HPC and AI for Prevision and Prediction of environmental phenomena (Day 1)

Universidad Carlos III de Madrid
April 20th and 21st, 2023

Prof. Raffaele Montella
University of Napoli "Parthenope"
https://raffaelemontella.it https://www.uniparthenope.it

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN
EuroHPC Joint Undertaking
ADMIRE malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

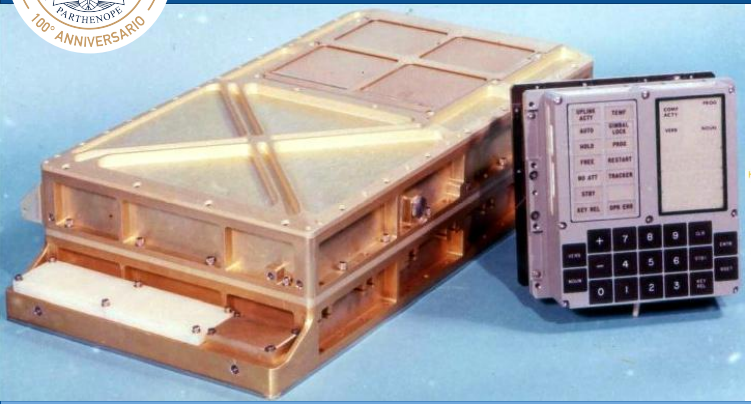
“...check the weather app…”

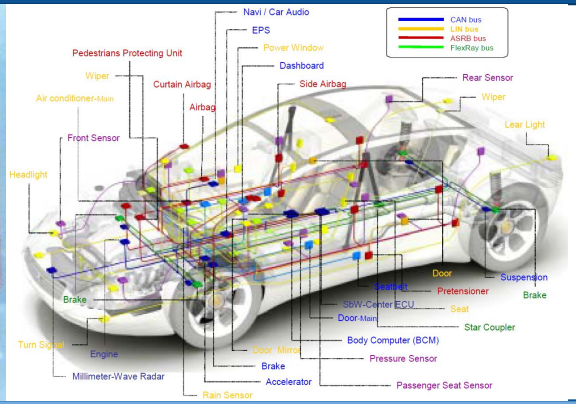
Performing a weather forecast is easy as 1, 2, 3:

an app **does** that in a matter of secs...

EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

“...and…





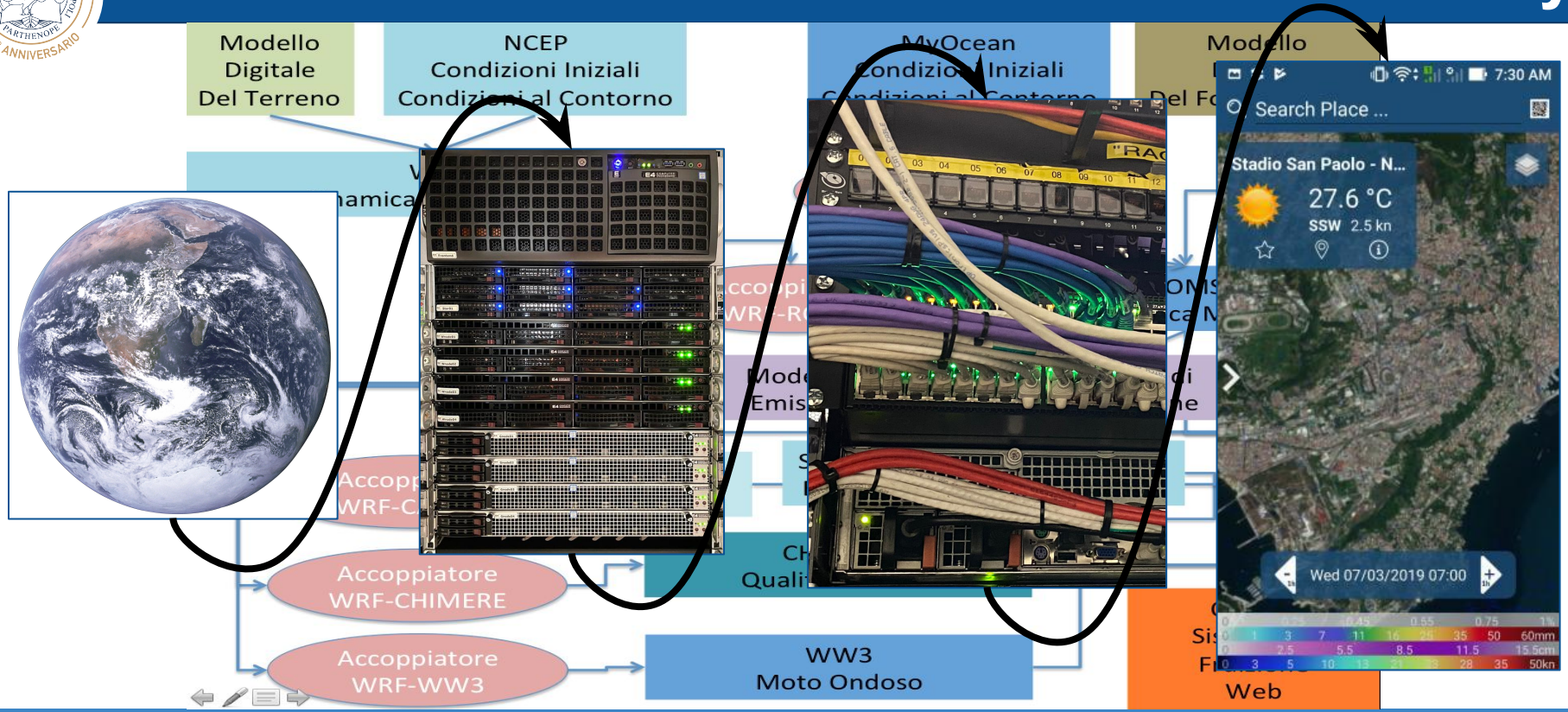
	Cray-2 Supercomputer	iPhone 8
GFLOPS	1.9	325.0
CPU Speed	244 MHz	2.39 GHz

Since 1999 each car is provided with more computing power than what has been used to send man to the Moon and back in 1969.

An iPhone 8 (2017) has more computing power than a 1985 Cray-2 supercomputing.

Nevertheless...

...it is not so easy!



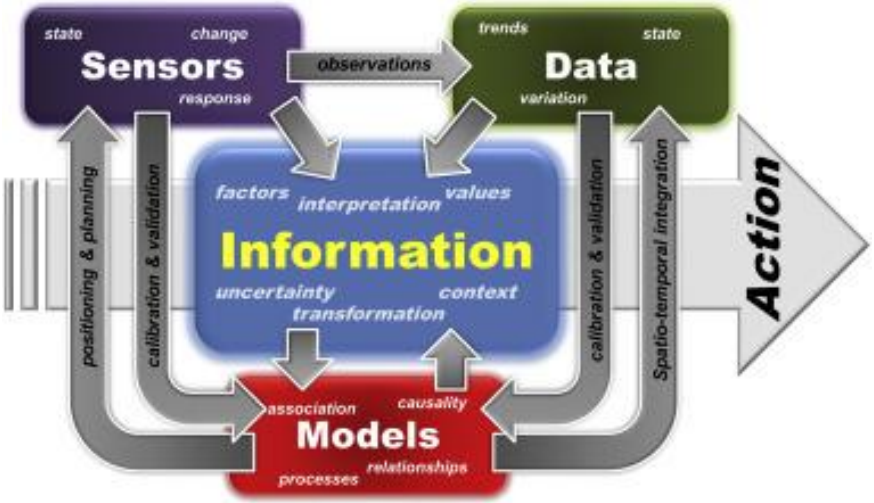
The app is the very last step, the iceberg peak of a complex workflow relying on science, technology, supercomputing and artificial intelligence.

Ice Breaking

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Introduction

- Environmental monitoring, simulations, forecasting, nowcasting and predictions are a strategic issue.
- The amount of collecting or produced data is skyrocketing.
- Reducing data dimension is needed for information extraction.

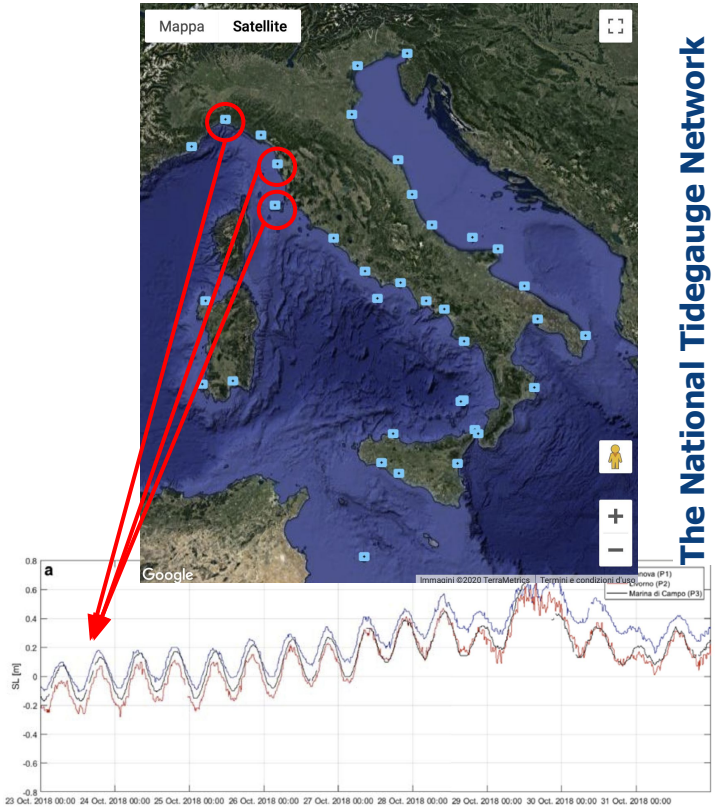


Reis, Stefan, et al. "Integrating modelling and smart sensors for environmental and human health." *Environmental Modelling & Software* 74 (2015): 238-246.

- Decision makers need for environmental data for natural hazard and accident responses, nature conservation and economic development.

Context definitions: “Monitoring”

- Monitoring: data sampling with a given space and time frequency.
  - Sensors aggregated in Instruments.
  - Point sensors (weather stations, air quality data loggers, buoys)
    - Wild, rural and urban zones.
    - Example: https://arrayofthings.github.io
  - Aerial sensors (satellites, airborne, shipborne, autonomous, radards).
  - Time series.
  - **Multidimensional georeferenced data.**

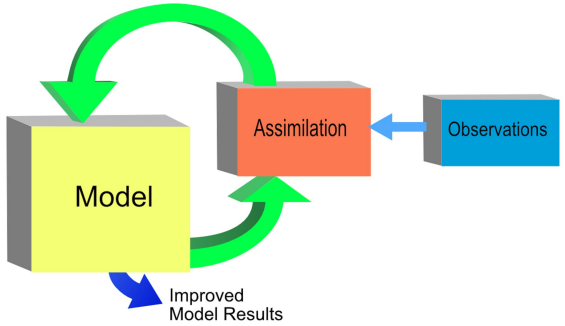


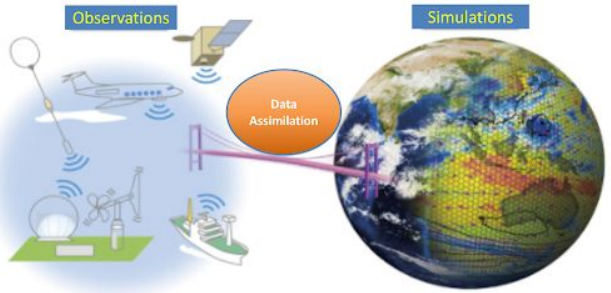
The National Tidegauge Network

EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Context definitions: “Simulations”

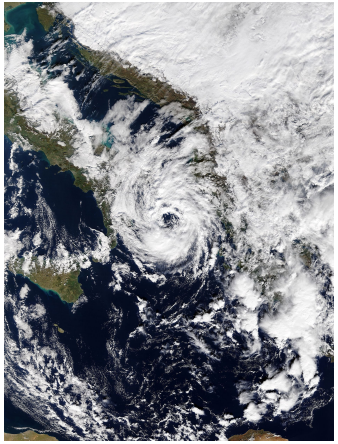
- Data produced by numerical models based on equations for physics, chemistry, …
  - The initial and boundary conditions are prepared from measured data from environmental monitoring.
  - Results affected by the initial and boundary conditions.
  - Compute intensive.
  - Wall clock time (usually) is not critical.
  - **Multidimensional georeferenced data.**
- Simulation of scenarios (past, future).



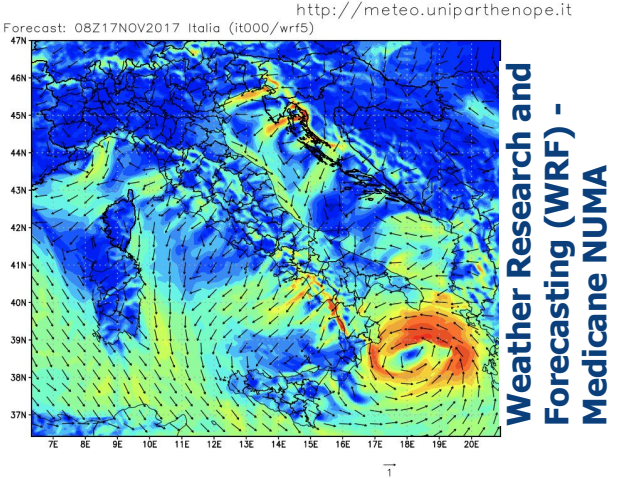


Context definitions: “Forecasting”

- Data produced by numerical models based on equations for physics, chemistry, …
  - The initial and boundary conditions are prepared from global/larger scale models..
  - Results strongly affected by the initial and boundary conditions.
  - Compute intensive.
  - Wall clock time is critical: the forecasts must be produced on time.
  - **Multidimensional georeferenced data.**
- Operational forecast (near future, days/weeks).



MODIS image captured by NASA’s Terra satellite



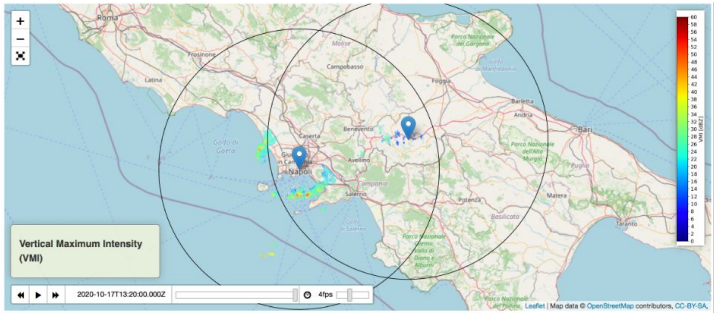
Weather Research and Forecasting (WRF) - Medicane NUMA

Context definitions: “Nowcasting”

- Data produced by geographically distributed sensor networks, …
  - It shares the technology with environmental monitoring.
  - Data is processed by humans/machine learning tools.
  - Data labelling is crucial.
  - Wall clock time is critical: data collection, processing, visualization.
  - **Multidimensional georeferenced data.**

- Operational forecasting (really near future, hours)

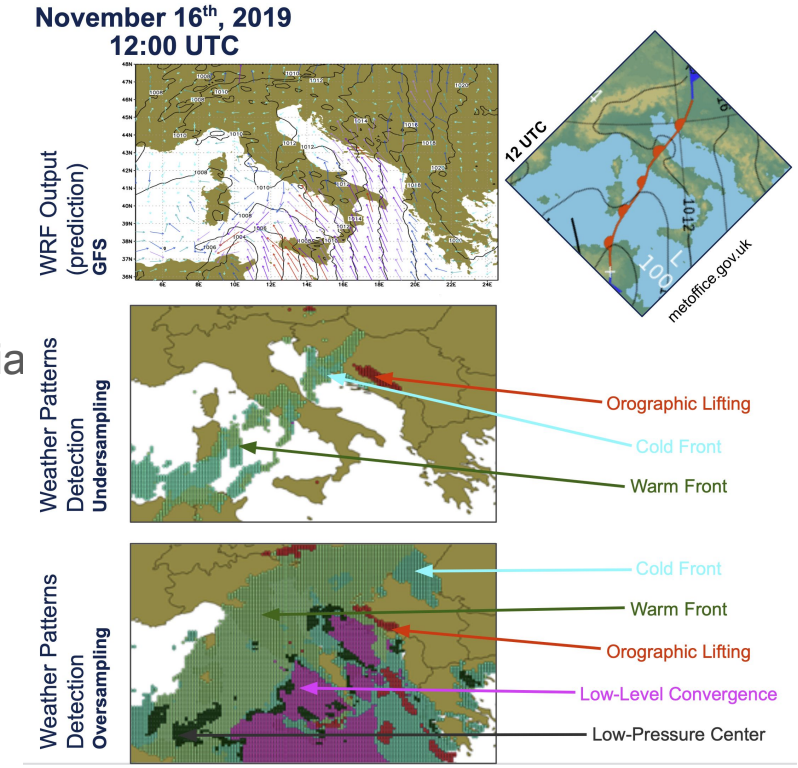
WR-10X Radar - meteo@uniparthenope



Context definitions: “Predictions”

- Data produced by machine learning algorithms.
  - Data sources:
    - Sensors and Instruments:
      →predictions from observed data.
    - Numerical models:
      →predictions as numerical models improvement.
  - Dataset preparation and data labelling are crucia
  - Diverse and different techniques.
  - Consume and produce **multidimensional georeferenced data.**

- Simulations and operational forecasting.



