WP5: Architectural Support for Theoretical and Experimental Physics Data Management on the Distributed CN infrastructure

WP5 Leader:
Daniele Spiga (INFN – sez. Perugia) & Elvira Rossi (Università Federico II di Napoli)

Contributors: INFN, UNIMIB, UNINA, ROMA1, UNITS, UNIBO, UNIPD, UNIFE

- *Support of the adaptation of existing applications on the data-lake distributed infrastructure, and via innovative computational models*
- *Competence center for the design, implementation and test of computing models*

Today’s Agenda

17:30 → 19:00 **Parallel WP5**

ZOOM https://cern.zoom.us/j/62234969655?pwd=RlhkWXlhRjJPY2Q4S3hHMXAvV2d4QT09

Join our Cloud HD Video Meeting
https://cern.zoom.us

Conveners: Daniele Spiga (Istituto Nazionale di Fisica Nucleare) , Elvira Rossi (Istituto Nazionale di Fisica Nucleare)

17:30 **WP5: Introduction and status**

Speakers: Daniele Spiga (Istituto Nazionale di Fisica Nucleare) , Elvira Rossi (Istituto Nazionale di Fisica Nucleare)

18:00 **WP1: Use cases/Requirements and summary of WP1 parallel**

Speaker: Mattia Bruno

18:15 **WP2: Use cases/Requirements and summary of WP2 parallel**

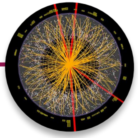
Speaker: Piergiulio Lenzi (Istituto Nazionale di Fisica Nucleare)

18:30 **WP3: Use cases/Requirements and summary of WP3 parallel**

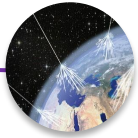
Speakers: Marco Landolfi, Paolo Natoli

Spoke 2 - Fundamental Research & Space Economy

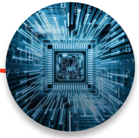
WP1: Design and development of science-driven tools and innovative algorithms for Theoretical Physics



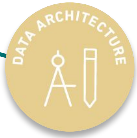
WP2: Design and development of science-driven tools and innovative algorithms for Experimental High Energy Physics



WP3:Design and development of science-driven tools and innovative algorithms for Experimental Astroparticle Physics and Gravitational Waves



WP4: Boosting the computational performance of Theoretical and Experimental Physics algorithms



WP5: Architectural Support for Theoretical and Experimental Physics Data Management on the Distributed CN infrastructure

WP6: Cross–domain Initiatives

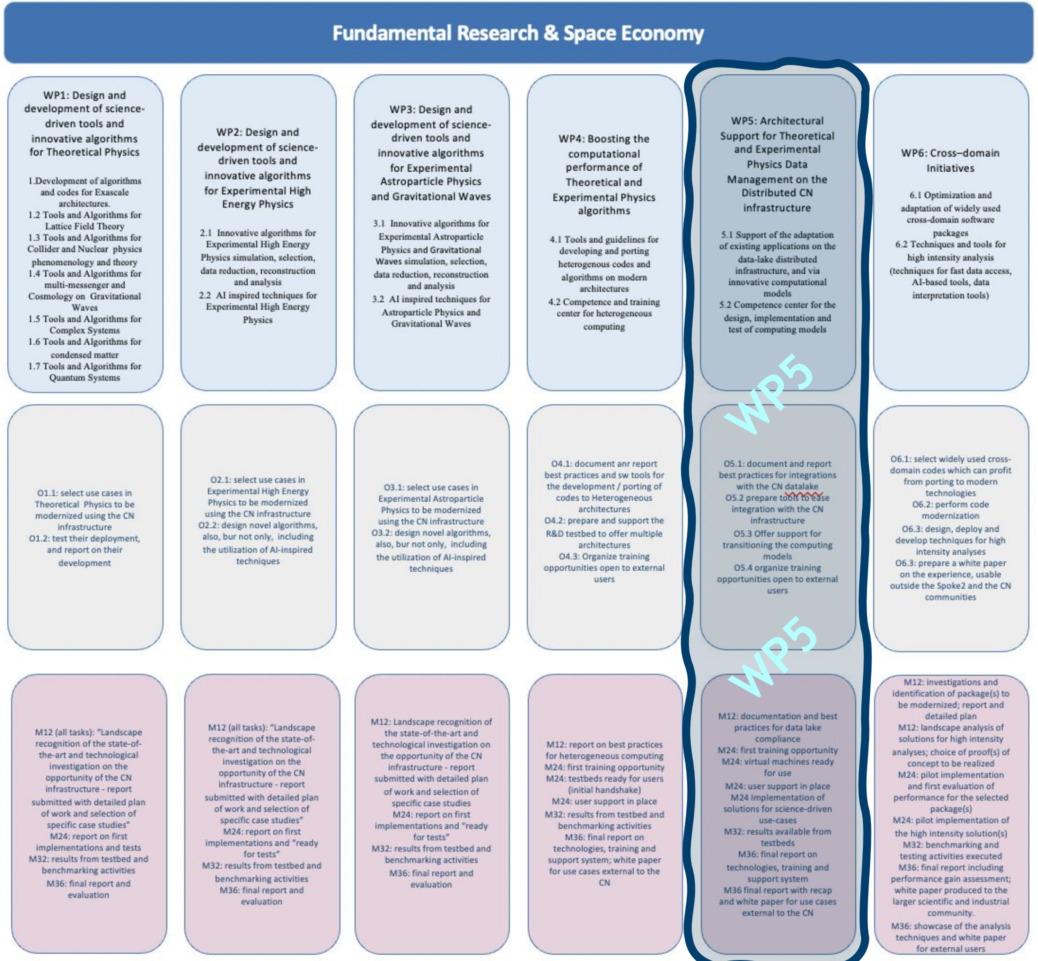
Spoke 2 - Fundamental Research & Space Economy

