Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Spoke 2 - FUNDAMENTAL RESEARCH & SPACE ECONOMY

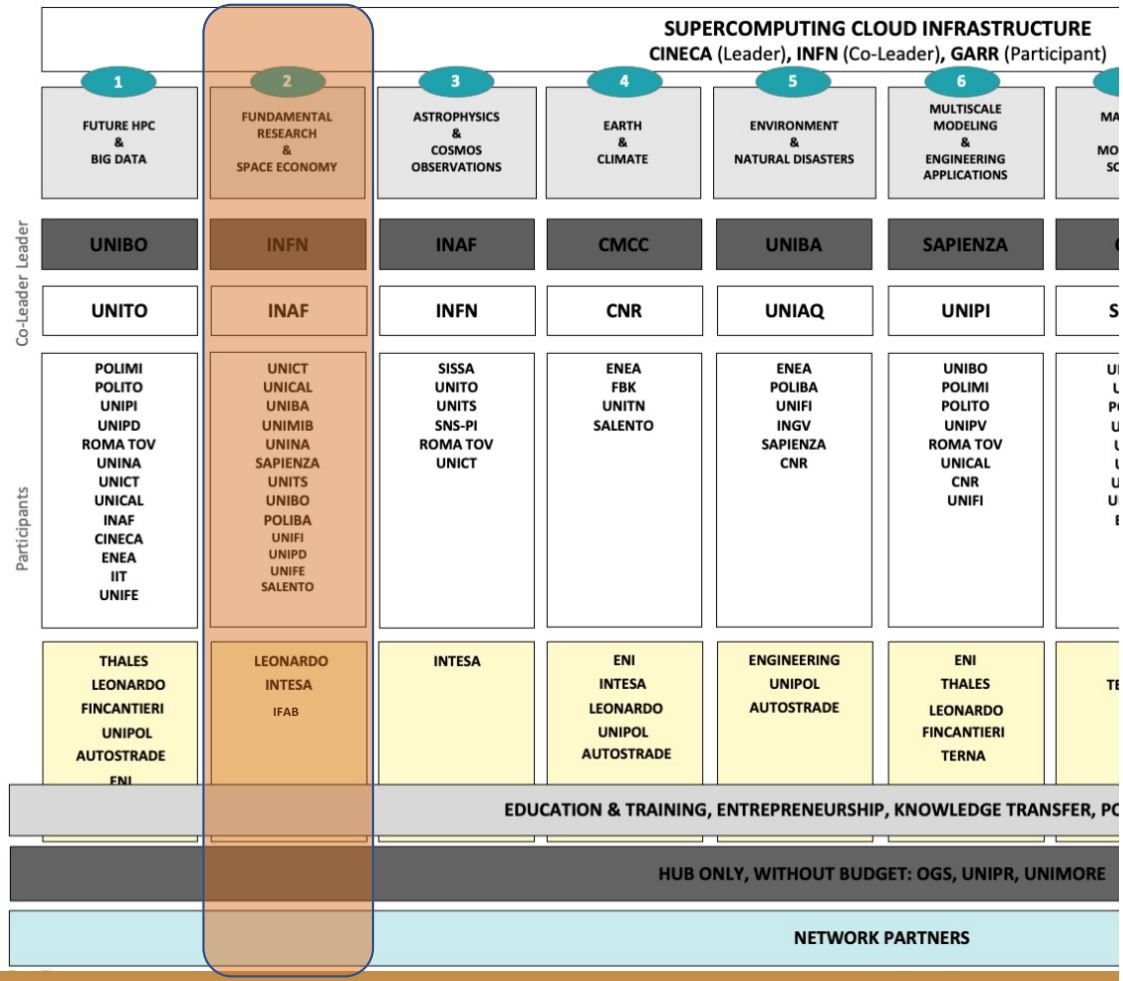
Sandra Malvezzi (INFN – MIB), Antonio Stamerra (INAF – OAR), Tommaso Boccali (INFN – PI)

Bologna, Kick-off Meeting, 25 Novembre 2022

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
INFN

Lo Spoke 2 in ICSC

- Spoke Leader: INFN
  - Sandra Malvezzi
  - Tommaso Boccali
- Spoke co-Leader: INAF
  - Antonio Stamerra
- 13 Università
- 3 privati da progetto



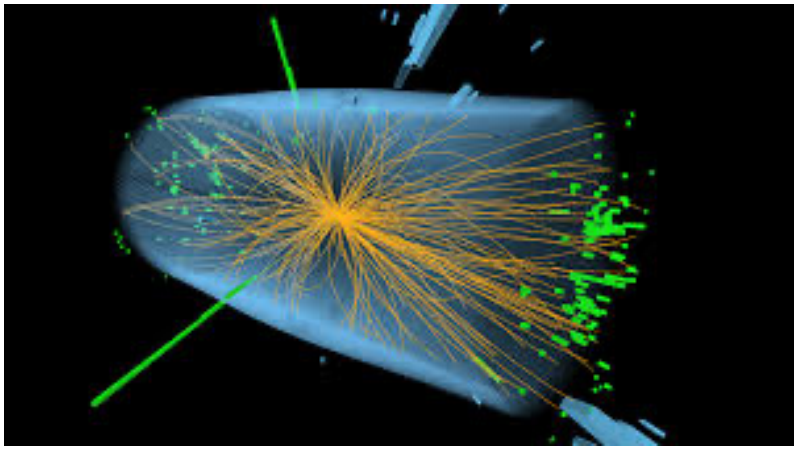
Finanziato dall'Unione europea
NextGenerationEU

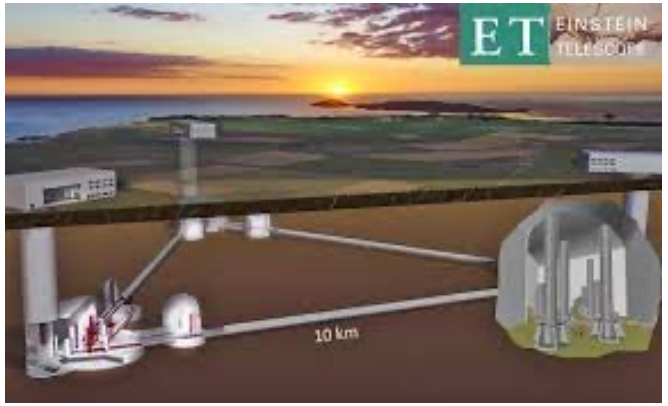
Ministero dell'Università e della Ricerca

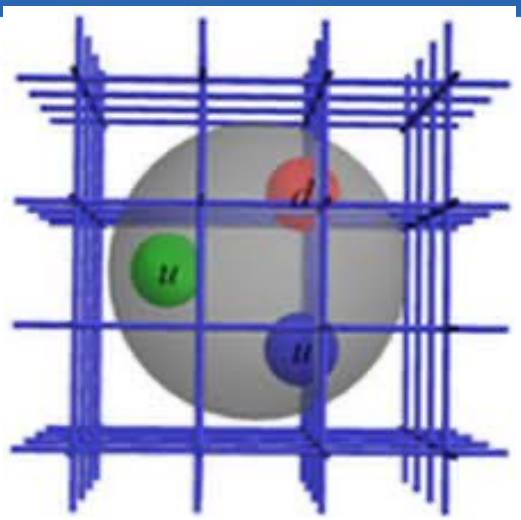
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

INFN

Quali sono le comunità scientifiche da cui veniamo?