EuroHPC Joint Undertaking
ADMIRE
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE
malleable data solutions for HPC

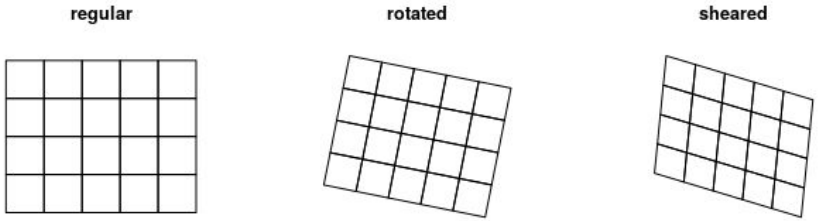
Multidimensional georeferenced data

- Data point dimensions:
  - Ensemble (different set of values for the same data point)
  - Time reference (gregorian, julian, epoch based)
  - Level (depth/height, pressure, *sigma*)
  - Latitude (any north/south reference)
  - Longitude (any west/east reference)

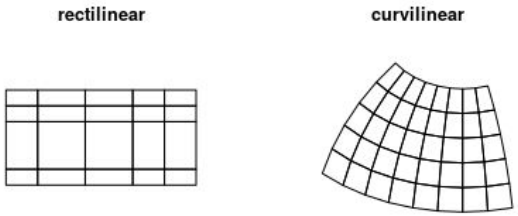
- Data point variables:
  - The values measured/calculated in the data point.
  - Have one, few or all data point dimensions (1D, 2D, …)
  - Attributes

Multidimensional georeferenced data

- Data distribution:
  - Regular grids
  - Not regular grids



- Geographical projections
  - None
  - Projected
  - EPSG4326



- Null values
  - Handled by custom labels

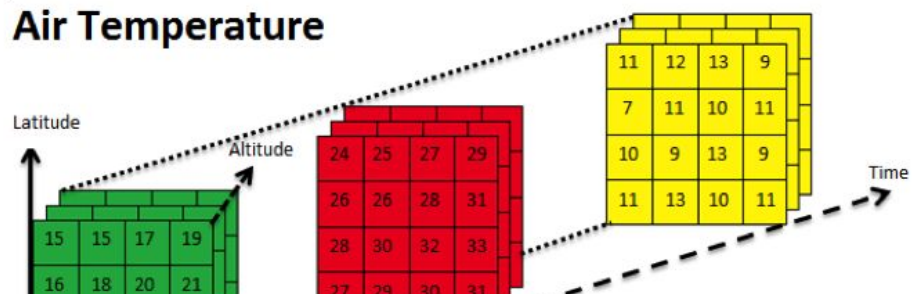
NetCDF: Network Common Data Form

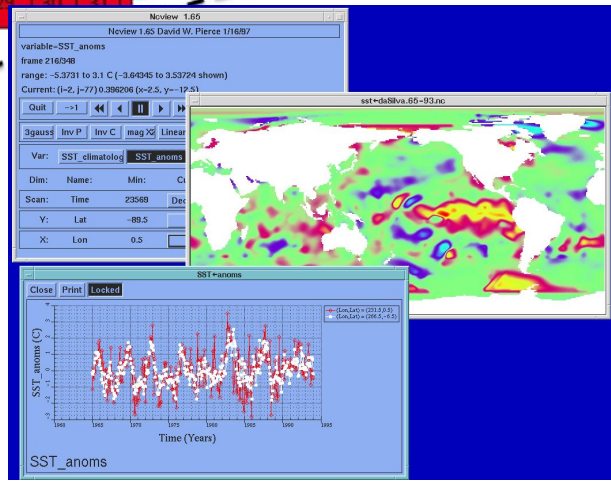
Definition:
"*NetCDF (Network Common Data Form) is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data*".

- Hosted by the Unidata program at the University Corporation for Atmospheric Research (UCAR).
- The format is an open standard.
- NetCDF Classic and 64-bit Offset Format are an international standard of the Open Geospatial Consortium.

NetCDF: Network Common Data Form

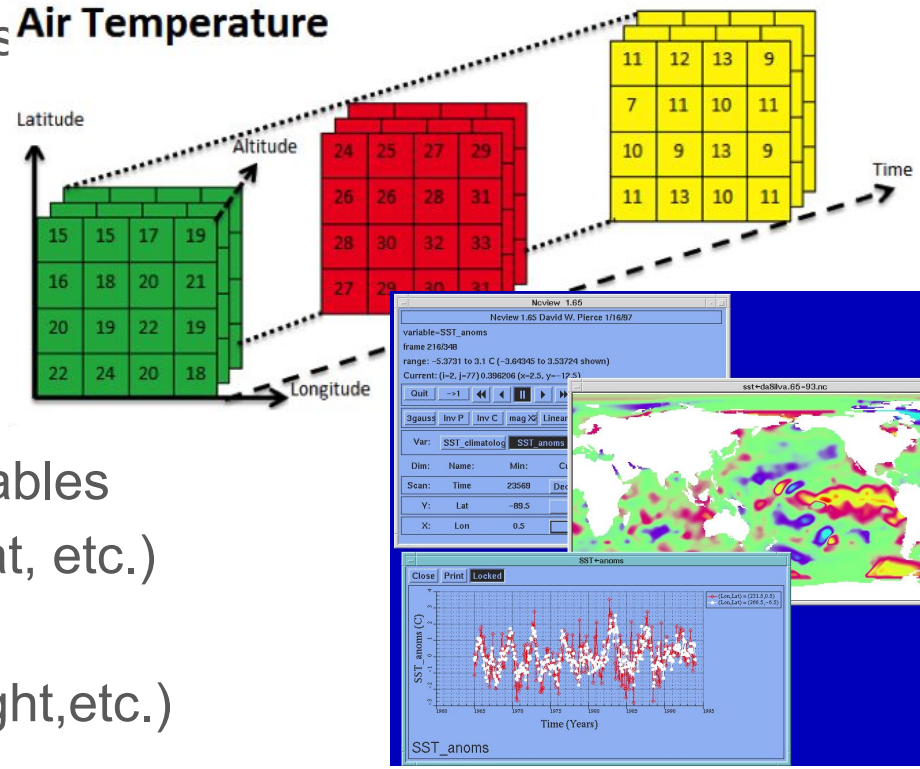
- Widespread file format for multidimensional environmental data.
- Supports unstructured, regular and curvilinear grids.
- Dimensions, variables and attributes.





NetCDF: Network Common Data Form

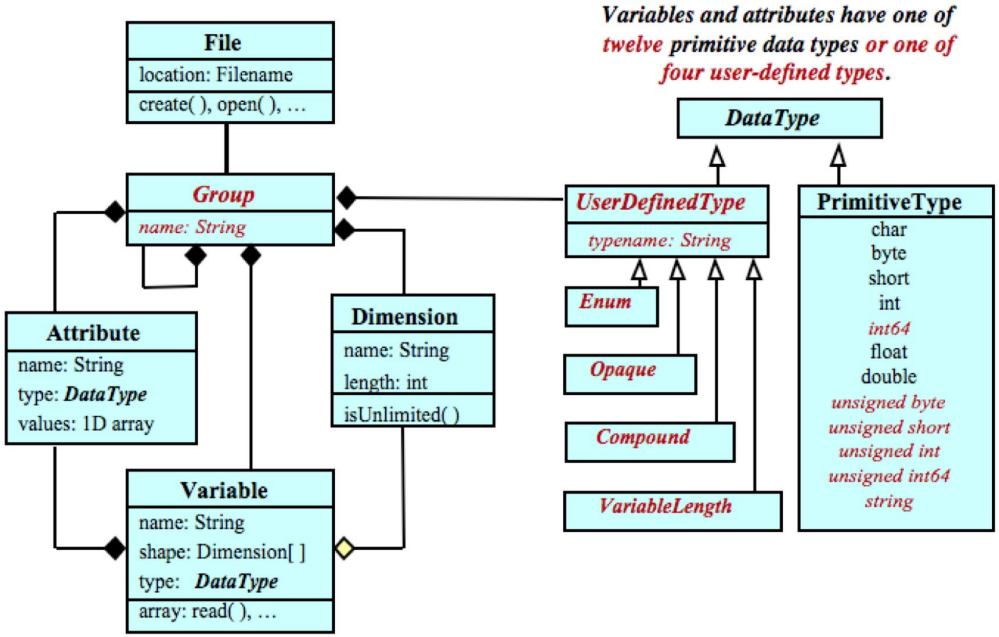
- Self descriptive and conventions
- Huge amount of data sources, libraries, and tools.
- Array oriented data format:
  - Multidimensional array variables
  - Variables are typed (int, float, etc.)
  - Coordinates for the dimensions(time,lat,lon,height,etc.)



NetCDF: Network Common Data Form

- Meta data for the variables (typed, including string)
- Global metadata
- Can take a“slice” (subdomain) from an array
- https://www.unidata.ucar.edu/software/netcdf/

NetCDF Structure



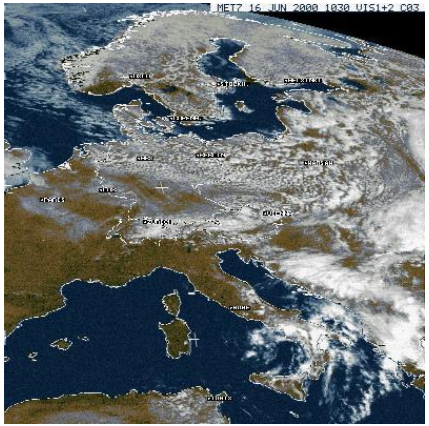
- https://www.unidata.ucar.edu/software/netcdf/docs/index.html

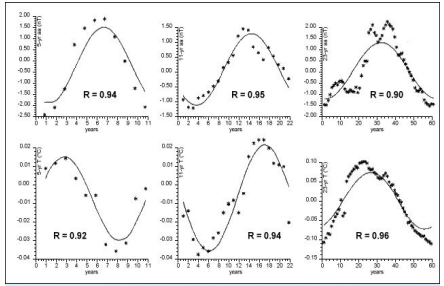
It came from supercomputers

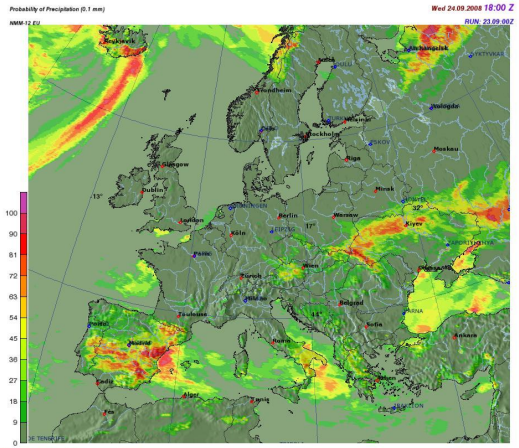
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

...focusing back on weather forecasts...

- How the weather forecast are produced today?
  - Observations (weather stations, …)
  - Remote sensing (satellites, …)
  - Statistics
  - Time series analysis
  - Numerical models





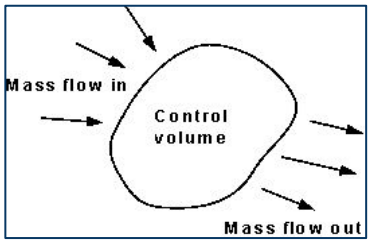


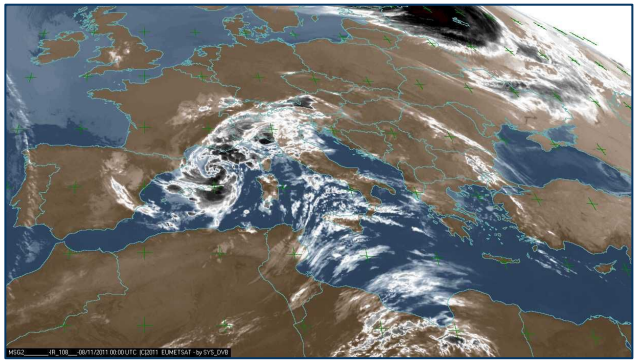
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

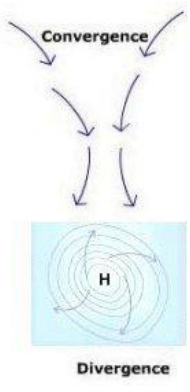
…making it simple at the max…

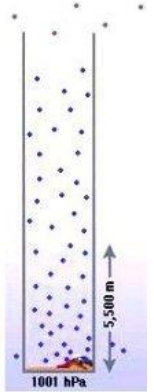
F=m*a
a=F
v/t=F
v/t=F(Δp,lat,g,a)

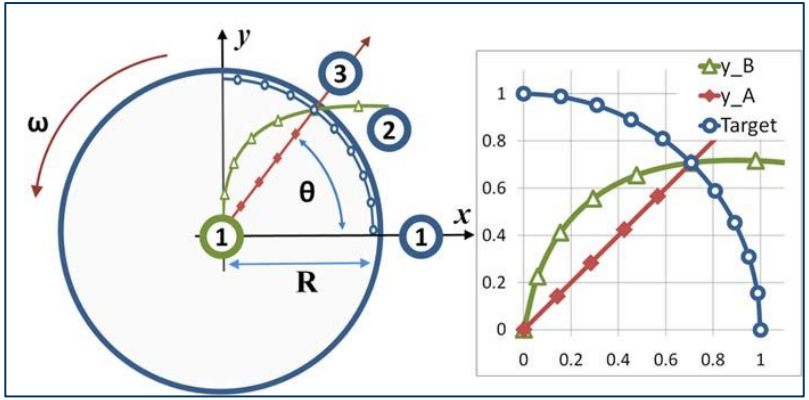
m=1, one air mass
Δp, how the air pressure is varying (pressure gradient)
lat, the Coriolis force is related to the latitude
g, the gravity acceleration
a, frictions





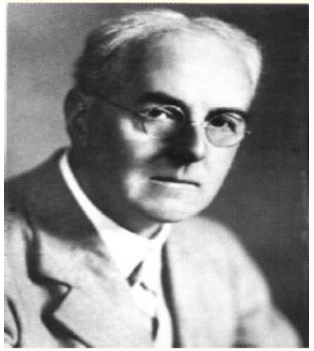


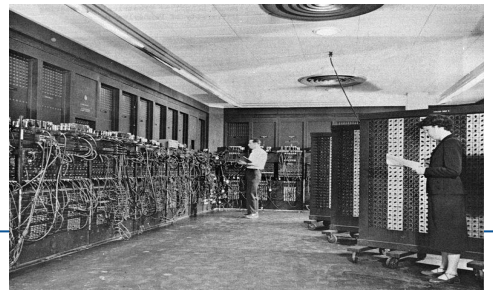




Let start from the beginning

- The british mathematician **Lewis Fry Richardson** proposed a numerical approach to weather predictions in **1922**.
- Richardson unfortunately **failed** because the **lack of computing power**.
- First numerical weather prediction: 1950 (USA).
- Weather team: Jule Charney, Philip Thompson, Larry Gates, Ragnar Fjörtoft.
- Numerical mathematician: **John von Neumann**.
- Digital computer: **ENIAC**.