Fundamental Research & Space Economy

WP1: Design and development of science-driven tools and innovative algorithms for Theoretical Physics 1.Development of algorithms and codes for Exascale architectures. 1.2 Tools and Algorithms for Lattice Field Theory 1.3 Tools and Algorithms for Collider and Nuclear physics phenomenology and theory 1.4 Tools and Algorithms for multi-messenger and Cosmology on Gravitational Waves 1.5 Tools and Algorithms for Complex Systems 1.6 Tools and Algorithms for condensed matter 1.7 Tools and Algorithms for Quantum Systems	**WP2: Design and development of science-driven tools and innovative algorithms for Experimental High Energy Physics** 2.1 Innovative algorithms for Experimental High Energy Physics simulation, selection, data reduction, reconstruction and analysis 2.2 AI inspired techniques for Experimental High Energy Physics	**WP3: Design and development of science-driven tools and innovative algorithms for Experimental Astroparticle Physics and Gravitational Waves** 3.1 Innovative algorithms for Experimental Astroparticle Physics and Gravitational Waves simulation, selection, data reduction, reconstruction and analysis 3.2 AI inspired techniques for Astroparticle Physics and Gravitational Waves	**WP4: Boosting the computational performance of Theoretical and Experimental Physics algorithms** 4.1 Tools and guidelines for developing and porting heterogenous codes and algorithms on modern architectures 4.2 Competence and training center for heterogeneous computing	**WP5: Architectural Support for Theoretical and Experimental Physics Data Management on the Distributed CN infrastructure** 5.1 Support of the adaptation of existing applications on the data-lake distributed infrastructure, and via innovative computational models 5.2 Competence center for the design, implementation and test of computing models	**WP6: Cross–domain Initiatives** 6.1 Optimization and adaptation of widely used cross-domain software packages 6.2 Techniques and tools for high intensity analysis (techniques for fast data access, AI-based tools, data interpretation tools)
O1.1: select use cases in Theoretical Physics to be modernized using the CN infrastructure O1.2: test their deployment, and report on their development	O2.1: select use cases in Experimental High Energy Physics to be modernized using the CN infrastructure O2.2: design novel algorithms, also, bur not only, including the utilization of AI-inspired techniques	O3.1: select use cases in Experimental Astroparticle Physics to be modernized using the CN infrastructure O3.2: design novel algorithms, also, bur not only, including the utilization of AI-inspired techniques	O4.1: document anr report best practices and sw tools for the development / porting of codes to Heterogeneous architectures O4.2: prepare and support the R&D testbed to offer multiple architectures O4.3: Organize training opportunities open to external users	O5.1: document and report best practices for integrations with the CN datalake O5.2 prepare tools to ease integration with the CN infrastructure O5.3 Offer support for transitioning the computing models O5.4 organize training opportunities open to external users	O6.1: select widely used cross-domain codes which can profit from porting to modern technologies O6.2: perform code modernization O6.3: design, deploy and develop techniques for high intensity analyses O6.3: prepare a white paper on the experience, usable outside the Spoke2 and the CN communities
M12 (all tasks): "Landscape recognition of the state-of-the-art and technological investigation on the opportunity of the CN infrastructure - report submitted with detailed plan of work and selection of specific case studies" M24: report on first implementations and tests M32: results from testbed and benchmarking activities M36: final report and evaluation	M12 (all tasks): "Landscape recognition of the state-of-the-art and technological investigation on the opportunity of the CN infrastructure - report submitted with detailed plan of work and selection of specific case studies" M24: report on first implementations and "ready for tests" M32: results from testbed and benchmarking activities M36: final report and evaluation	M12: Landscape recognition of the state-of-the-art and technological investigation on the opportunity of the CN infrastructure - report submitted with detailed plan of work and selection of specific case studies M24: report on first implementations and "ready for tests" M32: results from testbed and benchmarking activities M36: final report and evaluation	M12: report on best practices for heterogeneous computing M24: first training opportunity M24: testbeds ready for users (initial handshake) M24: user support in place M32: results from testbed and benchmarking activities M36: final report on technologies, training and support system; white paper for use cases external to the CN	M12: documentation and best practices for data lake compliance M24: first training opportunity M24: virtual machines ready for use M24: user support in place M24 Implementation of solutions for science-driven use-cases M32: results available from testbeds M36: final report on technologies, training and support system M36 final report with recap and white paper for use cases external to the CN	M12: investigations and identification of package(s) to be modernized; report and detailed plan M12: landscape analysis of solutions for high intensity analyses; choice of proof(s) of concept to be realized M24: pilot implementation and first evaluation of performance for the selected package(s) M24: pilot implementation of the high intensity solution(s) M32: benchmarking and testing activities executed M36: final report including performance gain assessment; white paper produced to the larger scientific and industrial community. M36: showcase of the analysis techniques and white paper for external users