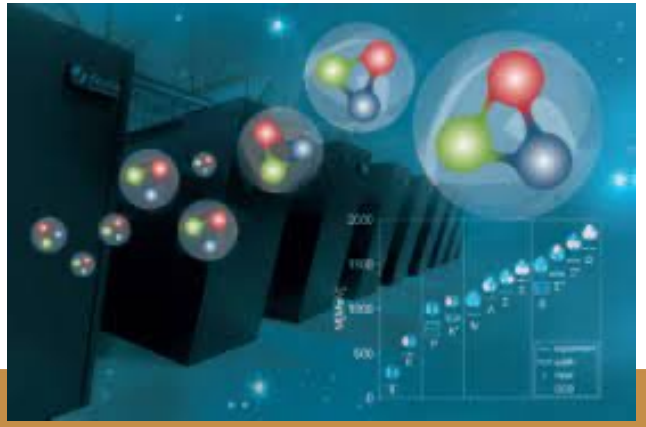






Le nostre necessità dal lato scientifico

- Anche se la stragrande maggioranza dei nostri affiliati è di estrazione "fisica", la realtà dei fatti è che nelle nostre attività abbiamo sempre più bisogno di capacità di calcolo avanzate, come dimensione e come complessità
- Esempi:
  - gli **esperimenti ai collisori** della generazione corrente hanno già superato la scala dell'**Exabyte** di dati raccolti, e li gestiscono mediante **> 100 centri di calcolo distribuiti** sul globo
  - Le simulazioni su Lattice sono fra i principali **clienti dei centri HPC**, e ne guidano la direzione di sviluppo
  - La quantità e la complessità (nonché il costo!) dei dati raccolti e delle soluzioni per processarlo hanno spinto il campo alla ricerca di **soluzioni avanzate di calcolo:**
    - **Il WEB, la Grid di calcolo di LHC, il test di soluzioni commerciali con HNSciCloud, la spinta per sistemi adatti a calcolo data-intensive in EuroHPC JU**
- Adesso, vediamo che altri domini (scientifici e non) si avvicinano a necessità simili alle nostre, e hanno conoscenze e idee da scambiare con noi
- **Lo spoke 2 →**



Cosa vogliamo fare? Dal progetto …

Spoke 2 - Fundamental Research & Space Economy

Science, and in particular science at the frontier of knowledge, is becoming more and more a computing intensive discipline. Current and next-generation experiments show processing and data needs comparable with the top global players and need a stack of solutions which are not typical of the curriculum of scientists. The trend has indeed started more than 15 years ago, with the development of solutions needed to satisfy the science of Collider Physics; since then, similar needs have been documented in other scientific domains, with Astroparticle physics showing by the end of the 2020s similar if not larger resource deployments. The activities in Spoke 2 "Fundamental Research and Space Economy" focus on boosting the science capabilities of current and future science initiatives, using the opportunities that PNRR in general and the National Centre for Big Data, HPC and Quantum Computing (CN) in particular offer in the next three years.

"share solutions between partners"

"use the CN infrastructure as enabling technology"

"no basic science, but computing FOR basic science"

GOALS AND OBJECTIVES

The activities in the "Fundamental Research & Space Economy" Spoke will be cast within the context of state-of-the-art research in basic science, and, in particular, of the domains of theoretical and experimental physics with accelerators and with space- and ground-based detectors for astroparticle physics and gravitational wave investigations. Within different time scales, all of these areas have or will have to face problems regarding the scaling and efficiency of computing infrastructures. These areas demand a scaling of the computing infrastructures and an improvement in efficiency. The Spoke intends to address these needs designing developing and testing solutions apt to the current and next-generation experiments, and fitting the contingent situation in Italy arising from the opportunities provided by the PNRR in general and the National Centre (CN) "Big Data, HPC and Quantum Computing".

In addition, the Spoke activities aim at demonstrating that the same set of solutions are of value in domains outside the bounds of basic science and of the specific research use cases. We plan to address at least two specific situations:

- the handling of data and processing in the context of the Space Economy Italian Strategy, including the handling and processing of data from the Mirror Copernicus program, by fostering the conditions to enable radically innovative services;
- the seeding of similar solutions in the productive context, with industry-research shared testbeds and proofs of concept; this will cover the productive domains in which data and processing resources need to be geographically dispersed, need a secure and granular solution for the data access, and experience computational problems with an unsustainable predicted future scaling (due to cost, efficiency, or performance).

The crucial aspects of its mission concern 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.

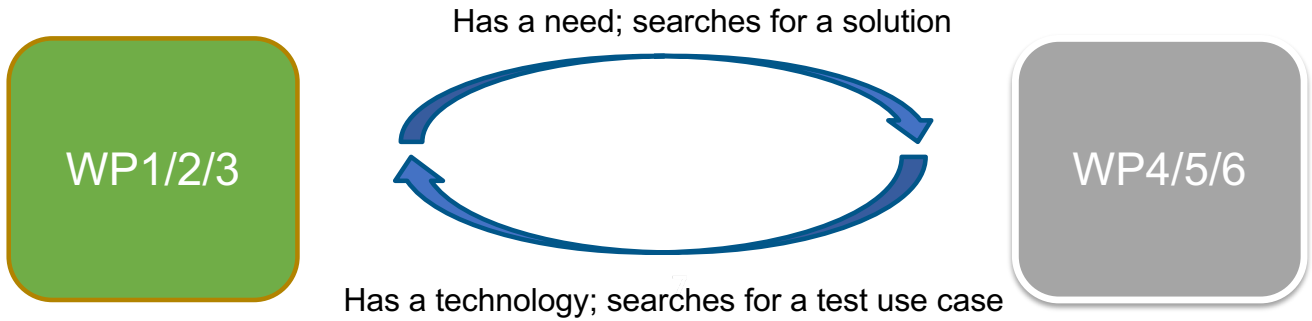
"design and test novel computing oriented solutions"

"go outside the academic environment; not to teach, but to share experience in both directions"

Some high level ideas / keywords

Suddivisione delle attività

- 6 Work Packages definiti; 3+3:
  - I primi 3 sono ”guidati da use cases (scientifici e industriali)”
    - “per fare questa misura di fisica, mi servirebbe un sistema che mi permetta di eseguire X operazioni su Y dati in Z secondi, usando al massimo W risorse”
  - I secondi 3 sono “guidati da tecnologie in nostro possesso / che vorremmo acquisire”
    - “l’affiliato X ha esperienza di sistemi di data management e accesso remoto al Terabit/s; a chi puo’ interessare?”
- Nella pratica, in questo modo ci aspettiamo che ogni use case (scientifico e industriale) possa afferire almeno a due WP



WP 1-2-3

- **WP2.1**: **Design and development of science-driven tools and innovative algorithms for Theoretical Physics**
  - ***Main Activities***: **development of algorithms, codes, and computational strategies** for the simulation of physical theories and models, towards pre-Exascale and Exascale architectures. Theoretical research projects in domains already using HPC solutions.
- **WP2.2**: **Design and development of science-driven tools and innovative algorithms for Experimental High Energy Physics**
  - ***Main Activities***: **selection, data reduction, simulation and reconstruction algorithms** (either via explicit programming or large scale Machine Learning solutions) for HEP experiments (LHC, Future Colliders, KEK, IHEP, neutrino experiments...), with applications ranging from innovative triggers to distributed analysis techniques.
- **WP2.3**: **Design and development of science-driven tools and innovative algorithms for Experimental Astroparticle Physics and Gravitational Waves**
  - ***Main Activities***: **data reduction, reconstruction and time cross-correlation algorithms, data selection and simulations** of astroparticle and gravitational waves experiments, tools for cross-correlations and pattern recognition in multi-messenger physics, including novel implementations using techniques like Machine Learning.

WP4

- **WP2.4**: **Boosting the computational performance of Theoretical and Experimental Physics algorithms**
  - *Main Activities*: **porting of applications to GPUs and heterogeneous architectures** (e.g., scalability of scientific codes and applications on **GPU/CPU** many-cores clusters, **local and remote offloading**, mission-critical algorithms on **FPGAs**, ...). The solutions and tools implemented during the project will be easily extendable to other scientific domains of the Centre and to the industrial partners in the Spoke; moreover, the personnel trained within the Centre will help to spread and boost the application of HPC methodologies to Italian academic and industrial fields, for a comprehensive advancement of the Italian system.

