NVIDIA AI TECHNOLOGY CENTER

Webinar for CINI labs, November 26th 2020

SPEAKERS

Frédéric Parienté
Engineering Mgr
NVAITC EMEA Director

Giuseppe Fiameni
Solutions Architect
NVAITC Italy Lead

Marco Rorro
AI consultant
CINECA

Massimiliano Guarrasi
HPC consultant
CINECA

David Prelogović
PhD student

Matteo Tomei
PhD student

AGENDA

NVAITC Explained (20 min)

Giuseppe Fiameni, Frédéric Pariente (NVIDIA)

CINECA resources and how to access them (20 min)

Massimiliano Guarrasi, Marco Rorro (CINECA)

Parameter estimation from 3D cosmic 21cm images with neural networks (10 min)

David Prelogović (Scuola Normale Superiore)

Distributed Deep Networks Training for Video Recognition (10 min)

Matteo Tomei (University of Modena and Reggio Emilia)

Q/A (30 min)


NVIDIA
NVAITC EXPLAINED

ITALY

Italy launched (March 2018) a national Artificial Intelligence program to facilitate the adoption of AI technologies in the Italian Public Administration to improve services to citizens and businesses. It entails nine main challenges:

- **Ethical**: Artificial Intelligence is always at the service of the citizen and not vice versa
- **Technological**: make existing and future technologies more similar to our way of relating to the world
- **Data**: must be of good quality, exempt as much as possible from biases produced by humans
- **Legal**: find out a well balance between the interests of the community and those of the individual
- **Public Sector**: Artificial Intelligence should support people and help them carry out their activities, without replacing them
- **Prevent Inequalities**: must be accessible to everyone and it must be of simple and immediate use
- **Measurement of the impact**: make the organizational processes more efficient and effective
- **Human being**: citizens and Institutions should be aware of the significant importance of these tools

https://libro-bianco-ia.readthedocs.io/en/latest/index.html

COLLABORATION FRAMEWORK

The Italian centre builds on the collaboration of three entities

consorzio interuniversitario nazionale per l'informatica

- Consortium of 41 universities in the fields of Computer Engineering, Computer science, and Information technologies. More that 1300 professors involved.
- Includes 11 labs spanning from Artificial Intelligence, Cyber Security, BigData to Digital Health.
- The AI Lab represents the consortium within the NVAITC.

CINECA

- Not-for-profit consortium, made up of the Italian Ministry of Education Universities and Research, 69 Italian universities, 11 Italian National Institutions.
- One of the largest HPC facilities in Europe and hosting entity of the Leonardo EuroHPC system.

Italy Forges AI Future in Partnership with NVIDIA

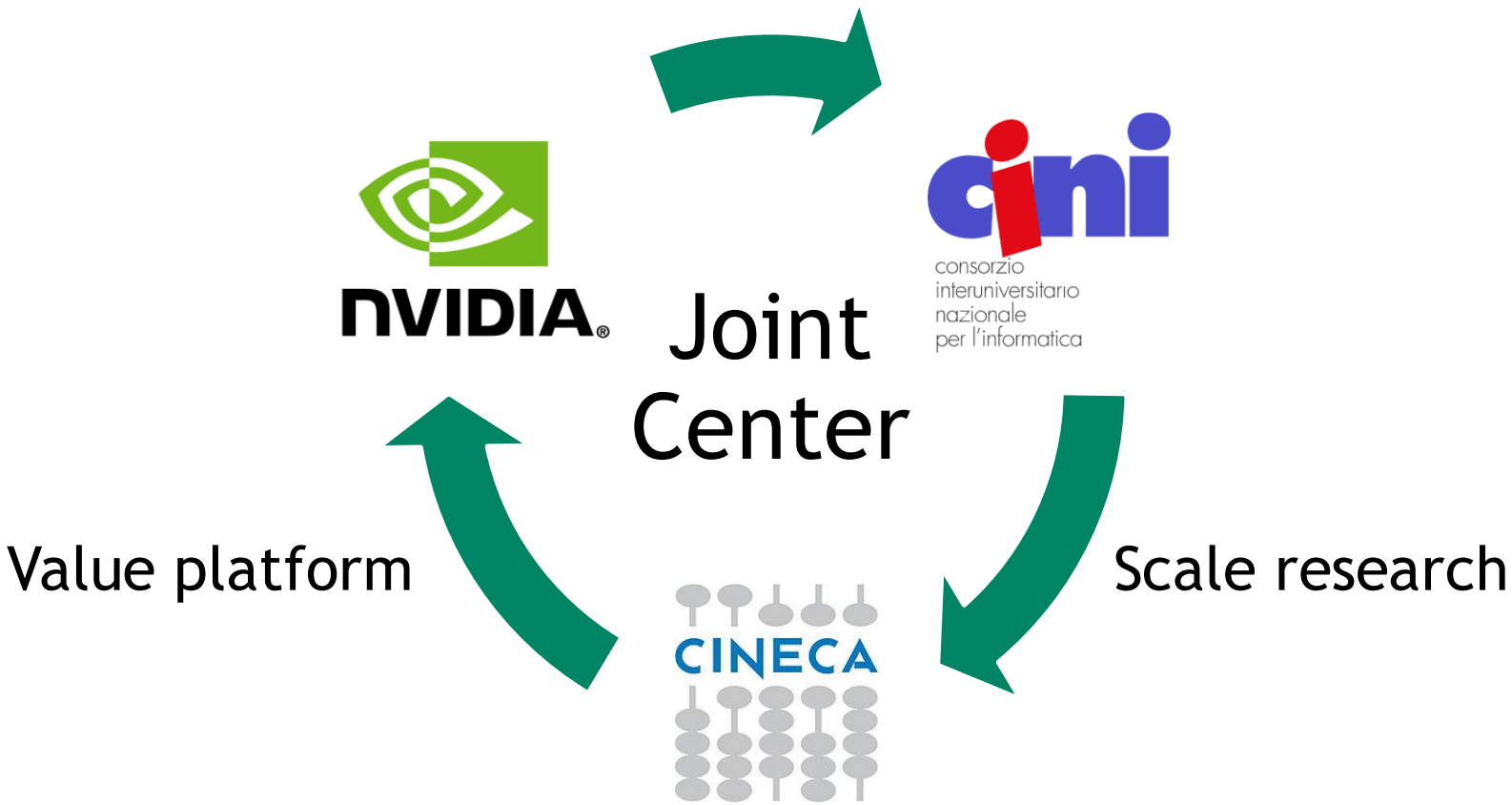
Collaboration begins with a research hub at AImageLab in Modena.

January 15, 2020 by FRÉDÉRIC PARIENTÉ

NVIDIA

MOTIVATION & VISION

Enable researcher



Value platform

Scale research

NVIDIA AI TECHNOLOGY CENTER (NVAITC)

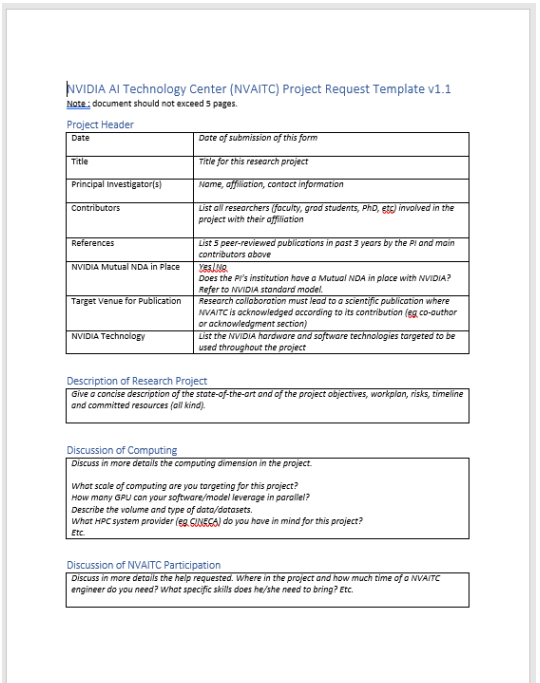
Catalyse AI transformation through research-centric integrated engagements

RULES OF ENGAGEMENTS

Research Intensity	5ish projects with noteworthy contribution (last authors) 1 "anchor" project with key contribution (first authors) 5ish projects with point support (acknowledgement)
Project selection	Based on NVAITC criteria (technology stack, publication, computing scale), an agreed-upon statement of work and approval by manning institution.
Compute Infrastructure	PI independently secures compute resources for the project. The center provides support to access CINECA.
Human Resources	NVIDIA scientists self-funded, 1 local + global team Italy may grow team to support more projects, NVIDIA can train/intern Invite collaborators for relevant industry opportunities
Funding	NVIDIA does not fund research under NVAITC, nor seeks funding. Boost research proposals with NVIDIA letters, NVAITC personnel, etc.
All-NVIDIA	Not a one-stop shop for all-NVIDIA enquiries Will do best effort to route requests to appropriate people at NVIDIA Inc. sales, start-up program, education, NVResearch, social media, etc