The Spoke 2 intends to address the needs of theoretical and experimental physics with accelerators, astroparticle physics with space- and ground-based detectors and gravitational wave investigation designing, developing and testing solutions apt to the current and next-generation experiments, and fitting the opportunities provided by the PNRR and the National Centre (CN) "Big Data, HPC and Quantum Computing".

Spoke 2 - Fundamental Research & Space Economy



Spoke 2 - objectives

The crucial aspects: the creation and/or optimization of algorithms and, in general, computing solutions capable of maximizing the potential physics output from experimental data and theoretical and phenomenological simulations, by using the tools made available by the Centre: e.g., heterogeneous and high-performance computing (via standard programming and AI-based solutions) and the ability to process large quantities of data beyond the capabilities of traditional methods.

A common denominator will be the utilization of more efficient strategies, reducing the computational costs and their power consumption footprint □ the project aims to be a key player in developing and testing solutions which ensure the sustainability of computing for the next generation of scientific experiments.

All the activities in Spoke 2 will be executed in strict collaboration with the major players in the respective scientific domains; it is thus important to maintain active and frequent communication channels with them.

WP5: Architectural Support for Theoretical and Experimental Physics Data Management on the Distributed CN infrastructure

Proposes:

- **5.1 Support of the adaptation of existing applications on the data-lake distributed infrastructure, and via innovative computational models**
- **5.2 Competence center for the design, implementation and test of computing models**

Objectives:

- ✔ **O5.1: document and report best practices for integrations with the CN datalake**
- ✔ **O5.2 prepare tools to ease integration with the CN infrastructure**
- ✔ **O5.3 Offer support for transitioning the computing models**
- ✔ **O5.4 organize training opportunities open to external users**

Milestones:

- ✔ **M12 (12 months): documentation and best practices for data lake compliance**
- ✔ **M24 (24 months): first training opportunity; virtual machines ready for use; user support in place; Implementation of solutions for science-driven use-cases**
- ✔ **M32 (32 months): results available from testbeds**
- ✔ **M36 (36 months): final report on technologies, training and support system; final report with recap and white paper for use cases external to the CN**