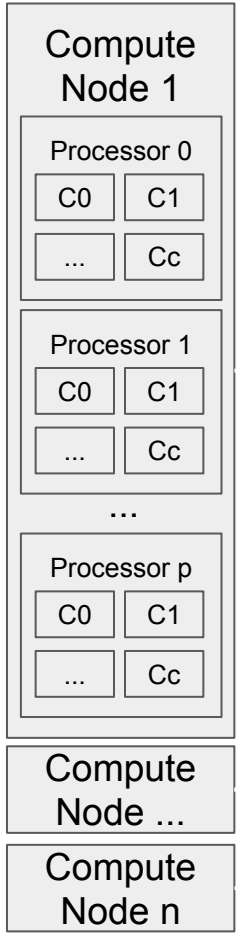


Computing Power

- A **supercomputer** is a computer with a high level of performance as compared to a general-purpose computer.
- Computing power is measured in floating-point operations per second (FLOPS).
- A desktop computer has performance in the range of hundreds of gigaFLOPS to tens of teraFLOPS.
- Today supercomputers can perform over 10^{17} FLOPS (a hundred quadrillion FLOPS, 100 PFLOPS)

High Performance Computing

- Supercomputers performs supercomputing or **High Performance Computing** (HPC)
- Today HPC systems are build by tens to hundreds of off the shelf computing nodes.
- Each computing node has more processors (CPUs).
- Each processor has many computing cores.
- Computing nodes are connected each other by high performance networks (from 1Gbit to hundreds Gbits)



- Computing nodes leverage on storage resources.
- Today computing nodes can be powered by accelerators.

- **General Purpose Graphic Processing Units:** high performance processors designed to speed up graphic operations, but then for HPCs.
- NVIDIA, Compute Unified Device Architecture (CUDA)

https://developer.nvidia.com/blog/cuda-refresher-reviewing-the-origins-of-gpu-computing/

High Performance Computing

- Supercomputers challenge to score better and better on http://top500.org

- The most part of supercomputers use Linux as operating system

- **Applications:**
  - quantum mechanics
  - **weather forecasting** & climate research
  - oil and gas exploration
  - molecular modeling
  - vaccines and drugs
  - cancer research
  - physical simulations
  - cryptanalysis

High Performance Computing

- High-end supercomputers cost hundred thousands dollars to billions.
- Alternatives:
  - **High Performance Cloud Computing:** using on demand public cloud resources to create a virtual HPC system to be used just for the needed computation.
  - **DIY “Beowulf”:** supercomputers made using off the shelf hardware components (desktops, servers, single board computers)

Focusing back on weather forecasts

- The needed parts checklist (tentative):

  ☑ A supercomputer.

  ☑ An operating system.

    - A software implementing the weather numerical model.

- A weather numerical model software implements a solver algorithm for the atmosphere state equations.

- Weather Research and Forecasting model is one of more used model for both research and operational services.

Weather Research and Forecasting

- The Weather Research and Forecasting (WRF) Model is a next-generation mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications.

- Web portal: https://www.mmm.ucar.edu/weather-research-and-forecasting-model

- Available as open source on github: https://github.com/wrf-model

Focusing back on weather forecasts

- The needed parts checklist (tentative):

  ☑ A supercomputer.

  ☑ An operating system (Linux).

  ☑ Weather forecasting model (WRF).

- The weather forecasting model needs atmosphere initial conditions produced by a global forecasting system:
  - European Centre for Medium-Range Weather Forecasts (ECMWF)
  - Global Forecasting System (NOAA, NCEP)

Computing Domains

- Numerical weather models solve the atmosphere state equations on a 3D grid.

- A **domain** is a geographical matrix covering the earth area on which the weather forecast is focused on.

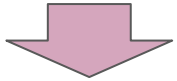
- Domains are characterized by:
  - A geographic center
  - A horizontal extension in terms of kilometers
  - A horizontal resolution (the size of each matrix cell in the real world)
  - A vertical resolution

The Mediterranean Area

300 x 200 cells, resolution 25 Km

Computing Domains

- High resolution initial conditions
- High resolution land digital models

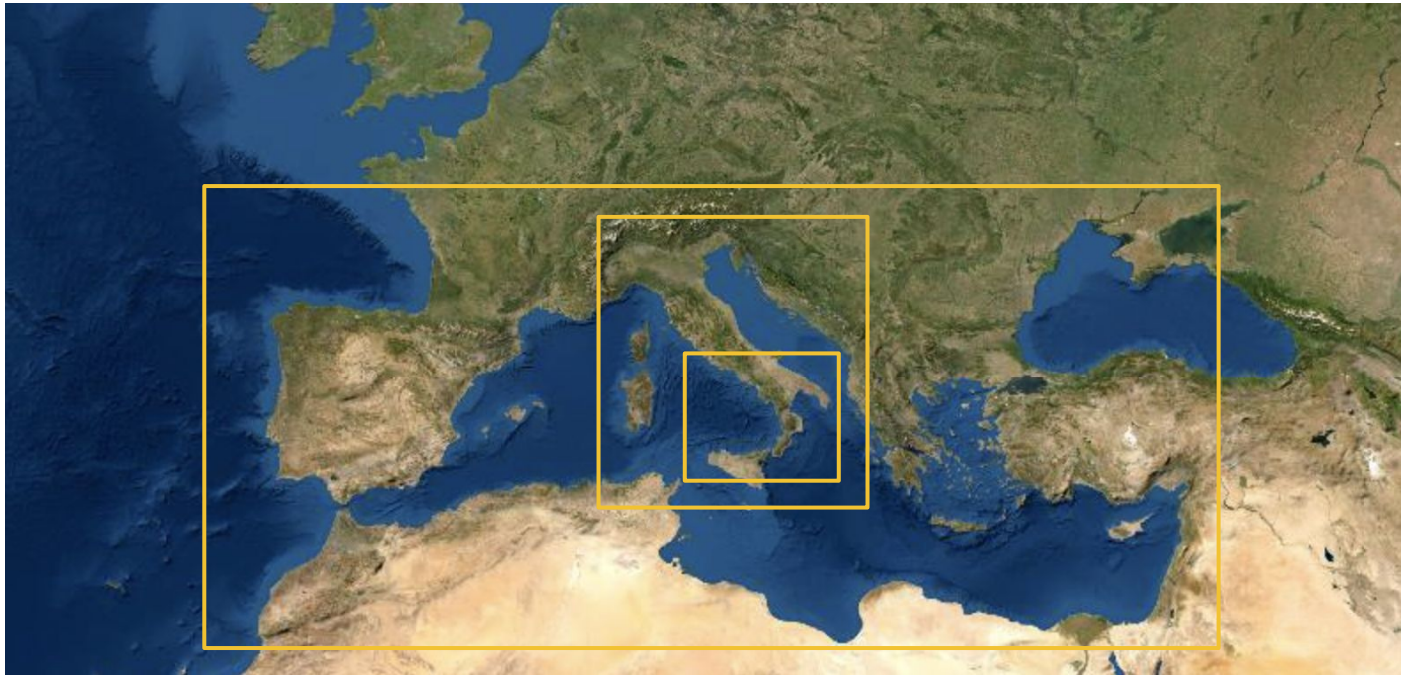


High resolution domains

More computing power needed

- Nesting domains into domains creases locally the weather forecast resolution mitigating the need for more computing power.

The Mediterranean Area



300 x 200 cells, resolutions 25 Km, 5 Km, 1 Km.

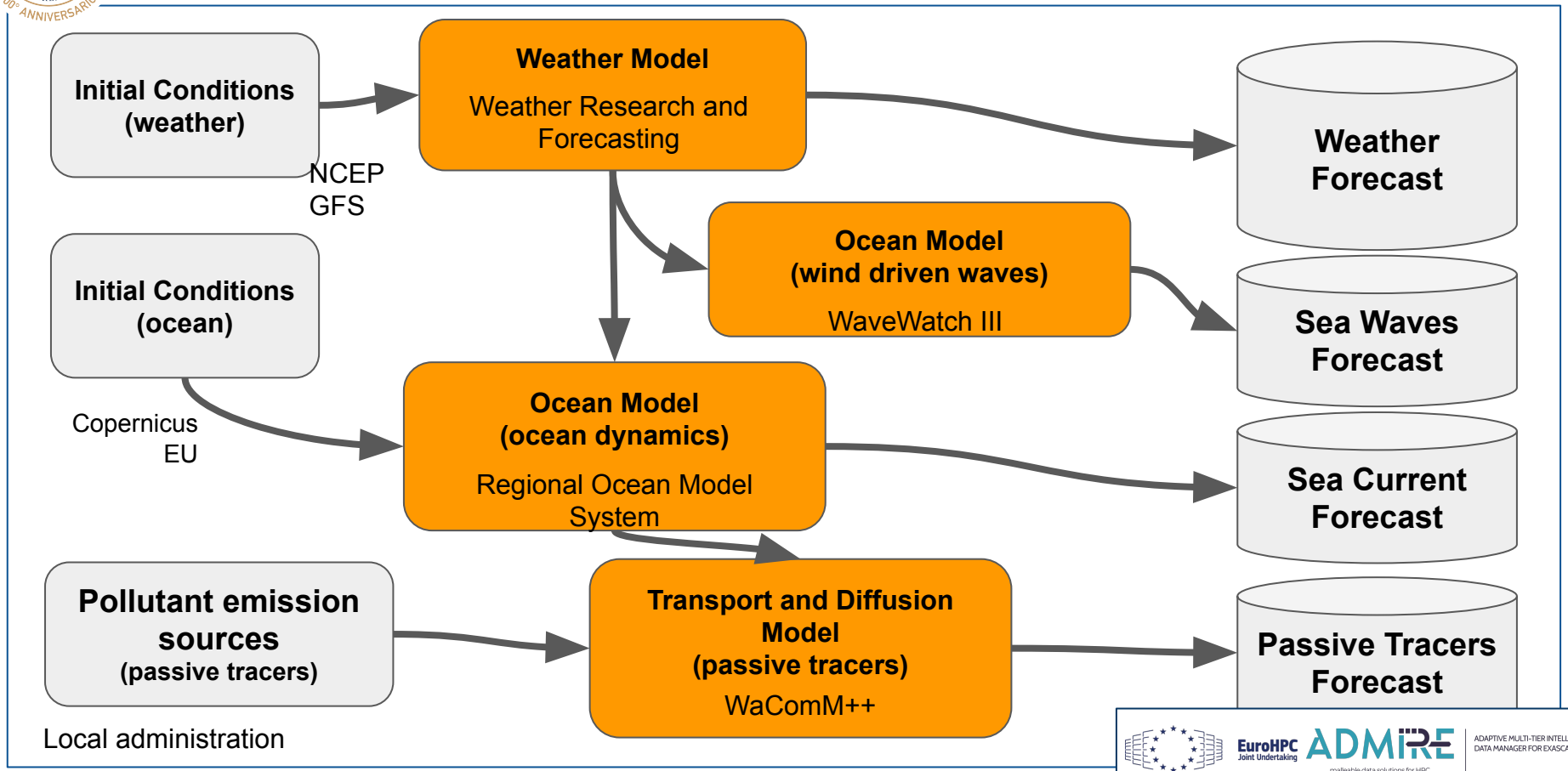
If you need weather forecast, who you gonna call?

Center for Marine/atmosphere Monitoring, Modelling, and Applications

Mission: *"deliver operational forecasts, predictions, and simulations for science, engineering, and management in the field of weather, ocean-dynamics, air-quality, and pollutant transport and diffusion at the sea."*

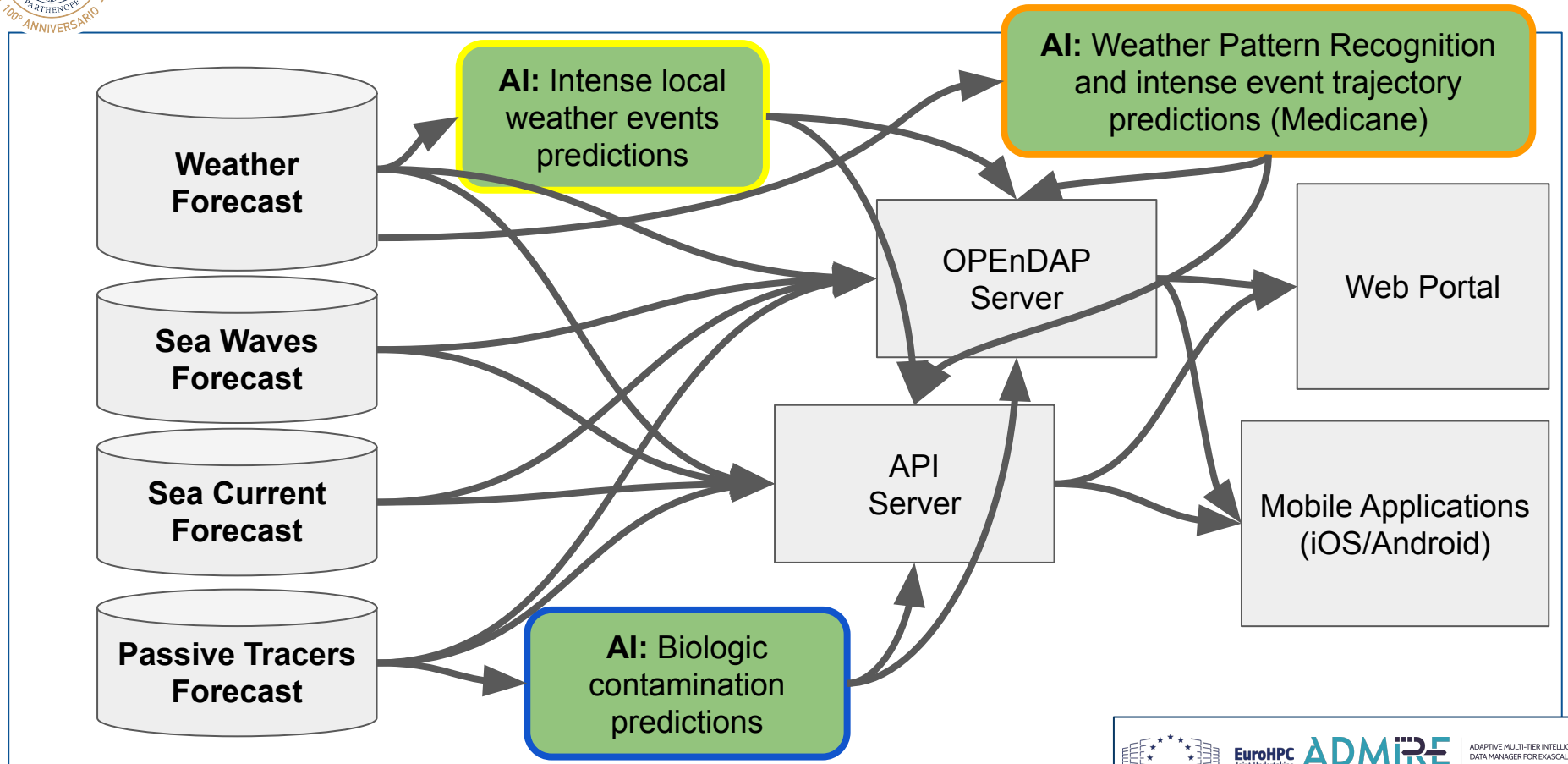
http://meteo.uniparthenope.it

From the initial conditions to the forecast



EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

From the forecast to the app (at last!)

