

Simulazione veloce in ATLAS basata su Machine Learning per il Run 3 di LHC e oltre

Federico A.G. Corchia
Università e INFN, Bologna

Workshop sul Calcolo nell'INFN
La Biodola, Portoferraio (LI)
26 maggio 2025

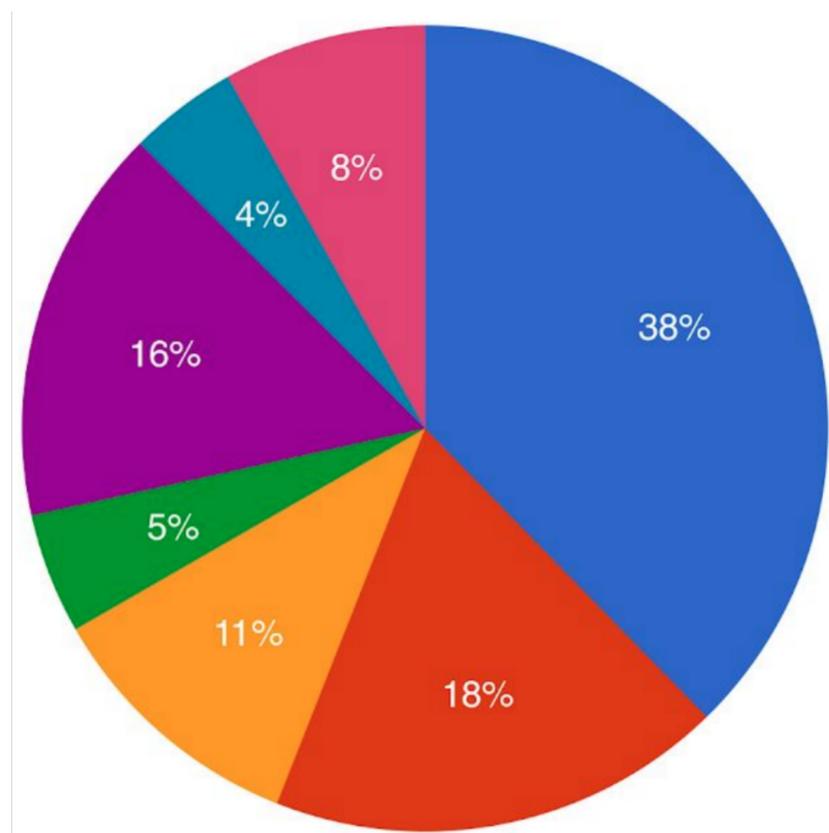


ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA



Introduzione

Simulazione dei calorimetri e carico computazionale



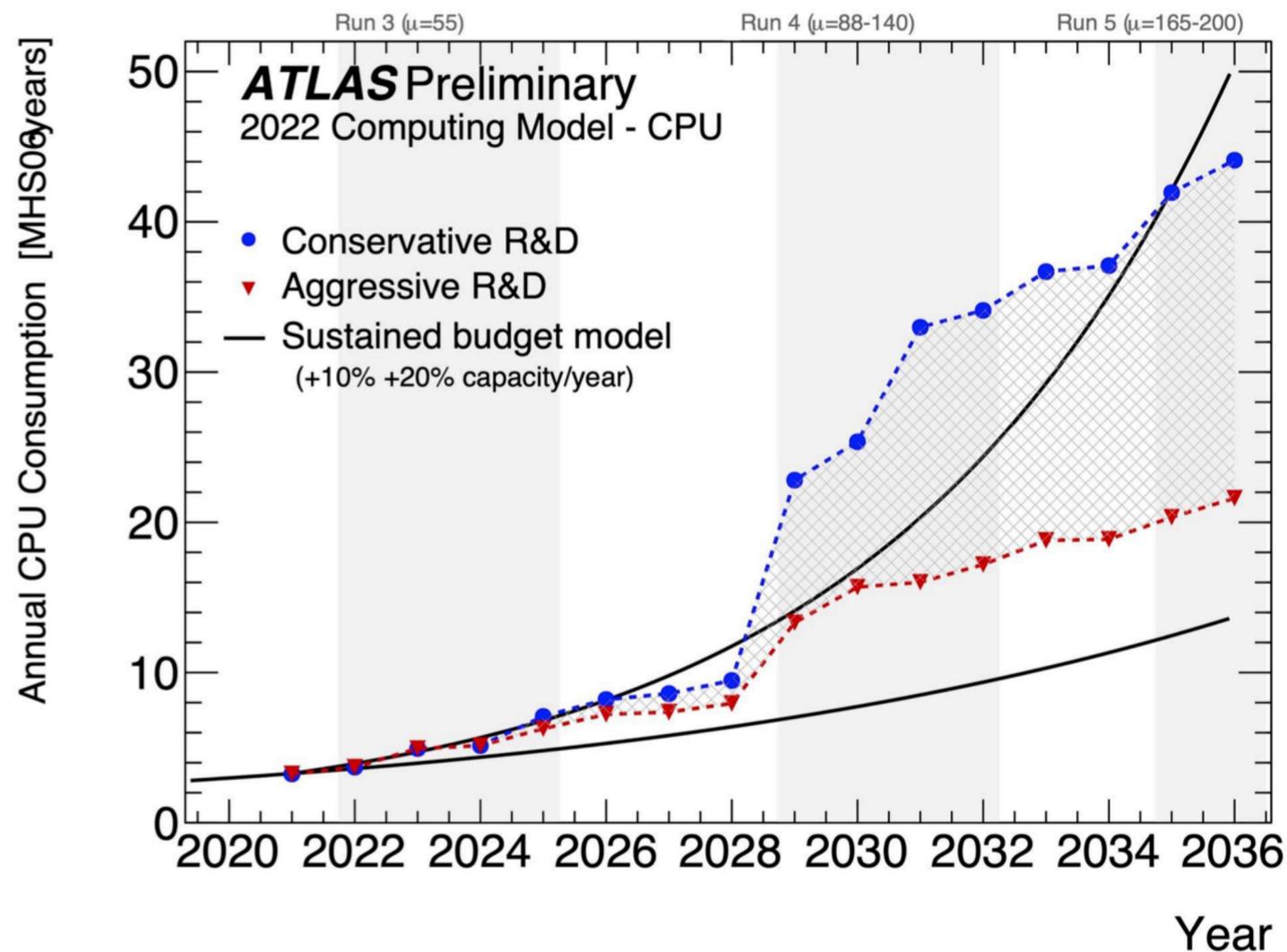
● MC simulation ● MC reconstruction ● MC event generation
● Analysis ● Group production ● Data processing
● Other

Consumo % di ore-CPU da parte delle attività di ATLAS
Coll. ATLAS, *ATLAS HL-LHC Computing Conceptual Design Report*,
[CERN-LHCC-2020-015](#) (2020)