WP5: Approach and workflow

Similar approach with the respect to WP4 and WP6:

- ***Planning and identification:*** landscape recognition for best solutions for the realization of heterogeneous and portable code (e.g. software frameworks, compilers, programming models, ...), for the integration of services into a data-lake infrastructure; cross domain software and services will be identified if appropriate. Moreover, solutions for handling user support, user fora, and training opportunities will be identified.
- ***Realization phase:*** phase in which the services and the support systems are put into place, at least in alpha/beta phase. These include the testbeds to be used for benchmarking of scientific and industrial solutions, the user support system, the training opportunities.
- ***Validation phase:*** phase in which experience on the supported services and codes are reported, to be used as a touch base before the end of the project.
- ***Wrap-up phase:*** phase in which results are reported for executed activities, and are disseminated via white papers for future and external use cases.

WP5: Approach and workflow

Similar approach with the respect to WP4 and WP6:

- ✔ ***Planning and identification:*** landscape recognition for best solutions for the realization of heterogeneous and portable code (e.g. software frameworks, compilers, programming models, ...), for the integration of services into a data-lake infrastructure; cross domain software and services will be identified if appropriate. Moreover, solutions for handling user support, user fora, and training opportunities will be identified.
- ✔ ***Realization phase:*** phase in which the services and the support systems are put into place, at least in alpha/beta phase. These include the testbeds to be used for benchmarking of scientific and industrial solutions, the user support system, the training opportunities.
- ✔ ***Validation phase:*** phase in which experience on the supported services and codes are reported, to be used as a touch base before the end of the project.
- ✔ ***Wrap-up phase:*** phase in which results are reported for executed activities, and are disseminated via white papers for future and external use cases.

Planning and identification phase is where we are now…

what we need and what we are starting to do is a recognition of:

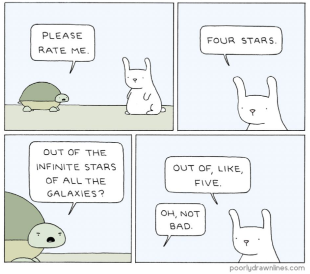
- person power and their knowledge to make the most of everyone's skills
- best solutions for the integration of services into a data-lake infrastructure
- design of possible user friendly solutions and users support

WP5: People

Contributors: INFN, UNIMIB, UNINA, ROMA1, UNITS, UNIBO, UNIPD, UNIFE

Local coordinators

WP5		
	Nome	Mail
Coordinatori	Elvira Rossi Daniele Spiga	elvira.rossi@unina.it daniele.spiga@pg.infn.it
Sede Affiliata		
INFN	Daniele Spiga	daniele.spiga@pg.infn.it
INAF		
UNIBA		
POLIBA		
UNIMIB	Mattia Bruno	Mattia.Bruno@cern.ch
UNINA	Elvira Rossi	elvira.rossi@unina.it
ROMA1		
UNITS	Andrea Bressan	andrea.bressan@ts.infn.it
UNIBO	Daniele Bonacorsi	daniele.bonacorsi@bo.infn.it
UNIFI		
UNICT		
UNIPD		
UniCal		
SALENTO		
UNIFE	Luca Tomassetti	luca.tomassetti@unife.it
IFAB		



