

# Big Data e Big Science

Alberto Di Meglio  
CERN openlab CTO

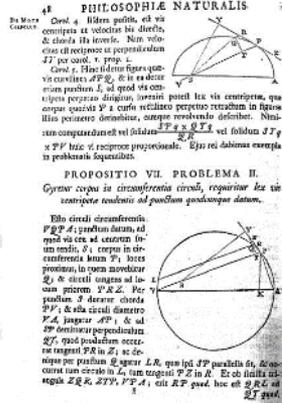


# Paradigmi della ricerca scientifica



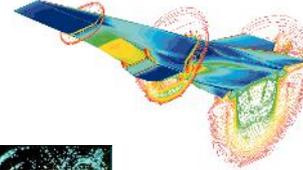
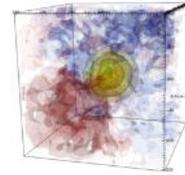
4000 anni fa

1 - Osservazione empirica



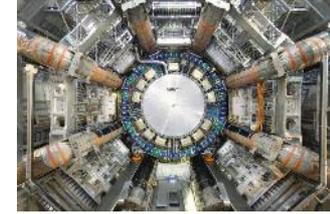
500 anni fa

2 - Generalizzazione Modelli matematici



~50 anni fa

3 - Simulazione Scienze computazionali



Oggi

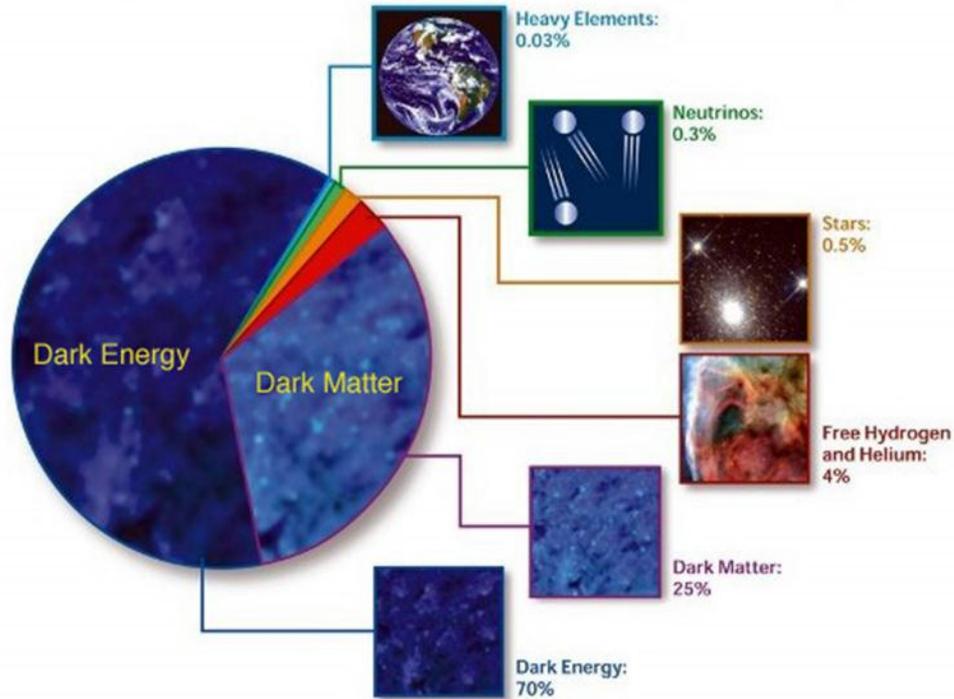
4 - Data-driven science eScience



# Scienza e collaborazioni internazionali



# Di cosa è fatto l'universo?

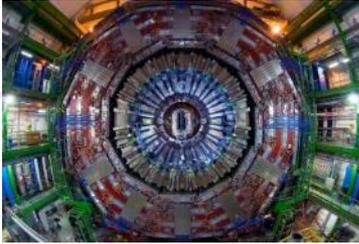


- Cosa dà alle particelle la loro massa?
- Come si può integrare la gravità in una teoria unificata?
- Perché c'è solo materia e non anti-materia nell'universo?
- Ci sono più dimensioni spazio-temporali delle 4 che conosciamo?
- Che cosa sono l'energia oscura e la materia oscura di cui è fatto il 95% dell'universo?

# Il Large Hadron Collider (LHC)



# Big Data !



I detectors: “microscopi” da 7000 tons  
150 milioni di sensori  
Generano dati 40 milioni di volte al secondo