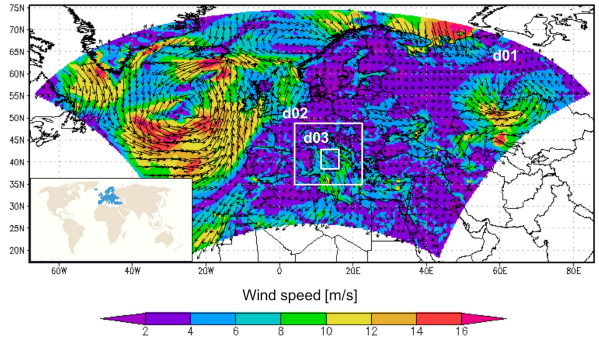


Computational Workflows

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Introduction, contextualization and motivations

- Any science field can be **now** considered as *computational science*.
- Workflow engines play a **primary role** in computational sciences.



- The Cloud technology provides **elastic** and virtually **infinite** computational **resources**.
- Workflow engines ensure computational experiment **reproducibility**.
- The **scientific approach** has been enriched by the use of computational **models** and data analysis.

2000s

*_Science + Computer Science = "*Computational Science*"...

Present Days

...any Science is "**computational**" by design.

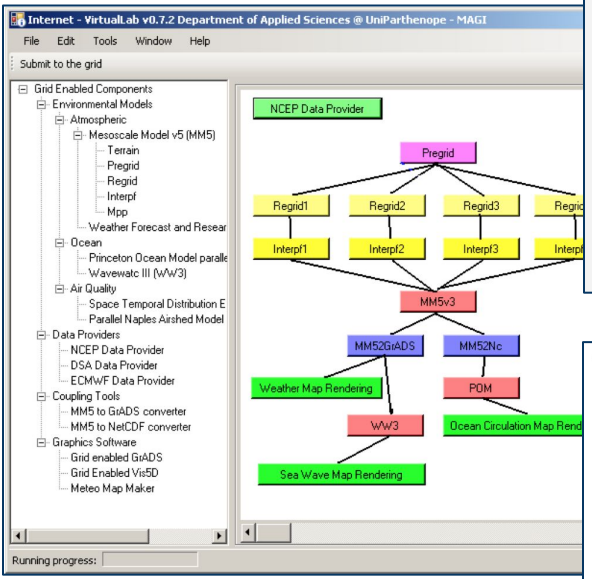
Outlook

...the "computational" is be **in** each Science **by default**.

EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

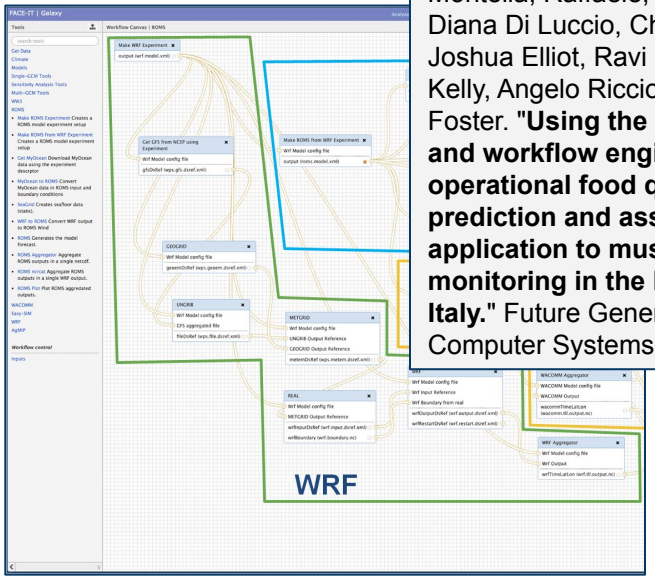
Introduction, contextualization and motivations

- Workflow engines for data-intensive science have existed since the beginning of the grid computing era.

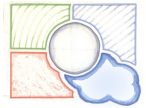


Ascione, Isabella, Giulio Giunta, Patrizio Mariani, Raffaele Montella, and Angelo Riccio. "**A grid computing based virtual laboratory for environmental simulations.**" In European Conference on Parallel Processing, pp. 1085-1094. Springer, Berlin, Heidelberg, **2006**.

```
<jfdl:jfs project="experiment01">
  <!-- Job definition -->
  <jfdl:jobs>
    <jfdl:job name="downloadConditions"
         target="dgric.uniparthenope.it"
         rsl="downloadConditions.rsl"/>

    <jfdl:job name="runMM5"
         target="dgbeobi.uniparthenope.it"
         rsl="runMM5.rsl"/>
```



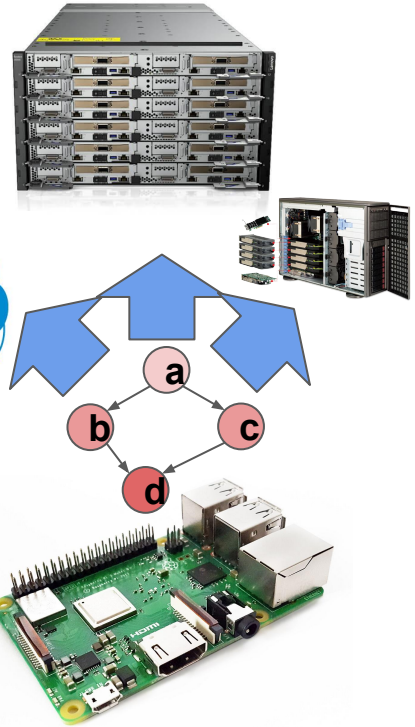
Montella, Raffaele, Alison Brizius, Diana Di Luccio, Cheryl Porter, Joshua Elliot, Ravi Madduri, David Kelly, Angelo Riccio, and Ian Foster. "**Using the FACE-IT portal and workflow engine for operational food quality prediction and assessment: An application to mussel farms monitoring in the Bay of Napoli, Italy.**" Future Generation Computer Systems (**2018**).



Workflows gained the role of first-class citizen in nowaday applications.
Ranging from mobile/embedded computing at the edge to large scale science.
...embedding workflows in regular applications could be the next big thing...

Introduction, contextualization and motivations

***General purpose**, or **dedicated** tools, that is the question.*



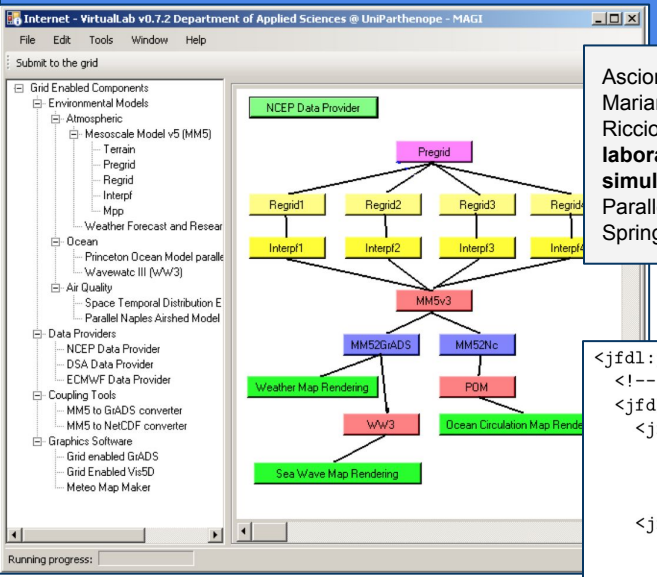
- Define workflows **programmatically** (production).
- Execute tasks in very **heterogeneous** environments.
- Use **multiple workflows** in the same applications, interacting each other if needed.
- **Move** data across **different domains** or, vice-versa, move the **computation close** to the data.

Really, do we need for a
“Yet Another Workflow Engine”!

So far…

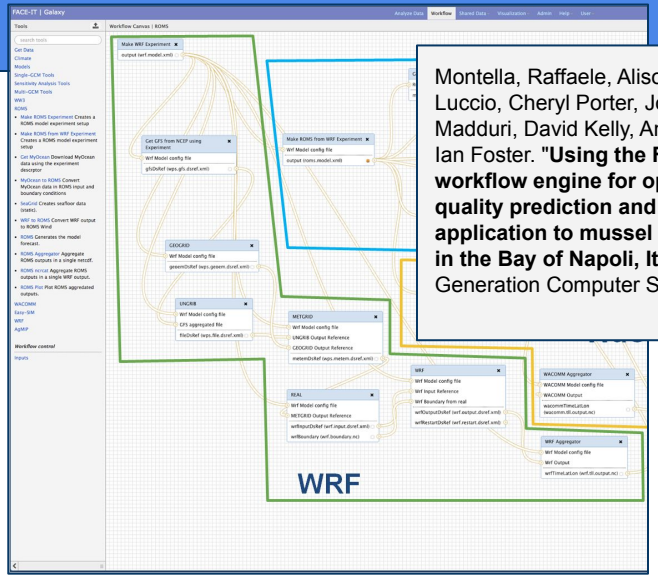
- Any science field can be **now** considered as *computational science*.
- Workflow engines have existed since the beginning of the grid computing era.

General purpose, *or* ***dedicated*** *tools, that is the question.*



Ascione, Isabella, Giulio Giunta, Patrizio Mariani, Raffaele Montella, and Angelo Riccio. "**A grid computing based virtual laboratory for environmental simulations.**" In European Conference on Parallel Processing, pp. 1085-1094. Springer, Berlin, Heidelberg, **2006**.

```
<jfdl:jfs project="experiment01">
  <!-- Job definition -->
  <jfdl:jobs>
    <jfdl:job name="downloadConditions"
         target="dgric.uniparthenope.it"
         rsl="downloadConditions.rsl"/>

    <jfdl:job name="runMM5"
         target="dgbeobi.uniparthenope.it"
         rsl="runMM5.rsl"/>
```



Montella, Raffaele, Alison Brizius, Diana Di Luccio, Cheryl Porter, Joshua Elliot, Ravi Madduri, David Kelly, Angelo Riccio, and Ian Foster. "**Using the FACE-IT portal and workflow engine for operational food quality prediction and assessment: An application to mussel farms monitoring in the Bay of Napoli, Italy.**" Future Generation Computer Systems (**2018**).

Galaxy PROJECT

Yes, we need for "Yet Another Workflow Engine" ?

Direct Acyclic Graphs as parallel jobs on anything

DagOn* is a production-oriented workflow engine:

- **Integration** in the Python environment.
- **Minimal** footprint for external software components execution.
- **Avoiding** any workflow engine **centered data management**.
- **Straightforward** definition of tasks:
  - Python scripts.
  - Web interaction.
  - External software components.
- **Execution** sites **independence**:
  - Local / scheduler (SLURM).
  - Containers (Docker).
  - Clouds (AWS, OpenStack, DigitalOcean).

DagOn*

Named after the Phoenician god-fish *Dagon* known by ancient Greeks as *Triton*.

NB: The * symbol is the wildcard for **anything**.

EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Programming Model

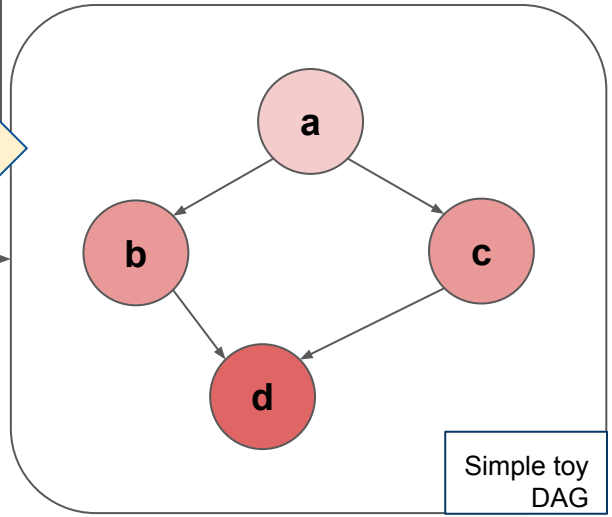
Python Script: "DagOn* Hello World App"

```
import dagon
…
workflow=Workflow("myapp",settings)
workflow.add_task(new Task("a","..."))
workflow.add_task(new Task("b","workflow:///a"))
workflow.add_task(new Task("c","workflow:///a"))
workflow.add_task(new Task("d","workflow:///b workflow:///c"))
workflow.run()
sys.exit(0)
```

Data Flow

Task Flow — Defined by task dependencies.

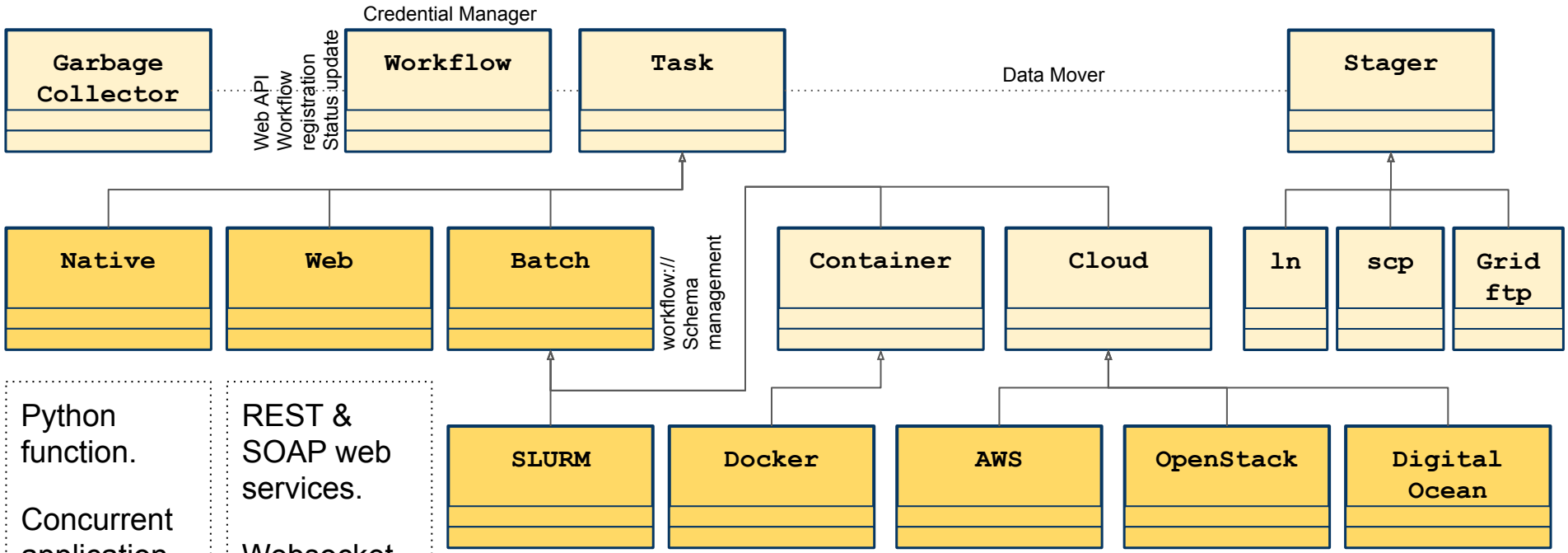
Data Flow — Defined by data dependencies.



- Dealing with actual data files instead of high-level defined datasets.
- Performing backward data references in order to create dependencies.
- Having more Workflow instances in the same Python application.

DagOn* has been designed starting from the desidered programming model.

Components



Python function.

Concurrent application thread.

REST & SOAP web services.

Websocket subscribe & publish.

DagOn Service.

External software.

SSH or scheduler.