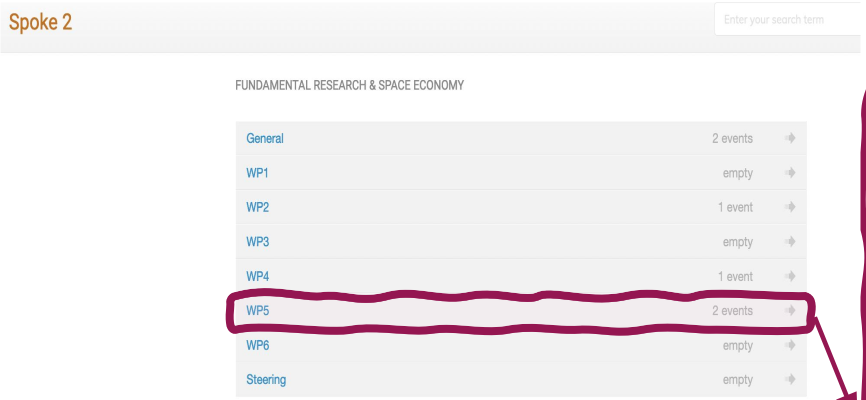
WP5 - People

- Alessandro De Salvo
- Andrea Bressan
- Antonio Stamerra
- Bernardino Spisso
- Daniele Bonacorsi
- Domenico Elia
- Eleonora Luppi
- Elvira Rossi
- Fabio Garufi
- Gianluca Sabella
- Giovanni Ianniruberto
- Giuseppe Milano
- Guido Russo (PO)
- Guido Russo (PU)
- Konstantinos Siettos
- Luca Tomasetti
- Lucia Silvestris
- Mattia Bruno
- Piergiulio Lenzi
- Sandra Malvezzi
- Simone Gennai
- Stefano Bagnasco

Very Very preliminary!!!
Only first declaration of interest not yet completed!

WP5: indico pages & email list

Spoke 2: https://agenda.infn.it/category/1774/

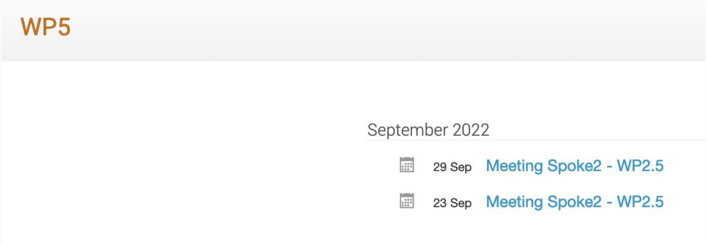


Email list

Please contact us if you are interested in contribut in WP5
email list: cn1-spoke2-wp5-all@lists.infn.it

a questo link trovate la pagina per iscriversi:
https://lists.infn.it/sympa/info/cn1-spoke2-wp5-all

WP5: https://agenda.infn.it/category/1781/



Two preliminary meetings with WP5 contacts to know each other and start to prepare the material for this kick off

Only WP5 contacts: Monthly/Biweekly/Weekly meetings as needed

For ALL WP5 enthusiasts: Monthly/Biweekly/Weekly meetings as needed will be organized

WP5: Un primo esercizio

WP5 Leader:
Daniele Spiga (INFN – sez. Perugia) & Elvira Rossi (Università Federico II di Napoli)

Un primo esercizio… “partiamo dal basso”

Abbiamo provato a farci un idea, in autonomia, basandoci sul materiale circolato durante la fase del “pre kick-off”.

- Una prima categorizzazione, a mo’ di esempio:

Hardware	Infrastrutture	Applicazioni e servizi	Tipi di Workflow	Distribuzione Software
CPU	batch (slurm)	Jupyter	pipeline automation	Container (docker/singula rity)
GPU	infinband	Spark	batch submission	
FPGA	Datalake	DB (NoSQL)	interattivo	
SSD/NVME				

NOTA: l’obiettivo qui è di stimolare la discussione iniziale ed evolvere (iterativamente). Riteniamo che il **processo di definizione dei requirements arrivi da WP1-3 (+4) o quantomeno in forte sinergia**.

Osservazioni e punti di discussione

Elemento centrale: probabilmente va definito, così da chiarire il gergo [**vedi dopo**]

Esempio di ottimi inputs: Vanno dettagliati e declinati, sottolineando:

- *pattern accesso dati, bookkeeping, servizi necessari, altre esigenze/richieste*

Hardware	Infrastrutture	Applicazioni e servizi	Tipi di Workflow	Distribuzione Software
CPU	batch (slurm)	Jupyter	pipeline automation	Container (docker/singularity)
GPU	infinband	Spark	batch submission	
FPGA	Datalake	DB (NoSQL)	interattivo	
SSD/NVME				

Definizione dei requisiti: Abbiamo una buona esperienza. **Capire i requisiti serve anche a definire modelli efficienti**

Quantificazione: **va definito "quanto serve".** Probabilmente con visione a breve e medio/lungo termine (i.e. ho bisogno di raggiungere una scala di N kcores..).

