Spoke2 Annual meeting

Monday, 18 December 2023 - Wednesday, 20 December 2023

CINECA

Book of Abstracts

Contents

Introduzione / 3

Saluti Fondazione

Stato dello Spoke 2 / 4

Saluti CINECA

Introduzione / 5

Research manager

Stato dello Spoke 2 / 6

Stato dello Spoke 2 - part 2

Corresponding Author: tommaso.boccali@pi.infn.it

Stato dello Spoke 2 / 7

Innovation Manager

Stato dello Spoke 2 / 8

Industrial Board

Stato dello Spoke 2 / 9

Stato Spoke 2 - FO

Corresponding Author: simona.petronici@pi.infn.it

Working packages / 10

WP1

Corresponding Authors: leonardo.cosmai@ba.infn.it, leonardo.giusti@mib.infn.it

Working packages / 11

WP2

Corresponding Authors: lenzip@fi.infn.it, alberto.annovi@pi.infn.it

Working packages / 12

WP3

Corresponding Authors: marco.landoni@brera.inaf.it, ntlpla@unife.it

Working packages / 13

WP4

Corresponding Authors: alexis.pompili@ba.infn.it, simone.gennai@mib.infn.it

Innovation grant projects / 14

ENI - PIML

Giorno preferito:

Innovation grant projects / 15

ENI - Predictive

Innovation grant projects / 16

Intesa - Fraud Detection

Innovation grant projects / 17

ENI - Ferrera Erbognone

Innovation grant projects / 18

buffer

Working packages / 19

WP6

Corresponding Author: antonio.stamerra@inaf.it

Working packages / 20

WP5

Corresponding Authors: elly@na.infn.it, daniele.spiga@pg.infn.it

Innovation grant projects / 21

Leonardo - IDL

Innovation grant projects / 22

Intesa - Agri

Innovation grant projects / 23

SOGEI / UnipolSAI - Hammon

Innovation grant projects / 24

Buffer

Lighning Talks / 25

The Milano-Bicocca FPGA Cluster

Corresponding Authors: simone.gennai@mib.infn.it, paolo.dini@mib.infn.it, francesco.brivio@mib.infn.it

The Milano-Bicocca FPGA cluster represents an incredible opportunity to explore different use cases for the Physics program of the CMS experiment upgrade foreseen for the High Luminosity LHC. In this talk I will describe the FPGA Cluster that will be built in Milano-Bicocca in the context of the ICSC Spoke-2 project, and describe one of the many physics use case that may benefit from an FPGA distributed analysis. In particular, this use case is represented by the search for the rare decay of one W boson into 3 pions using the Level-1 trigger scouting data designed for the Phase-2 upgrade of the CMS experiment.

Lightning talks - flash talks / 26

Lightning talk 2

Lightning talks - flash talks / 27

Lightning talk 1

Lightning talks - flash talks / 28

Lightning talk 1

Lightning talks - flash talks / 29

Lightning talk 1

Lightning talks - flash talks / 30

Lightning talk 1

Lightning talks - flash talks / 31

Lightning talk 1

Lightning talks - flash talks / 32

Lightning talk 1

Lightning talks - flash talks / 33

Lightning talk 1

Lightning talks - flash talks / 34

Lightning talk 1

Lightning talks - flash talks / 35

Lightning talk 1

Lightning talks - flash talks / 36

Lightning talk 1

Lighning Talks / 37

Hunting for WIMP Minimal Dark Matter and New Charged Particles at a Future Muon Collider

Corresponding Author: natascia.vignaroli@le.infn.it

A multi-TeV muon collider proves to be very efficient not only for the search for new heavy neutral particles, but also for the discovery of charged bosons of the W′ type. We find that, by analyzing the associated production with a Standard Model W, charged resonances can be probed directly up to multi-TeV mass values close to the collision energy, and for very small couplings with the SM fermions, of the order of 10
-10
times the SM weak coupling. This would mark an unprecedented level of sensitivity for a direct search. Furthermore, the channel offers a very efficient and alternative way to probe the WIMP scenario for the very special and compelling case of Minimal Dark Matter (MDM) in the 5-plet EW representation, by allowing the direct detection of the charged component of the MDM bound state. The reach on the WIMP 5-plet thermal target is found to be much higher than those of mono-X, missing-mass and disappearing tracks signatures.

Lighning Talks / 38

CI pipeline triggering analysis execution on INFN Analysis Facility

Corresponding Author: matteo.bartolini@fi.infn.it

High energy physics research has always leveraged bleeding edge
computing solutions. Now, as experiments turn their focus to high precision measurements, data
need to be understood and analyzed better than ever; this requires, among other things, new coding
paradigms that enforce reproducibility and improve usability.
In this talk I will show how it is possible to enable continuous integration with the CMS dataset
by using the gitlab CI and harnessing the computing resources made available by the INFN CMS
Analaysis Facility. In particular, I show that it is possible to integrate the submission of jobs to
HTCondor into the gitlab CI thus facilitating the handling of big datasets.
In this way analysts will be able to quickly run different tests on their data, perform different analyses
and, at the same time, keep tracks of all the changes made.

Lightning talks - flash talks / 39

CMS Level-1 trigger Data Scouting for online trigger-less data processing

Corresponding Author: sabrina.giorgetti@studenti.unipd.it

The Level-1 trigger Data Scouting (L1DS) is a novel data acquisition system under development
for the Phase-2 CMS detector at the High-Luminosity LHC (HL-LHC). Its purpose is to capture
and process L1 trigger information at the bunch crossing frequency of the LHC preceding the L1
accept. It has the potential for filterless detector diagnostics, luminosity studies and investigations
into signatures and processes that would be otherwise inaccessible or constrained by the standard
triggers. At LHC Run-3, a 40MHz Scouting demonstrator is being tested as proof of concept and
deep learning algorithms on FPGAs are being implemented to perform close-to real-time analysis.
The current status of the L1DS demonstrator will be present, along with the first data-taking results
and the developments of machine learning applications.