- La simulazione MC della risposta del rivelatore è una **grande sfida dal punto di vista computazionale** per gli esperimenti dell'LHC!
- **Gran parte** delle risorse richieste dalla simulazione del rivelatore è impegnata dalla **simulazione dei calorimetri** (~80% del consumo di CPU per processi $t\bar{t}$)

Introduzione

Simulazione dei calorimetri e carico computazionale futuro



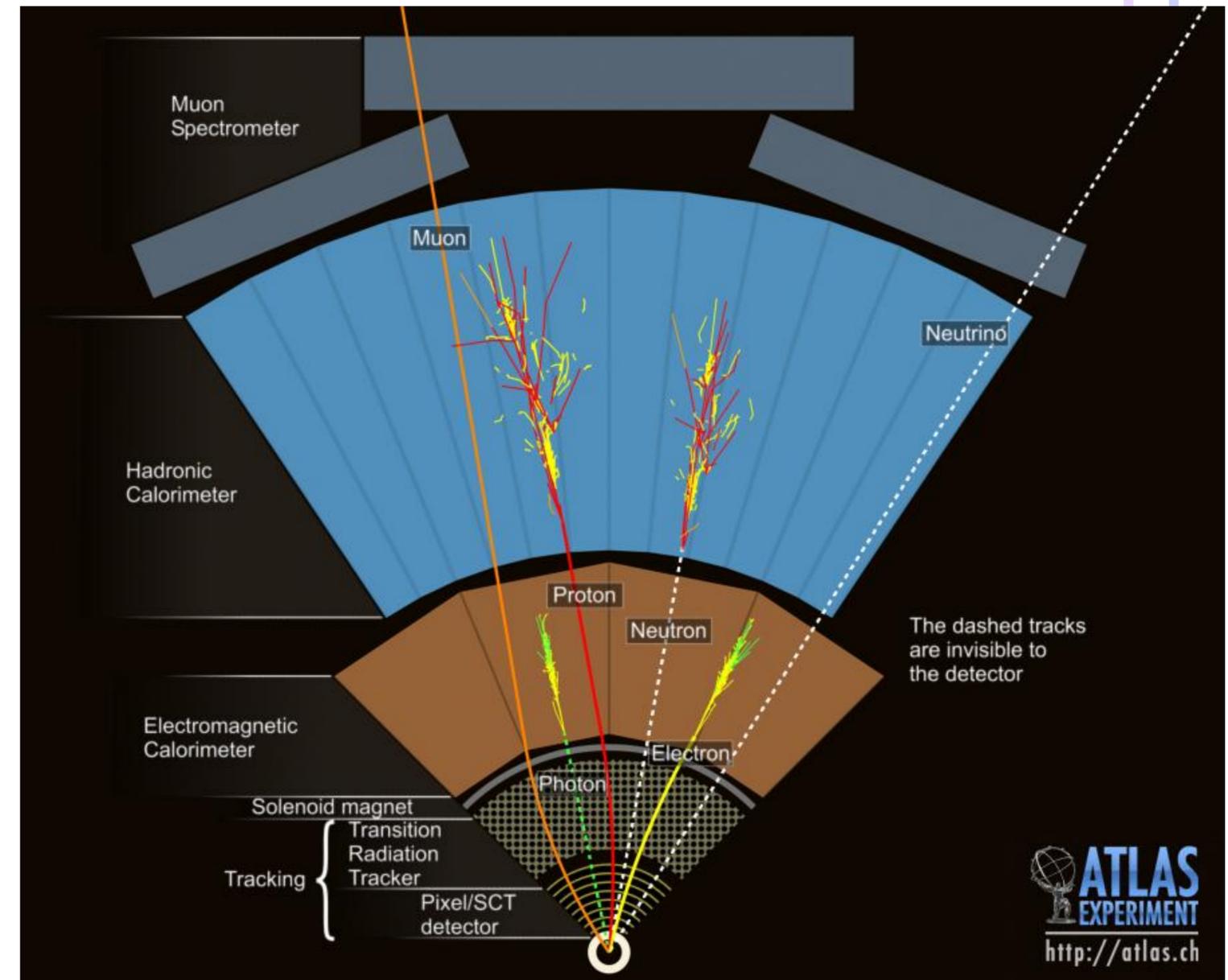
Coll. ATLAS, ATLAS Software and Computing HL-LHC Roadmap,
CERN-LHCC-2022-005 (2022)

- Si è previsto che il fabbisogno di risorse CPU **aumenterà** durante la Run 3 e in vista della Run 4 → necessario grande sforzo in ricerca e sviluppo per stare al passo!
- Servono soluzioni → **simulazione veloce!**

AtlFast3

La simulazione veloce in ATLAS

- **Sistemi di simulazione veloce:** capaci di simulare la risposta dei calorimetri più velocemente rispetto al sistema "tradizionale" di simulazione completa Geant4, ma mantenendo una buona accuratezza.
- **AtlFast3:** il sistema di simulazione veloce sviluppato in ATLAS.
 - Introdotto per la Run 2 e ulteriormente migliorato per la Run 3. Ora in produzione per la Run 3;
 - Sostituisce la lenta propagazione e le interazioni delle particelle incidenti dentro il volume del calorimetro con la generazione diretta dei depositi di energia, attraverso parametrizzazioni;
 - Impiega una geometria semplificata delle celle del calorimetro per semplificare la sua complessa struttura non omogenea.



cds.cern.ch/record/1096081

AtFast3

Componenti

- Impiega due approcci alla simulazione veloce:
 - **FastCaloSim V2**: approccio parametrico (v. prossima diapositiva);
 - **FastCaloGAN V2**: basato su **Reti Generative Avversarie** (*Generative Adversarial Network, GAN*).
- AtFast3 effettua la simulazione veloce attraverso uno dei due approcci, a seconda di quale dei due **restituisce la simulazione migliore** (cioè la più accurata rispetto a Geant4) per il tipo di particella incidente e per la sua energia.

Geant4 ancora usato per simulare tutte le particelle nell'Inner Detector, adroni a bassa energia nei calorimetri e muoni.

Muon Punchthrough (lo spray di particelle nello spettrometro dei muoni derivante da adroni ad alta energia che interagiscono tardi) modellato con un sistema basato su Reti Neurali Profonde (*Deep Neural Network, DNN*).