PROPOSAL FOR PROJECT SUBMISSION TO NVAITC

- PI submits proposal using template [1] to segreteria.aiis@consorzio-cini.it
- The local lead engineer (Giuseppe) is available for any input, suggestion, advise before submitting. PI is typically contacted for clarification and SoW settlement after submission.
- Giuseppe reviews proposal with help of fellow NVAITC engineers on a first-come first-serve basis. Review takes a couple of weeks.
- Evaluation is based on NVAITC criteria (target publication, technology stack and computing scale), rules of engagement (compact timeline, no compute, no funding, etc) and shared & realistic expectations (an agreed-upon SoW).
- This "call for proposals" remains open as long as we have the capacity to handle projects (process will be re-evaluated when victim of our own success :)).
- Access to computational resources is handled separately by CINECA via the ISCRA/PRACE programs (https://iscra.cineca.it/, http://prace-ri.eu).



NVIDIA AI Technology Center (NVAITC) Project Request Template v1.1

Note : document should not exceed 5 pages.

Project Header

Date	*Date of submission of this form*
Title	*Title for this research project*
Principal Investigator(s)	*Name, affiliation, contact information*
Contributors	*List all researchers (faculty, grad students, PhD, etc) involved in the project with their affiliation*
References	*List 5 peer-reviewed publications in past 3 years by the PI and main contributors above*
NVIDIA Mutual NDA in Place	*Yes|No* *Does the PI's institution have a Mutual NDA in place with NVIDIA? Refer to NVIDIA standard model.*
Target Venue for Publication	*Research collaboration must lead to a scientific publication where NVAITC is acknowledged according to its contribution (eg co-author or acknowledgment section)*
NVIDIA Technology	*List the NVIDIA hardware and software technologies targeted to be used throughout the project*

Description of Research Project

Give a concise description of the state-of-the-art and of the project objectives, workplan, risks, timeline and committed resources (all kind).

Discussion of Computing

Discuss in more details the computing dimension in the project.

What scale of computing are you targeting for this project?
How many GPU can your software/model leverage in parallel?
Describe the volume and type of data/datasets.
What HPC system provider (eg CINECA) do you have in mind for this project?
Etc.

Discussion of NVAITC Participation

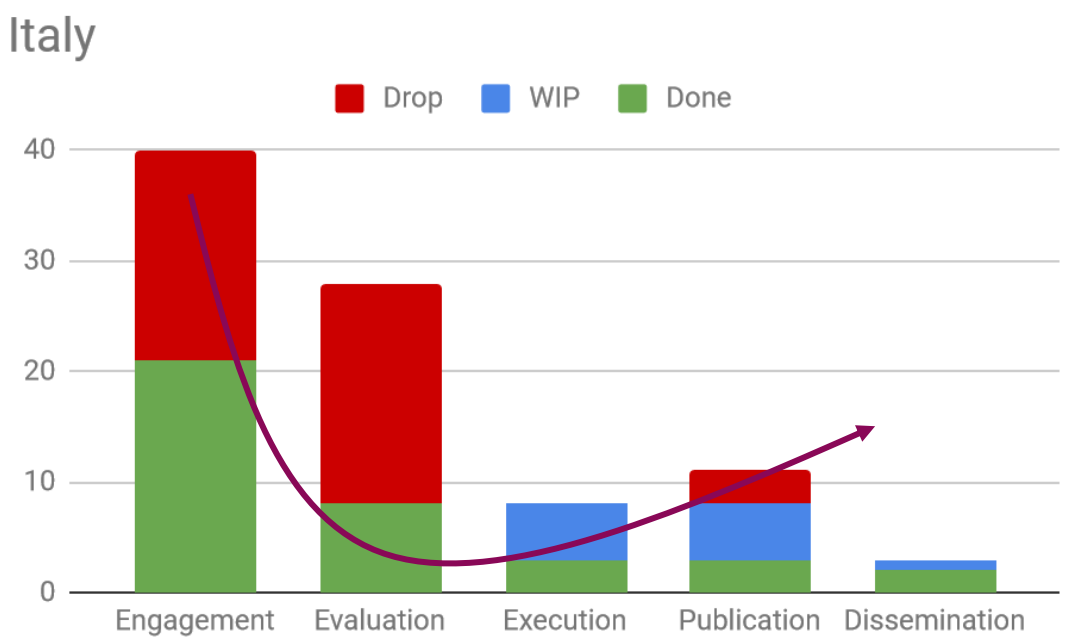
Discuss in more details the help requested. Where in the project and how much time of a NVAITC engineer do you need? What specific skills does he/she need to bring? Etc.

[1] https://www.consorzio-cini.it/index.php/it/labaiis-home/labaiis-bandi/1624-programma-nvidia-ai-nation-template-proposta-nvaitc

ITALY CENTER

Some statistics

	Nov 2020	June 2020
Awareness	455 (+210)	245
Training	490 (258)	232
PI engaged	39 (+5)	34
Institutions	21 (+2)	19
Project submitted	28 (+6)	22
Project accepted	8 (+5)	3
Project completed	3 (+3)	0
Publication	3 (+3)	0
Dissemination	2 (+2)	0



NVAITC PROJECTS

ITALY PROJECTS

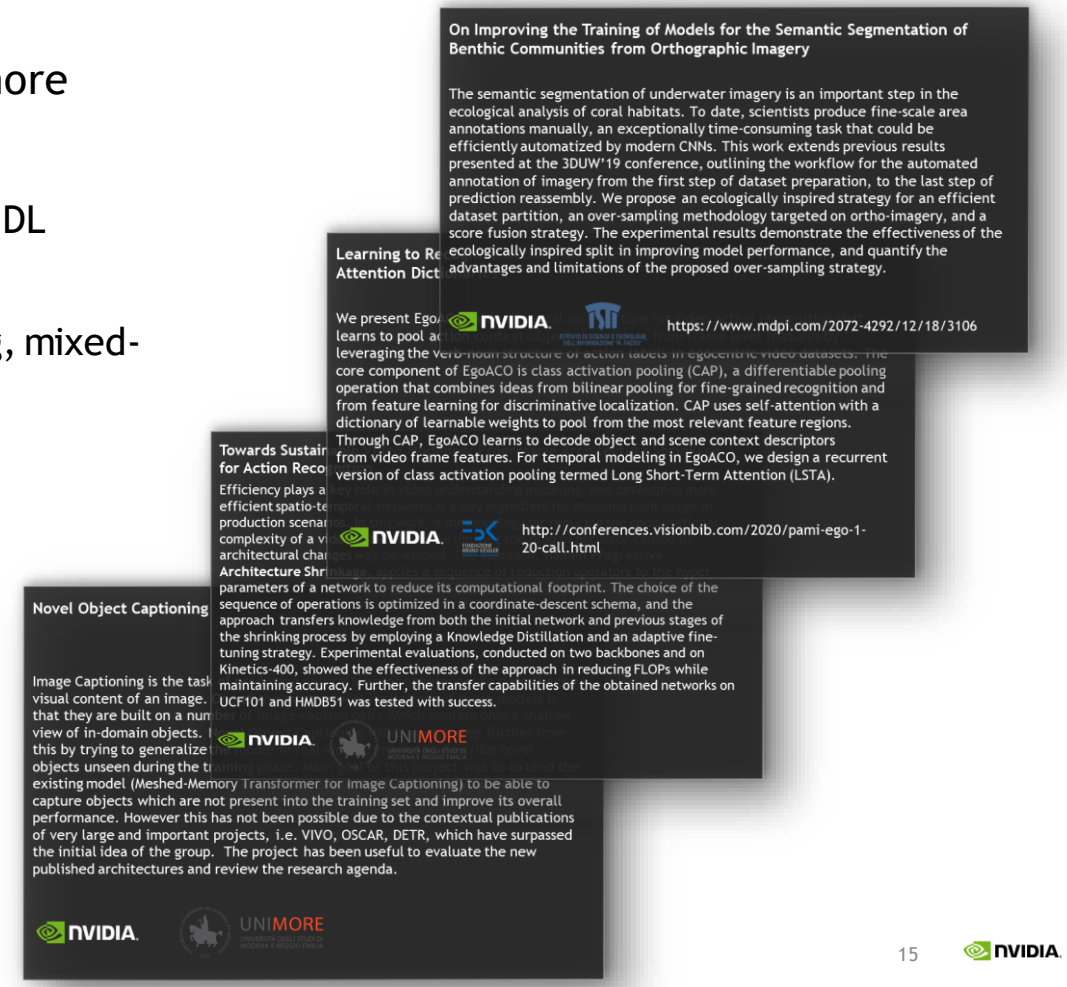
Title	Institution	
Inference of galaxy properties from radio images	Scuola Normale Superiore	CINECA
Graph-based learning of action recognition	University of Modena and Reggio Emilia	CINECA
Knowledge distillation applied to video classification	University of Modena and Reggio Emilia	
Grapevine Segmentation and Interpretation for Winter Pruning	Istituto Italiano di Tecnologia	
Novel Object Captioning (NOC)	University of Modena and Reggio Emilia	CINECA
Optimization of GSM and EgoACO models for video classification and representation learning	Fondazione Bruno Kesler	CINECA
Multi-view stereo semantic segmentation of 3D coral reefs	CNR-ISTI	
Point Cloud Domain Adaption via Forecasting	University of Trento	CINECA

CINECA = got access to CINECA resources through levering our collaboration.