Lightning talks - flash talks / 40

Developing and testing of a flexible and scalable high rate analysis platform

Corresponding Author: gianluca.sabella@na.infn.it

The talk will be focused to present the computing infrastructure under development as part of the
WP5 activities. The aim is to provide the users with an infrastructure that represents a tradeoff
between deployment speed-flexibility, resource efficiency and service performance, what we call
analysis facility.
In order to offer a general-purpose infrastructure, we leveraged container technology for running
applications and Kubernetes for orchestration. Within this framework, we tested a use case for the
analysis facility. This included the adoption of tools such as JupyterHub as an interactive web-based
development environment capable of managing multiple accesses; DASK as an open-source Python
library for parallel computing that can employ various batch systems (SLURM, HTC Condor, etc.)
or, if needed, be equipped with a native scheduler (our choice); ROOT as an object-oriented software
package; and optionally, S3 Object Storage for input/output handling
The near term goal of the activity is to automate the analysis facility deployable exploiting the ICSC
computing resources.

Lightning talks - flash talks / 41

Benchmark interactive analysis at future colliders

Corresponding Author: adelina.donofrio@na.infn.it

The challenges expected for the future colliders era are pushing to re think the HEP computing models at many levels.
A simple use case tested on the INFN analysis facility will be presented in the context of WP5, exploiting FCCee simulations.
The presented work will provide an overview of the main technologies involved and will describe the results of a first benchmark using IDEA detector concept.
One of the advantages of the above use case is the possibility to use it as a simple test for all the users willing to benefit from the WP5 infrastructure. Several metrics, from event throughput to resource consumption, will be shown to assess the reliability of the workflow using resources hosted at the INFN distributed analysis facility, in the framework of the thematic spoke "Fundamental Research and Space Economy"of the National Centre on HPC, Big Data and Quantum Computing (ICSC) project.

Lightning talks - flash talks / 42

Quasi interactive analysis of big data with high throughput - Initial steps and future perspectives

Corresponding Authors: francesco.giuseppe.gravili@le.infn.it, diotalevi@bo.infn.it

The ever-growing demand for fast processing of large datasets, as in the upcoming high-luminosity phases at the Large Hadron Collider (LHC), paves the way for innovative approaches. Leveraging the ICSC cloud DataLake model and integrating ongoing experiences in High Energy Physics (HEP) , a path towards an Analysis Facility (AF) is being forged. This new paradigm of data analysis moves from a batch-based to an interactive approach, based on a parallel and geographically distributed back-end.
This use case converges into a flagship activity (flagship UC2.2.2) of the Work Package 2 "Experimental High Energy Physics": firstly exploiting a testbed state-of-the-art prototype infrastructure developed by the CMS Collaboration, but foreseeing the adoption of a growing new infrastructure which integrates ICSC resources.

This lightning talk will report the national effort of porting different applications, ranging from quasi-interactive detector performance studies to physics data analysis from different scientific collaborations, offering a user-friendly interactive environment and adopting open-source industry standards. A survey of the ongoing activities will be provided, carried out mainly by newly hired Spoke2 personnel, and outlining the future steps of the flagship, in synergy with the architectural support defined in the Work Package 5.

Lightning talks - flash talks / 43

Facing the data-analysis challenge for the LISA mission

Corresponding Author: alice.spadaro97@gmail.com

Gravitational waves are perturbations of spacetime that propagate out through the Universe at the speed of light.
The Laser Interferometer Space Antenna (LISA) will be the first space-based observatory to survey the source-rich milliHertz band of the gravitational-wave spectrum.
LISA will revolutionize our understanding of the Universe, providing observations of astrophysical

sources ranging from Galactic white-dwarf binaries to mergers of massive black holes.

The signals from multiple gravitational-wave sources will be simultaneously present in the data, thus requiring a global data-analysis solution to efficiently isolate and characterize them.

In this talk I will review the challenges associated with analyzying the LISA data as well as the computational algorithms currently under development in order to maximize the scientific payoff of the LISA mission.

Lightning talks - flash talks / 44

Detecting vineyard diseases using high-resolution images acquired by airborne platforms

Authors: Alessia Tricomi[1]; Andrea Maino[2]; Enrico Chiarelli[2]; Fabio Mantovani[2]; Gioacchino Alex Anastasi[3]; Giuseppe Piparo[3]; Kassandra Giulia Cristina Raptis[2]; Matteo Albéri[2]; Michele Franceschi[2]

[1] *University of Catania, INFN Catania and Centro Siciliano di Fisica Nucleare e Struttura della Materia CSFNSM, Catania*

[2] *University of Ferrara & INFN Ferrara*

[3] *University of Catania & INFN Catania*

Corresponding Author: strati@fe.infn.it

Images acquired from aircrafts also integrate with satellite-based remote sensing allows for high-resolution data collection essential for ecosystems monitoring and risk management. This approach, combined with Artificial Intelligence (AI) algorithms serves as a reliable tool for the calibration and validation of satellite-derived data and ensures ground-truthing capabilities for more accurate data interpretation. In this talk, we present machine learning algorithms used to automatically analyze centimetric resolution images acquired by airborne experimental platforms for detecting vineyards diseases. The potential of the combined use of multispectral satellite imagery for symptom detection will also be discussed. In this framework, the use of high-performance computing resources is pivotal, accelerating image analysis for early symptoms detection and enabling early warning systems.