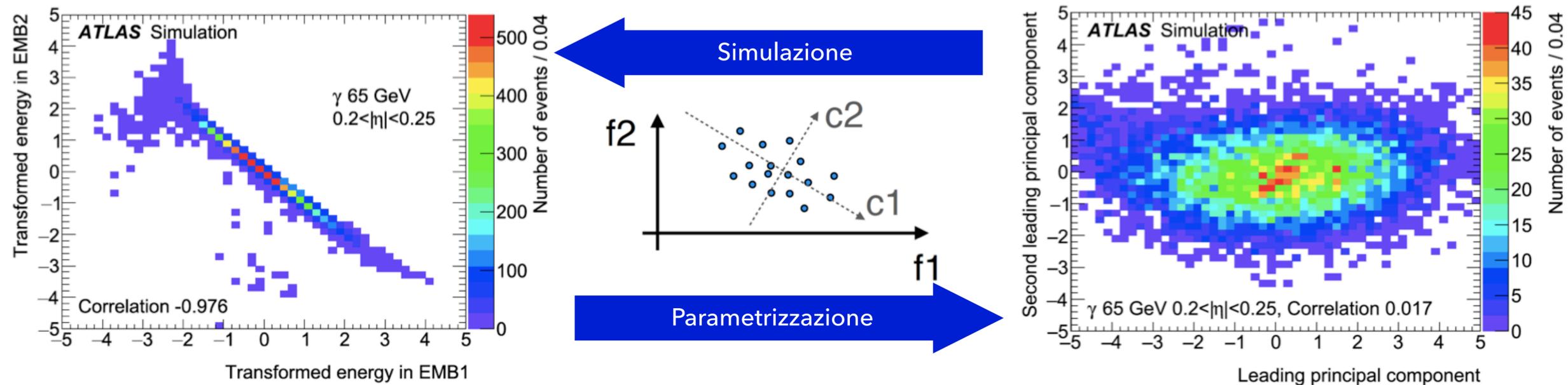
	Inner Detector	Calorimeters		Muon Spectrometer	
Electrons Photons	Geant4	FastCaloGAN V2 $E_{kin} < 8 \text{ GeV} \ \&\& \ \eta < 2.4,$ Except $[0.9 < \eta < 1.1, 1.35 < \eta < 1.5]$	FastCaloSim V2 $E_{kin} > 16 \text{ GeV} \ \&\& \ \eta < 2.4,$ All $E_{kin} \ \&\& \ [0.9 < \eta < 1.1, 1.35 < \eta < 1.5, \eta > 2.4]$		
Charged Pions Kaons		Geant4 Pions: $E_{kin} < 200 \text{ MeV}$ Other hadrons: $E_{kin} < 400 \text{ MeV}$	FastCaloSim V2 $E_{kin} < 4 \text{ GeV} \ \&\& \ \eta < 1.4,$ $E_{kin} < 1 \text{ GeV} \ \&\& \ \eta < 3.15$	FastCaloGAN V2 $E_{kin} > 8 \text{ GeV} \ \&\& \ \eta < 1.4,$ $E_{kin} > 2 \text{ GeV} \ \&\& \ 1.4 < \eta < 3.15,$ All $E_{kin} \ \&\& \ \eta > 3.15$	Muon Punchthrough + Geant4
Baryons		FastCaloGAN V2			
Muons		Geant4			

Coll. ATLAS, *Software and computing for Run 3 of the ATLAS experiment at the LHC*,
Eur. Phys. J. C **85**, 234 (2025)

FastCaloSim

- Modellizzazione **parametrizzata** impiegando campioni Geant4 di singoli γ , e^- e π^\pm ;
- Parametrizzazione in 17 bin in energia distanziati logaritmicamente da 64 MeV a 4 TeV e 100 bin in $|\eta|$ distanziati linearmente da 0 a 5;
- Parametrizzazione **separata** dello sviluppo longitudinale e laterale dello sciame;
- **Energia nei layer decorrelata** mediante *Principal Component Analysis* (PCA, v. immagine sotto);
- **Distribuzione dell'energia laterale parametrizzata** come funzioni di probabilità 2D.

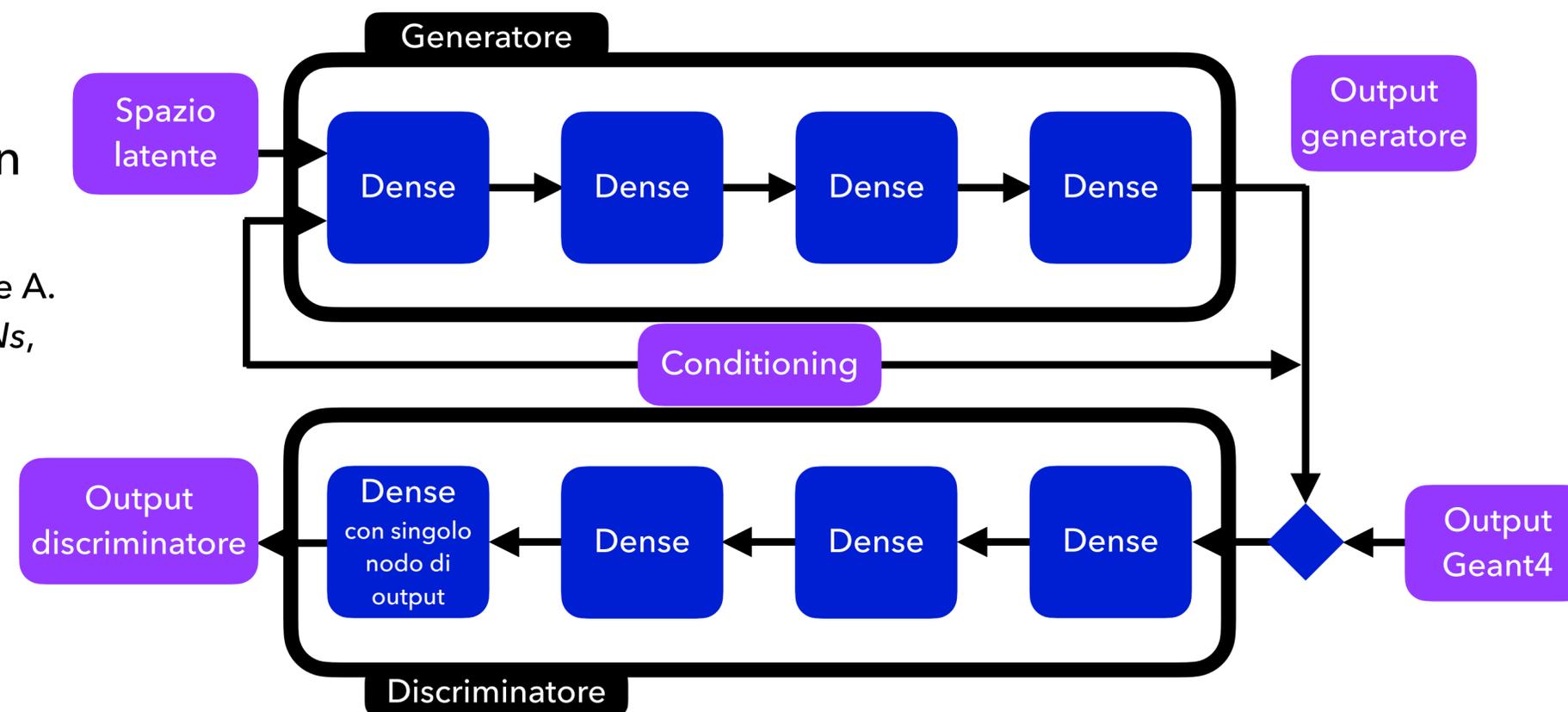
Coll. ATLAS, *AtFast3: The Next Generation of Fast Simulation in ATLAS*,
Comput Softw Big Sci **6**, 7 (2022)