In una frase:
ottimizzazione delle performance dei nostri algoritmi sul singolo nodo, anche mediante l'utilizzo di acceleratori

WP 5

- **WP2.5**: **Architectural Support for Theoretical and Experimental Physics Data Management on the Distributed CN infrastructure**
  - *Main Activities*: **support for the adaptation of existing applications on the data-lake distributed infrastructure, and via innovative computational models** (for example sharing of gauge configurations in lattice field theories, long- term data preservation, streaming access to data, tiered storage solutions, ...). The solutions implemented will be tailored to the needs of the scientific fields, easily extendible not only to the nearby scientific domains in the Centre, but also to all academic and industrial realities where needs to access distributed computing and large amounts of data exist. In particular, the industrial partners in the Spoke have expressed interest in using the same technologies for their specific use cases.

In una frase: scalare le performance da quelle del singolo nodo a quelle di sistemi (geograficamente) distribuiti, inclusi gli aspetti di *data management / movement / access*, e l' interfaccia al Datalake di ICSC

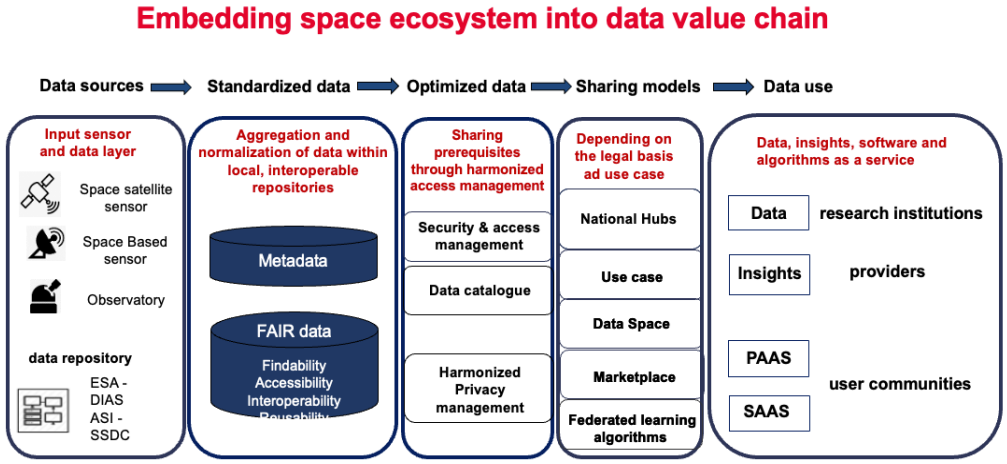
WP 6

- **WP2.6**: **Cross–domain Initiatives**
  - ***Main Activities***: **optimization and adaptation of widely used software packages on the national Centre infrastructure**, like Geant4 or FLUKA or generic high-performance techniques for data access/analysis; **statistical and AI-based tools; data-interpretations tools**. In the context of the Space Economy Italian Strategy, develop and deploy techniques to access, **analyse and process the data from the Mirror Copernicus program**, creating the conditions to enable radically innovative services. In particular, enable thorough and continuous observation programs for global and local processes, allowing external partners to operate a large variety of services, including the planning for emergencies, risks and resources.

Il WP più complesso; include lo studio di use cases di generale interesse come data analysis a alte prestazioni, a un aspetto specifico: è *l'hook* in ICSC per le attività di Space Economy

Le industrie

- Già elencate al momento del bando
  - **Leonardo**: interesse soprattutto lato Space Economy
  - **Intesa Sanpaolo**: molteplici punti di contatto
    - Data management, analisi dati da sorgenti eterogenee, AI …
- Discussioni post bando
  - **IFAB**: molteplici interessi; i più chiari al momento
    - **E4**: testbed tecnologici su soluzioni «non standard»
    - **BIP** e **Esteco**: data analysis e processing
  - In attesa di definizione:
    - **ENI**
    - **UnipolSAI**



I testbed

- Per noi la realizzazione di **testbed** è una parte essenziale dell'attività del 2o e 3o anno, e il modo principe di collaborare con l'industria
- Per noi il testbed è una piattaforma tecnologica, messa a disposizione dai partner Spoke 2 o mediante Open Calls o Innovation Grants o dallo Spoke 0 o dallo Spoke 1 ("Living Labs"), su cui sperimentare use cases accademici e industriali. Esempi (non finali):
  - Una piattaforma per test di **calcolo eterogeneo** (standard se da Spoke 0, meno se autoprocurata o da Spoke 1: ARM + GPU, GPU non Nvidia)
  - Una piattaforma per test di **calcolo distribuito** e porting di modelli su datalake
  - Una piattaforma ottimizzata per **ML training**
- Ovviamente l'ideale sarebbe condividerle ANCHE con altri spoke

Collaborazioni con altri Spoke

- Ancora in fase di studio; ci aspettiamo grosso contributo dalle Riunioni dello **Steering degli Spoke**
- Alcune Ovvie
  - **Spoke 0**: alcuni dei nostri use cases sono già pronti a USARE risorse → Spoke Allocation Board + Spoke 0; middleware datalake da integrare con WP5
  - **Spoke 3**: Spoke "sinergico". Scambio di persone a livello di Steering per armonizzare le attività, previste collaborazioni a molti livelli (attività "astroparticle", data management, …)
  - **Spoke 1**: sinergia a livello tecnologico: accesso a piattaforme di R&D, studio di framework per calcolo eterogeneo

Finanziato dall'Unione europea NextGenerationEU

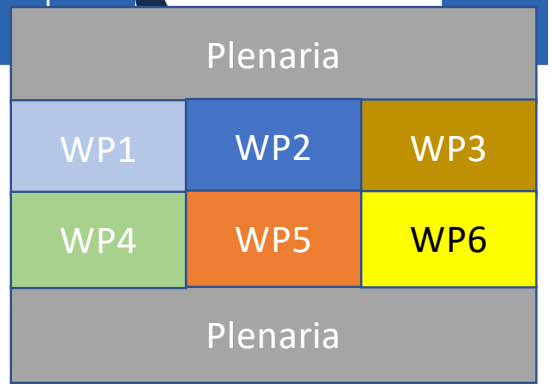
Ministero dell'Università e della Ricerca

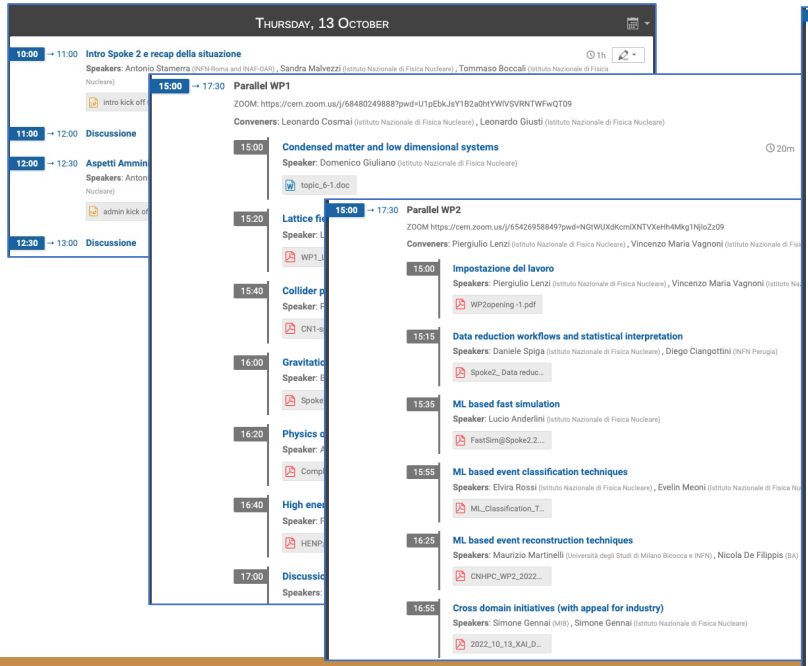
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