Teconolgie in nostro possesso: Sinergie con soluzioni disponibili ANCHE in progetti affini (INFN-Cloud/ML_INFN)

- **punto di forte interazione con Spoke0 (?!?)**

Verso cosa dobbiamo tendere: Un approccio “dall’alto”

Identificare e raccogliere delle User Stories ⇒ Sempre seguendo un processo dal basso abbiamo intercettato un paio di esempi:

- **Estratta da alcune note**
  - “Ci interessa che i risultati prodotti da una singola pipeline siano riproducibili ad esempio, e questo **è possibile solo utilizzando gli strumenti adatti**“
- **Estratta da alcune discussioni**
  - “Ci interessa poter utilizzare risorse HPC via slurm accedendo datasets prodotti altrove (un HPC diverso o altre risorse). Vorremmo che i dati prodotti viceversa fossero accessibili da altre facilities.
    Il nostro software è già pronto, per i dati non vorremmo spostare a mano via rsync”

Potrebbe aver senso un survey costruito cross-WP?

I DATI, lo storage, i trasferimenti…

I Dati e la loro gestione hanno un ruolo primario, tra l'altro parliamo di Datalake

- Analizzando il materiale disponibile si evince che **la parte di computing è "sbilanciata" verso l'accesso alle risorse di calcolo.** Su dati abbiamo ancora veramente pochi input da WP 1-2-3.

Gli aspetti che secondo noi andrebbero definiti e chiariti quanto prima sono, ad esempio:

- **Quantità di dati** che ci si aspetta di dover gestire
  - **Stato dell'arte:** dove stanno i dati oggi, come sono gestiti e di che tipo sono (*i.e. open data, embargoed, altro*)
- **Pattern di accesso** per i workflow identificati
- **Protocolli** di accesso e flessibilità dei software nell'adattarsi (*i.e. uso posix e il mio software può utilizzare solo posix*)
- **Cataloghi:** servono? se si quale è lo stato dell'arte?
- **Trasferimento Dati**: come anticipato ci si aspetta che questa funzionalità esista e sia "trasparente" all'utente. Quali sono le esigenze e i requisiti?

Alcune ipotesi operative: i testbeds per il compute

Sempre seguendo “l’approccio dal basso” abbiamo iniziato ad immaginare quali testbed potremmo proporre mettendoli a disposizione per i primi test:

- l’obiettivo principale sarebbe quello di raccogliere feedback
  - Utile per comprendere meglio le esigenze
  - Ma anche per concretizzare le esigenze in termini di servizi e di risorse

Facendo leva sulle sinergie possibili già oggi con i servizi e competenze disponibili in progetti scientificamente affini potremmo:

- Fornire soluzioni dedicate per batch con hardware specializzato (GPU)
- Creare ambienti per workflow interattivi
  - partendo da framework “dell’ecosistema python” si può testare la distribuzione e parallelizzazione su clusters (anche remoti/distribuiti)
- Testare soluzioni di offloading su risorse HPC
  - esempio di un Jupyter che si estende trasparentemente su cluster HPC

… testbed compute (cont)

Se mettiamo a frutto esperienze cloud con la possibilità di stabilire sinergie con progetti intorno a noi (e forse anche con altri Spoke!?!) ⇒ Potremmo proporre la costruzione di una soluzione per la gestione delle pipeline (vedi user story su riproducibilità)

Esistono soluzioni "cloud-native" che si sposano molto bene con piattaforme di gestione di servizi che già sono in uso su i.e. INFN-Cloud.

… e per i dati

Per quanto riguarda il Data Management potremmo isolare due macro aspetti:

1. la catalogazione dei dati (potrebbe essere community specific)
2. la gestione delle repliche e dei trasferimenti in ambiente distribuito: “data ingestion e data transfer“

In questo dominio un testbed da cui potremmo sicuramente partire è quello che si occupa della seconda:

- Abbiamo come riferimento un modello che già sappiamo può funzionare per più comunità (partiamo dagli output di ESCAPE)
- Identificare risorse disponibili oggi e a breve.
  - questo vale per tutti i testbeds, per i dati la questione è solo un po più delicata ovviamente

Alcuni punti aperti e recap

Una richiesta ricorrente è quella relativa all'accesso a risorse FPGA. Questo apre due aspetti da affrontare:

- Abbiamo già FPGA che possiamo mettere a disposizione con qualche soluzione
- Quale soluzione vogliamo pensare per consentire l'accesso a questo HW?
  - via cloud? bare metal? accesso container based?