FastCaloGAN

Concetti di base sulle GAN

- Basato su **GAN di Wasserstein** con gradient penalty
I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin e A. Courville, *Improved Training of Wasserstein GANs*, arXiv:1704.00028 [cs.LG] (2017)
- **Addestramento simultaneo di due reti neurali** (cioè la tipica struttura delle GAN):



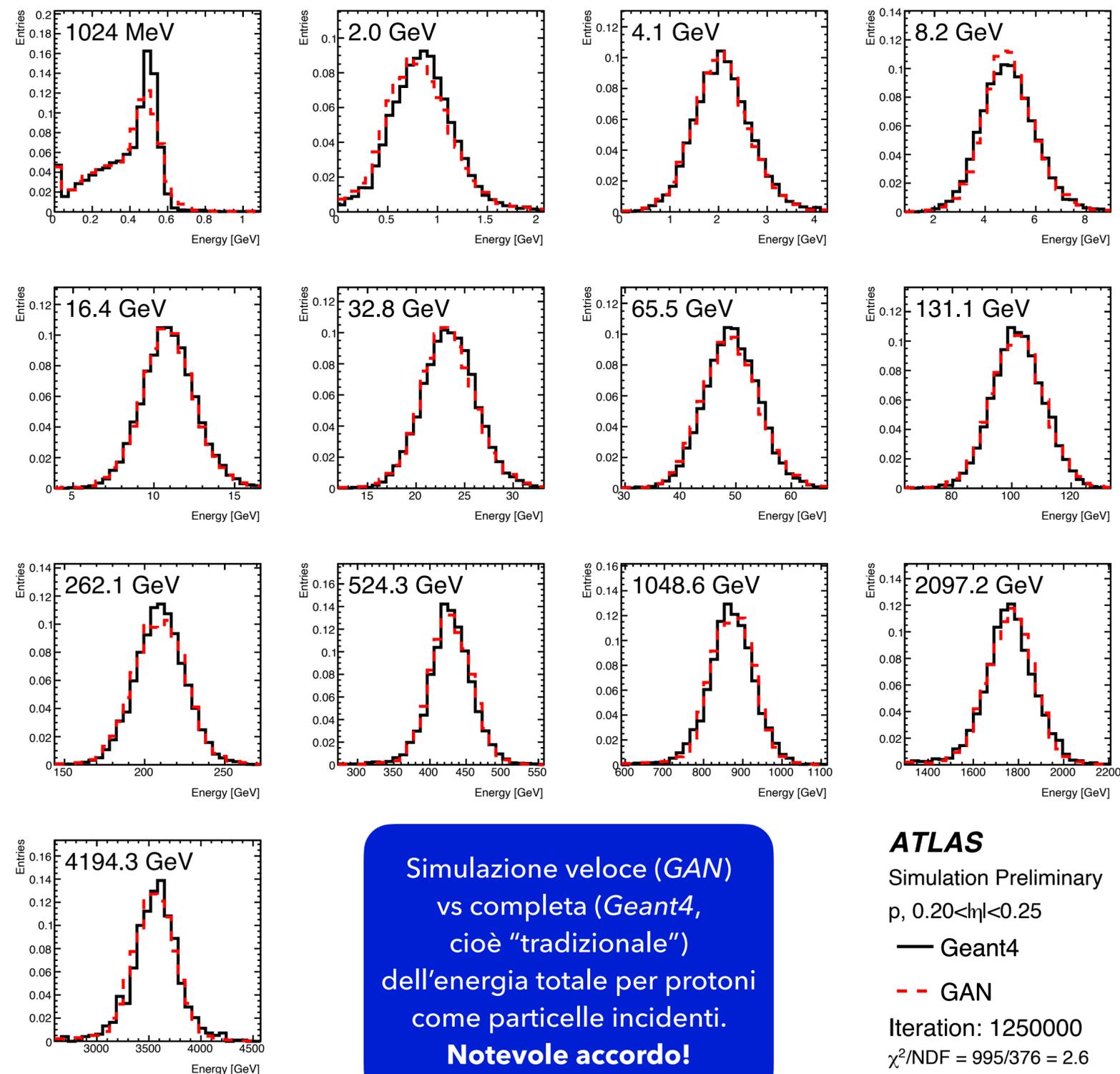
- **Generatore**: cerca di generare campioni il più simili possibile ai campioni di Geant4;
- **Discriminatore**: cerca di distinguere i campioni di Geant4 da quelli prodotti dal generatore.
- Quando raggiunge un **equilibrio** tra le due, FastCaloGAN simula gli sciami nei calorimetri molto più velocemente rispetto a Geant4, mantenendo una buona accuratezza.

FastCaloGAN

Ulteriori dettagli

- Addestramento con e^- , γ , p , π^\pm come particelle incidenti su ciascuno dei 100 bin in $|\eta|$, con conditioning sulla quantità di moto;
- Hit del calorimetro raggruppate in voxel (= bin 3D);
- L'architettura della rete, gli iperparametri e la granularità dei voxel sono stati ottimizzati.