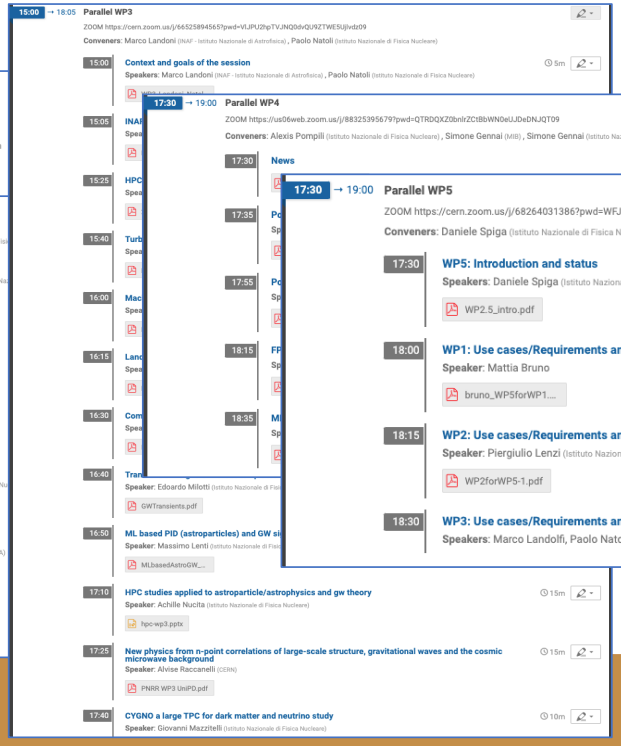
INFN

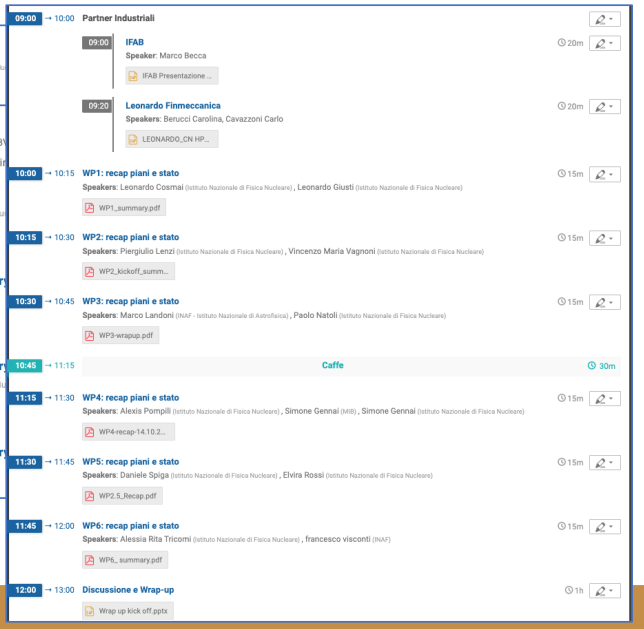
Stato delle attivita'

- Kick off «scientifico» dello Spoke 2 13-14 ottobre (link).
- 48 presentazioni fra plenarie e parallele





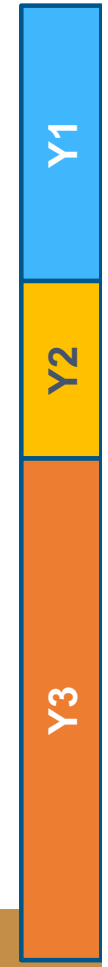




Il piano di attività

- Il piano di attività è definito in 4 fasi, allineate con le milestone di ICSC
- Buona parte del primo anno, a parte l'annosa questione del reclutamento, è dedicata alla selezione di use cases «**flagship**» da realizzare nel progetto
- Abbiamo convenuto con i WP Leaders di operare una prima identificazione per ~ Natale
  - Use cases + necessità tecnologiche + necessità di risorse

1. planning and identification: the first year of the project is dedicated to a landscape recognition for interesting state-of-the-art use case; its outcome is a work plan identifying the activities on which the core part of the project will be focusing - in particular, algorithms and services to be accomplished;
2. a realization phase, in which the actual development is performed via the staff/ hired personnel; the outcome is usable algorithms / services, documented (alpha/beta level) and ready to be tested on a larger scale;
3. a validation phase, in which the outcomes of the realization phase are verified in testbeds and proofs of concept, and benchmarked in order to assess their adherence to the specifications;
4. a wrap-up phase, in which results are analysed and consolidated in reports and white papers to be used as guidelines for similar use cases.



Highlights dai WP – WP1

- 7 research lines identificate
- Necessita' identificate su HPC/HTC per un totale di O(30) Mcoreh e PB-level storage

WP1@Spoke2: Research topics

a) Development of algorithms, codes and computational strategies for the simulation of physical theories and models,towards pre-Exascale and Exascale architectures. [L. Cosmai, L. Giusti]

b) Theoretical research projects in domains already using HPC solutions, such as:

b1) **Lattice field theory** (flavour physics, QCD phase diagrams, hadronic physics, interactions beyond the Standard Model, machine learning in quantum field theories, electromagnetic effects in hadronic processes) [L. Cosmai, L. Giusti]

b2) **Collider physics phenomenology** [P. Mastrolia]

b3) **Gravitational waves, cosmology and astroparticle physics** (neutron-star physics, primordial universe, dark matter and energy, neutrino physics) [B. Giacomazzo]

b4) **High energy nuclear physics** [F. Becattini]

b5) **Physics of complex systems** (fluid dynamics, disordered systems, quantitative biology) [A. Suma, G. Gonella]

b6) **Condensed matter in low dimensional systems** [D. Giuliano]

Possible interaction with WP4

Porting Parallel CPU codes to GPUs

Possible interaction with WP5

Data repositories (available to the public ?), and disk storage

WP1 - Lattice

I ricercatori c

Studio della
• Strategia c
QCD su ret
• Simulazion
• *Interazione*
configuraz

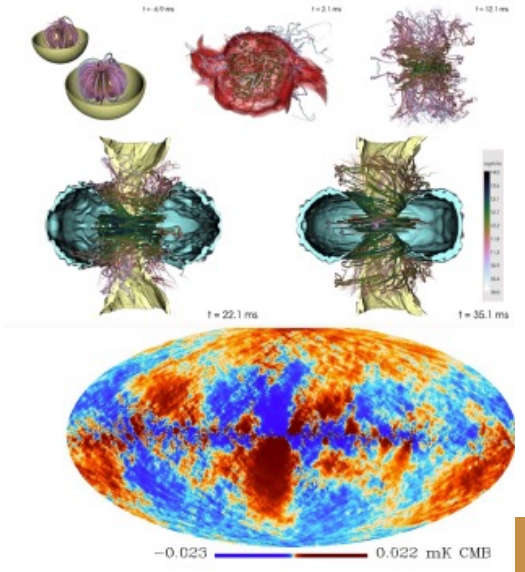
Simulazione
di gauge su r
• Algoritmi e
• Tools per l'

Dinamica del
• Strategia n
cromoelett
• Modulo pe
• *Interazione*
tensore cro

GW, Cosmology, and Astroparticle

Use cases (a synthesis)