Giorno preferito:

Lighning Talks / 45

Large Scale Simulations of Complex Systems

Corresponding Author: giuseppe.negro@ba.infn.it

Complex systems refer to a class of interconnected entities or components whose collective behavior cannot be easily deduced from the properties of their individual parts. These systems are characterized by a high degree of interdependence, non-linear relationships, and emergent properties that arise from the interactions and feedback loops among the elements. Our primary focus will be on the dynamics of active fluids and active particles. While our research foundation is built on 2D systems, the expansion to 3D scenarios represents a pivotal leap. Despite the computational challenges posed by this transition, it allows us to scrutinize the emergence of self-propelled directed motion due to the coupling between induced flow and topological defects depending on specific anchoring. In this talk I'll give some updated on the development, tests, and validation we have performed on an optimized parallel Lattice Boltzmann (LB) solver specifically designed for simulating 3D multiphase active droplets.

Lighning Talks / 46

BoGEMMS-HPC: development of Geant4 simulations in High-Performance Computing environments

Corresponding Authors: simone.lotti@iaps.inaf.it, fioretti@iasfbo.inaf.it, alex.ciabattoni@inaf.it

The Geant4 toolkit is a widely used particle transport code for the simulation of high energy space missions, enabling the evaluation of their performance and driving the instrument design. The evolving landscape of modern large-scale simulations provides a new challenge in managing the production of increasingly large datasets, along with memory and computing requirements, that leads to the design of next-generation simulation frameworks in the realm of High-Performance Computing (HPC). We propose the development of an open-source multi-threading (MT) and multi-node Geant4-based simulation pipeline with ad-hoc I/O interfaces (e.g. run-time input configuration, output databases) based on the Bologna Geant4 Multi-Mission Simulator (BoGEMMS), an astronomy-oriented Geant4-based application developed at INAF OAS. The BoGEMMS-HPC framework will allow user-friendly, multi-purpose simulations in HPC environments. The current release uses the already built-in MT library of Geant4, which distributes events (where each event is a primary particle with its secondaries) on different threads. The multi-node parallelism is instead handled by the G4MPI library, the only plug-and-play Geant4 interface with MPI currently available. The node parallelism is also at the level of primary events. BoGEMMS-HPC supports two distinct output data formats: FITS files and SQLite databases. The test case for BoGEMMS-HPC is the simulation of the anticoincidence system of COSI (Compton Spectrometer and Imager), a NASA Small Explorer satellite mission with launch planned in 2027. We present the BoGEMMS-HPC architecture, the preliminary verification results, and the development plans for the near future.

Lighning Talks / 47

Evolving High Rate Analysis infrastructure with seamless offloading on different type of providers

Corresponding Author: tommaso.tedeschi@pg.infn.it

The current High Rate Analysis platform that is being implemented offers a general purpose environment where analyzers can scale up computations within the size of the instantiated cluster that possibly scale within the provider. However, in order to handle potentially a huge amount of users with diverse use cases, we plan to evolve the general purpose infrastructure toward the offloading model. As a result the platform will be enabled to dynamically (on-demand) exploit all kinds of resources (HTC, HPC, Cloud) with almost no effort and transparently for the user. The technical plan is to adopt and enhance the existing prototype such as interLink [1] that generalizes the Virtual Kubelet concept and use. This is how we foresee the efficient use of the ICSC computing resources. In this presentation, the idea, the main concepts and possible work plan will be presented.

Lighning Talks / 48

Declarative paradigms for analysis description and implementation

Corresponding Authors: andrea.rizzi@cern.ch, alberto.annovi@pi.infn.it, paolo.mastrandrea@cern.ch

The software toolbox used for "big data" analysis in the last few years is changing fast. The adoption of approaches able to exploit the new hardware architectures plays a pivotal role in boosting data

processing speed, resources optimisation, analysis portability and analysis preservation.
The big scientific collaborations (ATLAS, CMS, LHCb, Alice, …) are devoting increasing resources to the development and implementation of bleeding-edge software technologies, pushing the reach of the single experiments and the whole HEP community.

The introduction of declarative paradigms in the analysis description and implementation is growing interest and support in the main collaborations. This approach can simplify and speed-up the analysis description phase, support the portability of an analysis among different datasets/experiments and strengthen the preservation of the results.
Furthermore this approach, providing a stronger decoupling between the analysis algorithm and back-end implementation, is a key element for present and future processing speed.

In the frame of the "Quasi interactive analysis of big data with high throughput" use case of the ICSC-S2-WP2 activity is ongoing for the development of a framework characterized by a declarative paradigm for the analysis description and able to operate on datasets from different experiments.
Starting from the existing NAIL (Natural Analysis Implementation Language - https://indico.cern.ch/event/769263/con
Python package, developed in the context of the CMS data analysis and devoted to the event processing, the activity is focusing on the development of general and effective interface able to support the data format of different experiments and on the extension of the declarative approach to the full analysis chain.

Lighning Talks / 49

Exotic Quark Decay in the 331 Model: LHC Prospects and Computational Techniques

Corresponding Authors: dario.melle@le.infn.it, paul.h.frampton@gmail.com, claudio.coriano@le.infn.it, gennaro.corcella@lnf.infn.it

We discuss a computational approch to the analysis of the 331 Model in Frampton's version with
. The model generalises the
weak sector of the Standard Model to
and it predicts exactly three generations through the cancellation of gauge anomalies within an intergenerational framework. One additional feature of the model is the presence of Bileptons, which are gauge bosons of charge
, due to the specific embedding of the
symmetry into the generalised electroweak structure. We show how to build a computational interface that allows to automatize and expand the analysis of this complex model. In particular we discuss its implementation in the code SARAH, the interface with SPheno and SSP, and perform the study of the decay of exotic quarks present in the model into Bilepons and ordinary quarks at the LHC using MadGraph5.
The analysis is contained in a paper in preparation.