ATLAS-SIM-2023-004

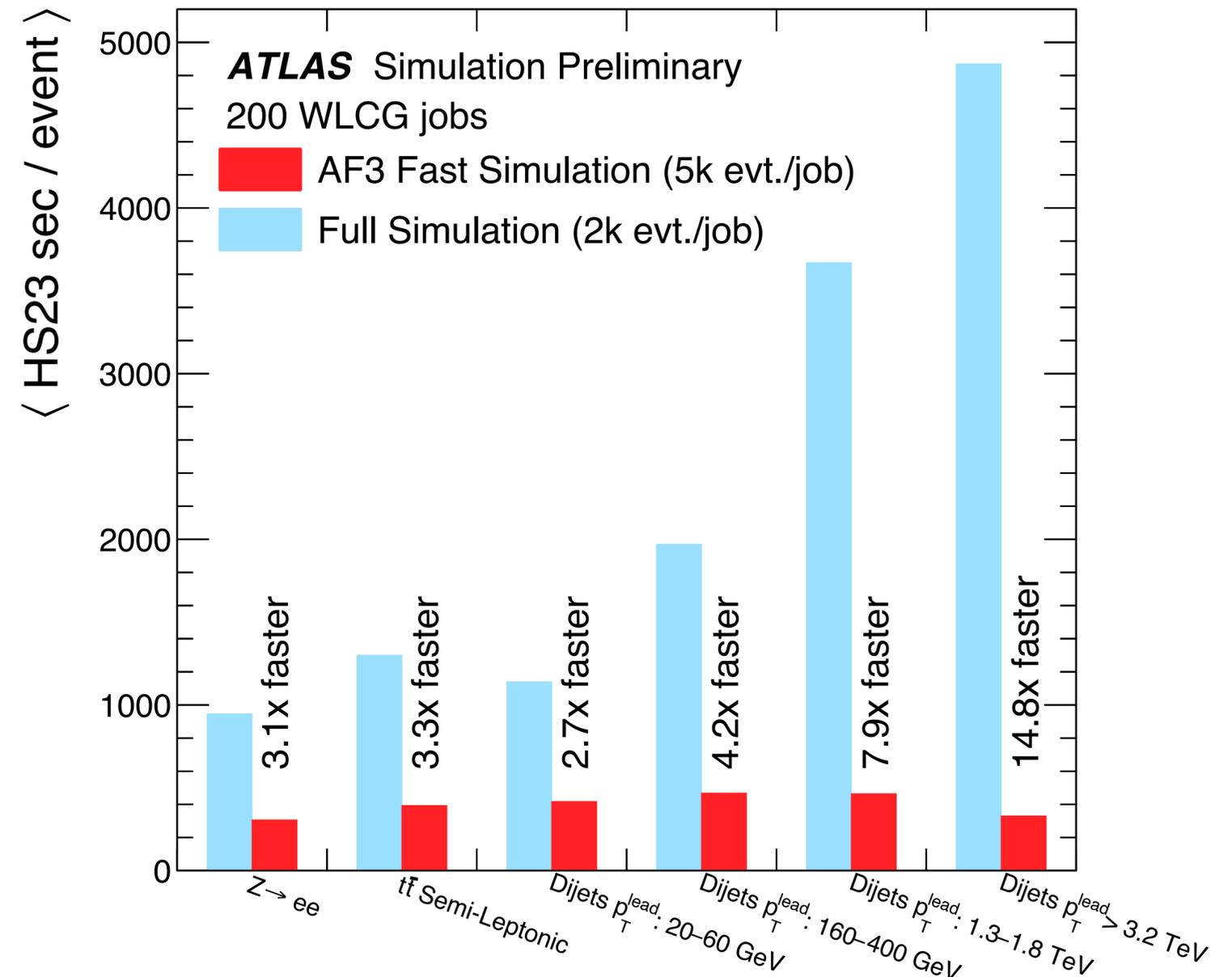


Prestazioni

Lato computazionale

- AtFast3 è da **3** (per eventi $Z \rightarrow ee$) a **15** (per eventi dijet ad alta p_T) **volte più veloce** rispetto a Geant4 (simulazione del rivelatore ATLAS della Run 3 per la presa dati del 2023)!

Tempo di simulazione in AtFast3
dominato dalla simulazione
completa dell'Inner Detector.



ATLAS-SIM-2023-005

Prestazioni

Lato fisico

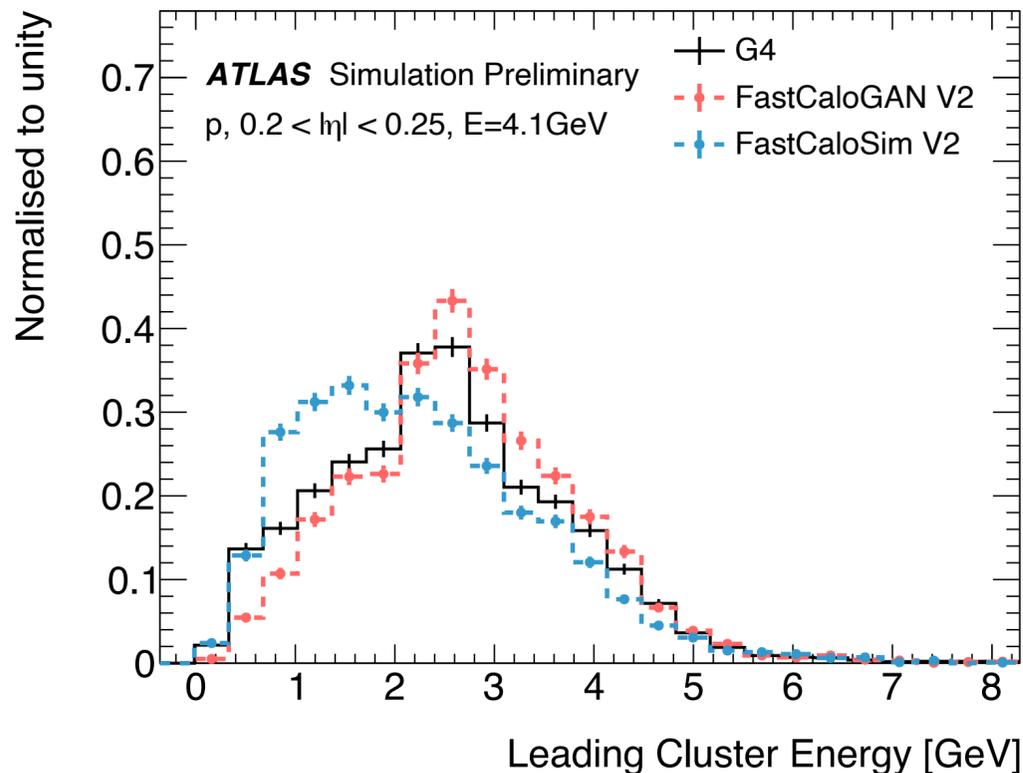
- Modellizzazione molto accurata dell'energia del leading cluster (con FastCaloGAN), così come del numero di costituenti per le variabili jet e substructure in eventi dijet (con l'approccio *ibrido*, cioè FastCaloSim e FastCaloGAN combinati);
- Miglioramento nelle prestazioni di fisica rispetto alle versioni precedenti;
- Per la maggior parte delle osservabili impiegate nelle analisi di fisica, AtlFast3 e Geant4 **hanno un accordo entro alcuni punti percentuali**;
- AtlFast3 può essere usato **per quasi ogni analisi** (non solo segnale ma anche fondo). Le misure ad altissima precisione potrebbero non essere in grado di impiegare AtlFast3, ma per la Run 4 **>90-95%** delle analisi dovrà impiegare la simulazione veloce, in quanto non ci sarà la capacità in CPU per permetterne la simulazione completa.