- Numerical Codes to study plasma in early universe
- Lattice and Markov Chain Monte Carlo Simulations of phase transitions
- Numerical Relativity Simulations of Compact Objects (including GRMHD, neutrinos, alternative gravity models)
- Numerical Algorithms to study neutrino oscillations and axion-photon conversions
- Cosmological simulations of the dynamics of bubbles of true vacua
- Computing the spectrum of gravitational waves in transplanckian collisions
- Simulating black hole formation environment (via SPH and N-body codes)
- Improving theoretical modelling in data analysis of large cosmological datasets
- Cosmological codes to calculate Large Scale Structure observables



Interaction with WP4

- Porting Parallel CPU codes to GPUs

Interaction with WP5

N/A

Institutions

UNICAL, UNIMIB, UNIBA, UNIFI, UNIFE, UNISAPIENZA, SALENTO, UNIPD

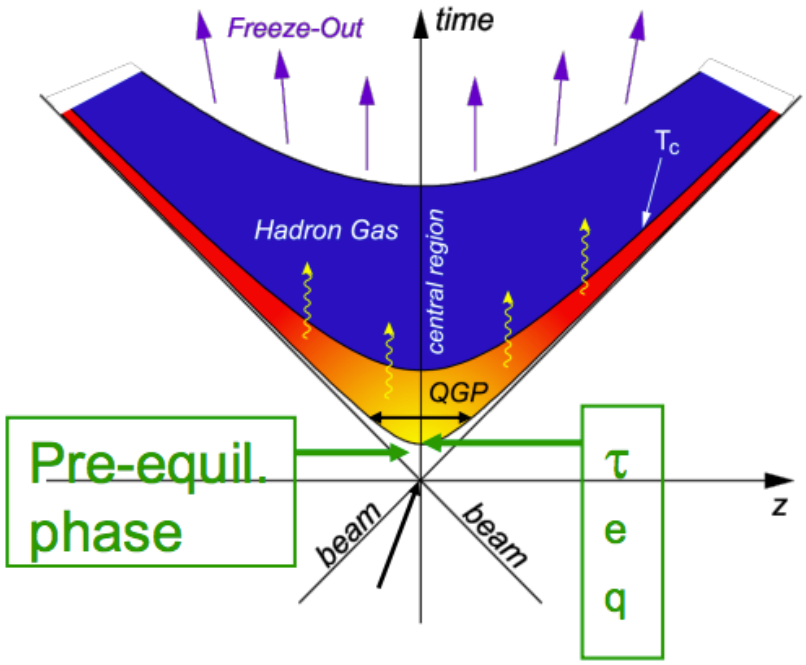
High energy nuclear physics

Use case
-Comple
-Emerge
-Fluid dy
-Medica
-Non-eq
-Model
-Bridgin
-Develo
Field Th
-Model

Institutio
UNIBA
UNIFE
UNINA

Use cases (a synthesis)

- Advanced relativistic hydrodynamics numerical codes including viscous terms to simulate QCD plasma formed in relativistic nuclear collisions;
- Relativistic kinetics numerical codes to study transport phenomena: heavy quark diffusion in the plasma etc.

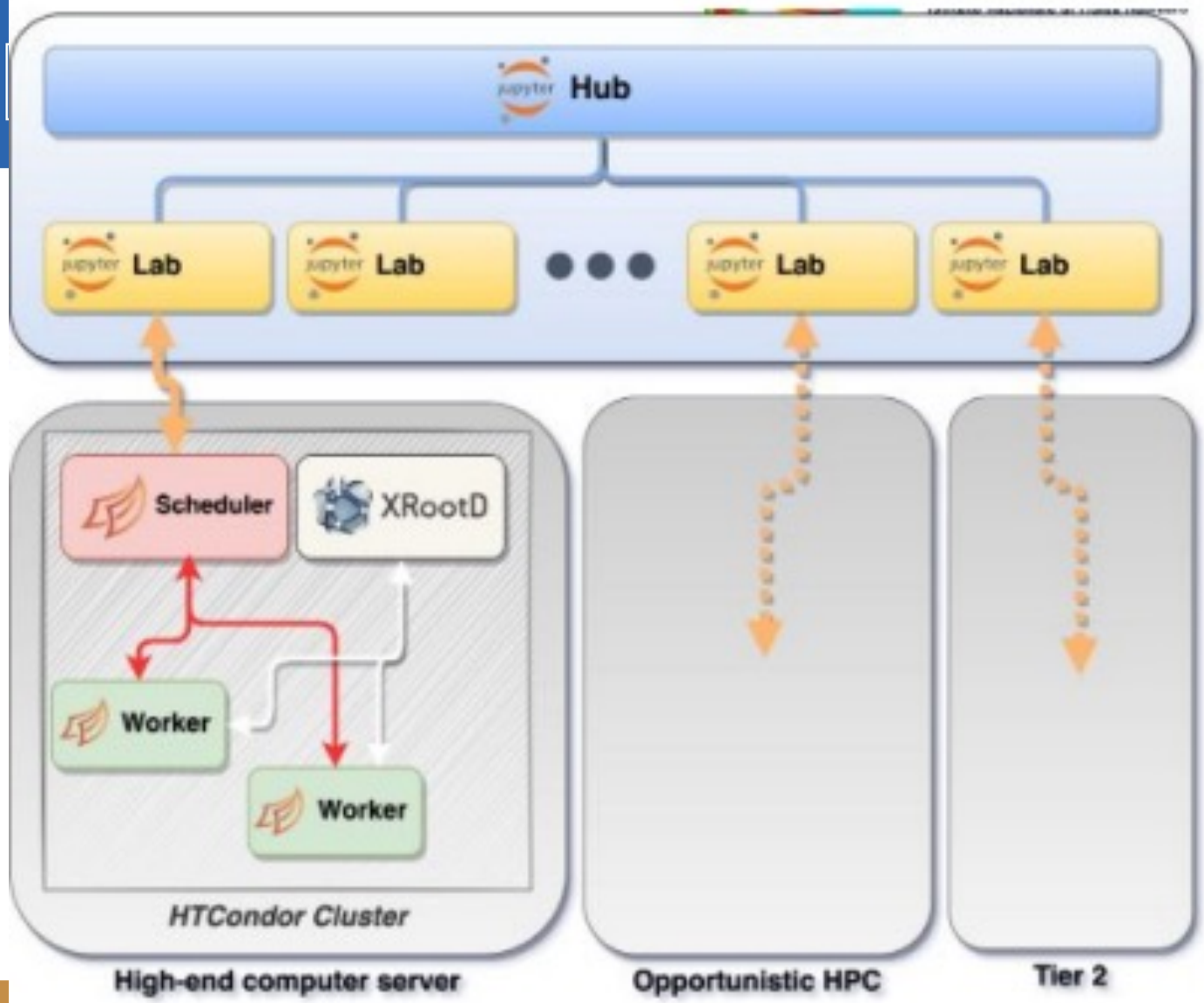


Institutions
UNIFI, UNICT

Institutions
SALENTO, UNICAL

Highlights dai WP – WP2

- 75 persone attualmente iscritte alle mailing list
- Necessità sia algoritmiche (migliorare le performance / tempi di esecuzione) → Machine Learning, porting su architetture eterogenee, sia infrastrutturali (analisi distribuita ad alto throughput)
  - → WP4, WP5, WP6
- Use cases da selezionare fra le piu' che abbondanti necessità della fisica ai collider
  - Simulazione veloce con metodo Monte Carlo
  - Algoritmi di ricostruzione sia offline sia real time
  - Algoritmi di analisi veloce ("come analizzare 1 PB di dati in 1 ora")



Jupyter notebooks (trasparentemente) distribuiti su macchine di tipologia diversa

- Nodi singoli
- Leonardo
- Data Center INFN

Di interesse di WP4 (acceleratori), WP5 (distribuzione di data / workload), WP6 (sistemi di analisi a alto throughput)

ML based event reconstruction

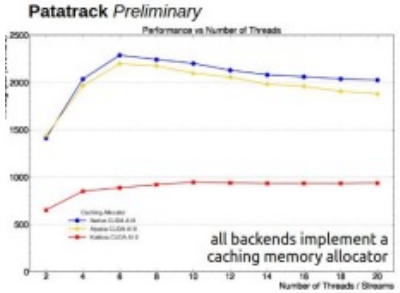
Maurizio Martinelli ha presentato un grande numero di attività che costituiscono un campione molto rappresentativo di tante iniziative in corso negli esperimenti.

