BP3-3 – HEROES Private Network

The HEORES infrastructure has been designed to have a single public point of access through the end-user interface.

All administrative interfaces are not exposed directly on the public internet and can be accessed by an administrator only after the administrator has authenticated into the project secure network via VPN.

Commenté [A9]: Rephrased. I think this is the ultimate purpose for this BP.

Commenté [A10]: Put reference at te end of the doc to the product documentation for encryption features

Commenté [A11R10]: Reference par added at the beginning of the doc., similar to the one in D3.1. I don't know how to make reference links work ☹

Commenté [A12]: Add detail about communication channels being mostly encrypted with SSL/TLS and certificates

Commenté [A13R12]: Added reference to certs & encryption for services

All API endpoints are similarly not exposed on the public internet but are accessible only through secured channels within the HEROES Private Network.

Furthermore, all endpoints in the network are subject to control and verification through certificate or key based authentication.

3.3.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Security and Identity management

3.3.2 HEROES module(s) impacted by the BP

- Module 1 – Identity Management
- Module 2 – Deployment & Integration
- Module 8 – Logging & Accounting
- Module 10 – User Interface & API

3.3.3 Relevance & expected benefits

This BP was adopted to minimize the surface of attack from malicious actors and ensure that all registered endpoints within the network can be trusted at any time.

This significantly reduces the vulnerability to attacks from public networks, greatly decreasing security risks.

3.3.4 Intended Outcome

This BP enables ensures the protection of information in HEROES networks and maintain the security of information transferred within HEROES and towards any external entity. The BP also follows the principle of minimizing the attack surface by limiting the publicly accessible endpoints of the HEROES platform to the minimum.

3.3.5 Evidence of implementation

This BP has been implemented in the backend tools used by the HEROES project for data management and identity management. This includes the back-end components (min.io[3], keycloak[4]) and the APIs and interfaces at the core of the HEROES platform. Further details on both the back-end components for endpoint verification and HEROES APIs identity checks are available in the project documentation (e.g. in D3.3 HEROES Initial Platform Mockup).

> **Commenté [A14]:** Link to the product documentation in reference
>
> **Commenté [A15R14]:** Added references

3.3.6 Limitations

The HEROES Private network does make use of the public internet for the integration of remote HPC infrastructure. Attack scenarios for this kind of integration, like "man-in-the-

middle" [5] for SSH connections to and from remote HPC infrastructures, need to be properly mitigated following specific security guidelines.

3.4 BP3-4 – Use of Universally Unique Identifiers (UUID) in the HEROES Platform

Each user, group and organization registered in the HEROES platform shall have a platform wide UUID assigned. This UUID is generated and stored in the backend software used in Module 1, keycloak.
The UUID shall be used on the HEROES platform to:
- identify data ownership on remote HPC infrastructures
- implement data isolation between users
- identify the ownership of API calls
- track action ownership through auditable logs

3.4.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Data management
- Security and Identity management
- System auditability

3.4.2 HEROES module(s) impacted by the BP

- All Modules

3.4.3 Relevance & expected benefits

This BP was adopted to ensure that each and every action performed by users while using the HEROES platform and its APIs is tracked and auditable. The BP also provides a way to implement data ownership and data isolation in situations where a single service user has to manage all user's data (for example in a typical remote HPC platform).

3.4.4 Intended Outcome

This BP enables:
- Protection of information and data isolation in remote HPC infrastructures
- Platform-wide user identification
- Platform-wide auditability of user actions
- Identification of ownership of all API calls

3.4.5 Evidence of implementation

This BP has been implemented in the backend tools used by the HEROES project for data and identity management. This includes the back-end components (min.io, keycloak) and the APIs and interfaces at the core of the HEROES platform. Further details on both the back-end components for endpoint verification and HEROES APIs identity checks are available in the project documentation (deliverables "D3.3_Initial Platform Mockup" and "D3.2_Deployment tools templates and scripts").

The BP has also been implemented in the HEROES Remote pseudo-Filesystem (HRFS) for managing user data ownership and isolation.

All API calls are logged with the UUID of the originator. The logs are then auditable from the log analysis console.

3.4.6 Limitations

Two UUIDs could conflict in theory. However, the probability is extremely small [6] and the HEROES platform is going to use a relatively small quantity of UUIDs.

3.5 BP3-5 – Main infrastructure scalability

The HEROES infrastructure and its modules should be scalable in all relevant dimensions as the number of users and integrated platforms grows or shrinks. This can be implemented by using cloud-native software and solutions as the building blocks of the HEROES infrastructure.

3.5.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Data management

3.5.2 HEROES module(s) impacted by the BP

- Module 1 – Identity Management
- Module 2 – Deployment & Integration
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 8 – Logging & Accounting

Commenté [A16]: Insert ref to the architectural docs and the script/deployment deliverable

3.5.3 Relevance & expected benefits

This BP was adopted to ensure that the HEROES architecture will be flexible and able to cope with variable performance requirements, following the growth of its user base.

3.5.4 Intended Outcome

Following this BP the HEROES infrastructure has been built using readily available cloud native software, which is often also Open Source software. This also provides the benefit of having a solid foundation on which to build the platform itself, since most of cloud-native software are sponsored or developed by large companies such as Google, RedHat, Amazon, etc.

3.5.5 Evidence of implementation

As an example, for Module 4 – Data Management a user may require an increase in storage space for him/herself (scale up) while an organization may require additional users with their own private storage space (scale out). The tool at the base of this module, min.io, provides both these features natively, and the way it was integrated within the HEROES infrastructure provides even more flexibility to each Organization enrolled in the platform (see D3.3 for more details).

3.5.6 Limitations

The HEROES architecture scalability is currently limited in a few modules. For example, there is a single entry point for users through the UI (Module 10). Such limitation will be areas of additional developments after the end of the project.

4 Best Practices for Work Package 4

4.1 BP4-1 – Platforms micro-benchmarks

The HEROES decision module requires a set of initial micro-benchmarks to be run on all HPC and Cloud platforms included in the marketplace. All new platforms shall run a set of selected benchmarks on all the different types of hardware made available in the marketplace. Those benchmarks will be used to predict the jobs characteristics (runtime, energy, cost) for new clusters for which we have no history.

4.1.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Decision module

4.1.2 HEROES module(s) impacted by the BP

- Module 2 - Deployment & Integration
- Module 3 – Energy Management & Optimisation
- Module 7 – Decision Module & Performance Metadata
- Module 9 – Cost Management

4.1.3 Relevance & expected benefits

This BP was adopted to ensure that the HEROES decision module provide accurate submission/placement recommendations and costs estimations. The projection of jobs runtime and energy consumption on various types of hardware (physical or virtual) relies on resources performance comparison for the selected workload. These micro-benchmarks allow the decision module to estimate the behavior of the workload on hardware on which the workload has never run before.

4.1.4 Intended Outcome

We expect the decision module to be able to identify the optimal cluster on which to run the jobs based on cost, runtime and energy consumption predictions. It will choose among all available clusters, the three best clusters (based on the objective to optimize, i.e. energy, time, cost or the three together/performance) and rank them. The identified clusters will be returned to the user along with runtime, cost and energy consumption associated to each one of them.

4.1.5 Evidence of implementation

Module 7 – Decision module: implements the algorithms (Multi-Criteria Decision Analysis) required for the decision process. Standard benchmarks (e.g., HPL, HPCG, STREAM…) were conducted on AWS and Azure, and can be reproduced on any HPC center.

4.1.6 Limitations

If no benchmark is available, the decision module won't be able to correctly compare the different hardware available and make a good performance/energy/cost placement decision. However, when the software has enough jobs history on the cluster, the benchmark will no longer be needed.

4.2 BP4-2 – Decision module predictors training

The HEROES decision module requires historical information on previously run jobs to train its predictors. The more data it has, the better the predictions will be. We usually recommend a history of 3-6 months.

Commenté [A17]: I think we have to add the energy also here. May be just:
The projection of jobs runtime and energy consumption on various …

The decision module will actually return (for max 3 clusters among all available):
- the ranking (for the best 3)
- the predicted cost, runtime and energy of the jobs on each of the 3 clusters

Commenté [A18]: I would say that this is a limitation at first. Because at some point we will have enough logs history to train the models, and we will be able to have the predictions without the need of using the benchs.

4.2.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Decision module

4.2.2 HEROES module(s) impacted by the BP

- Module 7 – Decision Module & Performance Metadata
- Module 8 – Logging & Accounting

4.2.3 Relevance & expected benefits

This BP was adopted to ensure that the HEROES decision module provides accurate submission/placement recommendations: the predictions on jobs characteristics (runtime, energy, costs…) are used during the decision process. Depending on the objective to optimize, the decision module will use a combination of multiple predictions to decide which are the best clusters on which to run the jobs. For example, if the user wants to optimize the time, the decision module will use two criteria to make a decision: the runtime and the slowdown (time to result / runtime).

4.2.4 Intended Outcome

Following this BP the HEROES decision module has been built using private or public historical workload archives (such as the Parallel Workload Archive - https://www.cs.huji.ac.il/labs/parallel/workload/). We expect the process and algorithms used in the decision module to be as robust and accurate as possible.

4.2.5 Evidence of implementation

Module 7 – Decision module: implements the algorithms required for the decision process. The tests that were conducted to validate their accuracy have been conducted on several datasets with at least 3 months of data.

4.2.6 Limitations

If no historical data is present, the decision module has to use the benchmarks to be able to compute jobs characteristics based on the predictions obtained on other clusters with history. If no benchmark has been run, the decision module will have to wait to have enough historical data to be able to predict jobs characteristics.

4.3 BP4-3 – Automatic generation and deployment of runtimes

The installation of HEROES components should be easily automated and require little to no manual configuration to set up. If all the required dependencies are installed beforehand, the detection of the necessary paths will be automatic. This has been achieved by implementing and improving the detection of the environment.

4.3.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

4.3.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management

4.3.3 Relevance & expected benefits

This BP was adopted to ensure that the HEROES components were easy to install and that their deployment could be easily automated, facilitating its adoption.

4.3.4 Intended Outcome

This BP enables standardization and automation of all the procedures required for the installation of HEROES components. By simplifying the procedure of installation, many man-made errors can be prevented, smoothing out the deployment of HEROES.

4.3.5 Evidence of implementation

This BP has been implemented in the HEROES project's components that must be installed by making the *./configure* command detect all dependencies without user input.

4.3.6 Limitations

If the necessary dependencies are not installed, the tool will try to work around it but some functionality may be missing.

4.4 BP4-4 – Dynamic initialization and runtime configurability

The HEROES runtime requires certain components to be installed or a certain level of privileges to be able to run with all functionalities. If any of those requirements are missing, it will dynamically try to find an available alternative or it will turn off the directly related

Commenté [A22]: Hacer que el configure sea 100% automatico y no tener que poner nada manualmente

Commenté [A23R22]: This will be managed automatically by the ansible deployment scripts.

Commenté [A24]: Idea: durante la inicializacion se detecta que hay y que no hay y se configura EAR en funcion de eso (politicas etc). Si EAR no esta instalado se crea una configuracion automatica (dummy) y el resto es 100% compatible

Commenté [A25R24]: Correct, this can be implemented as a "pre-script" within the workflow, right before launching the executable.

functionalities while keeping the other ones. Even if some of HEROES' components are not installed, the runtimes will try to provide as many functionalities as they can.

4.4.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

4.4.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 6 – Workflow & Job Management

4.4.3 Relevance & expected benefits

This BP was adopted to ensure that HEROES runtimes provide a minimum of functions when the system does not have all the necessary tools to provide full functionality, instead of providing no benefit at all, therefore always being ready to be deployed and do something.

4.4.4 Intended Outcome

The BP allows a minimum of functionality of HEROES runtimes even in systems where all prerequisites are not present instead of not working at all.

4.4.5 Evidence of implementation

This BP has been implemented in HEROES' runtimes by creating a baseline configuration with the minimum of functionalities in it and dynamically adding the rest if they are available.

4.4.6 Limitations

The limitations of the runtimes when some underlying tool is not available will still be there if no alternative is found.

4.5 BP4-5 – Utilization of standard and portable models for data sharing

Exporting data generated by HEROES components and runtimes shall be done with standard, extensible and portable formats to increase compatibility with external tools for data processing and visualization. The chosen format is CSV files, which are widely used and its a plain text format which is easy to modify, move or copy.

> **Commenté [A26]:** Utilizar formatos CSV para intercambio de datos (standard y portable, extensible). No tiene requerimientos en cuanto a herramientas externas y se puede mover ficheros con scp.

4.5.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms
- Data management
- Decision module

4.5.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 4 – Data Management
- Module 5 – Application and Container Management & Orchestration
- Module 6 – Workflow & Job Management
- Module 8 – Logging & Accounting

4.5.3 Relevance & expected benefits

By using a standardized format this BP allows HEROES runtimes and modules to easily interoperate with external tools that also use said format as well as increasing readability of the output data when looking directly at the file.

4.5.4 Intended Outcome

This BP enables high levels of compatibility with external tools by using a standard format. Furthermore, it increases the readability of the files by themselves as well as makes them portable by virtue of being a text file variant.

4.5.5 Evidence of implementation

This BP has been implemented by creating a data exporting system in HEROES runtimes and modules that create simple files in CSV format.

4.5.6 Limitations

This BP has no limitations. The format is widely used and, in the case where a necessary tool requires a different format, it is easy to read CSV files and then transform them.

4.6 BP4-6 – Extensibility and configurability at runtime to support new scenarios/architectures

HEROES modules need to be prepared for deployment in environments where necessary tools or other dependencies might be unavailable and needs to be able to proceed without them by either disabling the functionality or by having alternative ways to provide said

functionality. Furthermore, HEROES modules shall provide a method to provide inputs to them without the need of modifying the code already implemented.

4.6.1 Scope of application of the BP

- Integration with existing infrastructures
- Integration with cloud platforms

4.6.2 HEROES module(s) impacted by the BP

- Module 2 – Deployment & Integration
- Module 6 – Workflow & Job Management

4.6.3 Relevance & expected benefits

Implementing this BP allows HEROES to be deployed in any system regardless of its installed tools. Furthermore, it allows for extensibility of the modules, since they can load different subsystems depending on the environment, as well as accelerate the development of related systems by virtue of removing the need to recompile and redeploy HEROES to test their compatibility.

4.6.4 Intended Outcome

This BP prepares HEROES for deployment on any system as well as allowing it to be changed without necessitating reinstallation. Furthermore, by preparing it for deployment in "unprepared" systems it increases the robustness of the internal systems by providing sane defaults when a dependency is not present or not working properly.

4.6.5 Evidence of implementation

This BP has been implemented by creating start up checks to see whether certain systems are there or not, and in the case where they are not present a sane default has or an alternative have been implemented. Additionally, several HEROES components can be modified by setting or changing environment variables which allows for data input changes without having to modify the already operative code.

4.6.6 Limitations

This BP makes HEROES ready for deployment, but it may not be functionally ready or at least vacant in some ways without the necessary dependencies.