Prestazioni

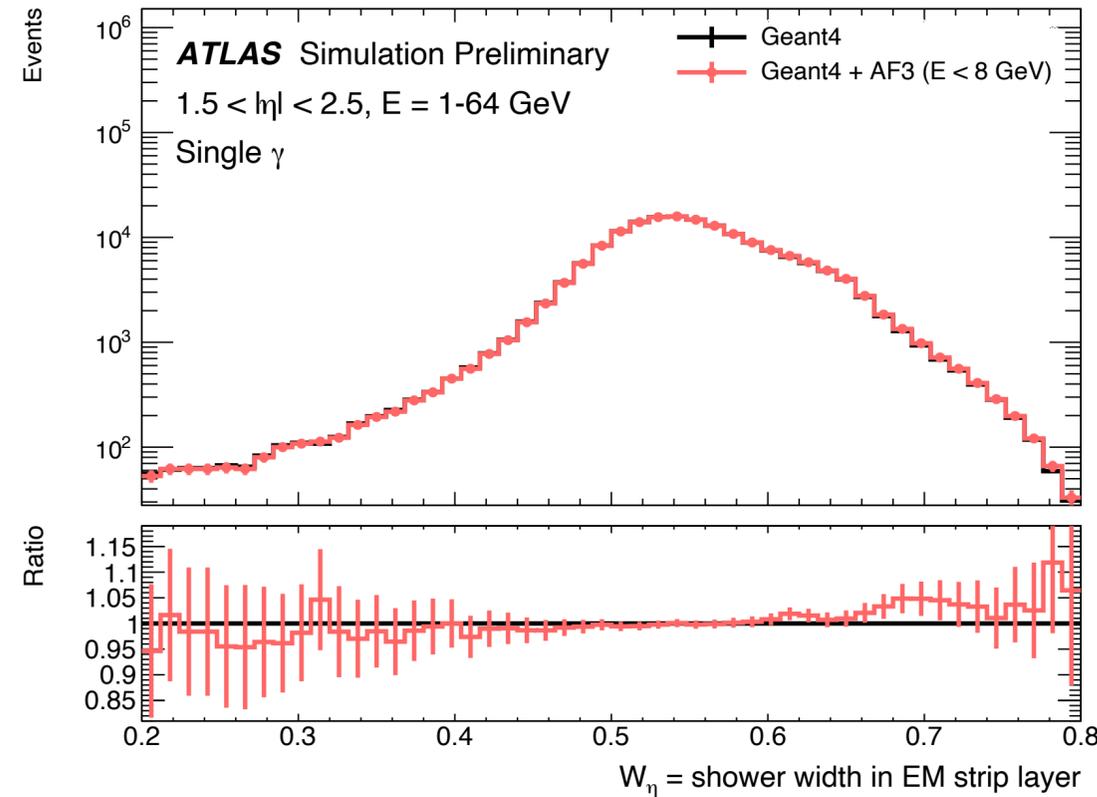
Lato fisico (continua)

Approccio ibrido = uso di FastCaloSim o FastCaloGAN in base a tipo di particella, energia ed η come da diapositiva 5 (invece di impiegare solo FastCaloSim o solo FastCaloGAN)

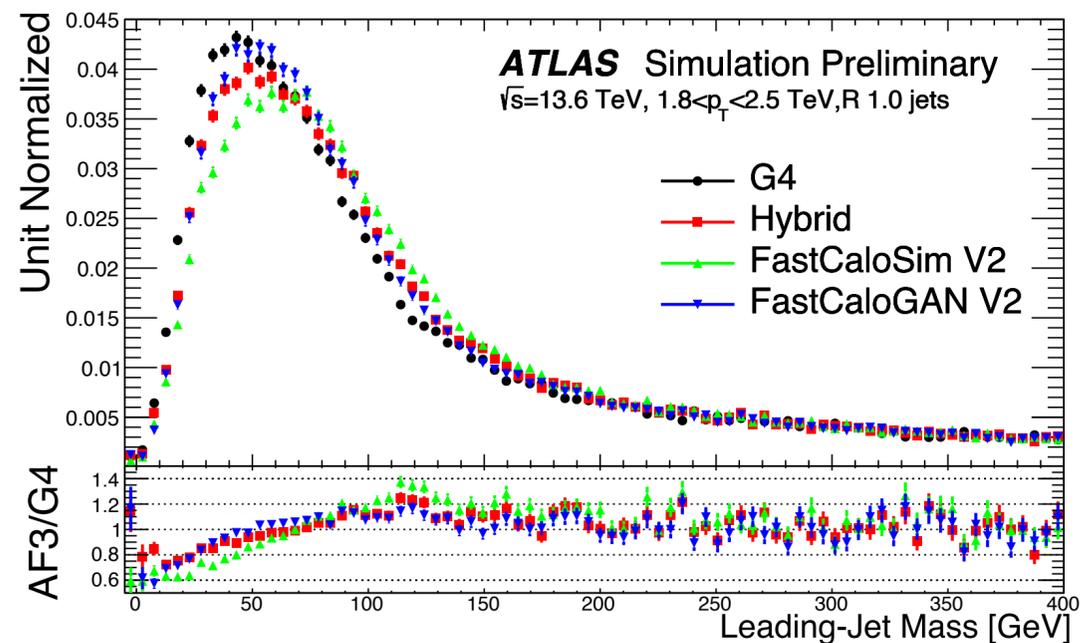
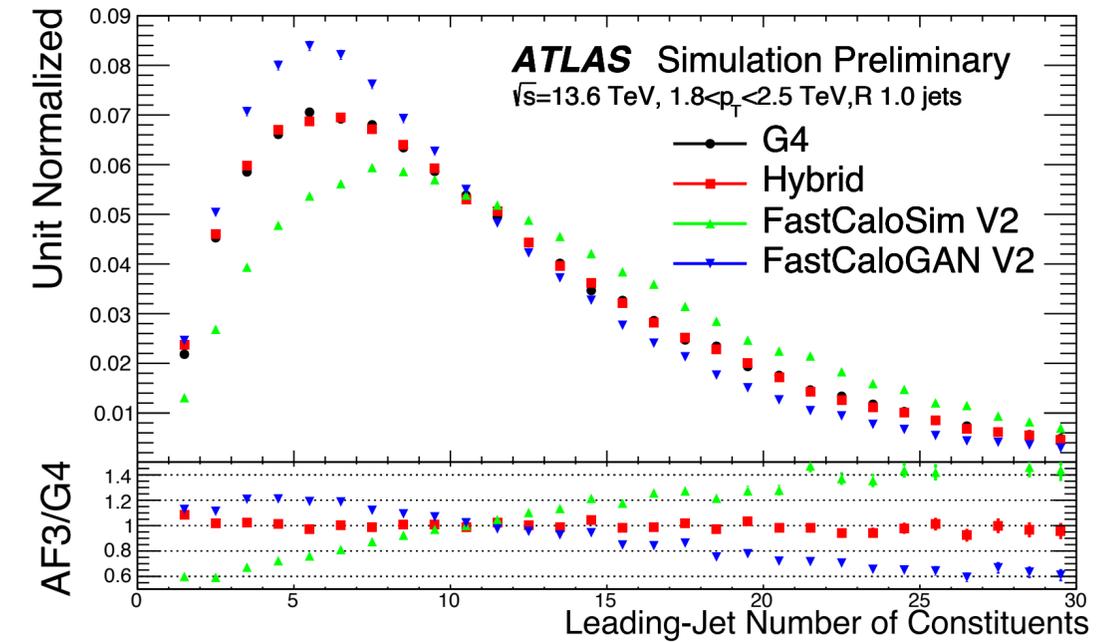
ATLAS-SIM-2023-004