Wrap-up and next actions / 50

Wrap-up e come continuare

51

The Milano-Bicocca FPGA Cluster

Authors: Francesco Brivio[1]; Paolo Dini[1]; Simone Gennai[2]

[1] *Istituto Nazionale di Fisica Nucleare*

[2] *MIB*

Corresponding Authors: simone.gennai@mib.infn.it, francesco.brivio@mib.infn.it, paolo.dini@mib.infn.it

The Milano-Bicocca FPGA cluster represents an incredible opportunity to explore different use cases for the Physics program of the CMS experiment upgrade foreseen for the High Luminosity LHC. In this talk I will describe the FPGA Cluster that will be built in Milano-Bicocca in the context of the ICSC Spoke-2 project, and describe one of the many physics use case that may benefit from an FPGA distributed analysis. In particular, this use case is represented by the search for the rare decay of one W boson into 3 pions using the Level-1 trigger scouting data designed for the Phase-2 upgrade of the CMS experiment.

Giorno preferito:

52

Hunting for WIMP Minimal Dark Matter and New Charged Particles at a Future Muon Collider

Author: Natascia Vignaroli[1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Author: natascia.vignaroli@le.infn.it

A multi-TeV muon collider proves to be very efficient not only for the search for new heavy neutral particles, but also for the discovery of charged bosons of the W′ type. We find that, by analyzing the associated production with a Standard Model W, charged resonances can be probed directly up to multi-TeV mass values close to the collision energy, and for very small couplings with the SM fermions, of the order of 10^{-3}-10^{-4} times the SM weak coupling. This would mark an unprecedented level of sensitivity for a direct search. Furthermore, the channel offers a very efficient and alternative way to probe the WIMP scenario for the very special and compelling case of Minimal Dark Matter (MDM) in the 5-plet EW representation, by allowing the direct detection of the charged component of the MDM bound state. The reach on the WIMP 5-plet thermal target is found to be much higher than those of mono-X, missing-mass and disappearing tracks signatures.

Giorno preferito:

53

CI pipeline triggering analysis execution on INFN Analysis Facility

Author: Matteo Bartolini[1]

[1] *INFN, Sez. Firenze*

Corresponding Author: matteo.bartolini@fi.infn.it

High energy physics research has always leveraged bleeding edge
computing solutions. Now, as experiments turn their focus to high precision measurements, data need to be understood and analyzed better than ever; this requires, among other things, new coding paradigms that enforce reproducibility and improve usability.

In this talk I will show how it is possible to enable continuous integration with the CMS dataset by using the gitlab CI and harnessing the computing resources made available by the INFN CMS Analaysis Facility. In particular, I show that it is possible to integrate the submission of jobs to HTCondor into the gitlab CI thus facilitating the handling of big datasets.

In this way analysts will be able to quickly run different tests on their data, perform different analyses and, at the same time, keep tracks of all the changes made.

Giorno preferito:

19 Dicembre Pomeriggio

54

CMS Level-1 trigger Data Scouting for online trigger-less data processing

Author: Sabrina Giorgetti[1]

[1] *Università di Padova e INFN*

Corresponding Author: sabrina.giorgetti@studenti.unipd.it

The Level-1 trigger Data Scouting (L1DS) is a novel data acquisition system under development for the Phase-2 CMS detector at the High-Luminosity LHC (HL-LHC). Its purpose is to capture and process L1 trigger information at the bunch crossing frequency of the LHC preceding the L1 accept. It has the potential for filterless detector diagnostics, luminosity studies and investigations into signatures and processes that would be otherwise inaccessible or constrained by the standard triggers. At LHC Run-3, a 40MHz Scouting demonstrator is being tested as proof of concept and deep learning algorithms on FPGAs are being implemented to perform close-to real-time analysis. The current status of the L1DS demonstrator will be present, along with the first data-taking results and the developments of machine learning applications.

Giorno preferito:

19 Dicembre Pomeriggio

55

Large Scale Simulations of Complex Systems

Author: Giuseppe Negro[1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Author: giuseppe.negro@ba.infn.it

Complex systems refer to a class of interconnected entities or components whose collective behavior cannot be easily deduced from the properties of their individual parts. These systems are characterized by a high degree of interdependence, non-linear relationships, and emergent properties that arise from the interactions and feedback loops among the elements. Our primary focus will be on the dynamics of active fluids and active particles. While our research foundation is built on 2D systems, the expansion to 3D scenarios represents a pivotal leap. Despite the computational challenges posed by this transition, it allows us to scrutinize the emergence of self-propelled directed motion due to the coupling between induced flow and topological defects depending on specific anchoring. In this talk I'll give some updated on the development, tests, and validation we have performed on an

optimized parallel Lattice Boltzmann (LB) solver specifically designed for simulating 3D multiphase active droplets.