Domanda per WP1-2-3: vedete qualcosa di utile nelle proposte discusse oggi?

Repetita: Per i Testbed servono servizi e risorse HW. Per le risorse HW abbiamo capito che esiste un Board che se ne occupa. Come funzionerà per i servizi?

- Noi proponiamo soluzioni come WP5 ma probabilmente è importante avere una connessione con Spoke0 o per i servizi pensiamo di agire in totale autonomia ?

Chiamata alle armi

Nei giorni scorsi abbiamo cercato di contattare persone che sappiamo svolgono attività scientificamente affini a quanto faremo in Wp5

Ovviamente non avremo intercettato tutti (banalmente avremo dimenticato…) per questo chiederemmo a tutti di aiutare e segnalarci interessati o possibili interessati

- esempio: se alla luce di quanto discusso già oggi, qualche parola vi ha fatto da trigger e avete in mente attività o persone “affini” contattate direttamente
    - elvira (elvira.rossi@na.infn.it)
    - daniele (daniele.spiga@pg.infn.it)

o comunque fate riferimento alla ML: **cn1-spoke2-wp5-all@lists.infn.it**

Backup

WP5: Approach and workflow

Similar approach with the respect to WP4 and WP6:

- ✔ ***Planning and identification:*** landscape recognition for best solutions for the realization of heterogeneous and portable code (e.g. software frameworks, compilers, programming models, ...), for the integration of services into a data-lake infrastructure; cross domain software and services will be identified if appropriate. Moreover, solutions for handling user support, user fora, and training opportunities will be identified.
- ✔ ***Realization phase:*** phase in which the services and the support systems are put into place, at least in alpha/beta phase. These include the testbeds to be used for benchmarking of scientific and industrial solutions, the user support system, the training opportunities.
- ✔ ***Validation phase:*** phase in which experience on the supported services and codes are reported, to be used as a touch base before the end of the project.
- ✔ ***Wrap-up phase:*** phase in which results are reported for executed activities, and are disseminated via white papers for future and external use cases.

Realization and Validation phase:

High integration with the industrial partners:

- industrial partners can provide testbed platforms, by offering their infrastructure in-kind, or procuring in their centres new R&D platforms not available in the CN;
- industrial partners are expected to test their typical-use cases on the same platforms, under the paradigm that "data are data", and once one can abstract from the specific domain, the technology to treat them is similar. This second part is expected to siphon experience and technologies between the academic and productive sectors, and vice versa.

The testbed and benchmarking ("validation") phase will be executed partially via the "innovation grants" available via the project, in which the solutions developed in alpha/beta level will be tested together with industrial partners, on hardware either provided by the CN, or acquired via the same grants.

WP5: Approach and workflow

Realization and Validation phase:

- **Shared testbeds and proof of concept**, using the infrastructure of the CB and other R&D platforms procured as needed, in order to test the processing of large, dispersed and heterogeneous data sources. Solution to be tested tentatively include:
  - Test on processing advantage using smart/dumb caches for remote/local data;
  - test using tiered storage systems (from tape to rotating disks to solid state disks);
  - test on remote streming vs lazy download vs caches w or w/o prefetch.

  - **Use cases:**
    - processing of O(10-100) TB of data from collider experiments, for example using workflows such as data processing in HEP;
    - typical data-intensive-use cases from companies, such as the processing of agricultural data, as suggested by Intesa Sanpaolo.

- Metrics to be considered for the testbeds are (at least) the processing efficiency / total time and power consumption using the various storage solutions, and the cost and scalability of the storage systems.
- All the activities will be executed in strict collaboration with the major players in the respective scientific domains; it is thus important to maintain active and frequent communication channels with them.
- **Open Calls:** number of support and ancillary services, like a web portal, a ticketing / support system, and help in organizing activities like benchmarking and training. We intend to use a part of the “Open Calls” to this purpose, selecting professionalities from companies or academic institutions with experience on the subjects from previous projects.