ATLAS-SIMU-2024-08



ATLAS-SIM-2023-004



ATLAS-SIM-2023-004

ATLAS-SIM-2023-004



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Prospettive future

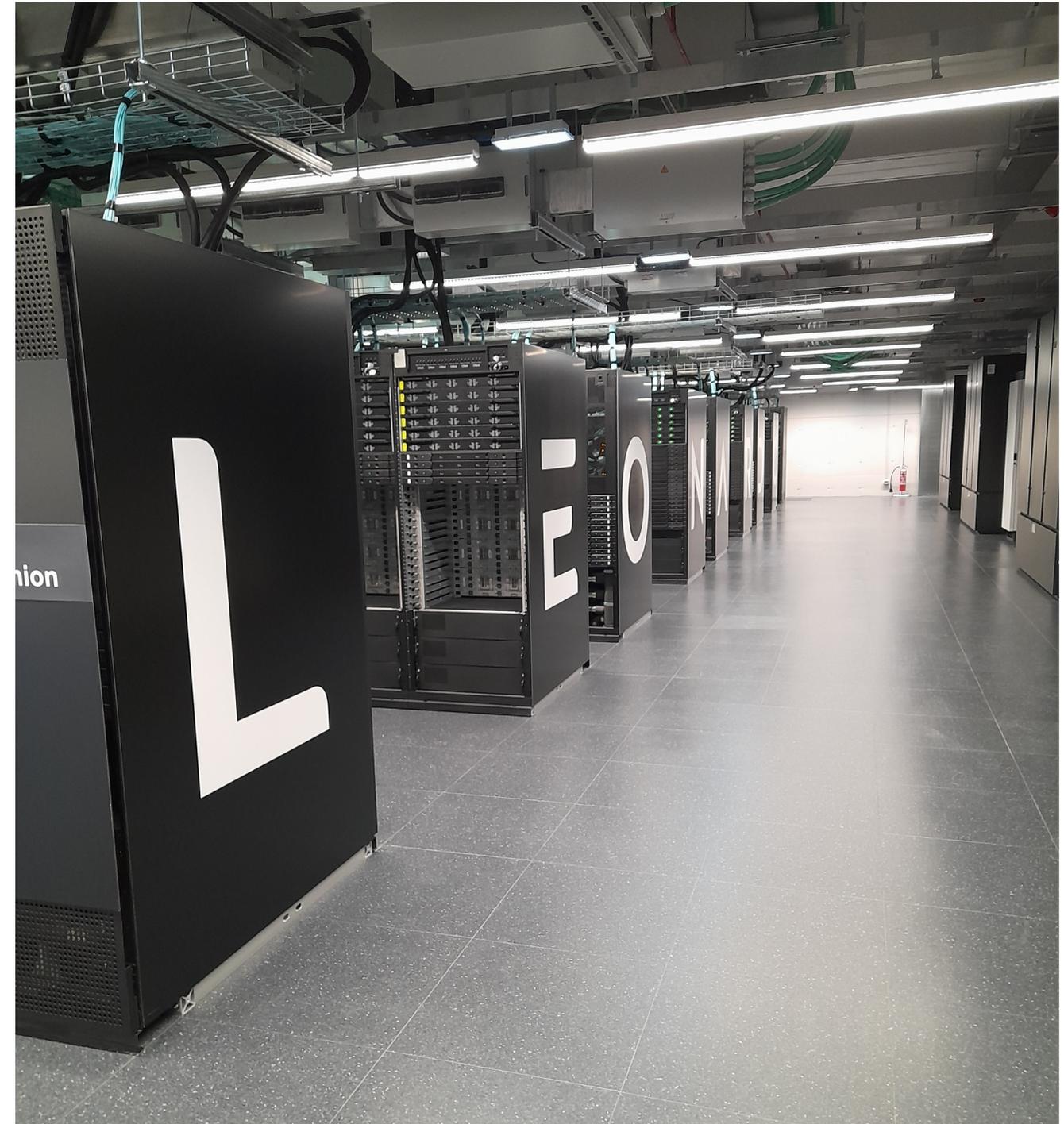
Verso la Run 4

- Ora che AtlFast3 è in produzione per la Run 3, è in corso lo sviluppo in vista della parte finale della Run 3 e la Run 4:
 - test su voxelizzazione **migliorata** e più fine che riduce i bias dovuti alla geometria del calorimetro;
 - ricerca su modelli di **diffusione** (basati su **transformer** - *CaloDiT*) e **Reti Neurali Invertibili** (*Invertible Neural Network, INN*). Se producono buone simulazioni potremo aggiungerli come ulteriori componenti ad AtlFast3 (o addirittura sostituire FastCaloSim e/o FastCaloGAN)!
- I risultati attuali mostrano che **le GAN hanno ancora potenziale** e che competono bene con le INN, le quali a loro volta **producono risultati molto buoni**.

Prospettive future

Verso la Run 4 (continua)

- L'addestramento delle GAN e la relativa ricerca avvengono non solo sulle risorse CERN (LX BATCH) ma anche sul **supercomputer Leonardo** del CINECA a Bologna;
- Il 9° supercomputer più potente nella classifica TOP500 (novembre 2024);
- RHEL 8.7, 4992 nodi di calcolo di cui 3456 con (per ciascun nodo) 4 GPU NVIDIA A100 (64 GB), 32 core e 512 GB di RAM.