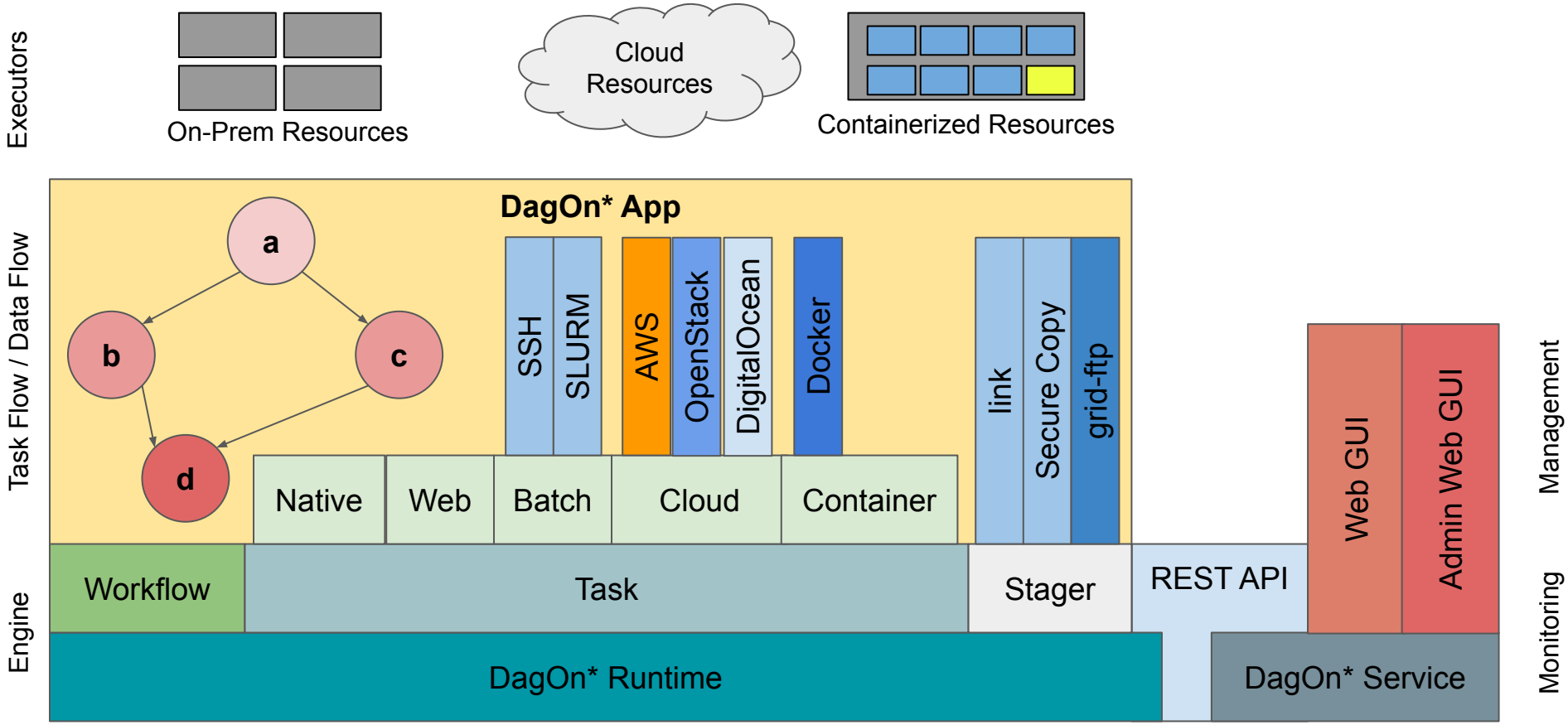
Task embedded in container.

Task embedded in a virtual machine.

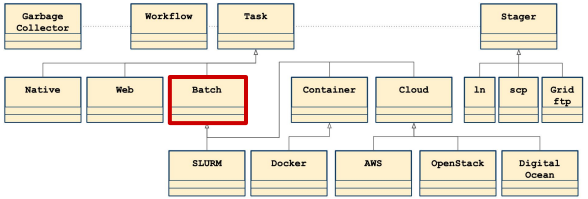
Instance status management.

Cloud independent.

Architecture



The workflow:// schema



The **Batch** component takes charge of the management of data dependencies using the **workflow:// schema**.

workflow://workflow_unique_name**/**task_unique_name**/**

The schema label

The workflow unique name
An UUID could be used
If empty means “current workflow”

The task unique name
Can be dynamically generated by the Python script when the workflow is created programmatically.

Task scratch directory root

Workflow

Stager

- Local
- Shared File System
- Remote scratch directory on physical machine, virtual instance or container

One Of...

ln

scp

Grid ftp

out

in

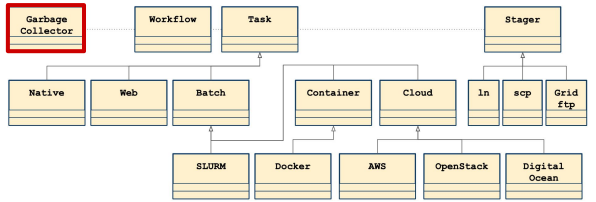
External software **sandbox**

workflow:///task_a/results.csv

Scratch directory

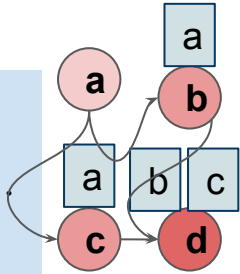
The garbage collector

- Tracks the storage and computational resources allocated during tasks execution.
- Proceeds to dispose them when no longer needed.



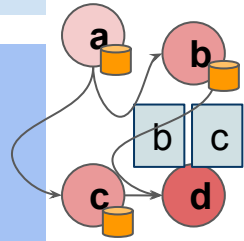
Make Dependencies

```
For each batch task in the <workflow> ...
     For each workflow://<workflow>/<task>/ reference in the task command line ...
          Increment the number of reference to <task>
```



On Task Finish

```
For each workflow://<workflow>/<task>/ reference in the task command line ...
     Decrement the number of reference to <task>
     If the number of reference to <task> is 0, clean up the involved resource
```

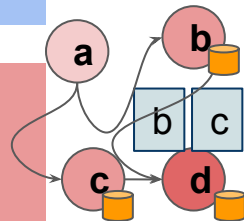


Clean Up

Local, remote or shared file system:
Remove the scratch directory.

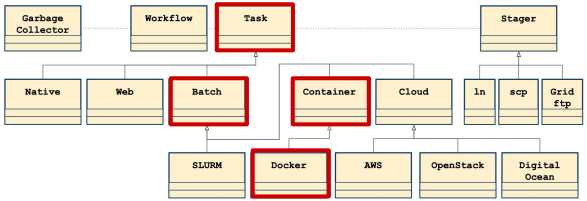
Virtual machine instance:
Stop the instance.

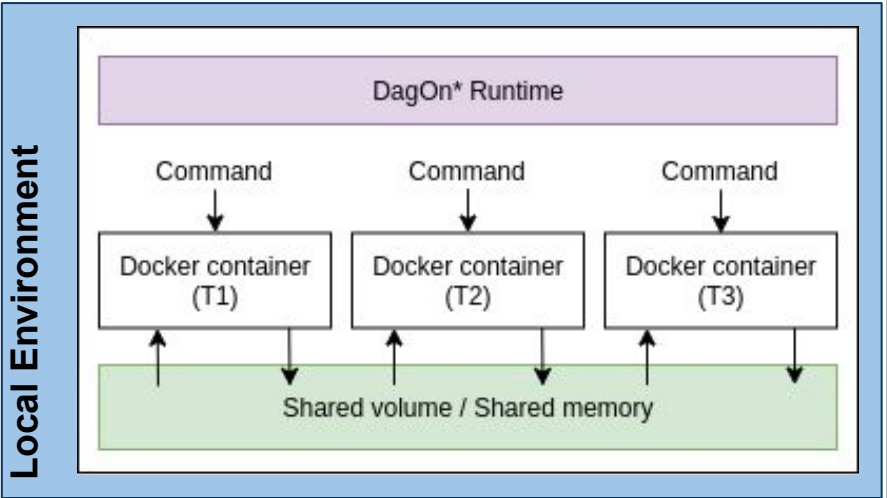
Container:
Stop the container.

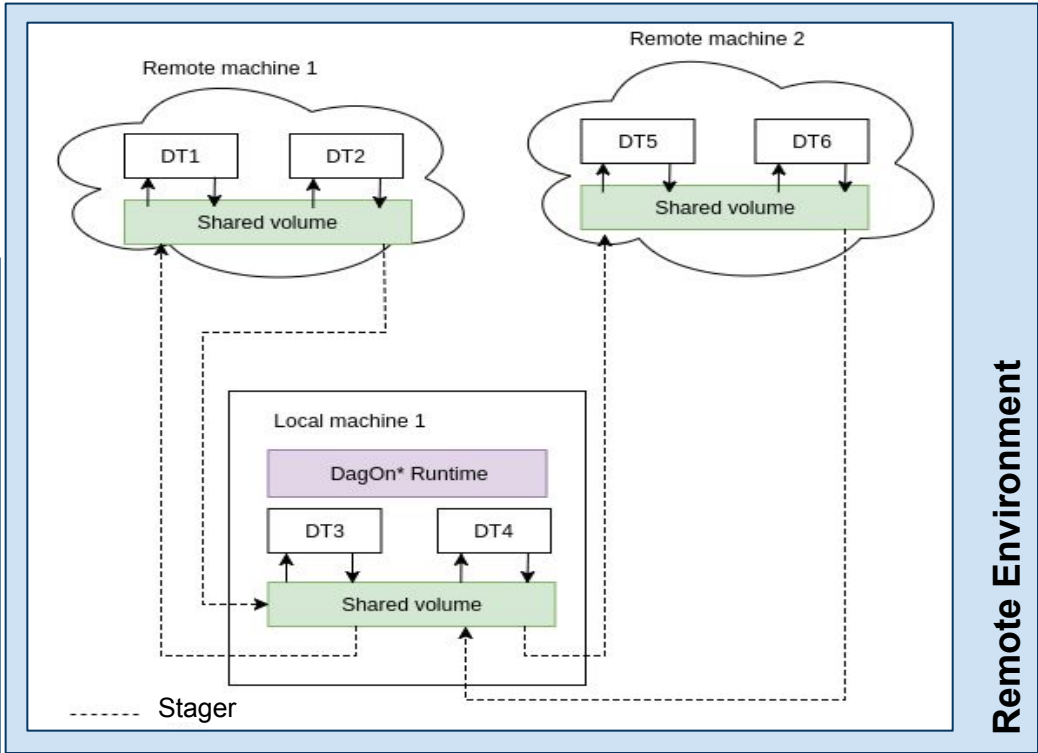


Container tasks

- Deployable in any Docker machine.
- Share a volume with the operating system host file system.
- If the tasks are on the same machine, the data transfer is done using shared memory.
- In a remote environment, data is copied to the volume shared between containers.

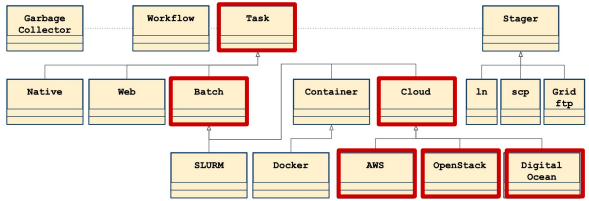


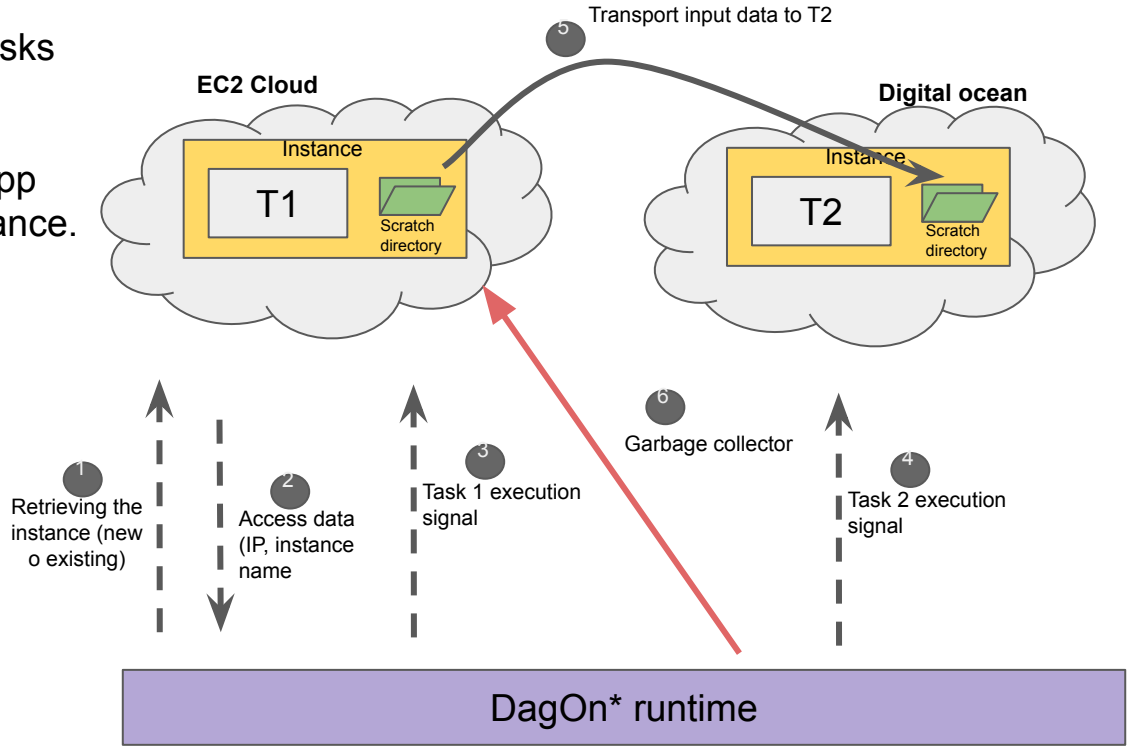




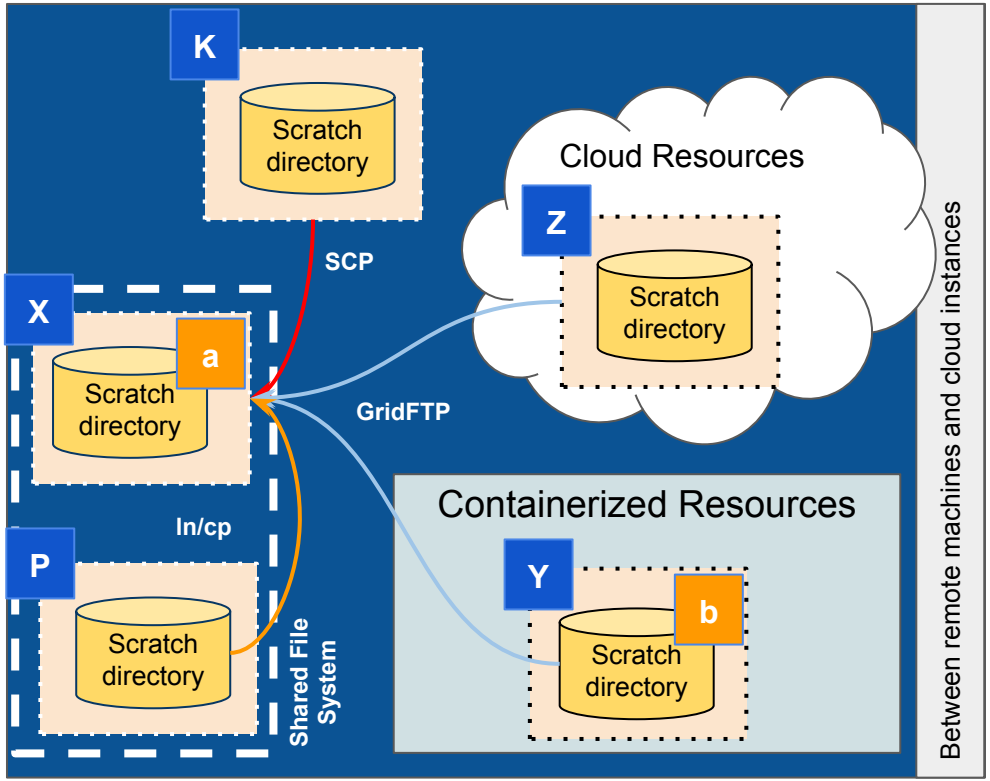
Cloud tasks

- Deployable in private, public and hybrid clouds.
- Define programmatically the flavour and configuration of the instance.
- Interoperable with other types of tasks (batch, containers, etc).
- SSH is used to make the DagOn app controlling the virtual machine instance.
- Data is transferred between tasks using the Stager component.
- Leverage on Apache Libcloud
- Tested with:
  - AWS
  - OpenStack
  - Digital Ocean
  - **Google Cloud**