MAIN AREAS OF CONTRIBUTION

What we do in practice

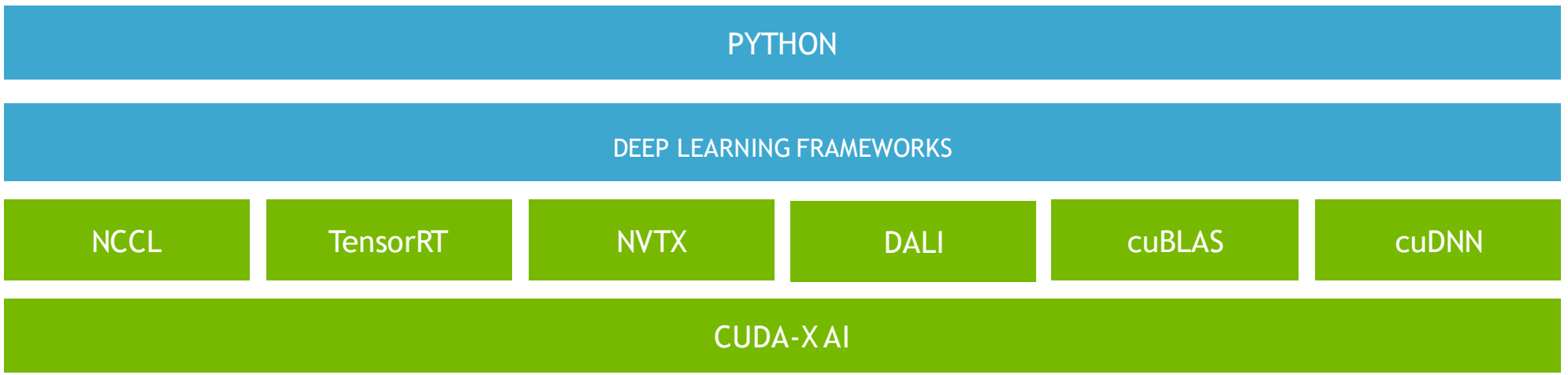
- Enable academics at all levels to do their own research more efficiently
  - **Adoption of DL/ML frameworks** (NVIDIA heavily contributes to DL frameworks development)
  - **Technology selection and optimization (**efficient data loading, mixed-precision, inference)
  - **Model architectural choices**
  - **Software development**
  - **Performance optimization and tuning through profiling**
  - **Workload scaling on multiple GPUs/nodes**
  - **Discussion on research studies**
  - **Training**
  - **Support to access HPC resources**



NVAITC TOOLKIT

Educational Code Base

- Code base to showcase interoperability of CUDA-X AI software stack in multi-GPU environments
- Provide researchers a reference framework to build new projects on
- Available on official NVIDIA github organization
  - **http://github.com/nvidia/nvaitc-toolkit**



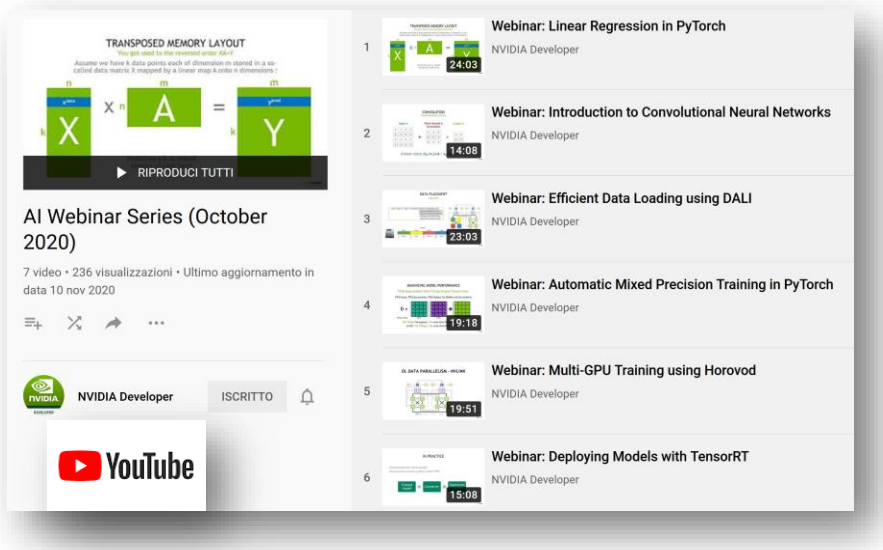
NVAITC WEBINAR SERIES

Now available online

AI Webinar Series on Deep Learning for CINI AIIS Labs - June 29th/July 3rd 2020

The goal of this webinar series is to explore the fundamentals of deep learning by building and training neural networks, optimizing data loading and performance through mixed-precision and parallelization, and deploying your trained model in production for inference. You will learn how to design, train, optimize, profile and deploy a deep neural network using NVIDIA technologies. Each session is split.

Date	Topic and slides	
Session 0	Linear Regression in Pytorch - Christian Hundt	Video
Session 1	Convolutional Neural Networks (from slide 45) - Christian Hundt	Video
Session 2	Efficient Data Loading using DALI - Giuseppe Fiameni	Video
Session 3	Mixed Precision Training using Apex - Paul Graham	Video
Session 4	Multi-GPU Training using Horovod - Gunter Roeth	Video
Session 5	Deploying Models with TensorRT - Niki Loppi	Video
Session 6	Profiling with NVTX - Giuseppe Fiameni	Video