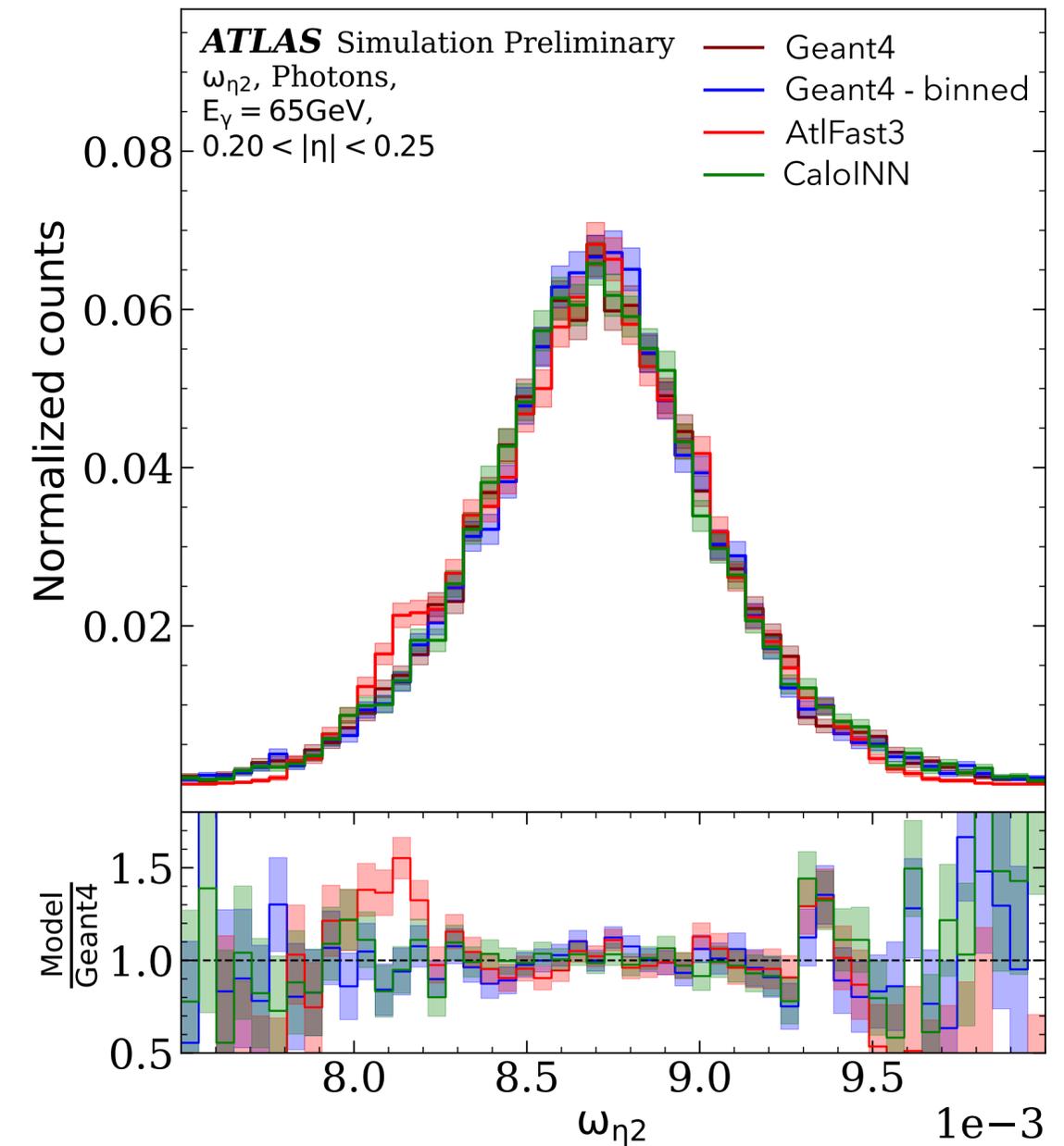
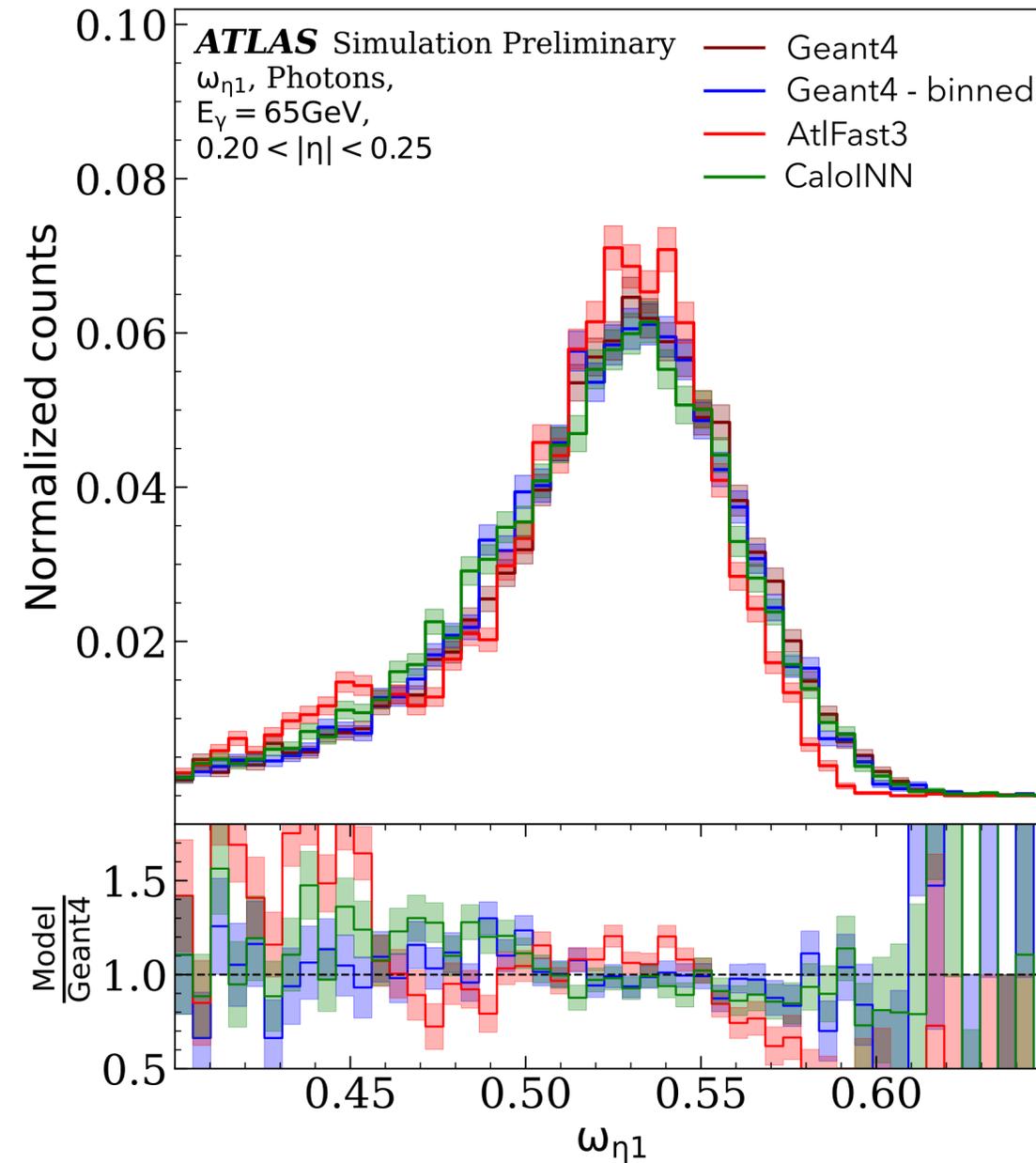
Prospettive future

CaloINN

ATLAS-SIMU-2024-10

La voxelizzazione è stata completamente riottimizzata e si è visto che la vecchia voxelizzazione introduce rilevanti artefatti fisici. Le INN (*CaloINN*) riproducono accuratamente le distribuzioni di Geant4 con le voxelizzazioni ottimizzate.

$\omega_{\eta 1}$ = shower width laterale ricostruita nello strip layer del calorimetro di ATLAS
 $\omega_{\eta 2}$ = shower width laterale ricostruita nel layer EMB2 del calorimetro di ATLAS



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Conclusioni

- AtlFast3 è il sistema di simulazione veloce in ATLAS allo stato dell'arte, capace di simulare un'ampia gamma di processi fisici con alta precisione.
- AtlFast3 per la Run 3 raggiunge un'alta precisione in molti casi con un'accelerazione nelle prestazioni computazionali di 3-15 volte.
- Essenziale per rispondere al fabbisogno computazionale delle future run dell'LHC, così come le esigenze di accuratezza nella modellizzazione fisica.

Grazie per l'attenzione