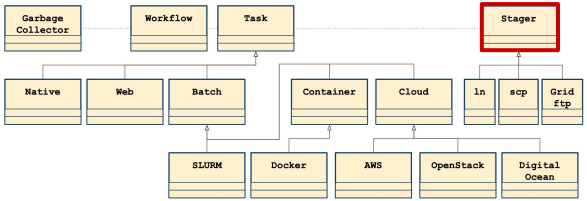


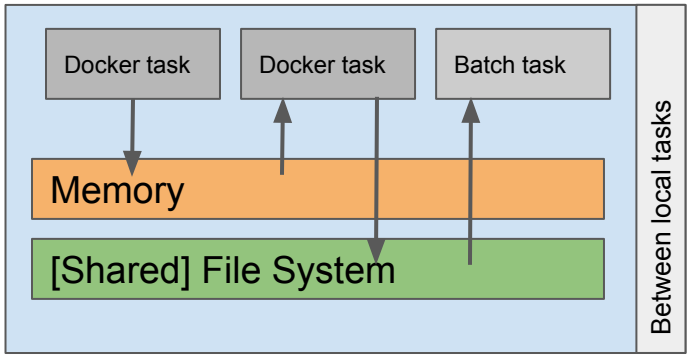


Staging




```
globus-url-copy -vb -p 4 gsiftp://X/tmp/a/f1 gsiftp://Y/tmp/b/f2
```



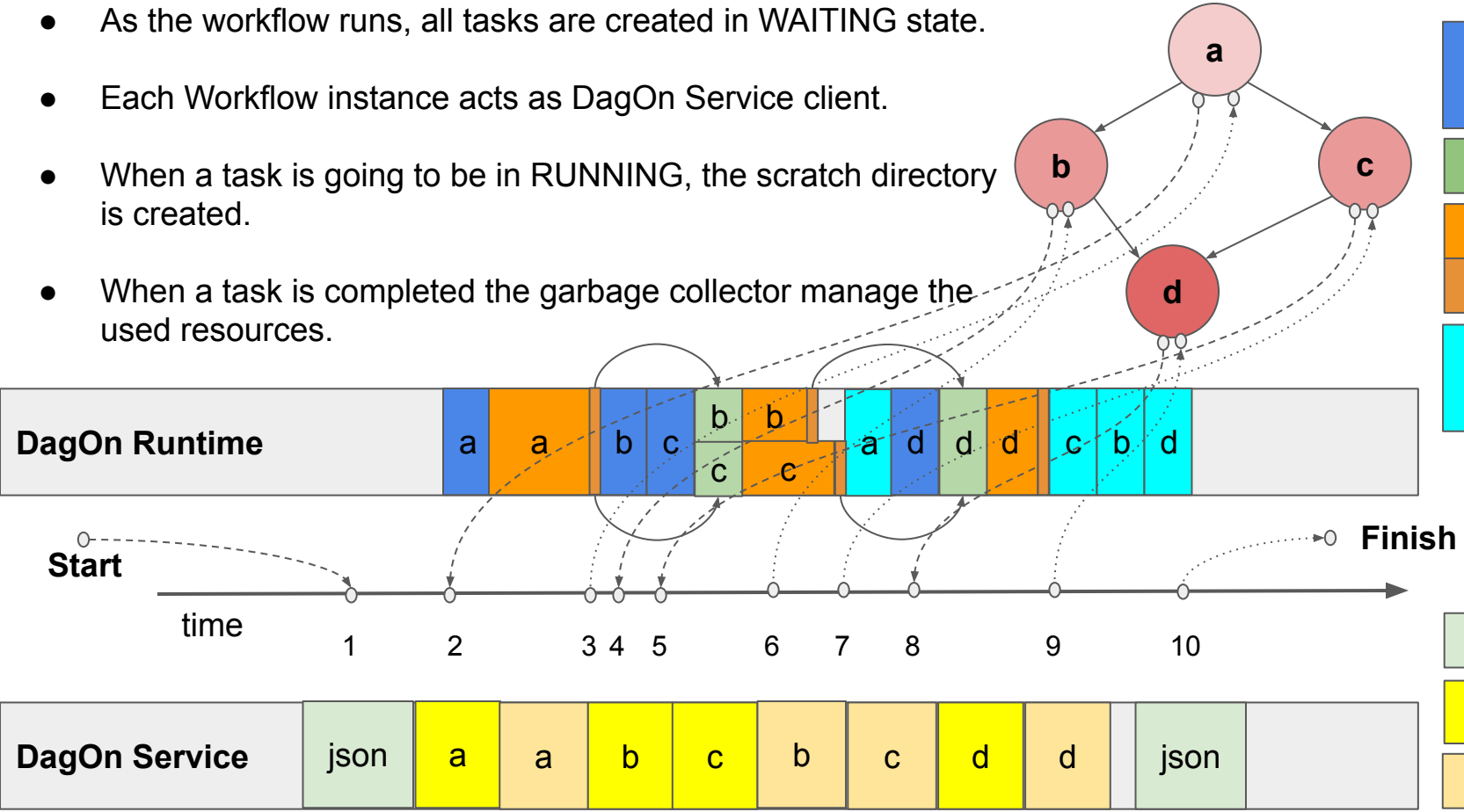


- Manages the data movement between all type of tasks.
- Fallback strategy:
  a. GridFTP
  b. Secure Copy
- Local tasks: memory, [shared] file system.

Globus Connect Server

Application lifecycle

- As the workflow runs, all tasks are created in WAITING state.
- Each Workflow instance acts as DagOn Service client.
- When a task is going to be in RUNNING, the scratch directory is created.
- When a task is completed the garbage collector manage the used resources.



Create scratch dir

Stage-out

Run

Completed

Remove scratch dir

Publish

Register

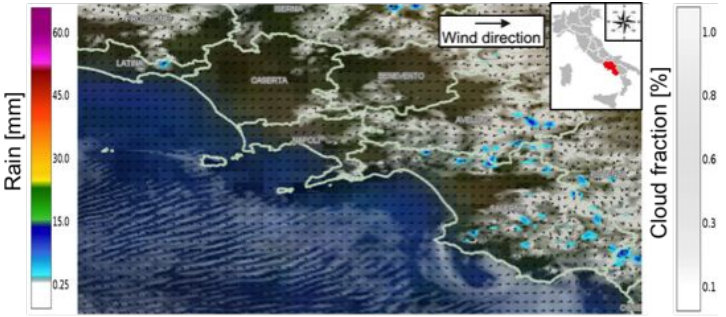
Unregister

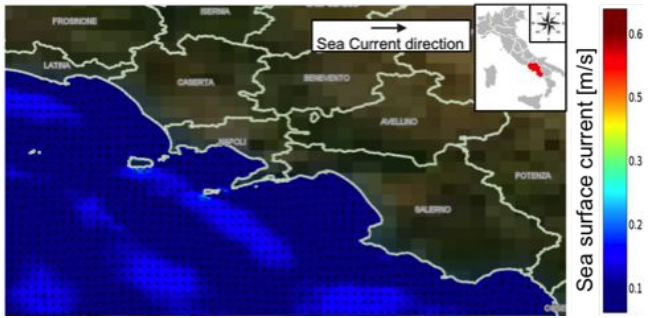
Operational weather forecast apps

Operational sea/weather forecast on the European regions with higher resolution on the Southern Italy area.

Applications

1. ROMS and WaComM for **mussels** farms **quality assessment** and management.
2. WaveWatch III for **coastal flooding** forecast in case of extreme sea storm events.
3. CALMET/CALPUFF for **wildfire** smoke tracing.



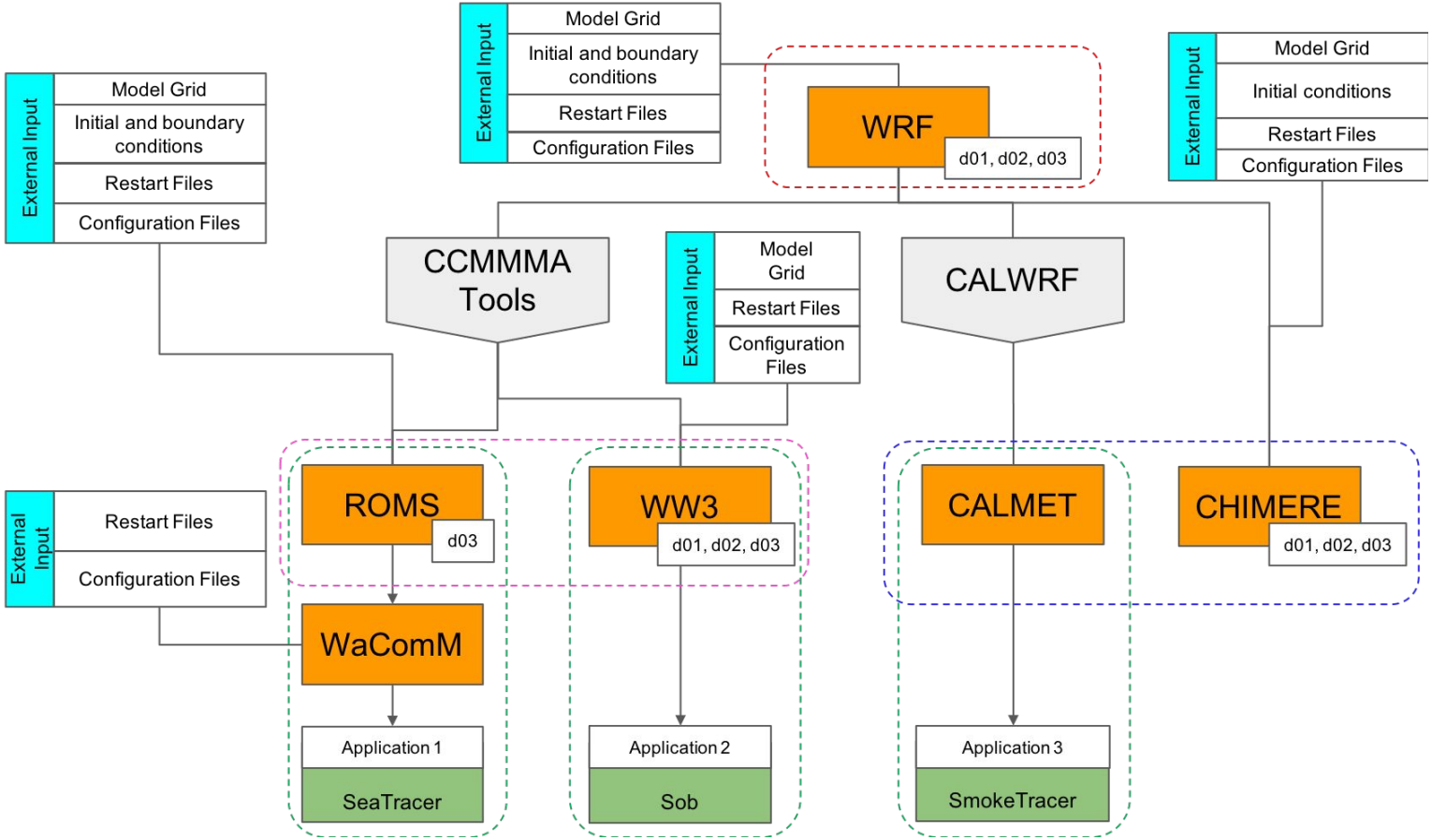


Model	Domain	Δt [sec.]	Grid dimension [N. of cells]			Grid spacing [m]	Output File Dimension [Gb] (1-hour of simulation)	N. of output file for each run (7 days)	N. of MPI processes for each Δt	Execution time [mm:ss] (1-hour of simulation)
			West-East	South-North	Bottom-Top					
WRF	d01	150	230	209	28	25000	0.11	169	96 (12 core*8 nodes)	05:05
	d02	30	361	336		5000	0.29			
	d03	6	301	306		1000	0.22			
ROMS	d03	30	1375	1021	30	200	2.28		80 (10 core*4 nodes)	07:12

Models, problem size and wall clock time.

Operational weather forecast apps

ROMS and WaComM for **mussels** farms **quality assessment** and management (MytiluSE app).



Use case: ADMIRE Environmental App

- Part of the DagOn* App running in production at http://meteo.uniparthenope.it
- The workflow is defined dynamically.
- Tasks are defined as Batch/Slurm ones.
- Once all the tasks have been added, the workflow is run.

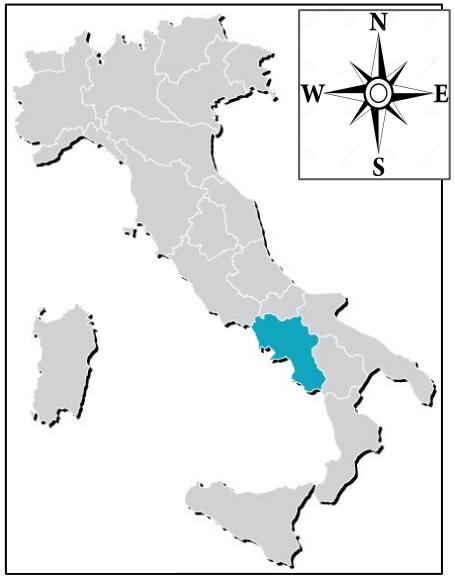
```
  # Create the orchestration workflow
  workflow=Workflow("roms",config)

  # Some beauty logging
  workflow.logger.info("initialization date: %s",i_date)

  # The myocean2roms task executed using Slurm
  taskMyocean2roms=batch.Slurm("myocean2roms",command_dir_base+"/myocean2roms "+i_date+" "+gridFilename+" "+domainId)

  # The wrf2roms task executed using Slurm
  taskWrf2roms=batch.Slurm("wrf2roms",command_dir_base+"/wrf2roms "+i_date+" "+str(hours)+" "+gridFilename+" "+domainId+" "+wrfDom
ainId)

  # The roms task executed using Slurm
  taskRoms=batch.Slurm("roms",command_dir_base+"/roms "+i_date+" "+gridFilename+" "+domainId+" workflow://myocean2roms/ini-"+domai
nId+".nc workflow://myocean2roms/bry-"+domainId+".nc workflow:///wrf2roms/wind-"+domainId+".nc")

  # The wrf2roms task executed using Slurm
  taskPublishRomsOutput=batch.Slurm("publishRomsOutput",command_dir_base+"/publishRomsOutput "+i_date+"00 "+i_date+"00 rms3 workfl
ow://roms/output/ocean_his_")

  # add tasks to the workflow
  workflow.add_task(taskMyocean2roms)
  workflow.add_task(taskWrf2roms)
  workflow.add_task(taskRoms)
  workflow.add_task(taskPublishRomsOutput)
```

```
frontend: 20181108Z00 data download completed!
Running workflow...
2018-11-08 21:35:25,829 root         INFO     initialization date: 20181108Z00
2018-11-08 21:35:25,830 root         DEBUG    Running workflow: roms
2018-11-08 21:35:25,831 root         DEBUG    myocean2roms: Status.WAITING
2018-11-08 21:35:25,831 root         DEBUG    myocean2roms: Status.RUNNING
2018-11-08 21:35:25,831 root         DEBUG    myocean2roms: Executing...
2018-11-08 21:35:25,832 root         DEBUG    wrf2roms: Status.WAITING
2018-11-08 21:35:25,832 root         DEBUG    wrf2roms: Status.RUNNING
2018-11-08 21:35:25,832 root         DEBUG    roms: Status.WAITING
2018-11-08 21:35:25,833 root         DEBUG    wrf2roms: Executing...
2018-11-08 21:35:25,833 root         DEBUG    publishRomsOutput: Status.WAITING
2018-11-08 21:35:25,848 root         DEBUG    myocean2roms: Scratch directory: /home/ccmmma/dev/yawe/tmp//1541712925832-myocean2roms
2018-11-08 21:35:25,856 root         DEBUG    wrf2roms: Scratch directory: /home/ccmmma/dev/yawe/tmp//1541712925833-wrf2roms
```

Local queue status (HPC scenario)

```
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
 4193       gpu    wrf_6   ccmmma  R    1:09:57      8 node[1,3-5,9-12]
 4195     hicpu wrf2roms   ccmmma  R       3:01      1 node13
 4196     hicpu myocean2   ccmmma  R       3:01      1 node14
```