Non richieste gravose in termini di risorse al CN. Infrastruttura centralizzata un beneficio

Trigger infrastructure

Anomaly detection strategies for trigger con uso di FPGA

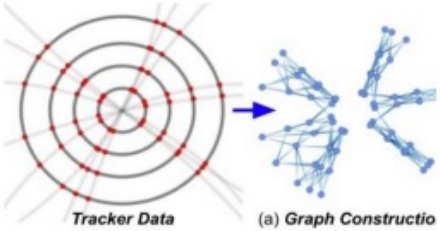
Porting eterogeneo del codice di trigger su GPU



Vertexing e tracking

Graph NN per pattern recognition

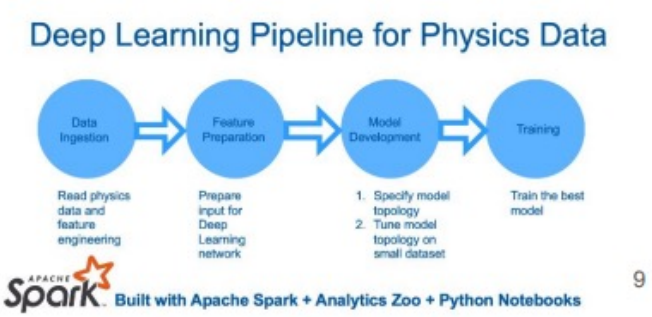
NN per 4D vertexing



Particle ID

Varie iniziative nel contesto di esperimenti attuali e futuri

Pipeline di deep learning



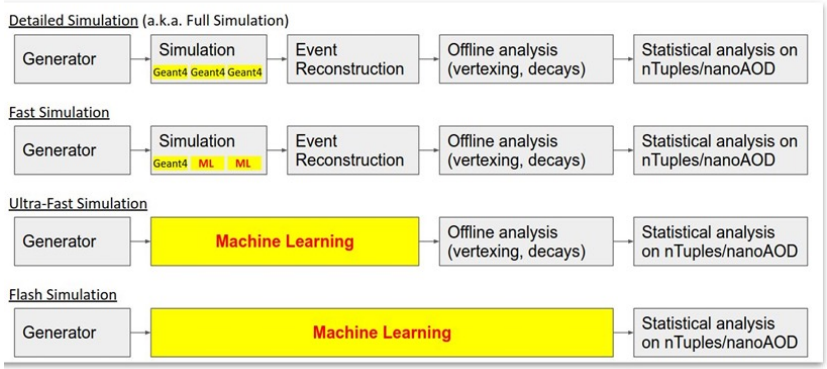
9

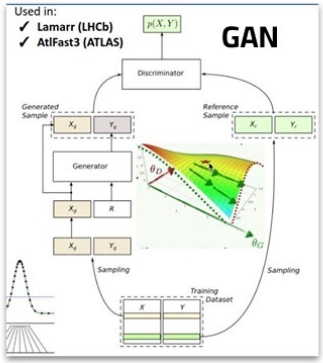
Uso di tecnologie sw e hardware (→ testbeds) per training di pipeline ML a larga scala, data intensive

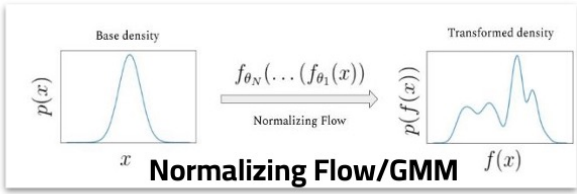
Di interesse di WP4 (acceleratori), WP5 (distribuzione di data / workload)

ML based Simulation - Fast, *Faster*, *Flash* (I)

Lucio Anderlini ha discusso le iniziative legate alla ML-based fast simulation in Italia







Forte coinvolgimento di gruppi italiani in LHCb, ATLAS, CMS

Attività essenzialmente **production ready**

Massive GPU resources needed in **bursts of time**, to shorten the development cycle.

ICSC un'occasione anche per migliorare il coordinamento nazionale su questi temi

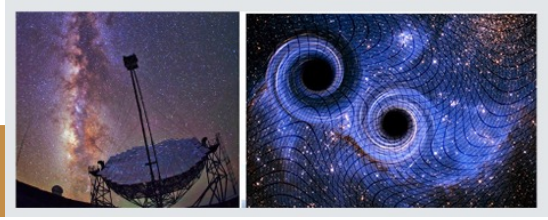
6

- Da simulazioni dettagliate (per «noi» vuol dire Geant 4) a sistemi *GAN-based* per generazione veloce
  - Da minuti a msec
  - Cambio infrastrutturale: da GRID per il processing, a HPC per il training

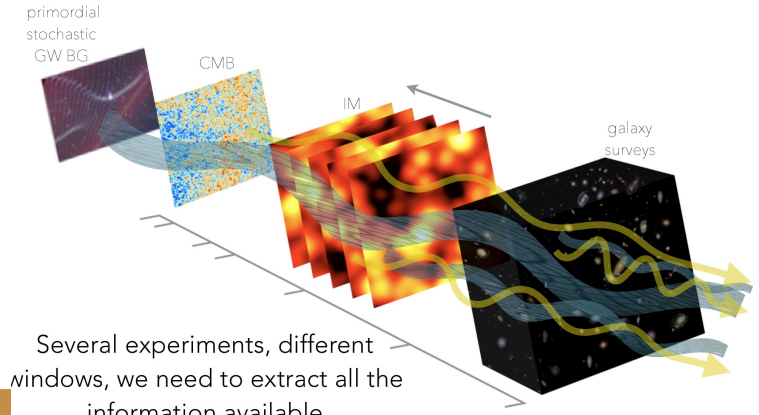
Highlights dai WP – WP3

- **Use cases in esame classificati come**

- IACT/CTA (Cherenkov) + astroparticle detection (e.g. CALET, JEM-EUSO, HERD, …)
- Geodetic satellites (LAGEOS/LARES)
- Pipeline for satellite data analysis (atroparticle/X/Gamma) in “heteregenous” synergy (e.g GW)
- Cosmology/CMB/early Universe (LiteBIRD, CMB-S4, Euclid, SKA, …)
- Turbolence with application on atmo- and helio-spheres + Compact obeject simulation and merging



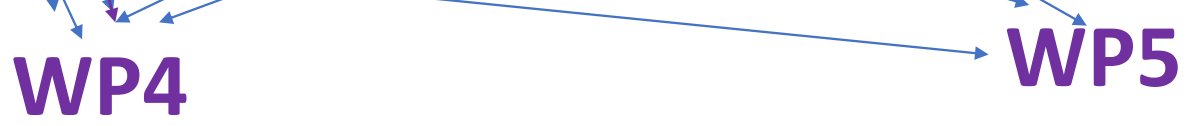
- GW:
  - Lisa/Virgo/Kagra
  - LISA
  - ET
- Multimessenger astrophysics
- Some interest in“near astrophysics”, e.g. solar sytem
- DarkMatter + neutrino (CYGNO, DAMPE, Hyper Kamiokande



Several experiments, different windows, we need to extract all the information available