Giorno preferito:

20 Dicembre Mattina

56

BoGEMMS-HPC: development of Geant4 simulations in High-Performance Computing environments (Flagship UC 2.3.6)

Authors: Alex Ciabattoni[1]; Valentina Fioretti[2]; Simone Lotti[3]

[1] *Università di Bologna, INAF OAS Bologna*

[2] *INAF OAS Bologna*

[3] *INAF/IAPS*

Corresponding Authors: valentina.fioretti@inaf.it, alex.ciabattoni@inaf.it, simone.lotti@iaps.inaf.it

The Geant4 toolkit is a widely used particle transport code for the simulation of high energy space missions, enabling the evaluation of their performance and driving the instrument design. The evolving landscape of modern large-scale simulations provides a new challenge in managing the production of increasingly large datasets, along with memory and computing requirements, that leads to the design of next-generation simulation frameworks in the realm of High-Performance Computing (HPC). We propose the development of an open-source multi-threading (MT) and multi-node Geant4-based simulation pipeline with ad-hoc I/O interfaces (e.g. run-time input configuration, output databases) based on the Bologna Geant4 Multi-Mission Simulator (BoGEMMS), an astronomy-oriented Geant4-based application developed at INAF OAS. The BoGEMMS-HPC framework will allow user-friendly, multi-purpose simulations in HPC environments. The current release uses the already built-in MT library of Geant4, which distributes events (where each event is a primary particle with its secondaries) on different threads. The multi-node parallelism is instead handled by the G4MPI library, the only plug-and-play Geant4 interface with MPI currently available. The node parallelism is also at the level of primary events. BoGEMMS-HPC supports two distinct output data formats: FITS files and SQLite databases. The test case for BoGEMMS-HPC is the simulation of the anticoincidence system of COSI (Compton Spectrometer and Imager), a NASA Small Explorer satellite mission with launch planned in 2027. We present the BoGEMMS-HPC architecture, the preliminary verification results, and the development plans for the near future.

Giorno preferito:

57

Evolving High Rate Analysis infrastructure with seamless offloading on different type of providers

Author: Tommaso Tedeschi[1]

[1] *Università e INFN Perugia*

Corresponding Author: tommaso.tedeschi@pg.infn.it

The current High Rate Analysis platform that is being implemented offers a general purpose environment where analyzers can scale up computations within the size of the instantiated cluster that possibly scale within the provider. However, in order to handle potentially a huge amount of users

with diverse use cases, we plan to evolve the general purpose infrastructure toward the offloading model. As a result the platform will be enabled to dynamically (on-demand) exploit all kinds of resources (HTC, HPC, Cloud) with almost no effort and transparently for the user. The technical plan is to adopt and enhance the existing prototype such as interLink [1] that generalizes the Virtual Kubelet concept and use. This is how we foresee the efficient use of the ICSC computing resources. In this presentation, the idea, the main concepts and possible work plan will be presented.

[1] https://github.com/interTwin-eu/interLink/tree/main

Giorno preferito:

58

Developing and testing of a flexible and scalable high rate analysis platform

Author: Gianluca Sabella[1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Author: gianluca.sabella@na.infn.it

The talk will be focused to present the computing infrastructure under development as part of the WP5 activities. The aim is to provide the users with an infrastructure that represents a tradeoff between deployment speed-flexibility, resource efficiency and service performance, what we call analysis facility.
In order to offer a general-purpose infrastructure, we leveraged container technology for running applications and Kubernetes for orchestration. Within this framework, we tested a use case for the analysis facility. This included the adoption of tools such as JupyterHub as an interactive web-based development environment capable of managing multiple accesses; DASK as an open-source Python library for parallel computing that can employ various batch systems (SLURM, HTC Condor, etc.) or, if needed, be equipped with a native scheduler (our choice); ROOT as an object-oriented software package; and optionally, S3 Object Storage for input/output handling
The near term goal of the activity is to automate the analysis facility deployable exploiting the ICSC computing resources.

Giorno preferito:

59

Benchmark interactive analysis at future colliders

Author: Adelina D'Onofrio[1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Author: adelina.donofrio@na.infn.it

The challenges expected for the future colliders era are pushing to re think the HEP computing models at many levels.
A simple use case tested on the INFN analysis facility will be presented in the context of WP5, exploiting FCCee simulations.
The presented work will provide an overview of the main technologies involved and will describe the results of a first benchmark using IDEA detector concept.
One of the advantages of the above use case is the possibility to use it as a simple test for all the users willing to benefit from the WP5 infrastructure. Several metrics, from event throughput to resource consumption, will be shown to assess the reliability of the workflow using resources hosted at the

INFN distributed analysis facility, in the framework of the thematic spoke "Fundamental Research and Space Economy"of the National Centre on HPC, Big Data and Quantum Computing (ICSC) project.

Giorno preferito:

19 Dicembre Pomeriggio

60

Quasi interactive analysis of big data with high throughput - Initial steps and future perspectives

Author: Tommaso Diotalevi[1]

Co-author: Francesco Giuseppe Gravili [2]

[1] *Università e INFN, Bologna*

[2] *Istituto Nazionale di Fisica Nucleare*

Corresponding Authors: diotalevi@bo.infn.it, francesco.giuseppe.gravili@le.infn.it

The ever-growing demand for fast processing of large datasets, as in the upcoming high-luminosity phases at the Large Hadron Collider (LHC), paves the way for innovative approaches. Leveraging the ICSC cloud DataLake model and integrating ongoing experiences in High Energy Physics (HEP) , a path towards an Analysis Facility (AF) is being forged. This new paradigm of data analysis moves from a batch-based to an interactive approach, based on a parallel and geographically distributed back-end.
This use case converges into a flagship activity (flagship UC2.2.2) of the Work Package 2 "Experimental High Energy Physics": firstly exploiting a testbed state-of-the-art prototype infrastructure developed by the CMS Collaboration, but foreseeing the adoption of a growing new infrastructure which integrates ICSC resources.

This lightning talk will report the national effort of porting different applications, ranging from quasi-interactive detector performance studies to physics data analysis from different scientific collaborations, offering a user-friendly interactive environment and adopting open-source industry standards. A survey of the ongoing activities will be provided, carried out mainly by newly hired Spoke2 personnel, and outlining the future steps of the flagship, in synergy with the architectural support defined in the Work Package 5.