https://www.consorzio-cini.it/index.php/en/labaiis-home/labaiis-nvaitc

HOW TO APPROACH THE CENTER

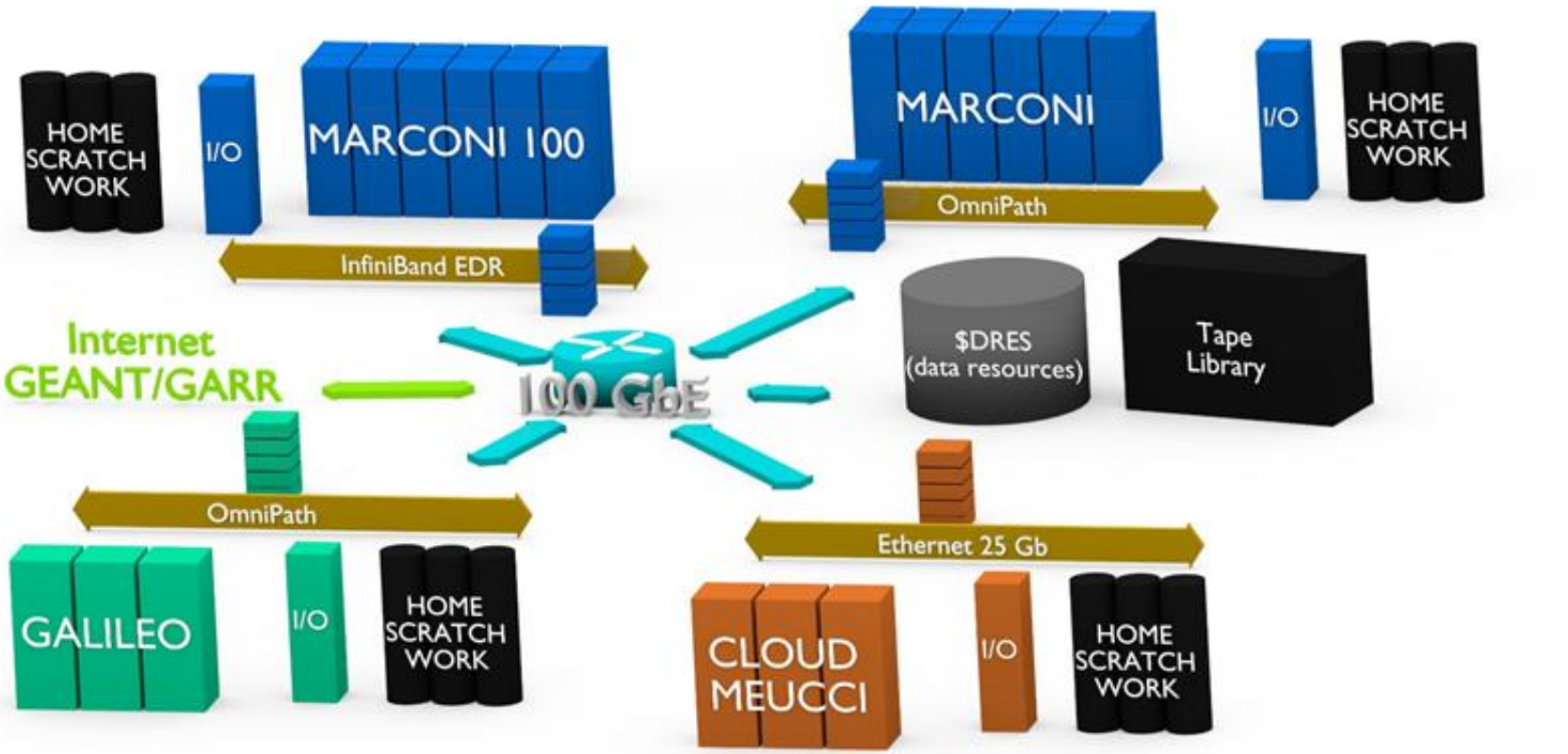
In few steps

1. Access the online AI Webinars Series and try the NVAITC toolkit
    - *The webinars include code snippets taken from the toolkit*
2. Define a project proposal in collaboration with the local lead engineer (Giuseppe)
3. Submit a project proposal whenever it is fine with you.
    - *Continuous open call, no deadlines*
4. Secure computational resources leveraging internal capacity or interacting with CINECA.
    - *The next CINECA call (ISCRA B) opens on Dec the 1st. This is good period of the year to consider submitting a proposal to NVAITC too*

CINECA RESOURCES AND HOW TO ACCESS THEM

Marco Rorro m.rorro@cineca.it - Massimiliano Guarrasi m.guarrasi@cineca.it

CINECA Infrastructure



CINECA Roadmap

2016/2017	2020	2020	2021/22
Marconi	Marconi100	DGX	Leonardo
Intel Skylake and KNL architecture (3188 + 3600 nodes) 8 PF + 6,3 PF	IBM Power9 (980 nodes) and NVIDIA Volta GPUs (4 per node) 32 PF	DGX-A100 (3 nodes) and NVIDIA Ampere GPUs (8 per node) 0.5 PF	Atos XH2000 (3500 nodes) and NVIDIA Ampere next-gen GPUs (4 per node) >250 PF

Marconi100

- Model: IBM Power AC922 (Whiterspoon)
- Racks: 55 total (49 compute)
- Nodes: 980
- Processors: 2x16 cores IBM POWER9 AC922 at 2.6(3.1) GHz
- Accelerators: 4 x NVIDIA Volta V100 GPUs/node, Nvlink 2.0, 16GB
- Cores: 32 cores/node, Hyperthreading x4
- RAM: 256 GB/node (242 usable)
- Local disk: 1.6TB NVMe
- Internal Network: Mellanox IB EDR DragonFly++
- Disk Space: 8PB raw GPFS storage

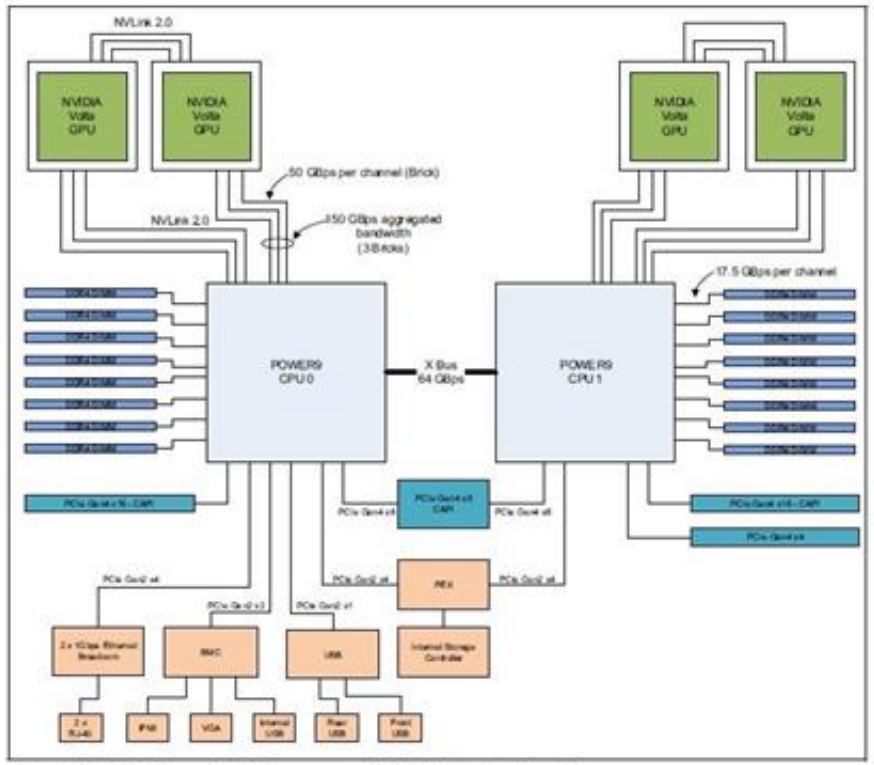
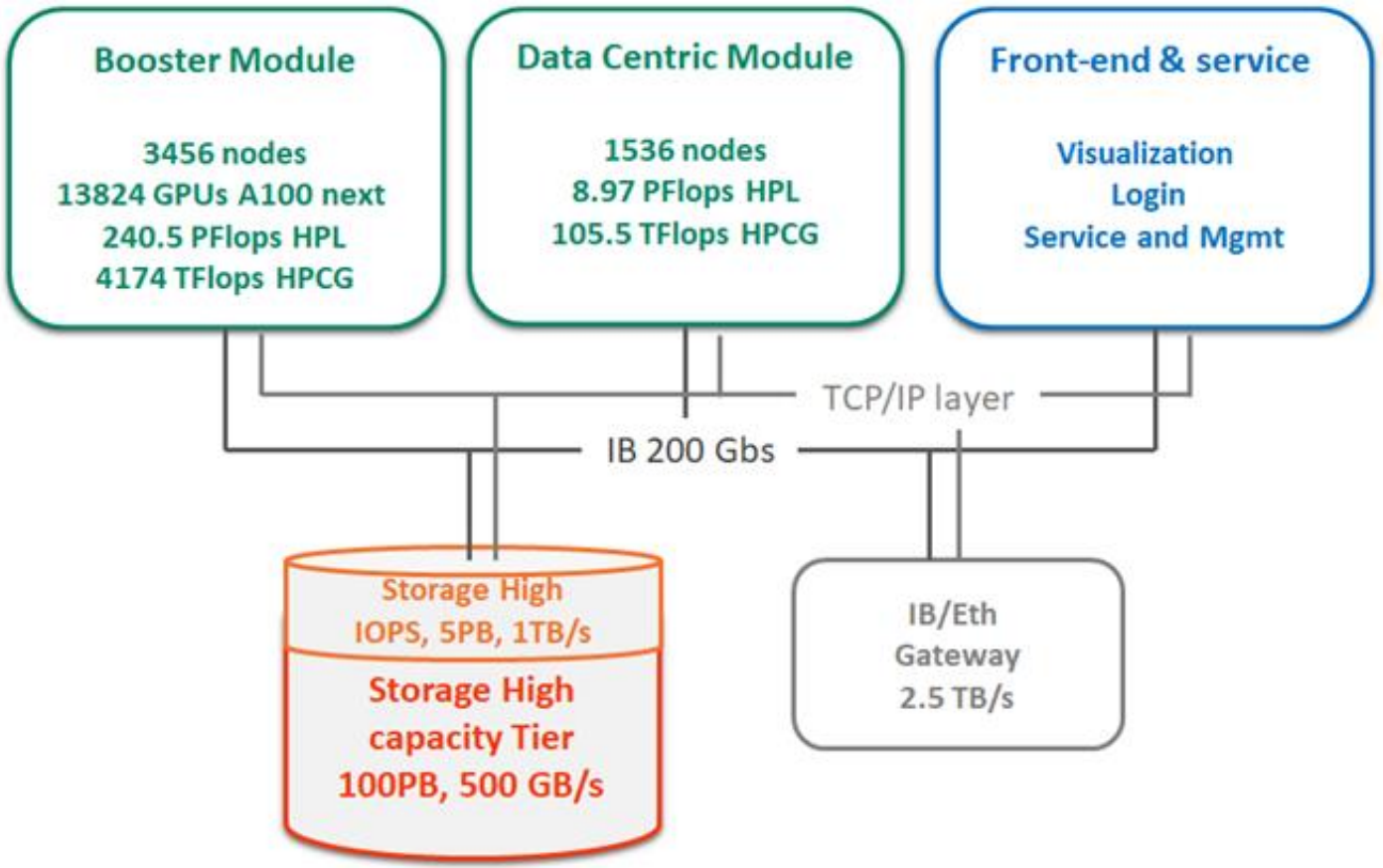