WP3 – le necessità tecniche

- Very preliminary assessment of use cases involved:
  - Event search and/or classification (Cherenkov, GW, DM, …)
    - Largely GPU based
    - A few requests for FPGA
    - ML techniques often proposed
  - Simulation/reduction pipelines for cosmological surveys
    - Strongly CPU/HPC based
    - Some I/O + bandwith requests •
- Request for CPU/GPU flops for existing codes (not necessarily connected to algorithm development)
- Code development:
  - Cross-analysis of datasets, joint analysis of heterogenous data
  - Machine learning development appears "almost everywhere"
  - Porting existing codes on GPU: wide spectrum of readiness
- Requests for improving data accessibility



Highlights dai WP – WP4

Qualche considerazione preliminare al lavoro che ci aspetta

- GPU
  - Ci sono già in corso diverse attività sia teoriche che sperimentali, al momento l'interesse maggiore e' lo sviluppo di framework platform agnostic
    - qui la maggiore esperienze viene dalla parte sperimentale, dove l'effort in questa direzione e' iniziato da tempo
    - serve rafforzare l'expertise italiana e fare formazione diffusa sul porting di algoritmi
    - al di la' del porting degli algoritmi c'e' poi da capire il fine tuning di diversi parametri per ottimizzare la memoria e l'esecuzione
      - ovviamente questo puo' essere molto project dependent
- FPGA
  - Area con interesse crescente negli ultimi tempi, e **molto varia per quanto riguarda progetti e anche tipologia di HW**
    - attività principalmente lato sperimentale con forte liason a tematiche legate a trigger selection e reconstruction
  - Probabilmente l'area dove c'e' meno expertise e quindi anche dove serve maggiormente la formazione di nuovo person-power
    - c'e' una conoscenza diffusa dei vari codici per la scrittura di firmware, manca forse la conoscenza per il setup di **cluster di FPGA** su media-grande scala (forse più indicato per WP5)
    - esistono vari testbed già funzionanti per alcuni progetti ma su scala ridotta, in altri casi si usano risorse messe a disposizione dai vari esperimenti (principalmente al CERN).
- Algoritmi ML
  - Largamente in uso nella nostra comunità, con sempre maggiore interesse per deployment su acceleratori vari
  - Per quanto riguarda il WP4, i punti di maggiore interesse potrebbero essere
    - ottimizzazione delle performance su GPU (eliminare i "colli di bottiglia" come il trasferimento dei dati o tuning di hyper-parameters
    - sviluppare know-how per il porting a FPGA (molti usano HLS4ML che e' in fase di completa riscrittura, ma sembrerebbe che nuovi tool siano in via di sviluppo)
  - per quanto riguarda la formazione (hackaton, workshop, repository di esempi e mini tutorial) significativa e' l'esperienza già accumulata nell'ambito del **progetto ML_INFN**: sarebbe opportuno approfittarne il piu' possibile evitando di duplicare lavoro già fatto.

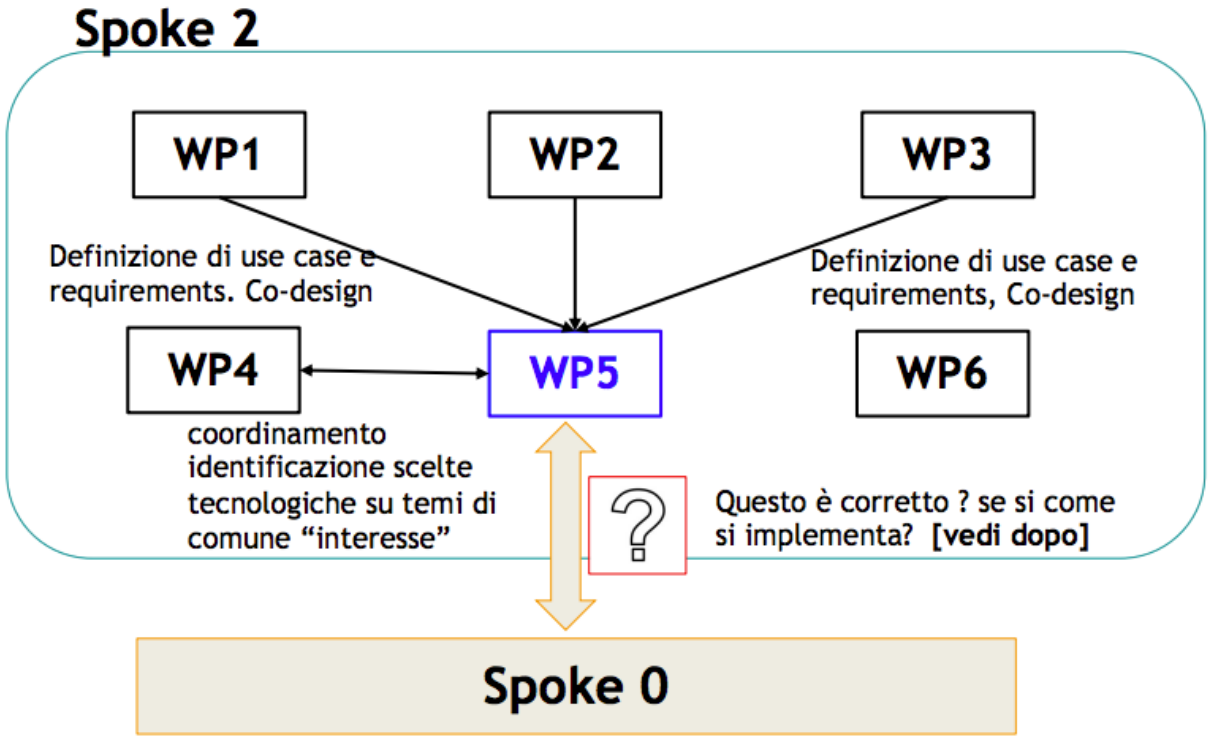
5

Esperienze in Exp HEP e Th HEP complementari; forte investimento della parte Exp in framework come **Alpaka** e **Kokkos**

Sistemi sia real-time sia offline; forte interesse nel deployment di reti ML su FPGA per inferenza

Realizzazione di piattaforme ML ottimizzate, basate su cloud per velocizzare R&D

Highlights dai WP – WP5



WP5: People

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

Coordinatori

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			

INAF ha manifestato interesse per temi comuni e probabilmente indicheranno un contact in WP5

WP5 - People

Alessandro De Salvo	Giovanni Ianniruberto
Andrea Bressan	Giuseppe Della Ricca
Andrea Contu	Giuseppe Milano
Antonio Stamerra	Guido Russo (PO)
Konstantinos Siettos	Guido Russo (RU)
Daniele Bonacorsi	Piergiulio Lenzi
Domenico Elia	Luca Tomasetti
Eleonora Luppi	Lucia Silvestris
Fabio Garufi	Mattia Bruno
Francesco Noferini	Michele Pavone
Gianluca Sabella	Nadia Rega
	Sandra Malvezzi

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

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 (?!?)**

Qualche idea preliminare sui testbed

Sulla base delle informazioni circolate finora e discusse ieri noi vediamo interessante partire con la costruzione di alcuni testbed che ad alto livello riassumiamo qui sotto:

- **Spostamento dati:** questo è un requirement discusso in modo preliminare. Automatizzare lo spostamento dati tra più endpoint (incluso HPC …) sarebbe un valore aggiunto:
  - abbiamo modelli di riferimento
  - abbiamo soluzioni da cui partire (non strettamente in ambito nazionale…)
  - vediamo questo come un esempio di una attività di interesse anche inter-spoke
- **Abilitazione dell'accesso a risorse HPC / HTC / Cloud**:
  - conosciamo soluzioni prototipali da cui evolvere
  - potrebbe essere utile per testare offloading di workflow in modo trasparente all'utente finale
  - potrebbe essere utile per abilitare analisi interattive su grosse moli di dati
  - potrebbe essere utile per fare esercizi con GPUs (interesse con WP4)
- **Soluzioni per la gestione delle pipeline (à la CI/CD):**
  - forte interazione con tecniche cloud-native
  - varie esperienze note intorno a noi
  - probabilmente di interesse trasversale (probabile interesse anche per WP4)