Giorno preferito:

19 Dicembre Pomeriggio

61

Facing the data-analysis challenge for the LISA mission

Author: Alice Spadaro[1]

[1] *Università di MIlano-Bicocca*

Corresponding Author: alice.spadaro97@gmail.com

Gravitational waves are perturbations of spacetime that propagate out through the Universe at the speed of light.

The Laser Interferometer Space Antenna (LISA) will be the first space-based observatory to survey the source-rich milliHertz band of the gravitational-wave spectrum.

LISA will revolutionize our understanding of the Universe, providing observations of astrophysical sources ranging from Galactic white-dwarf binaries to mergers of massive black holes.

The signals from multiple gravitational-wave sources will be simultaneously present in the data, thus requiring a global data-analysis solution to efficiently isolate and characterize them.

In this talk I will review the challenges associated with analyzying the LISA data as well as the computational algorithms currently under development in order to maximize the scientific payoff of the LISA mission.

Giorno preferito:

19 Dicembre Pomeriggio

62

Detecting vineyard diseases using high-resolution images acquired by airborne platforms

Authors: Virginia Strati[1]; Matteo Albéri[2]; Enrico Chiarelli[2]; Michele Franceschi[2]; Andrea Maino[2]; Fabio Mantovani[2]; Kassandra Giulia Cristina Raptis[2]; Giuseppe Piparo[3]; Alessia Tricomi[4]; Gioacchino Alex Anastasi[3]

[1] *Istituto Nazionale di Fisica Nucleare*

[2] *University of Ferrara & INFN Ferrara*

[3] *University of Catania & INFN Catania*

[4] *University of Catania, INFN Catania and Centro Siciliano di Fisica Nucleare e Struttura della Materia CSFNSM, Catania*

Corresponding Author: strati@fe.infn.it

Images acquired from aircrafts also integrate with satellite-based remote sensing allows for high-resolution data collection essential for ecosystems monitoring and risk management. This approach, combined with Artificial Intelligence (AI) algorithms serves as a reliable tool for the calibration and validation of satellite-derived data and ensures ground-truthing capabilities for more accurate data interpretation. In this talk, we present machine learning algorithms used to automatically analyze centimetric resolution images acquired by airborne experimental platforms for detecting vineyards diseases. The potential of the combined use of multispectral satellite imagery for symptom detection will also be discussed. In this framework, the use of high-performance computing resources is pivotal, accelerating image analysis for early symptoms detection and enabling early warning systems.

Giorno preferito:

19 Dicembre Pomeriggio

63

Advanced Calculus for Precision Physics

Authors: Jonathan Ronca[1]; Manoj Kumar Mandal[2]; Pierpaolo Mastrolia[2]; William J. Torres Bobadilla[3]

[1] *Roma Tre university*

[2] *Istituto Nazionale di Fisica Nucleare*

[3] *Max-Planck-Institute for Physics*

Corresponding Authors: manojkumar.mandal@pd.infn.it, pierpaolo.mastrolia@pd.infn.it, ronca@roma3.infn.it, torres@mpp.mpg.de

The Advanced Calculus for Precision Physics (ACPP) use case is an initiative dedicated to enhance the computation of scattering amplitudes and cross-sections to support the phenomenology analyses for prospects of detection and observations of new physics events in advanced collider physics programs (such as CERN and Fermilab) and gravitational wave detectors (including LIGO-VIRGO-KAGRA, ET, and LISA). In this presentation, I will highlight our ongoing research and development efforts focused on creating software and employing techniques to improve the efficiency of evaluating multi-loop scattering amplitudes. This work aims to contribute significantly to deciphering the physics encoded in data collected by collider and gravitational wave detectors.

Giorno preferito:

20 Dicembre Mattina

64

Inference of cosmological and astrophysical population properties from gravitational wave observations with and without electromagnetic counterparts

Author: Matteo Tagliazucchi[1]

[1] *University of Bologna*

Corresponding Author: matteo.tagliazucchi2@unibo.it

Gravitational waves (GWs) from compact binary coalescences can be used as a new and independent cosmological probe if external binary redshift information is injected into the inference process. Methods for incorporating redshift information range from direct detection of electromagnetic counterparts ("bright sirens") to statistical inference of binary redshift using a catalog of possible hosts or spectral features in the source-frame mass distribution of the GW population when GW events are not followed by an EM event ("dark sirens").
In this talk I will present two pipelines, CHIMERA and icarogw2.0, which combine all these methods within a hierarchical Bayesian framework to fully exploit multi-messenger information and constrain both cosmological and GW population parameters. I will then describe their general workflow, the main computational bottlenecks and the activities we are carrying out for the flagship use case UC2.3.3 within the Spoke 2 - WP3 to improve the performance of the hierarchical inference, as well as the mock data challenge between the two codes.

Giorno preferito:

65

Deterministic Detection of Photovoltaic Panels in Aerial Images

Author: Emiliano Alessio Tramontana[None]

Co-authors: Daniele Marletta ; Alessandro Midolo

Corresponding Author: tramontana@dmi.unict.it

Automatically detecting the area of photovoltaic panels in images gives the possibility to forecast and plan the green energy production in a community. Most existing approaches for panel detection resort to machine learning to analyse images and find the photovoltaic panels. However, each geographical area is likely to have its own sorrounding/background colours and panel colours, as the former depend on the latitude, and the latter on the materials used. Then, a specific set of annotated images and a further training phase are needed, therefore increasing the amount of time (and cost) needed for training. We propose a deterministic approach that extracts a range of significant colours for photovoltaic panels in images. This consists of the most frequent panel colours different from the surrounding parts, for the given conditions of light exposure in annotated images. Then, by reckoning pixel density and comparable levels of light, the colours in other images are compared with the extracted range to reveal panels. The approach is fast and results produced were highly accurate, independently of the shape of panels.