Figure 2-5 The Power AC922 server model GTH logical system diagram

DGX

- Model: Nvidia DGX-A100
- Architecture: Linux Cluster
- Nodes: 3
- Processors: 2 x 64-cores AMD EPYC (ROME) processors 3,4 GHz
  - Cores: 128 cores/node
- RAM: 1TB + 15 TB NVME Gen4 per node
- Accelerators:
- 8 GPU Nvidia A100 (40 GB HBM2 DRAM memory per GPU) per node
- 9 Mellanox ConnectX-6 VPI HDR InfiniBand
- 200 Gb Ethernet
- Peak performance single node:
  - 5 petaFLOPS AI
  - 10 petaOPS INT8
  - 156 teraFLOPS FP32

Leonardo



- Based on Atos XH2000 platform technology
- NVIDIA Mellanox HDR 200Gb/s InfiniBand
- 4 x NVIDIA Ampere next-gen GPUs and an Intel CPU per node
- 14000 NVIDIA Ampere GPUs
- 100 PB DDN Storage
- 10 exaflops of FP16 AI performance

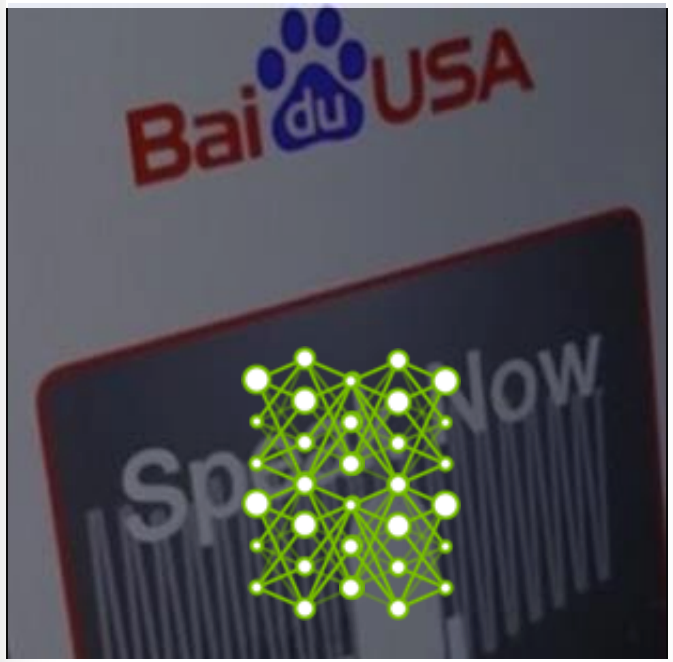
NEURAL NETWORK COMPLEXITY IS EXPLODING

To Tackle Increasingly Complex Challenges

7 ExaFLOPS
60 Million Parameters

2015 – Microsoft ResNet
Superhuman Image Recognition

20 ExaFLOPS
300 Million Parameters

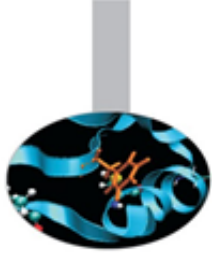


2016 – Baidu Deep Speech 2
Superhuman Voice Recognition

100 ExaFLOPS
8700 Million Parameters

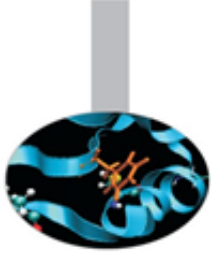
2017 – Google Neural Machine Translation
Near Human Language Translation

CINECA & AI



- Exascale systems are being designed for HPC, AI, HPDA
- HPC is at the forefront of R&D for important use cases
  - Precision medicine
  - Automated driving systems
  - Fraud and anomaly detection
  - Affinity marketing
  - Business intelligence
  - Cyber security
  - IoT/Smart cities

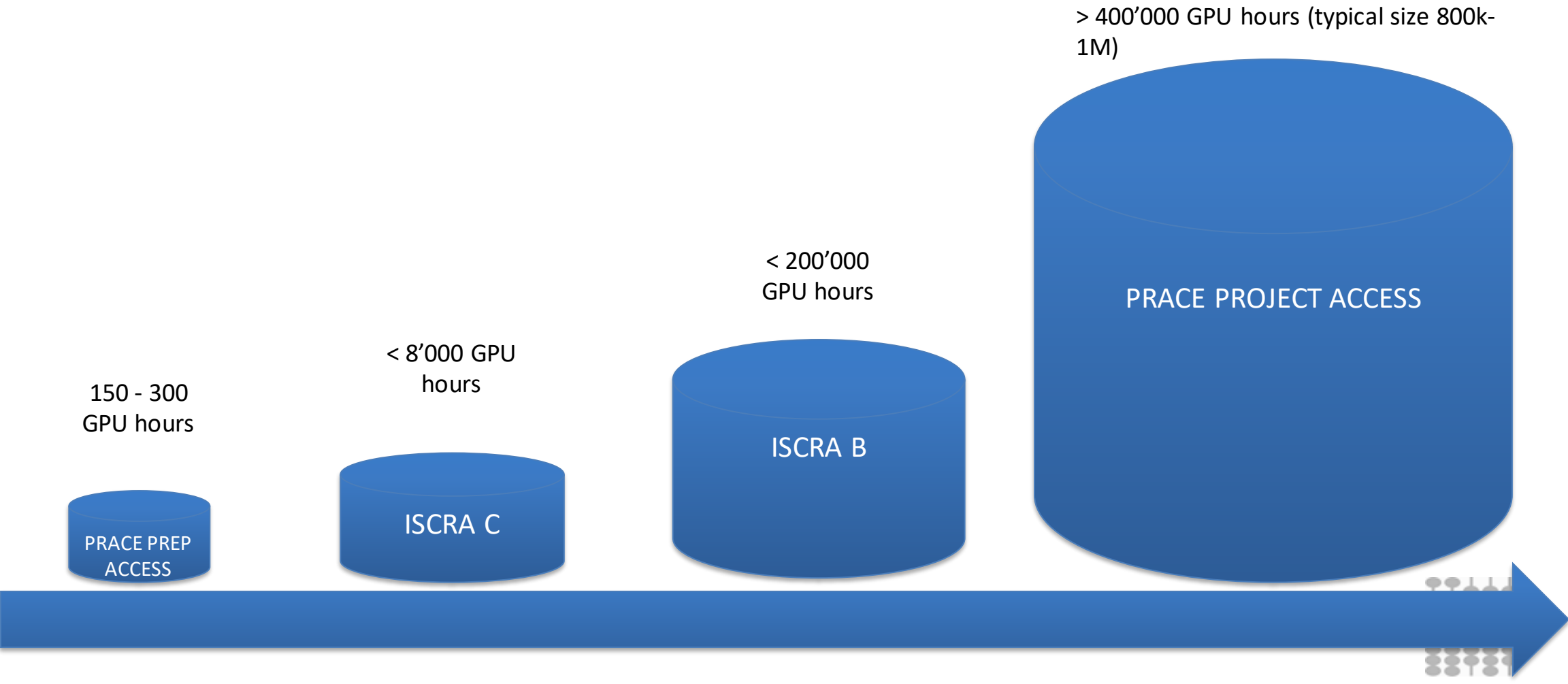
HPC access programmes



Italy

Europe

Projects dimension

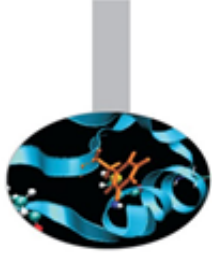


HPC offer in Italy

Italy

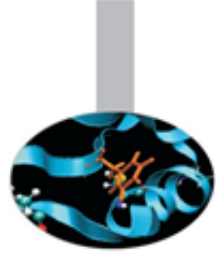
Submission and informations are available at:
https://www.hpc.cineca.it/services/iscra