Prime espressioni di interesse:

- Attività sinergiche a temi di ML → UniBO
  - reti neurali parametriche per HEP, ML a servizio (training/inferenza) …
  - NOTA: questo è da discutere in dettaglio e sicuramente con WP2 (attività cross WP)

Inoltre abbiamo intercettato aspetti come

- Accesso risorse (calcolo e dati) su HPC (CINECA)
- pipeline ML, riproducibilità e accesso a GPU (FPGA)

Ovviamente molti spunti sono arrivati grazie ai report dei WP 1-2-3 (grazie dello sforzo fatto per le sintesi live)

- Il materiale che è stato raccolto ora va digerito e sistematizzato (i.e. identificare elementi comuni)

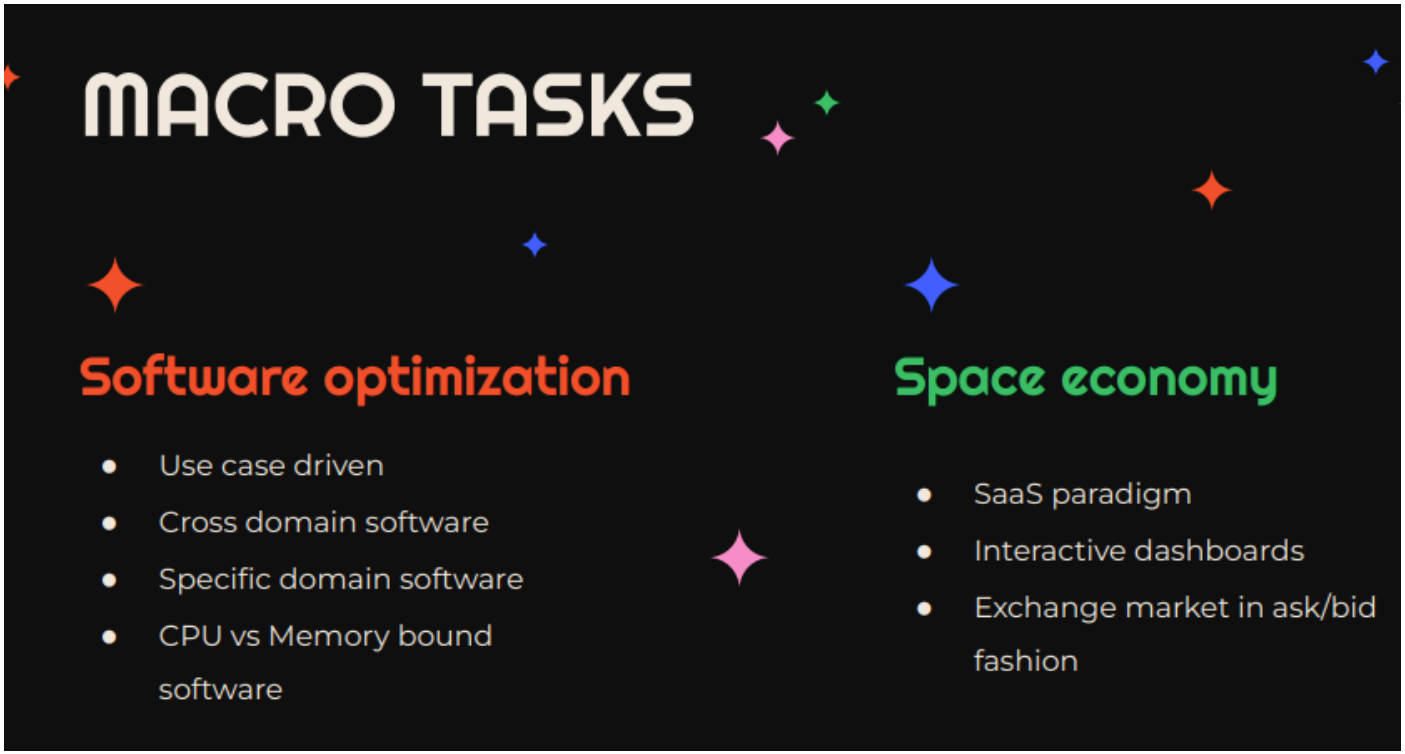
Alcune osservazioni "a caldo":

- **Pensare ai testbeds sembra una buona idea** → Dobbiamo definire le priorità in base agli interessi
- Oltre l'implementazione dobbiamo pensare al **supporto utenti/formazione** → Dobbiamo pensare come si implementa (effort)
- C'è una prima convergenza sul fatto che oltre le risorse HW il **dispiegamento dei servizi richiede un coordinamento cross spoke** → Da Discuter
- L'ipotesi di **effettuare un survey orientato agli aspetti "tecnologici"** ha risuonato durante la discussione. → Dobbiamo decidere se è uno strumento "eccessivo" o meno. Vogliamo/dobbiamo coordinarlo? i.e. alcune cose potrebbero essere di interesse comune con WP4
- **A quali quantità di risorse possiamo ambire…** → Non poniamoci limiti in fase di proposte/discussioni
- Abbiamo osservato una **ricorrente richiesta di FPGA** →qui vediamo un aspetto di interazione con

→ Survey lanciato, in comune con WP4!

Highlights dai WP – WP6

Approcci con le collaborazioni che guidano sw di ampio uso , come **Geant4** e **Corsika**, per proporre evoluzione di specifiche parti dei pacchetti (usando tecnologia «WP4», per esempio)



Interessi «interni» ai Membri CN, ma ci aspettiamo un forte impulso con la collaborazione con il tavolo italiano di Space Economy

Exec summary dai WPs

Ambito Scientifico ⟷ **Ambito Tecnologico**

- Siamo nella fase attiva di raccolta degli use case (*bottom-up*) e della loro sistematizzazione (*top-down*) in macro use cases, per quello che riguarda
  - Tematiche di ricerca
  - Necessità di calcolo per la «soluzione»
- Dall'altro lato, quello tecnologico:
  - Raccolta delle tecnologie «nel portfolio» dei partner (in una seconda fase del CN, mediante comunicazione fra gli spokes)
  - Analisi degli use cases e match con le tecnologie
  - Proposta di testbed tecnologici su cui testare le soluzioni

Una piccola digressione: stato del personale

- 195 staff hanno dato disponibilità a partecipare alle attività dello Spoke
- I 15 affiliati "accademici" hanno in parte cominciato le attivita' di recruitment
- Ad oggi, risultano banditi / in fase finale di bando **32** posizioni fra (R)TD e PhD
  - Ma non tutti hanno bandito / non tutti hanno bandito tutto
  - C'è genuina curiosità di vedere come andranno queste procedure, vista la "concorrenza" di tante altre, PNRR e non

Conclusioni

- Le attivita' di Spoke 2 si inseriscono in linee di ricerca **ben consolidate** e, se realizzate con successo, possono dare un **impulso reale** ai domini scientifici di riferimento
- I nostri domini di riferimento hanno ampia e variegata esperienza in soluzione tecnologiche, che devono però devono essere validate sulla scala del CN; questo non vuol dire che non siamo intenzionati a valutare soluzioni da altri domini!
- Le attività hanno un'importante **impronta tecnologica**, secondo le problematiche classiche dei big data e del loro processing
  - Tanti dati, tanto processing, dati eterogenei, dati distanti, necessità di *real* time
- … per cui siamo sicuri di poter instaurare proficue collaborazioni con le realtà industriali, condividendo soluzioni e idee