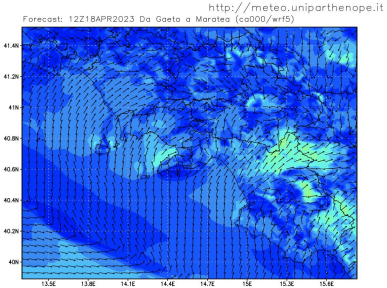
Use case: ADMIRE Environmental App

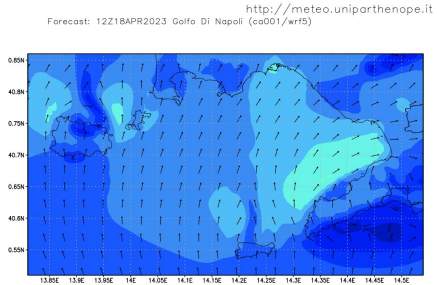


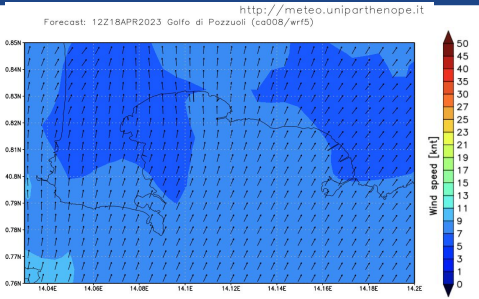
April 18th, 2023
12:00 UTC

Simulation: 169h
Computing: ~6h

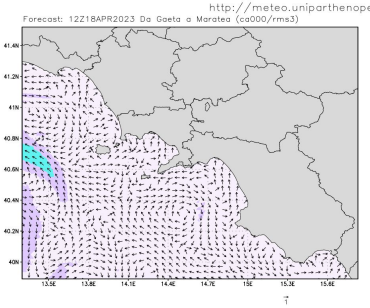
10m Wind speed and direction

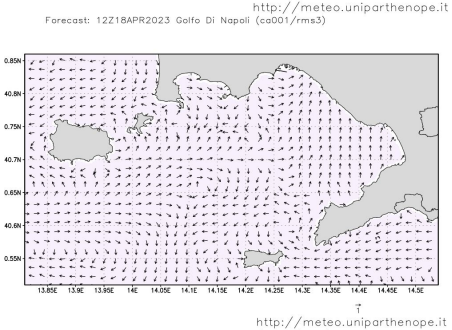


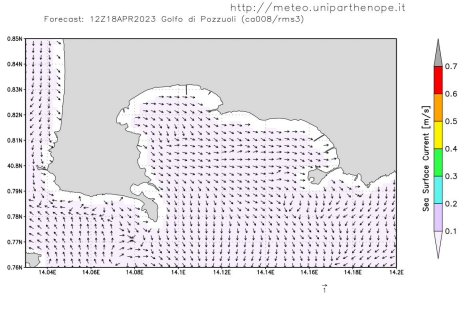




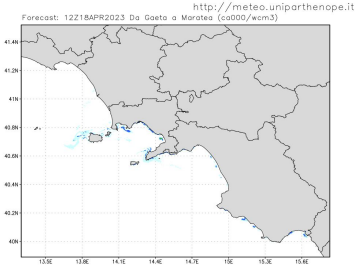
Sea Surface Current

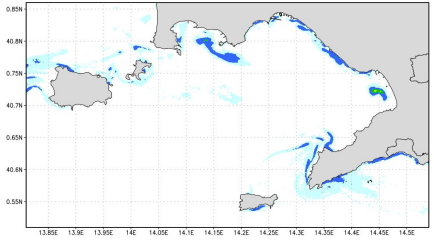


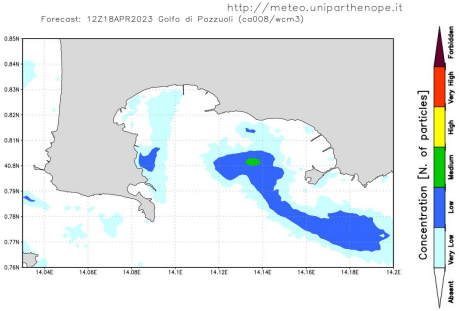




Forecasted sea pollution





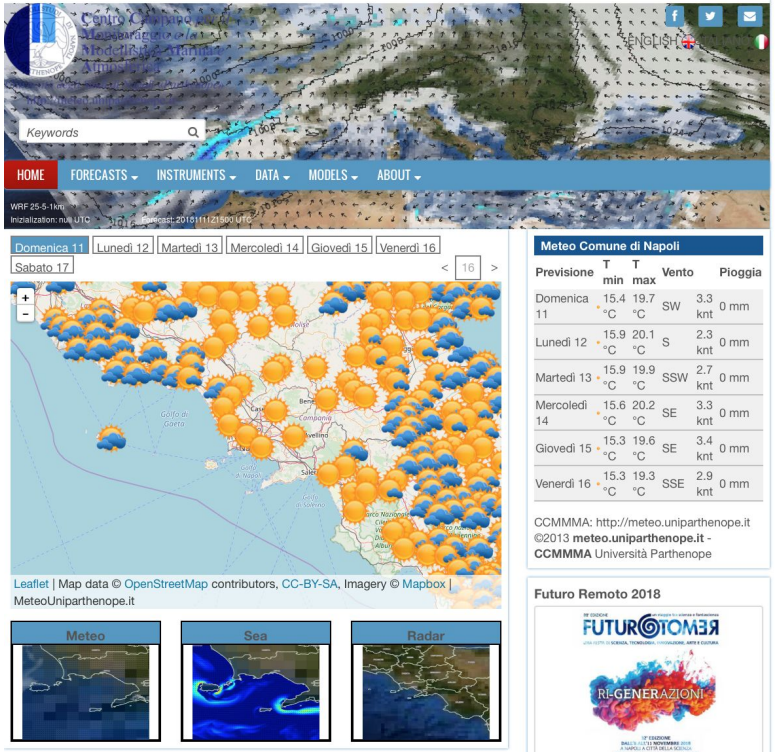


Related Work

	http://swift-lang.org	http://parsl-project.org	https://galaxyproject.org	
	swift	Parsl	Galaxy PROJECT	
Workflow creation	Scripting	Python scripts Annotations	Intuitive Web Interface	Python scripts and JSON files **Object Oriented**
Representation	Graphs Data structure	Dependent task graph Futures	Direct Acyclic Graphs Data structure	Direct Acyclic Graphs, **workflow:// schema**
Data exchange	Files	Python object serialization	Datatype/Dataset Shared File System	Files Shared memory objects, Shared File System, Secure Copy, Grid-Ftp **Garbage collector**
Deployment	Local, HPC, cloud	Local, HPC, containers	Local, HPC, cloud	Local, HPC, containers, **cloud**
Disadvantages	The access to the filesystem could generate a bottleneck.	Can just manage local files and files accessible by Globus.	Requires the deployment of several different software packages.	Still work in progress. Quantitative performance evaluation on experimental use cases.

Conclusion

- Python based tool for data intensive scientific workflows targeting **production applications**.
- DagOn* programming model enables the developer to **embed workflows in** already existing Python **scientific applications**.
- Peculiar features as the dependencies management via the **workflow://** schema and the ***garbage collector***.
- Tasks can be executed **locally**, using a local **scheduler**, on local or remote **containers** and on public, private or hybrid **clouds**.



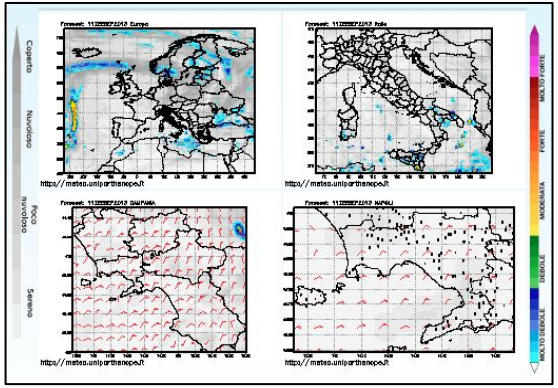
http://meteo.uniparthenope.it

Real world use case application for **weather and marine forecasts.**

Wrapping up and hands on

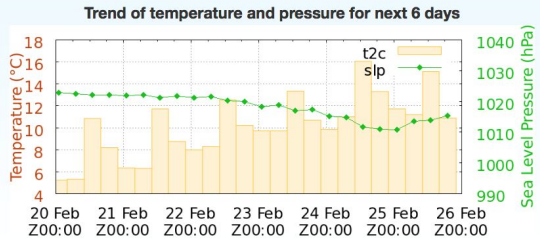
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Weather forecast products



Meteo Comune di Napoli

Previsione	T min	T max	Vento		Pioggia
Lunedì 20	5°C	12°C	NNW	2 knt	-
Martedì 21	6°C	12°C	WNW	1 knt	-
Mercoledì 22	9°C	13°C	SSE	4 knt	-
Giovedì 23	9°C	14°C	SSW	4 knt	0.2 mm
Venerdì 24	11°C	16°C	SSE	9 knt	0.2 mm
Sabato 25	9°C	14°C	SSW	4 knt	-



Places: Regione CAMPANIA

UTC: 26/11/2020 09:00

meteo uniparthenope

Home Forecast Strumenti Dati Modelli Progetti Servizi Info

Regione Campania - Previsioni meteorologiche per giovedì 26 novembre 2020

Regione Campania - Previsioni meteorologiche per giovedì 26 novembre 2020 *armando.rocco* Wed, 11/25/2020 - 19:45

meteo uniparthenope

Previsioni meteorologiche per giovedì 26 novembre 2020

L'ampia area di alta pressione presente sull'Europa orientale e sul Mediterraneo garantirà condizioni di stabilità atmosferica sulle aree centro-settentrionali della Penisola; pertanto, sulla regione Campania si prevedono condizioni di cielo prevalentemente poco nuvoloso.

Giovedì 26 novembre 2020: di notte cielo nuvoloso o poco nuvoloso su Irpinia e alto Sannio; prevalentemente sereno altrove. In mattinata ampie schiarite interesseranno i settori centro-orientali della Regione, mentre la provincia di Caserta, l'area metropolitana di Napoli e l'entroterra cilentano saranno interessate da un breve passaggio di nubi alte. A partire dal tardo pomeriggio, è atteso un graduale aumento della nuvolosità su Irpinia, Sannio e basso Cilento; un passaggio di nubi alte e sottili è previsto, invece, sui settori nord-occidentali della regione, in estensione su tutta la fascia costiera. **Temperature:** minime in diminuzione sulle aree interne e sulle principali pianure, con valori compresi rispettivamente tra 3°C e 7°C e tra 7°C e 9°C; in lieve diminuzione lungo la fascia costiera e comprese tra 9°C e 12°C. Massime in lieve aumento e comprese tra 12°C e 15°C sulle aree interne, pressoché stazionarie sulle principali pianure e lungo la fascia costiera con valori compresi tra 16°C e 19°C. **Venti:** molto debole o calmo sui settori interni ed a regime variabile; debole o al più moderato dai quadranti orientali lungo la fascia costiera durante le ore notturne. Dal tardo pomeriggio, vento molto debole o calmo ed a regime variabile su tutta la regione. **Mare:** durante l'intero arco della giornata, quasi calmo sotto costa, calmo a largo.

Previsioni per la città di Napoli: cielo sereno durante le ore notturne; velato durante il prosieguo della giornata. **Temperature:** minime in lieve diminuzione e comprese tra 9°C e 11°C; massime pressoché stazionarie e comprese tra 16°C e 18°C. **Vento:** debole dai quadranti orientali sino alle tarde ore pomeridiane, quando il vento spirerà molto debole ed a regime variabile. **Mare:** quasi calmo durante tutta la giornata.

Armando Rocco & Francesco Costa

Type

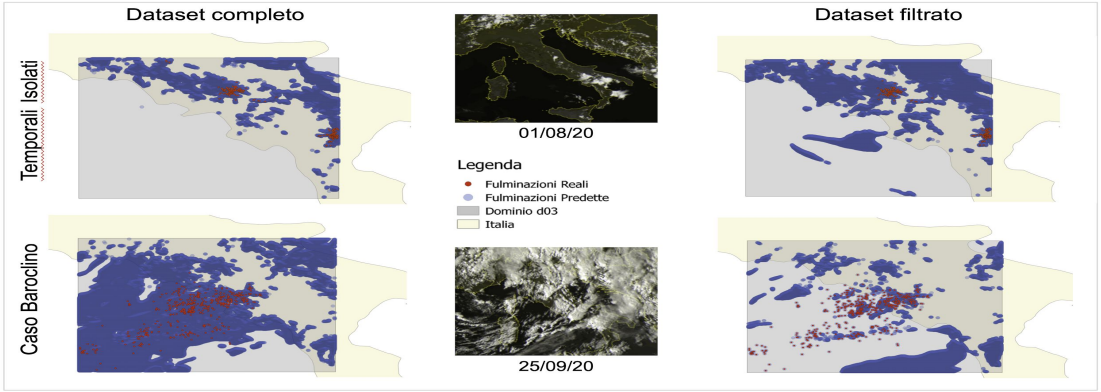
Weather Forecast

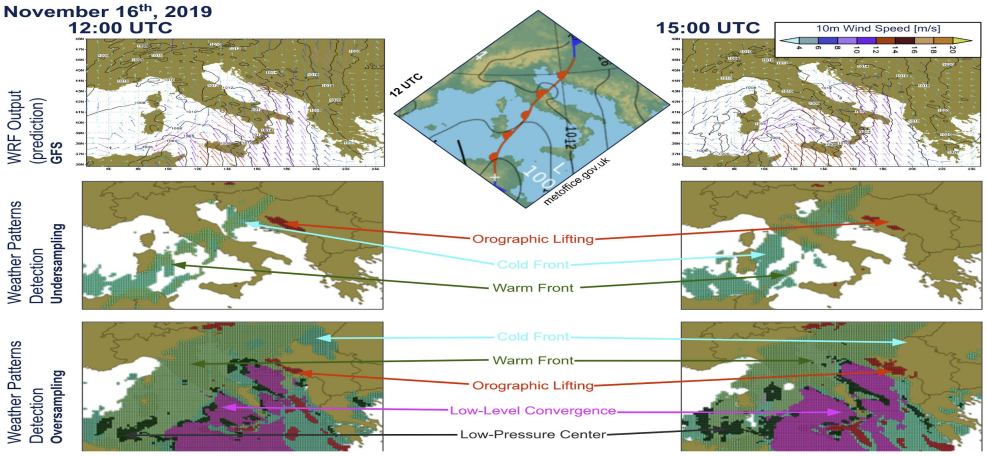
I dati e le mappe da essi generati sono frutto di elaborazioni di modelli fisico-matematici che operano in modo automatico senza intervento umano. I prodotti sono forniti così come sono. L'interpretazione deve essere svolta da personale esperto a proprio rischio. Le informazioni riportate non possono essere considerate sostitutive dei documenti ufficiali; l'orario indicato è in ora UTC. In nessun caso staff di meteo@uniparthenope, il CMMMA, il Dipartimento di Scienze e Tecnologie,Università degli Studi di Napoli "Parthenope" e tutte le persone coinvolte nelle relative attività di ricerca e produzione sono responsabili per danni a persone o cose.

EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Artificial Intelligence based products

Lightning predictions at sub-grid resolution.



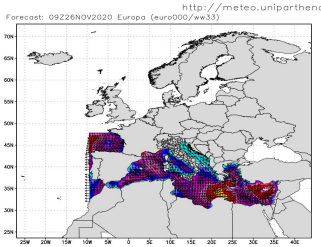


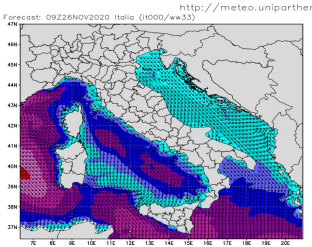
Weather Pattern recognition.