National Resources - ISCRA



- The aim of ISCRA is to ensure adequate support to Italian scientists and engineers
- The allocation is per year (data for 2020):
  - of ~3M NODE hours (12M GPU Hours) on MARCONI100
  - of ~1.5M NODE hours (54M CORE hours) on Galileo
  - of ~26K NODE hours (210k GPU hours, i.e. 3.3M CORE hours) on DGX
- The access is by
  - Online submission of proposals
  - Applications must be submitted in English
- The proposal are scientifically evaluated by international reviewers and technically evaluated by Cineca experts
- Applications and codes are evaluated on the basis of their computational readiness

National Resources - ISCRA



Two types of call (B and C) available for accessing Marconi:

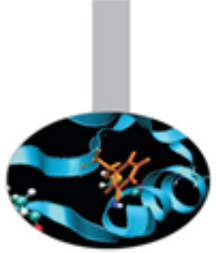
Class B: Standard Projects; two calls / year

- up to **50'000 NODE hours** on the **MARCONI100** machine
- No resources will be available on the Galileo machine for the next calls, but if necessary an allocation on this machine could be considered anyway.
- Duration: 12 months
- The Call is now open, next deadline: 15 January 2021 and 31 June 2021

Class C: Small Projects; continuous submission, 10 selections per year (only technical evaluation)

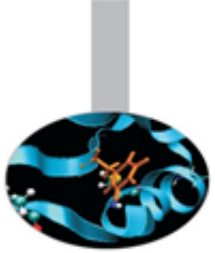
- On **MARCONI100**: up to **8.000 GPU hours** (2000 NODE hours)
- On **Galileo**: up to 5.000 NODE hours (CPU only)
- On **DGX**: no strict limitations are imposed. We assume:
  - Up to **3'000 GPU hours** (~ 390 NODE hours, i.e ~50'000 core hours) for a Development & Benchmark project
  - Up to **12'500 GPU hours** (~ 1'560 NODE hours, i.e 200'000 core hours) for a Production proposal
- Duration: 9 months
- **Trial**: on demand -> superc@cineca.it

CINECA and NVAITC



- Inside the new ISCRA C calls one of the admitted special focus is devoted to AI & ML applications. A special attention will be given to this class of proposal. This means:
  - Dedicated resources (DGX)
  - Faster evaluation (if needed)
  - Less strict requirements
- Huge amount of resources for AI projects will be available thanks to ISCRA B calls. We will give support in preparing the applications

HPC offer in Europe



CINECA

Europe

PRACE

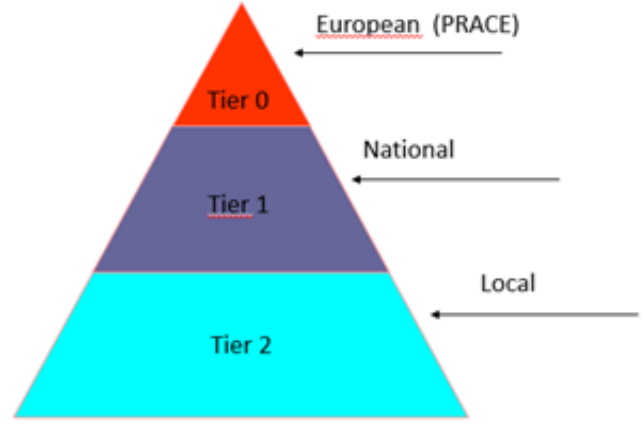
HPC Europa 3

FENIX
RESEARCH INFRASTRUCTURE

PRACE

- The mission of PRACE is to enable high impact scientific discovery and engineering research and development across all disciplines to enhance European competitiveness for the benefit of society
- PRACE provides world class computing and data management resources and services through a peer review process
- PRACE is established as an international non-profit association (of 25 countries) with its seat in Brussel
- Five Hosting Members (France, Germany, Italy, Spain, Switzerland) provide Multi-PFlop/s Tier-0 Systems

http://www.prace-ri.eu/



PRACE offers access to Tier 0 (and Tier 1) systems

CINECA

PRACE calls

- **Preparatory Access**
  - Intended for preliminary resource use required to prepare proposals for Project Access
  - Simplified form to access limited resources
  - Technical review
- **Project Access**
  - Intended for individual researchers and research groups including multi-national research groups
  - Technical and Scientific review
- **Shape Access**
  - to help SMEs benefit from the expertise and knowledge developed within the PRACE RI.
  - Intended for SME with the potential of using HPC

Joliot Curie, (GENCI@CEA France)

Marconi100 (CINECA, Italy)

Mare Nostrum (BSC, Spain)

HAWK (GCS@HLRS, Germany)

Piz Daint, (ETH Zurich/CSCS, Switzerland)

Juwels + Juwels Booster (GCS@FZJ, Germany)

SuperMUC NG (GCS@LRZ, Germany)

Conclusions

- We can help you to
  - exploit architectures
  - ask resources and use them
  - optimize and scale applications

- Useful links
  - http://www.hpc.cineca.it/services/iscra
  - https://prace-ri.eu/hpc-access/
  - http://www.hpc-europa.eu/
  - https://fenix-ri.eu/
- Useful contacts
  - superc@cineca.it
  - iscra@cineca.it
  - prace-tech@cineca.it
  - peer-review@prace-ri.eu
  - staff@hpc-europa.org

Parameter estimation from 3D cosmic 21cm images with neural networks

David Prelogović

PhD @ SNS Pisa

Supervisor: prof. Andrei Albert Mesinger

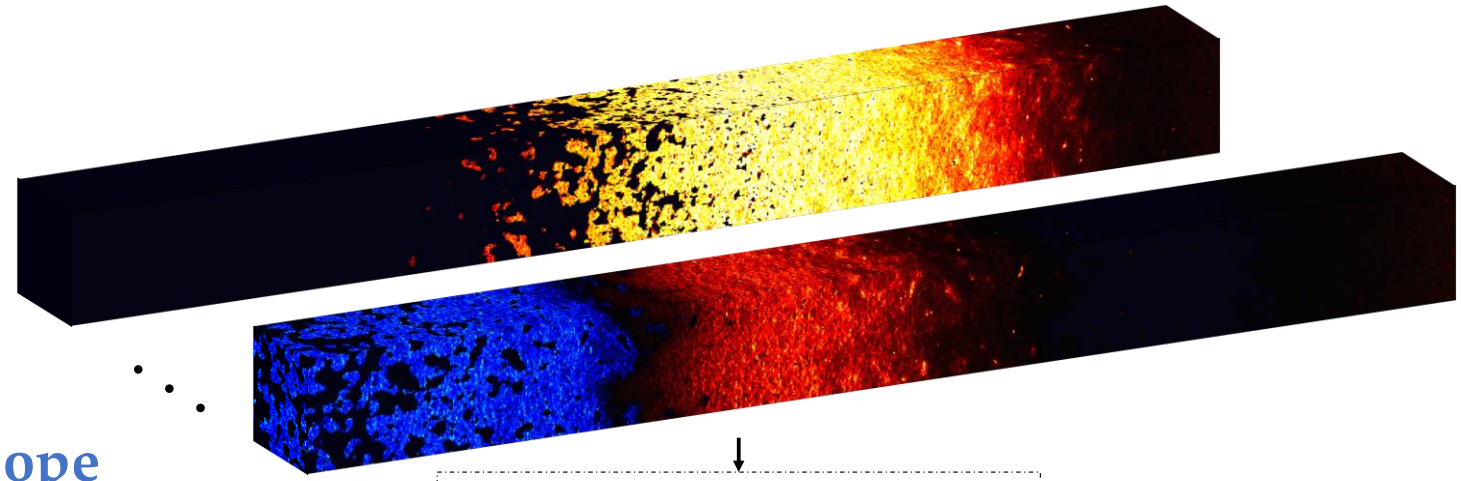
NVAITC collaboration: Giuseppe Fiameni

SCUOLA NORMALE SUPERIORE

Prelogović et.al. in prep.