Giorno preferito:

19 Dicembre Pomeriggio

66

Declarative paradigms for analysis description and implementation

Authors: Alberto Annovi[1]; Andrea Rizzi[2]; Paolo Mastrandrea[1]; Tommaso Boccali[1]

[1] *Istituto Nazionale di Fisica Nucleare*

[2] *INFN Pisa*

Corresponding Authors: andrea.rizzi@cern.ch, tommaso.boccali@pi.infn.it, alberto.annovi@pi.infn.it, paolo.mastrandrea@cern.ch

The software toolbox used for "big data" analysis in the last few years is changing fast. The adoption of approaches able to exploit the new hardware architectures plays a pivotal role in boosting data processing speed, resources optimisation, analysis portability and analysis preservation.
The big scientific collaborations (ATLAS, CMS, LHCb, Alice, …) are devoting increasing resources to the development and implementation of bleeding-edge software technologies, pushing the reach of the single experiments and the whole HEP community.

The introduction of declarative paradigms in the analysis description and implementation is growing interest and support in the main collaborations. This approach can simplify and speed-up the analysis description phase, support the portability of an analysis among different datasets/experiments and strengthen the preservation of the results.
Furthermore this approach, providing a stronger decoupling between the analysis algorithm and back-end implementation, is a key element for present and future processing speed.

In the frame of the "Quasi interactive analysis of big data with high throughput" use case of the ICSC-S2-WP2 activity is ongoing for the development of a framework characterized by a declarative paradigm for the analysis description and able to operate on datasets from different experiments.
Starting from the existing NAIL (Natural Analysis Implementation Language - https://indico.cern.ch/event/769263/con Python package, developed in the context of the CMS data analysis and devoted to the event processing, the activity is focusing on the development of general and effective interface able to support the data format of different experiments and on the extension of the declarative approach to the full analysis chain.

Giorno preferito:

67

Exotic Quark Decay in the 331 Model: LHC Prospects and Computational Techniques

Authors: Claudio Coriano'[1]; Dario Melle[2]; Gennaro Corcella[3]; Paul Frampton[4]

[1] *LE*

[2] *INFN Lecce*

[3] *Istituto Nazionale di Fisica Nucleare*

[4] *University of Salento*

Corresponding Authors: gennaro.corcella@lnf.infn.it, dario.melle@le.infn.it, claudio.coriano@le.infn.it, paul.h.frampton@gmail.com

We discuss a computational approch to the analysis of the 331 Model in Frampton's version with $\beta = \sqrt{3}$. The model generalises the $SU(2)_L \times U(1)_Y$ weak sector of the Standard Model to $SU(3)_L \times U(1)_X$ and it predicts exactly three generations through the cancellation of gauge anomalies within an inter-generational framework. One additional feature of the model is the presence of Bileptons, which are gauge bosons of charge ± 2, due to the specific embedding of the $U(1)_{em}$ symmetry into the generalised electroweak structure. We show how to build a computational interface that allows to automatize and expand the analysis of this complex model. In particular we discuss its implementation in the code SARAH, the interface with SPheno and SSP, and perform the study of the decay of exotic quarks present in the model into Bilepons and ordinary quarks at the LHC using MadGraph5.
The analysis is contained in a paper in preparation.

Giorno preferito:

68

Simulating lattice QCD at high temperatures

Author: Pietro Rescigno[1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Author: p.rescigno1@campus.unimib.it

Numerical simulations of lattice Quantum ChromoDynamics offer a non perturbative approach from first principles to compute the properties of the theory of strong interactions. The design of efficient algorithms and the increasing computing power of latest and future generation HPC systems allow to push simulations to more interesting (and challenging) regimes. Within the context of the National Center for High Performance Computing, I will present some work-in-progress results from my Ph.D. project, which concerns the simulation of QCD in the very high temperature regime.

Giorno preferito:

69

Algorithm optimization to improve continuous gravitational-wave searches

Author: Lorenzo Pierini[1]

Co-authors: Pia Astone [1]; Marco Serra [1]; Stefano Dal Pra [1]; Cristiano Palomba [1]

[1] *Istituto Nazionale di Fisica Nucleare*

Corresponding Authors: stefano.dalpra@cnaf.infn.it, pia.astone@roma1.infn.it, marco.serra@roma1.infn.it, cristiano.palomba@roma1.infn.it, lorenzo.pierini@roma1.infn.it

The recent observation of gravitational waves from merging binary systems of compact astrophysical objects has opened a new window to explore the Universe. A strong effort is still ongoing to detect signals from different sources, like rotating isolated neutron stars, which are expected to produce continuous, persistent gravitational waves. In this talk, I will show that those searches are typically computationally bounded and that the optimization of the algorithm can lead to a direct improvement of the search sensitivity. In this optic, I will present the work we are doing as a flagship use case of the Spoke 2 WorkPackage 3.

Giorno preferito:

19 Dicembre Pomeriggio

Lightning talks - flash talks / 70

Inference of cosmological and astrophysical population properties from gravitational wave observations with and without electromagnetic counterparts

Corresponding Author: matteo.tagliazucchi2@unibo.it

Gravitational waves (GWs) from compact binary coalescences can be used as a new and independent cosmological probe if external binary redshift information is injected into the inference process. Methods for incorporating redshift information range from direct detection of electromagnetic counterparts ("bright sirens") to statistical inference of binary redshift using a catalog of possible hosts or spectral features in the source-frame mass distribution of the GW population when GW events are not followed by an EM event ("dark sirens").
In this talk I will present two pipelines, CHIMERA and icarogw2.0, which combine all these methods within a hierarchical Bayesian framework to fully exploit multi-messenger information and constrain both cosmological and GW population parameters. I will then describe their general workflow, the main computational bottlenecks and the activities we are carrying out for the flagship use case UC2.3.3 within the Spoke 2 - WP3 to improve the performance of the hierarchical inference, as well as the mock data challenge between the two codes.