→ **Peta Bytes / sec !**



I Filtri di Basso Livello (Triggers)  
100,000 selezioni al secondo

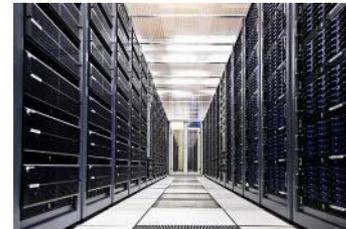
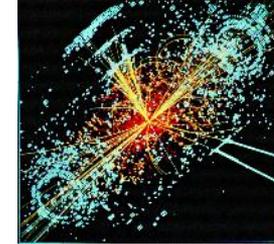
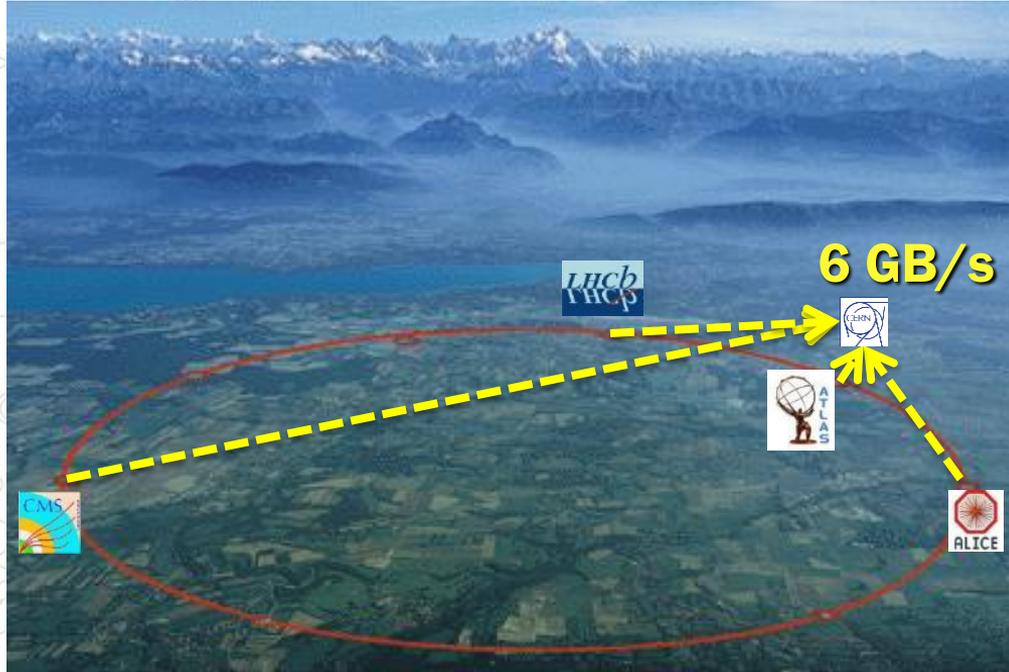
→ **Tera Bytes / sec !**



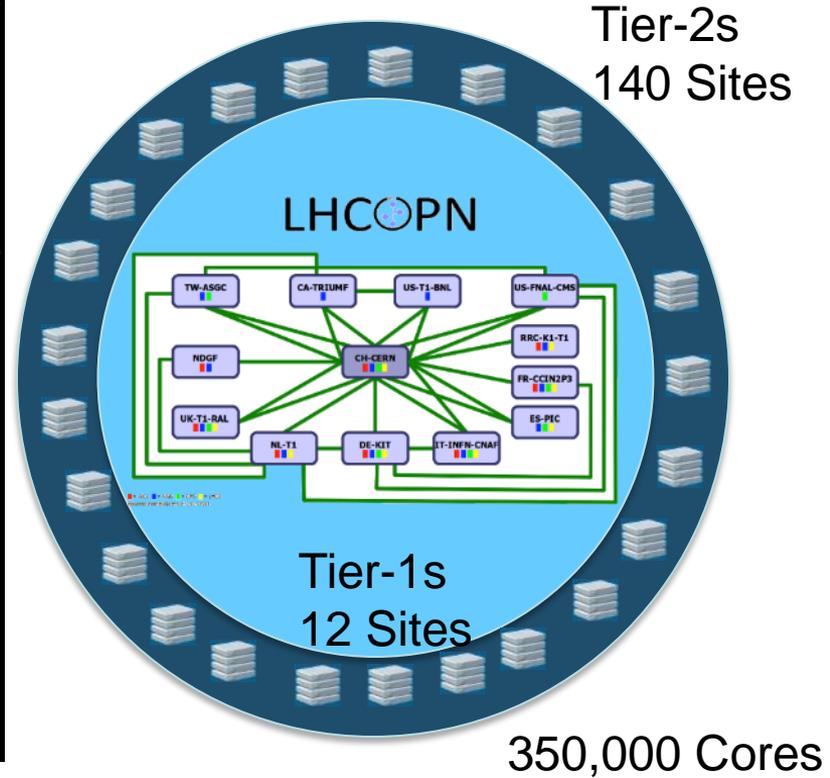
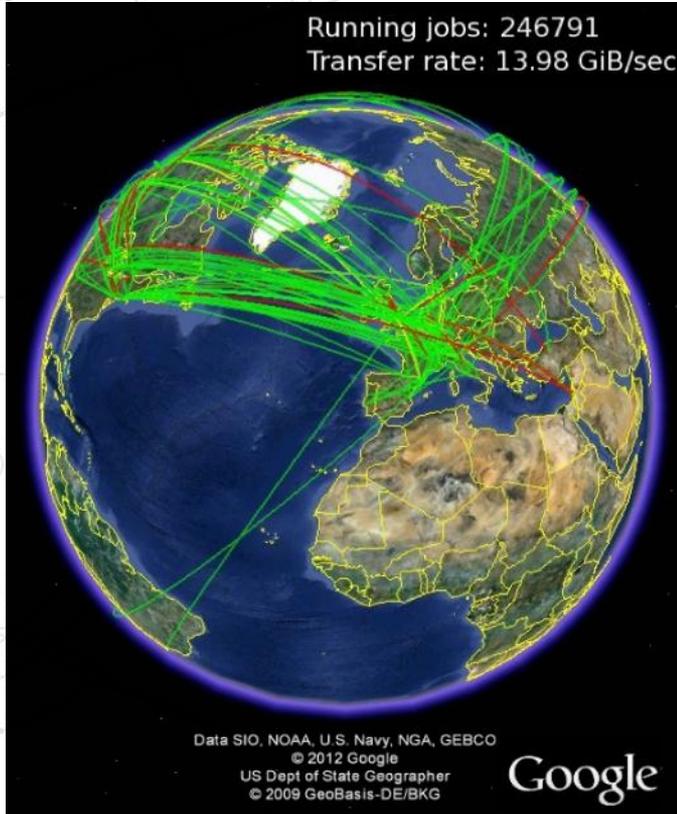
I Filtri di Alto Livello (HLT/Filters)  
100 selezioni al secondo