Cosmic 21cm signal

- Patterns reflect properties of the first non-visible galaxies



21cmFAST simulations

Radiotelescope simulator

Observational Pipeline

Regressor (NN)

Astrophysical and Cosmological parameters

SCUOLA NORMALE SUPERIORE

Project resources

- Hyperparameter search ~ 1000
- Multiple architectures and datasets

- CSCS PizDaint – 300 000 GPU hours – P100

- Cineca Marconi100 – small C call, ~ 2000 GPU hours – V100

SCUOLA NORMALE SUPERIORE

DL Pipeline

Dataset(s) Creation	Architecture Construction	Hyperparameter Optimization	Training Procedure	Trained Models Management

• Fits on GPU?
• Fits in RAM?
• Dataset format
• Data aug.
  • On fly?
	• Frameworks
	• Layer types
	• Activations and random initializers
	• Mixed precision
	• Transfer learning?
		• Grid search / bayes opt.
		• Which parameters to tune
		• Flexible architecture objects
			• Manual or built-in?
			  • workers
			• Parallelization
			• Bottlenecks
			• Optimization for GPU type
			• Installing lib.
			  • containers
				• Database?
				• Retrieving pred.
				• Model comparison
				• Network visualization
				  • CAM

• DALI • Tfrecord	• Tensorflow/ Pytorch…		• Horovod • NSIGHT	• Tensorboard

SCUOLA NORMALE SUPERIORE

DL Pipeline

Dataset(s) Creation	Architecture Construction	Hyperparameter Optimization	Training Procedure	Trained Models Management

• Fits on GPU?
• Fits in RAM?
• Dataset format
• Data aug.
  • On fly?
	• Frameworks
	• Layer types
	• Activations and random initializers
	• Mixed precision
	• Transfer learning?
		• Grid search / bayes opt.
		• Which parameters to tune
		• Flexible architecture objects
			• Manual or built-in?
			  • workers
			• Parallelization
			• Bottlenecks
			• Optimization for GPU type
			• Installing lib.
			  • containers
				• Database?
				• Retrieving pred.
				• Model comparison
				• Network visualization
				  • CAM

• DALI • Tfrecord	• Tensorflow/ Pytorch…		• Horovod • NSIGHT	• Tensorboard

*In red the task the center helped on

SCUOLA NORMALE SUPERIORE

PyTorch

AImage Lab

Distributed Deep Networks Training for Video Recognition

Matteo Tomei
AImageLab
UNIMORE
matteo.tomei@unimore.it

NVIDIA

CINECA

Distributed Deep Networks Training for Video Recognition

AImage Lab

Video Data

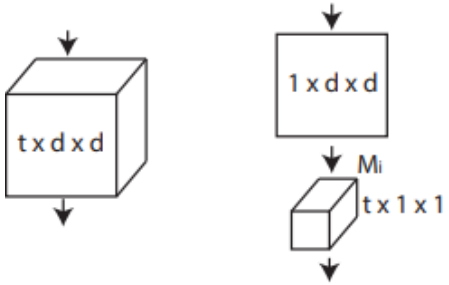
Kinetics-400 [1]
- 400 human action classes
- About 240.000 training videos
- About 20.000 validation videos

UCF-101 [2]
- 101 action classes
- 3 splits
- About 10.000 training videos
- About 3.000 test videos

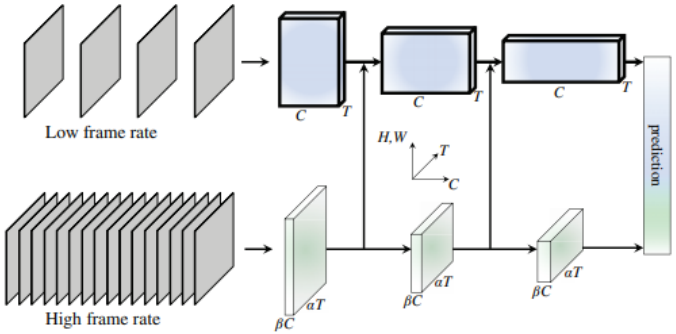
HMDB-51 [3]
- 51 action classes
- 3 splits
- About 6.000 videos

Video Backbones

R(2+1)D Network [4]



SlowFast Network [5]

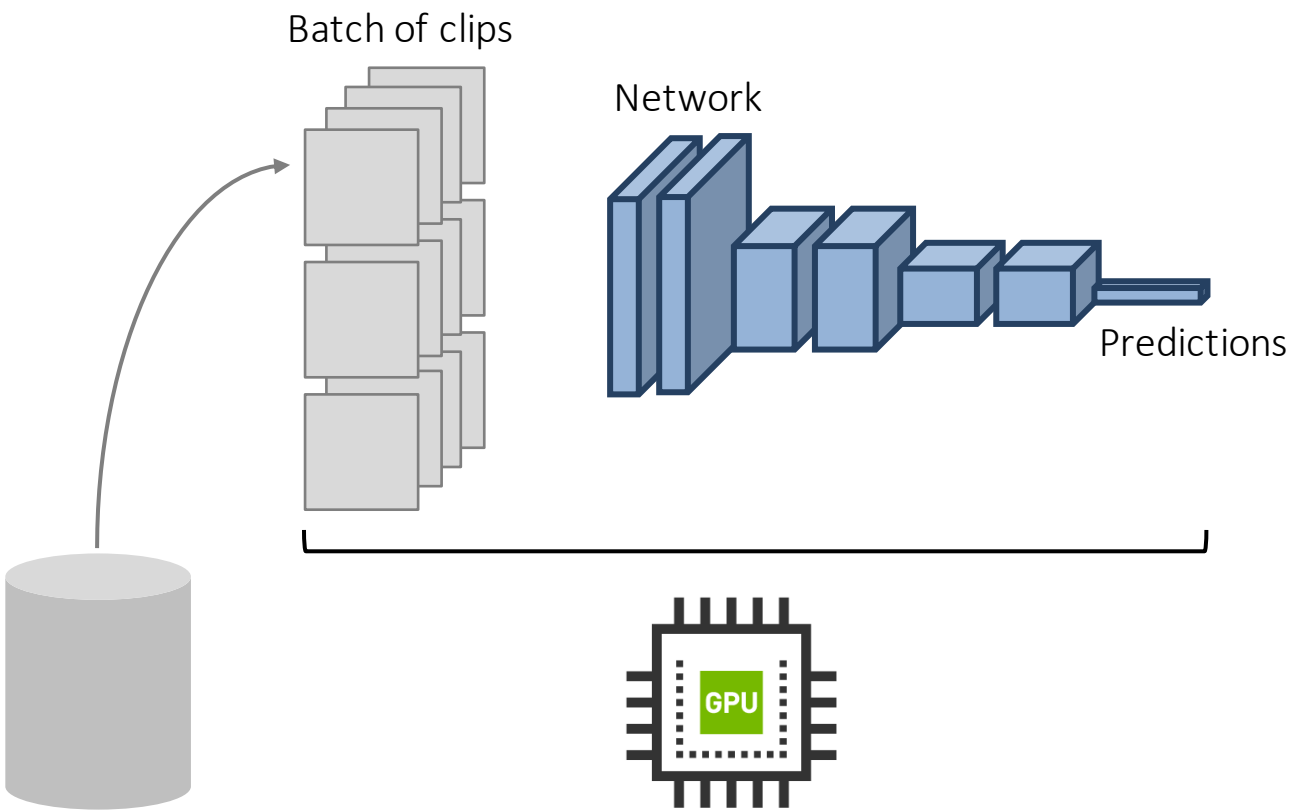


HW Resources

CINECA MARCONI100 accelerated cluster
- 980 nodes
- 4 x 16GB Volta NVIDIA GPUs per node
- 32 cores per node
- 256 GB RAM per node

[1] Will Kay, Joao Carreira, Karen Simonyan, Brian Zhang, Chloe Hillier, Sudheendra Vijayanarasimhan, Fabio Viola, Tim Green, Trevor Back, Paul Natsev, et al. The kinetics human action video dataset. arXiv preprint arXiv:1705.06950, 2017.
[2] Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah. Ucf101: A dataset of 101 human actions classes from videos in the wild. arXiv preprint arXiv:1212.0402, 2012.
[3] Hildegard Kuehne, Hueihan Jhuang, Est´ıbaliz Garrote, Tomaso Poggio, and Thomas Serre. Hmdb: a large video database for human motion recognition. In Proc. ICCV, 2011.
[4] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri. A closer look at spatiotemporal convolutions for action recognition. In Proc. CVPR, 2018.
[5] Christoph Feichtenhofer, Haoqi Fan, Jitendra Malik, and Kaiming He. Slowfast networks for video recognition. In Proc. ICCV, 2019.