Lighning Talks / 71

Advanced Calculus for Precision Physics

Corresponding Authors: torres@mpp.mpg.de, ronca@roma3.infn.it, pierpaolo.mastrolia@pd.infn.it, manojkumar.mandal@pd.infn.it

The Advanced Calculus for Precision Physics (ACPP) use case is an initiative dedicated to enhance the computation of scattering amplitudes and cross-sections to support the phenomenology analyses for prospects of detection and observations of new physics events in advanced collider physics programs (such as CERN and Fermilab) and gravitational wave detectors (including LIGO-VIRGO-KAGRA, ET, and LISA). In this presentation, I will highlight our ongoing research and development efforts focused on creating software and employing techniques to improve the efficiency of evaluating multi-loop scattering amplitudes. This work aims to contribute significantly to deciphering the physics encoded in data collected by collider and gravitational wave detectors.

Lightning talks - flash talks / 72

Deterministic Detection of Photovoltaic Panels in Aerial Images

Corresponding Author: tramontana@dmi.unict.it

Automatically detecting the area of photovoltaic panels in images gives the possibility to forecast and plan the green energy production in a community. Most existing approaches for panel detection resort to machine learning to analyse images and find the photovoltaic panels. However, each geographical area is likely to have its own sorrounding/background colours and panel colours, as the former depend on the latitude, and the latter on the materials used. Then, a specific set of annotated images and a further training phase are needed, therefore increasing the amount of time (and cost) needed for training. We propose a deterministic approach that extracts a range of significant colours for photovoltaic panels in images. This consists of the most frequent panel colours different from the surrounding parts, for the given conditions of light exposure in annotated images. Then, by reckoning pixel density and comparable levels of light, the colours in other images are compared with the extracted range to reveal panels. The approach is fast and results produced were highly accurate, independently of the shape of panels.

Lightning talks - flash talks / 73

Algorithm optimization to improve continuous gravitational-wave searches

Corresponding Authors: cristiano.palomba@roma1.infn.it, stefano.dalpra@cnaf.infn.it, marco.serra@roma1.infn.it, pia.astone@roma1.infn.it, lorenzo.pierini@roma1.infn.it

The recent observation of gravitational waves from merging binary systems of compact astrophysical objects has opened a new window to explore the Universe. A strong effort is still ongoing to detect signals from different sources, like rotating isolated neutron stars, which are expected to produce continuous, persistent gravitational waves. In this talk, I will show that those searches are typically computationally bounded and that the optimization of the algorithm can lead to a direct improvement of the search sensitivity. In this optic, I will present the work we are doing as a flagship use case of the Spoke 2 WorkPackage 3.

Lightning talks - flash talks / 74

Simulating lattice QCD at high temperatures

Corresponding Author: p.rescigno1@campus.unimib.it

Numerical simulations of lattice Quantum ChromoDynamics offer a non perturbative approach from first principles to compute the properties of the theory of strong interactions. The design of efficient algorithms and the increasing computing power of latest and future generation HPC systems allow to push simulations to more interesting (and challenging) regimes. Within the context of the National Center for High Performance Computing, I will present some work-in-progress results from my Ph.D. project, which concerns the simulation of QCD in the very high temperature regime.

Lighning Talks / 76

Boosting unmodeled searches of gravitational wave transients

Corresponding Author: giacomo.principe@inaf.it

A strong effort is ongoing to develop and improve unmodeled methods for detecting generic GW signals, including those for which we still miss precise modelling, such as the burst events produced by supernovae, magnetar flares, fast radio bursts..
In this context, coherent WaveBurst (cWB) is currently the most efficient and utilised burst pipeline in the LVK Collaboration. cWB is based on a wavelet decomposition of the GW strain signals from the detectors in the worldwide GW interferometer network, and on a global likelihood for the source localisation.
Because of its non-stationarity and non-Gaussianity, the background noise must be empirically estimated by repeating any successful search many times while effectively "switching off"the GW source, making the process computational burden.
In this talk, I will present the current cWB pipeline and the plan for the optimisation of its algorithm.

Introduzione / 77

Stato Spoke 2 - parte 1

Corresponding Author: tommaso.boccali@pi.infn.it

78

Cena presso il Pizzikotto

indirizzo: https://maps.app.goo.gl/kGTFPhov724kjdoy5

Lighning Talks / 79

Machine Learning Algorithms for Multi-Messenger Astroparticle Physics

Corresponding Author: nicola.calabria@na.infn.it

Machine Learning algorithms bring a new opportunity of investigation and analysis of phenomena in the context of Astroparticle Physics and multi-messenger Astrophysics. We focus here on Water Cherenkov detectors, such as Super-Kamiokande and Hyper-Kamiokande, which offer a low noise environment ideal for the study of neutrinos from astrophysical sources, but also for the detection of rare events, like proton decay, for which an accurate knowledge of neutrino interaction events is critical since they play the part of an unavoidable background.
On the Super-Kamiokande side we want to introduce Machine Learning based reconstruction algorithms, along with the existing reconstruction algorithm based on Likelihood Maximization (fiTQun), to improve reconstruction performance, such as particle counting and particle identification, with focus on proton decay analysis as a benchmark. On the Hyper-Kamiokande side we want to study the improvement of reconstrution processing time, motivated by the increase of the number of photosensors.