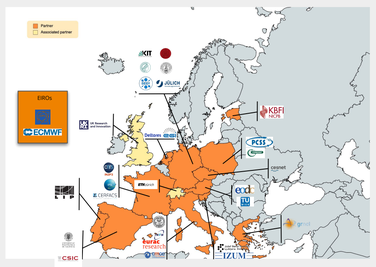
HPC, HTC and Cloud: converging toward a seamless computing federation with interLink

Daniele Spiga (INFN-Perugia), Diego Ciangottini (INFN-Perugia), Tommaso Boccali (INFN-Pisa), Andrej Filipcic (JSI), Andrea Manzi (EGI), Ahmed Shiraz Memon (Juelich), Teo Prica (IZUM), Tommaso Tedeschi (INFN-Perugia), Giacomo Surace (INFN-Perugia)



29 **Participants**, including 1 affiliated entity and 2 associated partners

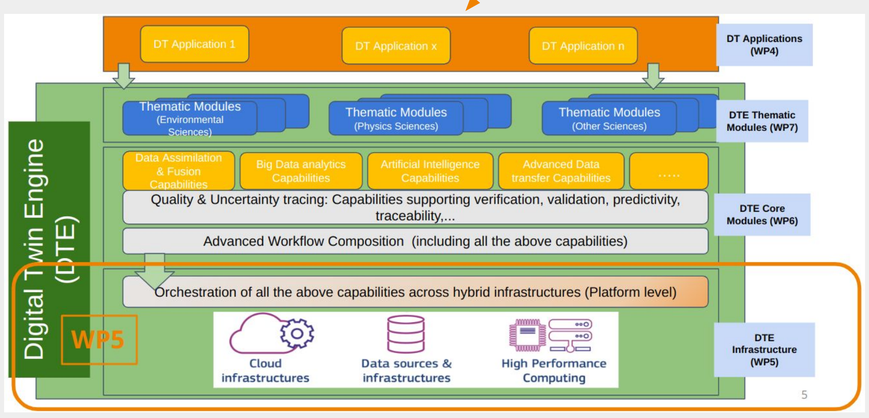
Consortium at a glance

10 Providers	11 Technology providers	14 Community representants
cloud, HTC , HPC resources and access to Quantum systems	delivering the DTE infrastructure and horizontal capabilities	from 5 scientific areas; requirements and developing DT applications and thematic modules

1.09.22 - 31.08.25 | Budget 11,7 M euro

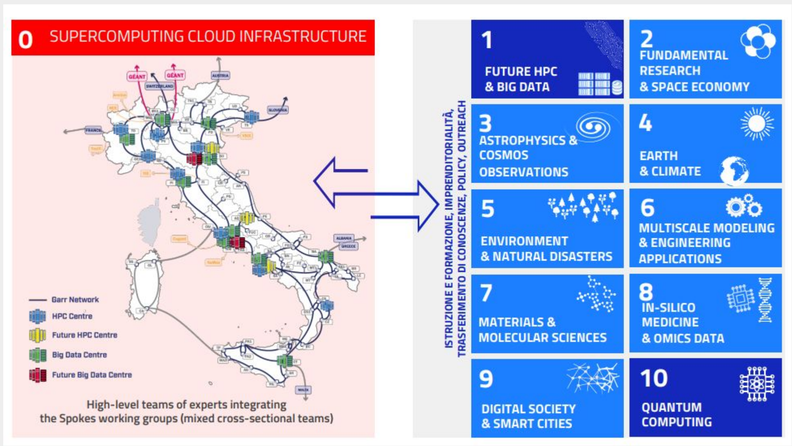
In the era of digital twins a federated system capable of integrating High-Performance Computing (HPC), High-Throughput Computing (HTC), and Cloud computing can provide a robust and versatile platform for creating, managing, and optimizing Digital Twin applications. **One of the most critical problems involve the logistics of wide-area with multi stage workflows that move back and forth across multiple resource providers.**

interTwin co-designs and implements the prototype of an interdisciplinary Digital Twin Engine (DTE) – an open source platform based on open standards that offers the capability to integrate with application-specific Digital Twins (DTs). Its functional specifications and implementation are based on a co-designed interoperability framework and conceptual model of a DT for research – **the DTE blueprint architecture.** The ambition of interTwin is to create consensus on a common approach to the implementation of DTs that is applicable across the whole spectrum of scientific disciplines that will facilitate developments and interoperability across different DTs.