→ **Giga Bytes / sec !**

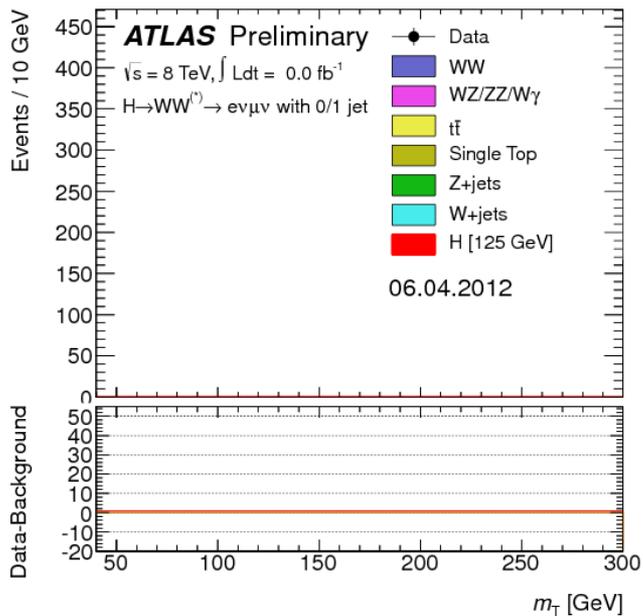
# Stoccaggio, Ricostruzione, Simulazione, Distribuzione



# Worldwide LHC Computing Grid



# Grandi collaborazioni, grandi scoperte



# Evoluzione dell'LHC

2009 2010 2011 2011 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 ... 2030?



LHC startup  
900 GeV

Phase-0 Upgrade  
(design energy,  
nominal luminosity)

Phase-1 Upgrade  
(design energy,  
design luminosity)

x 3

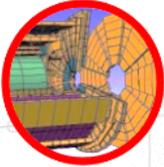
x 6

x 3

# Aree di ricerca in Information Technology



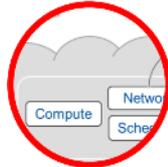
Acquisizione dati e selezione



Piattaforme di calcolo, analisi dati, simulazione



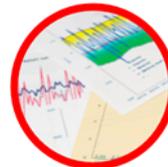
Stoccaggio dati e conservazione a lungo termine



Gestione delle risorse di calcolo (cloud)



Reti e comunicazione



Data analytics

# Nuove professioni

**Programmazione  
processori  
multicore,  
processori  
grafici (GPU),  
multithreaded  
software**

**Ingegneri  
Elettronici &  
Software**

**Piattaforme di  
analisi dati,  
statistica,  
matematica,  
presentazione e  
visualizzazione,  
(familiarità con  
la fisica delle  
particelle)**

**Data Scientist**

**Applicazioni  
della fisica alla  
medicina  
(acceleratori per  
uso medico)**

**Applicazioni  
multidisciplinari**



## GET IN TOUCH [www.cern.ch/openlab](http://www.cern.ch/openlab)

### EXECUTIVE CONTACT

Bob Jones, Head of CERN openlab  
[bob.jones@cern.ch](mailto:bob.jones@cern.ch)

### TECHNICAL CONTACTS

CERN openlab Chief Technology Office  
Alberto Di Meglio, [alberto.di.meglio@cern.ch](mailto:alberto.di.meglio@cern.ch)  
Sverre Jarp, [sverre.jarp@cern.ch](mailto:sverre.jarp@cern.ch)

### COMMUNICATION CONTACT

Mélissa Gaillard, CERN openlab Communications Officer  
[melissa.gaillard@cern.ch](mailto:melissa.gaillard@cern.ch)

This work is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License. It includes photos, models and videos courtesy of CERN and uses contents provided by CERN and CERN openlab