Distributed Deep Networks Training for Video Recognition



Training details:

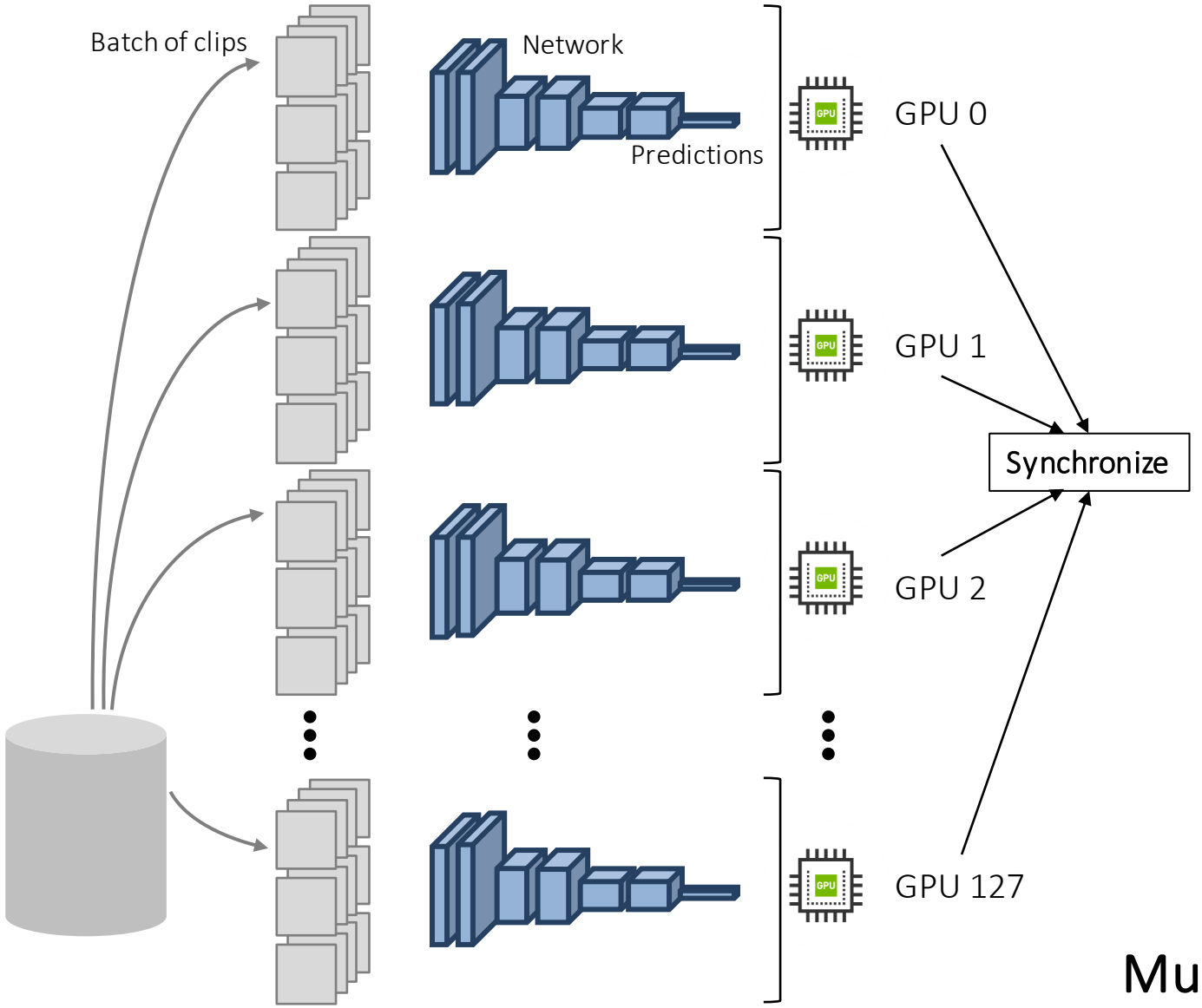
- SlowFast-4x16-R50 Network [1]
- Kinetics-400 dataset
- 64x256x256 clip shape
- Batch Size: 16 clips
- 45 epochs training
- Single 16GB NVIDIA V100 GPU

Estimated Training Time: ~ 68 days

Single-Process Single-GPU

[1] Christoph Feichtenhofer, Haoqi Fan, Jitendra Malik, and Kaiming He. Slowfast networks for video recognition. In Proc. ICCV, 2019.

Distributed Deep Networks Training for Video Recognition



Training details:

- SlowFast-4x16-R50 Network
- Kinetics-400 dataset
- 64x256x256 clip shape
- Batch Size: 16 clips per GPU
- 45 epochs training
- 128 16GB NVIDIA V100 GPU

Estimated Training Time: ~ 16 hours

Multiple-Processes Multiple-GPUs

Distributed Deep Networks Training for Video Recognition

AImage Lab

Computational Resources used so far:

- 75K GPU-hours (~8.5 years using 1 GPU)
- 800K CPU-hours (~91 years using 1 CPU)
- 64 to 128 nodes (with 4 16GB V100 NVIDIA GPUs each)

Main technologies:

- PyTorch
- Torch DistributedDataParallel
- NVIDIA Apex Amp (Automatic Mixed Precision)

Support from the NVIDIA Technology Centre:

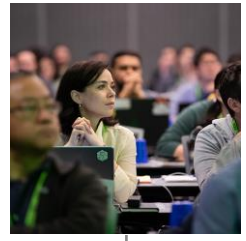
- SOTA review for model architecture
- Test execution and optimization
- Joint publication

PyTorch

NVIDIA


NVIDIA
EVENTS

EVENTS



Feb 2020
OpenACC Course at CINECA

Mar 2020
GTC Digital

Apr/May 2020
Online seminars On Multi-GPU training

May 2020
CINECA MARCONI-100 3920 GPUs

May 2020
NVAITC explained

June/July 2020
AI Webinar Series

Oct/Nov 2020
GTC Digital

Jan 2021
ICPR 2020

CADL2020

International Workshop on Computational Aspects of Deep Learning
January 11, 2021 (http://cadl.it)

Two invited speakers confirmed

Tom Gibbs - NVIDIA
Dr. Gibbs is currently responsible for strategy and implementation of programs to enable and promote developers to take full advantage of NVIDIA technology. Tom brings over 30 years of experience in HPC, and has applications expertise in industries ranging from Physics, Aerospace, Healthcare, Life Sciences, Energy and Financial Services.
Winner of the Gordon Bell Special Price in HPC 2020.

Marco Rorro - CINECA
Degree in Mathematics and PhD in Applied Mathematics at the University of Rome "La Sapienza".Technical consultant and developer within the Middleware for HPC services group of the SCAI department.

NVIDIA GTC

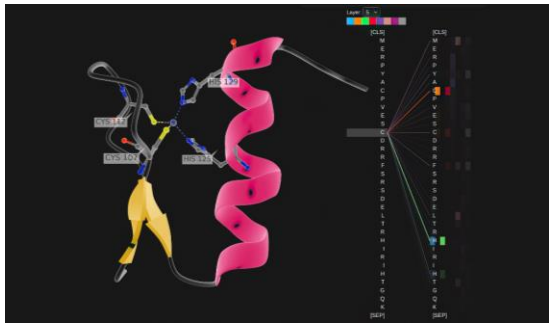
SHARE YOUR LIFE'S WORK AT GTC 2021

JOIN A GLOBAL COMMUNITY OF BRILLIANT MINDS ONLINE NEXT MARCH

NVIDIA's GTC brings together a global community of developers, researchers, engineers, and innovators with the common goal of sharing achievements while discovering new technologies and tools that drive change around the globe.

If you work with any of our GPUs, DPUs, or software offerings is making a difference, submit a talk or poster to join us online in March.

March 15 — 25, 2021
Submit your ideas at www.nvidia.com/gtc




NVIDIA
Q/A

Q/A

Q: How to give an estimation to Cineca about the amount of computational resources that might be needed by the project?

A: Use existing benchmark. Request for a trial access for testing and scalability study. Do an extrapolation starting from results using local resources. NVIDIA and Cineca are willing to help.

Q: Are interfaces like Google Collab available at Cineca?

A: NVIDIA and Cineca are collecting requirements from the AI community and working to implement them. For instance, Cineca provides a mechanism to spawn Jupyter Notebook on the Marconi-100 system.

Q: Does NVAITC also provide support to startups?

A: NVIDIA supports startups through the NVIDIA inception program (https://www.nvidia.com/en-us/deep-learning-ai/startups/)

AImage Lab
University of Modena and Reggio Emilia

SCUOLA NORMALE SUPERIORE

CINECA

cini
consorzio interuniversitario nazionale per l'informatica

NVIDIA