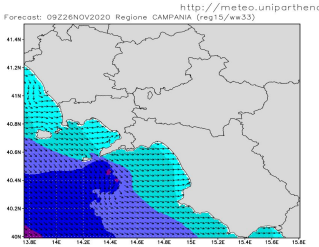
EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

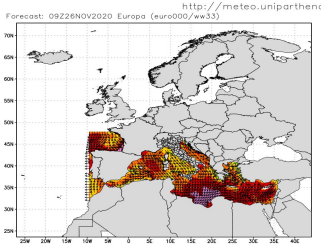
Ocean forecast products

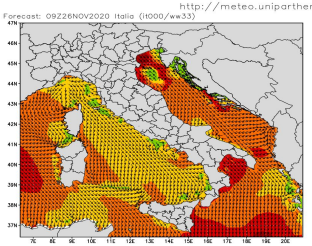
Sea Waves

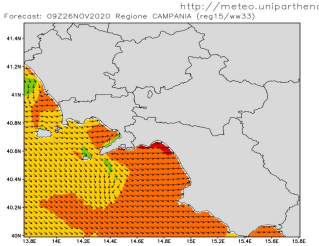




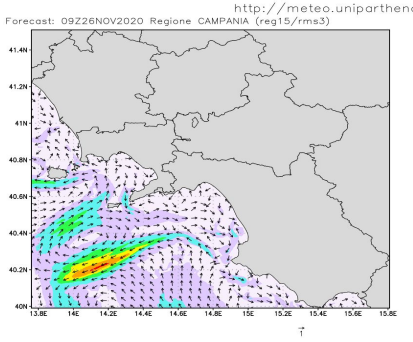




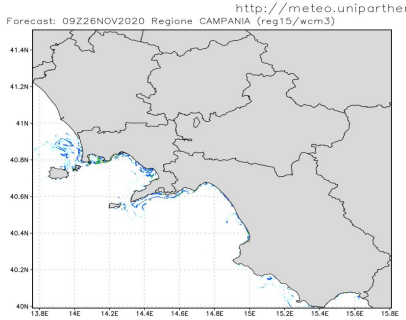




Sea Current



Passive Tracers



EuroHPC Joint Undertaking
ADMIRE
malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Weather application program interface

API 1.0

[Base URL: /]
http://193.205.230.6/swagger.json

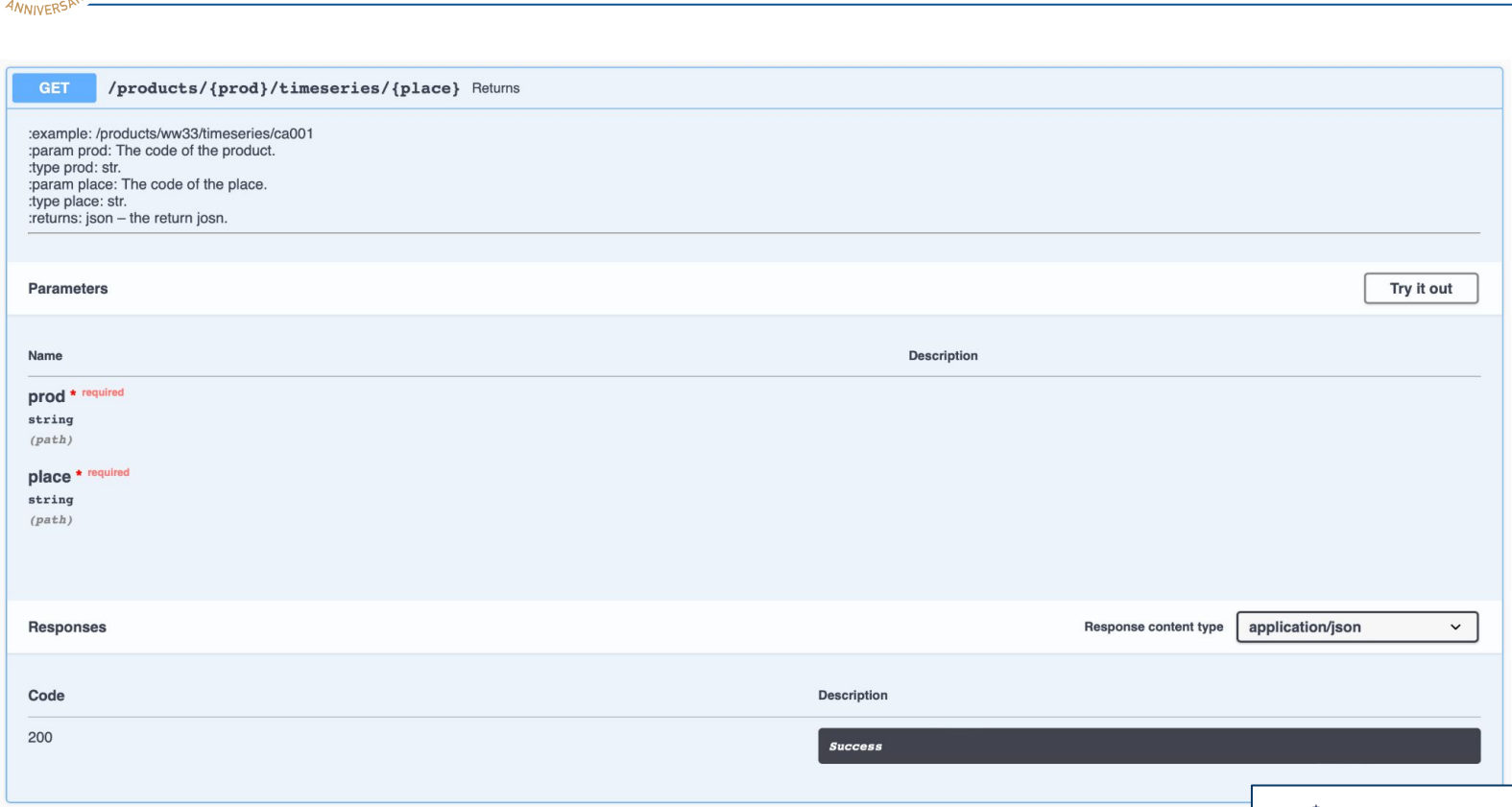
default Default namespace

GET	/apps/owm/{prod}/{placeprefix}/{z}/{x}/{y}.geojson	Returns
GET	/apps/sais/risk	Returns
GET	/apps/sais/risk/ondameters	Returns
GET	/apps/sais/risk/transects	Returns
GET	/apps/sais/risk/transects/{tid}	Returns
GET	/box/today/{place}	Returns
GET	/legal/disclaimer	Returns the Discaimer
GET	/legal/privacy	Returns the Privacy
GET	/places/search/byboundingbox/{minLatitude}/{minLongitude}/{maxLatitude}/{maxLongitude}	Returns
GET	/places/search/bycoords/{latitude}/{longitude}	Returns

http://api.meteo.uniparthenope.it

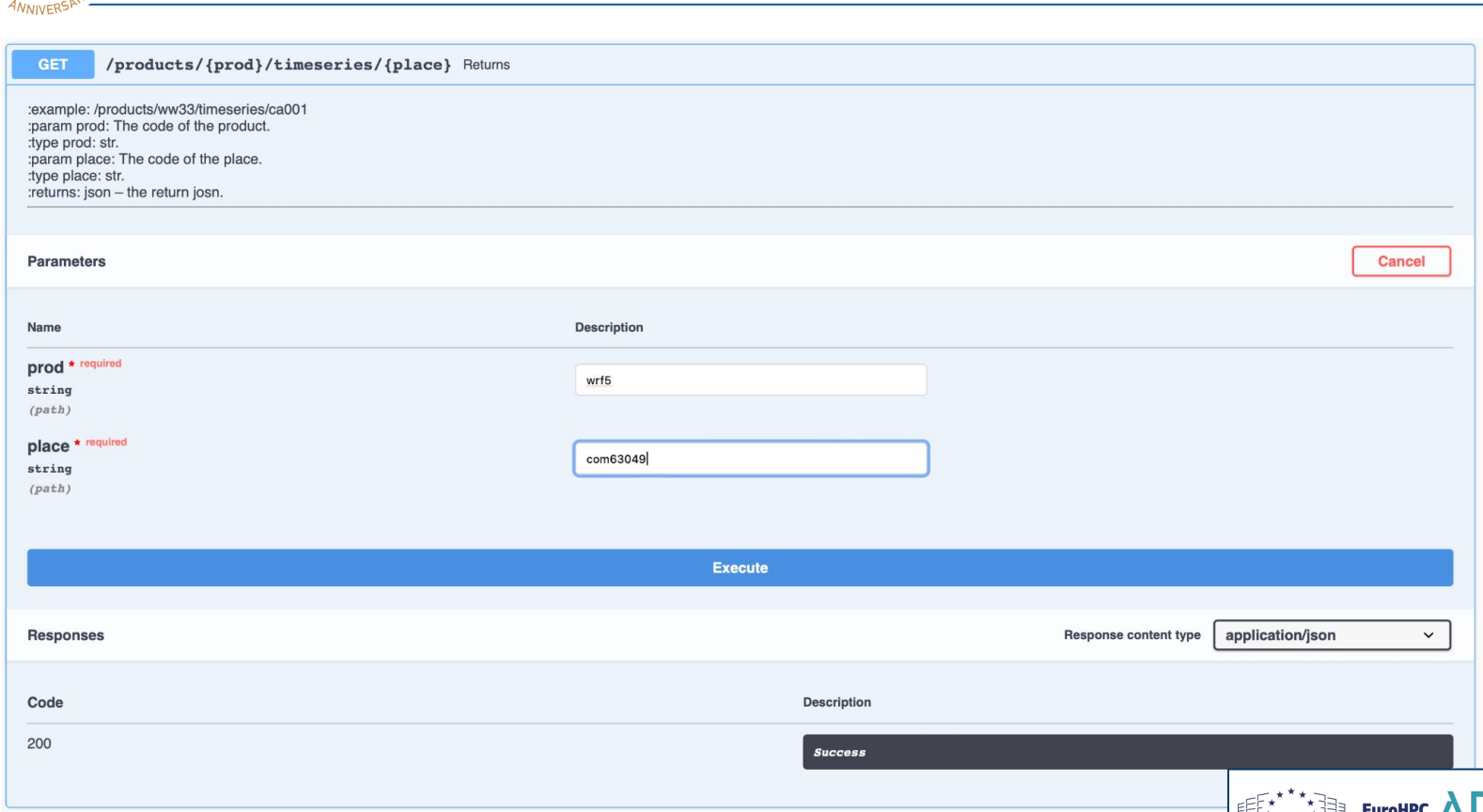
EuroHPC Joint Undertaking ADMIRE malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Weather application program interface



http://api.meteo.uniparthenope.it

Weather application program interface



http://api.meteo.uniparthenope.it

Weather application program interface

Responses — Response content type: application/json

Curl

```
curl -X GET "http://193.205.230.6/products/wrf5/timeseries/com63049" -H "accept: application/json"
```

Request URL

```
http://193.205.230.6/products/wrf5/timeseries/com63049
```

Server response

Code	Details
200	Response body


```
{
  "result": "ok",
  "timeseries": [
    {
      "clf": 0.99,
      "crd": 0,
      "crh": 0,
      "dateTime": "20211206Z0000",
      "dwd10": 29.4,
      "dws10": 0.2,
      "iDate": "20211205Z00",
      "icon": "cloudy5_night.png",
      "link": "product=wrf5&place=com63049&date=20211206Z0000",
      "rh2": 80,
      "rh300": 27.28,
      "rh500": 69.71,
      "rh700": 92.51,
      "rh850": 98.26,
      "rh925": 82.59,
      "rh950": 79.72,
      "slp": 1000.5,
      "swe": 0,
      "t2c": 9,
      "tc300": -52.74,
      "tc500": -28.31,
      "tc700": -9.01,
      "tc850": -0.02,
      "tc925": 5.23,
      "tc950": 6.86,
      "text": {
        "en": "Covered",
        "it": "Coperto"
      },
```


http://api.meteo.uniparthenope.it

EuroHPC Joint Undertaking ADMIRE malleable data solutions for HPC
ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE

Weather application program interface


```
{
  "result": "ok",
  "timeseries": [
    {
      "clf": 0.99,
      "crd": 0,
      "crh": 0,
      "dateTime": "20211206Z0000",
      "dwd10": 29.4,
      "dws10": 0.2,
      "iDate": "20211205Z00",
      "icon": "cloudy5_night.png",
      "link": "product=wrf5&place=com63049&date=20211206Z0000",
      "rh2": 80,
      "rh300": 27.28,
      "rh500": 69.71,
      "rh700": 92.51,
      "rh850": 98.26,
      "rh925": 82.59,
      "rh950": 79.72,
      "slp": 1000.5,
      "swe": 0,
      "t2c": 9,
      "tc300": -52.74,
      "tc500": -28.31,
      "tc700": -9.01,
      "tc850": -0.02,
      "tc925": 5.23,
      "tc950": 6.86,
      "text": {
        "en": "Covered",
        "it": "Coperto"
      },
```

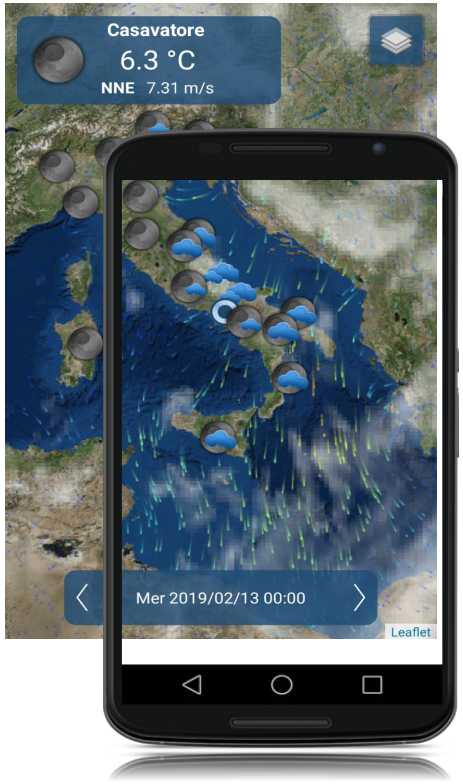

http://api.meteo.uniparthenope.it

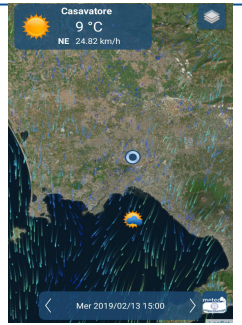
Weather application program interface

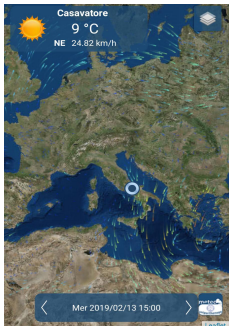

```
      "text": {
        "en": "Covered",
        "it": "Coperto"
      },
      "u10m": -2.22,
      "u300": 22.02,
      "u500": 11.45,
      "u700": 3.24,
      "u850": 1.38,
      "u925": -3.12,
      "u950": -3.96,
      "v10m": -0.52,
      "v300": 30.98,
      "v500": 12.94,
      "v700": 6.74,
      "v850": 1.29,
      "v925": -1.2,
      "v950": -0.58,
      "wchill": 9,
      "wd10": 95,
      "winds": "E",
      "ws10": 2.9,
      "ws10b": 2,
      "ws10k": 10.4,
      "ws10n": 5.6
    },
```

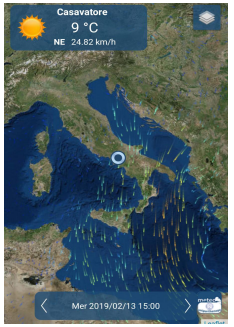

http://api.meteo.uniparthenope.it

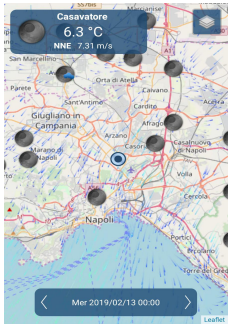
Eventually, the Weather App!











Download on the App Store

ANDROID APP ON Google play

EuroHPC Joint Undertaking

ADMIRE malleable data solutions for HPC

ADAPTIVE MULTI-TIER INTELLIGENT DATA MANAGER FOR EXASCALE