THE CONTEXT

- Build a world-class supercomputing cloud infrastructure to store, manage and process all the produced data
- Set up centers of excellence with teams of high-level experts to develop domain applications
- Set up strong links between Academia, Industry and Public Administration
- Train the next generation of data scientists and managers to become experts in the digital transition
- Implement structural measures for innovation and for dissemination



The ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, conducts R&D, nationally and internationally, for innovation in high-performance computing, simulations, and big data analytics. This aim is pursued through a state-of-the-art infrastructure for high-performance computing and big data management, which leverages existing resources and integrates emerging technologies.

Our Challenge

The main goal is to provide software solutions to enable resources provisioning on a wide range of compute providers, primarily will deal with:
customised runtime environment at user level
deployment of specific network configuration on the nodes, where/if needed
Edge services to enable federation and orchestration layer communication

- **Federate a set of highly heterogeneous and disparate providers**
  - **enabling a "transparent payload offloading"**
- **Flexible extension of existing computing clusters to remote resources**
  - **E.g. HPC, specialized hardware...**

Our vision

We envision a model where such a challenge can be addressed enabling a "transparent offloading" of containerized payloads using the Kubernetes API primitives, enabling a transparent access to any number of external hardware machines and type of backends. Thus we created the interLink project, an open source extension to the concept of Virtual-Kubelet with a design that aims for a common abstraction over heterogeneous and distributed backends. The primary goal is to have HPC centers exploitable with native Kubernetes APIs with an effort close to zero from all the stakeholders' standpoint.

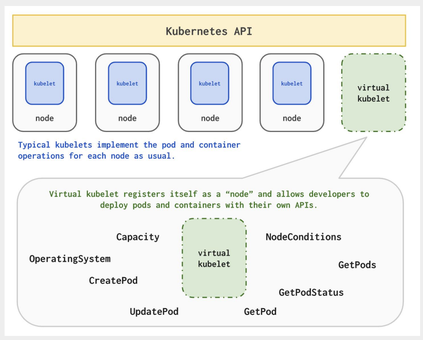
- **Extend the container orchestration de-facto standard (K8s) to support offloading under the hood**
  - **With little to no knowledge required from the end user perspective**
- **What about making a Kubernetes cluster send your containers/pod to a "virtual" node that seamlessly takes care of your application lifecycle on a remote server/hpc batch queue?**
  - **While exposing the very SAME experience of running a pod on the cloud resources**
    - Meaning that the API layer exposed is the "normal" set of K8s apis

THE PROJECT

We extended the VK solutions with a first draft of a generic API layer for delegating pod execution on ANY remote backend
k8s POD requests are digested through the API layer (e.g. deployed on an HPC edge) into batch job execution of a container.

Virtual kubelet (VK):
"Open-source Kubernetes kubelet implementation that masquerades as a kubelet. This allows Kubernetes nodes to be backed by Virtual Kubelet providers"

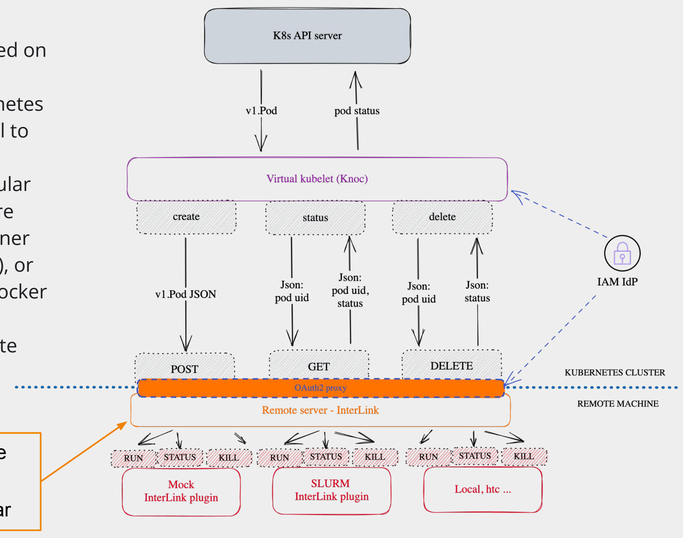
N.B. It can be imagined as a translation layer:
- It takes your pod and executes it wherever



interLink

- A Kubernetes Virtual Node: based on the VirtualKubelet technology. Translating request for a kubernetes pod execution into a remote call to the interLink API server.
- The interLink API server: a modular and pluggable REST server where you can create your own Container manager plugin (called sidecars), or use the existing ones: remote docker execution on a remote host, singularity Container on a remote SLURM batch system.

- Oauth2 proxy: authN with IAM and authZ configurable on aud and groups
- "Digests" and manipulates calls from VK to the sidecar



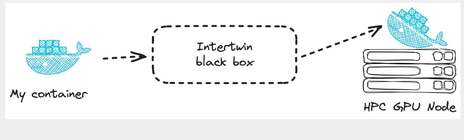
ARCHITECTURE

Current Status and Future Directions

A first version of interLink software that enables the offloading system, has been prototyped and made available through testbeds:

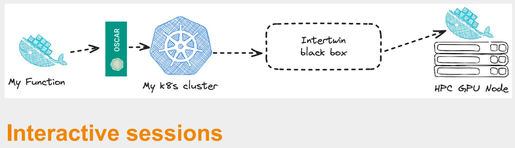
Simple process (POD) execution

To create a simple container to be scheduled on a remote Slurm batch on an HPC center



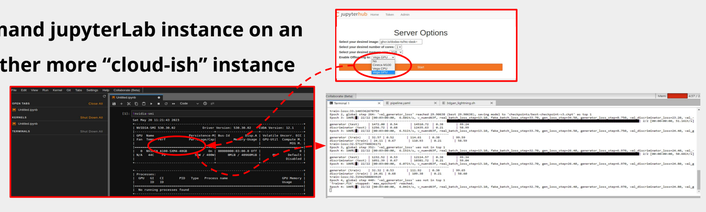
Serverless workflows

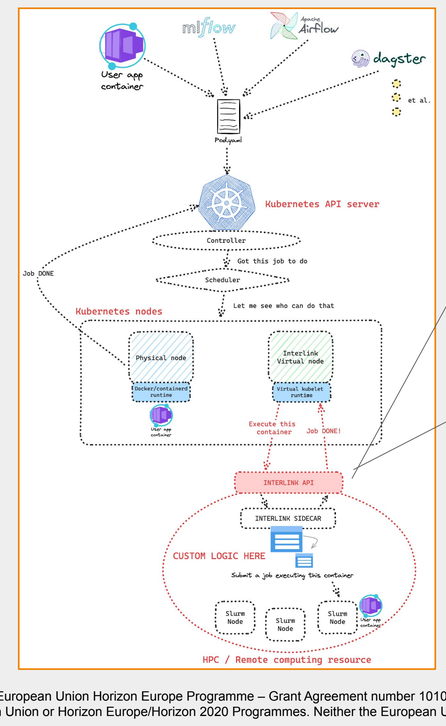
To execute payload in response of an external trigger being it a storage event or a web server call



Interactive sessions

Spawning on-demand jupyterLab instance on an HPC along with other more "cloud-ish" instance on K8s.





Vega
- **Deployed interLink Slurm layer on a VM on the edge of the HPC**

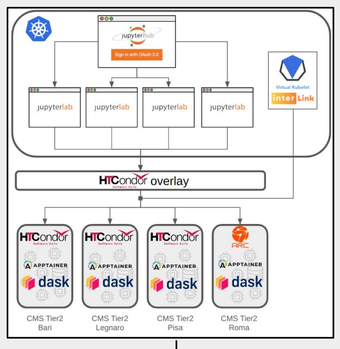
Juelich
- **Deployed "remote docker run" interLink on a login node**
- **Developed a custom plugin to offload containers on UNICORE**

PSNC
- **Deploying a network segregated setup**

HTC / Grid Sites

Grid Tier2 Italian Federation
- **ReCaS Bari, Rome, Legnaro and Pisa**
- **Central interLink deployed to remotely access all the Tier2. HTCondor and ARC Plugin/Sidecar**

KBFI
- **Deployed interLink layer to enable the offloading to the HTC Site. ARC Sidecar used**



Adopted solution to offload LHC (quasi-)interactive analysis via Dask Distributed
See https://indico.cern.ch/event/1330797/contributions/5796597/ contribution on High Rate analysis

HPC and Euro HPC

Summary and next steps

Integrating heterogeneous resources via interLink approach might represent a game changer in our scientific context. We foresee further developments in order to accommodate new use cases both in the context of ML/AI and also on pre-processing and interactive data analysis, a key point under discussion in many domains included in the High Energy Physics. Moreover, being interLink a community agnostic solution, it can be integrated in several Virtual Research Environments emerging nowadays.

Further testbeds are also foreseen in order to gather feedback from the resources providers perspectives. So far the experience made with EuroHPC centers thanks to the interTwin project was a key to reach the current level of maturity.

This work is partially supported byf the interTwin project (https://www.intertwin.eu). The interTwin project is funded by the European Union Horizon Europe Programme – Grant Agreement number 101058386. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or Horizon Europe/Horizon 2020 Programmes. Neither the European Union nor the granting authorities can be held responsible for them.

"This work is partially supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU".