High Throughput Computing and Its Impact on the Society
The Higgs Discovery and Particle Physics

The African School of Fundamental Physics 2022

December 7, 2022

Dr. Jaehoon Yu
Department of Physics
University of Texas at Arlington

Outline

- Who am I and how am I related to ASP?
- Introduction
- The problem
- A solution using the Computing Grid
- What HTC did for a Nobel winning discovery
- Conclusions

- My full name
- Lived in Sou
  - I take freed
  - Obtained B
    Korean Arr
- Joined the P
  obtained Ph.
  - Ph.D. thesi
    prototyping
    data analys
  - All my 3 ch
- 1st postdoc a
  postdoc at F
  building the
- Fermilab stat

orship"S"

ery seriously

service in the

Who am I? – 2

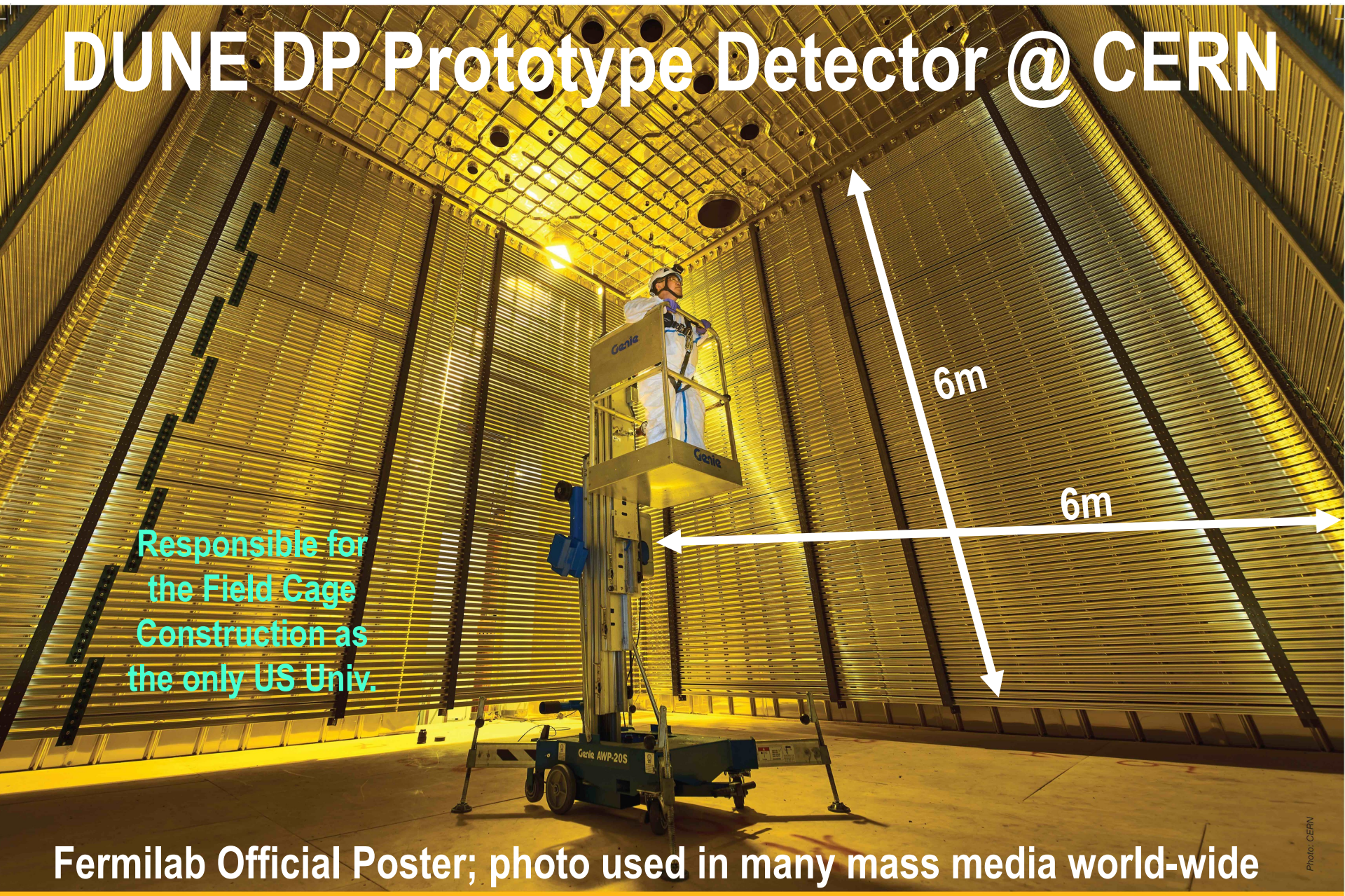
- Professor at U. Texas Arlington (2001 – present)
  - Led the design and implementation of D-Zero computing grid
    - Led the group on discovery of Higgs in WW final states
  - Led International Linear Collider detector R&D beam testing
  - Joined ATLAS @ LHC 2005 and led the grid computing user services
  - Led a subgroup in LHC Higgs Cross section working group
  - Contributed to 2012 Higgs discovery (see the TV interview) and the subsequent precision property measurements
  - Moved to neutrino experiment and created and leading the Beyond the Standard Model physics group till 2021 (1st ever in the community!)
  - Constructed two DUNE field cages (2018&2022) for Prototypes @CERN
  - To construct half the FC and the whole FC for first two 17,000t modules
  - Leading the technical design of the 2nd 17,000t module HV system

Where is USA?


ARCTIC OCEAN
Greenland (DENMARK)
Franz Josef Land (RUSSIA)
Svalbard (NORWAY)
Novaya Zemlya
RUSSIA
Alaska (U.S.)
CANADA
UNITED STATES
NORTH PACIFIC OCEAN
NORTH ATLANTIC OCEAN
SWEDEN
FINLAND
NORWAY
UNITED KINGDOM
POLAND
UKRAINE
FRANCE
SPAIN
MOROCCO
ALGERIA
LIBYA
EGYPT
TURKEY
IRAQ
IRAN
KAZAKHSTAN
MONGOLIA
CHINA
JAPAN
INDIA
SAUDI ARABIA
MAURITANIA
MALI
NIGER
CHAD
SUDAN
NIGERIA
ETHIOPIA
SOMALIA
MEXICO
VENEZUELA
COLOMBIA
PERU
BRAZIL
BOLIVIA
CHILE
ARGENTINA
ANGOLA
PHILIPPINES
MALAYSIA
INDONESIA
PAPUA NEW GUINEA
AUSTRALIA
NEW ZEALAND
KIRIBATI
French Polynesia (FR.)
INDIAN OCEAN
SOUTH ATLANTIC OCEAN
French Southern and Antarctic Lands (FR.)
SOUTHERN OCEAN
ANTARCTICA
The World
POLITICAL
Miller Projection

Where is Texas?

DUNE DP Prototype Detector @ CERN

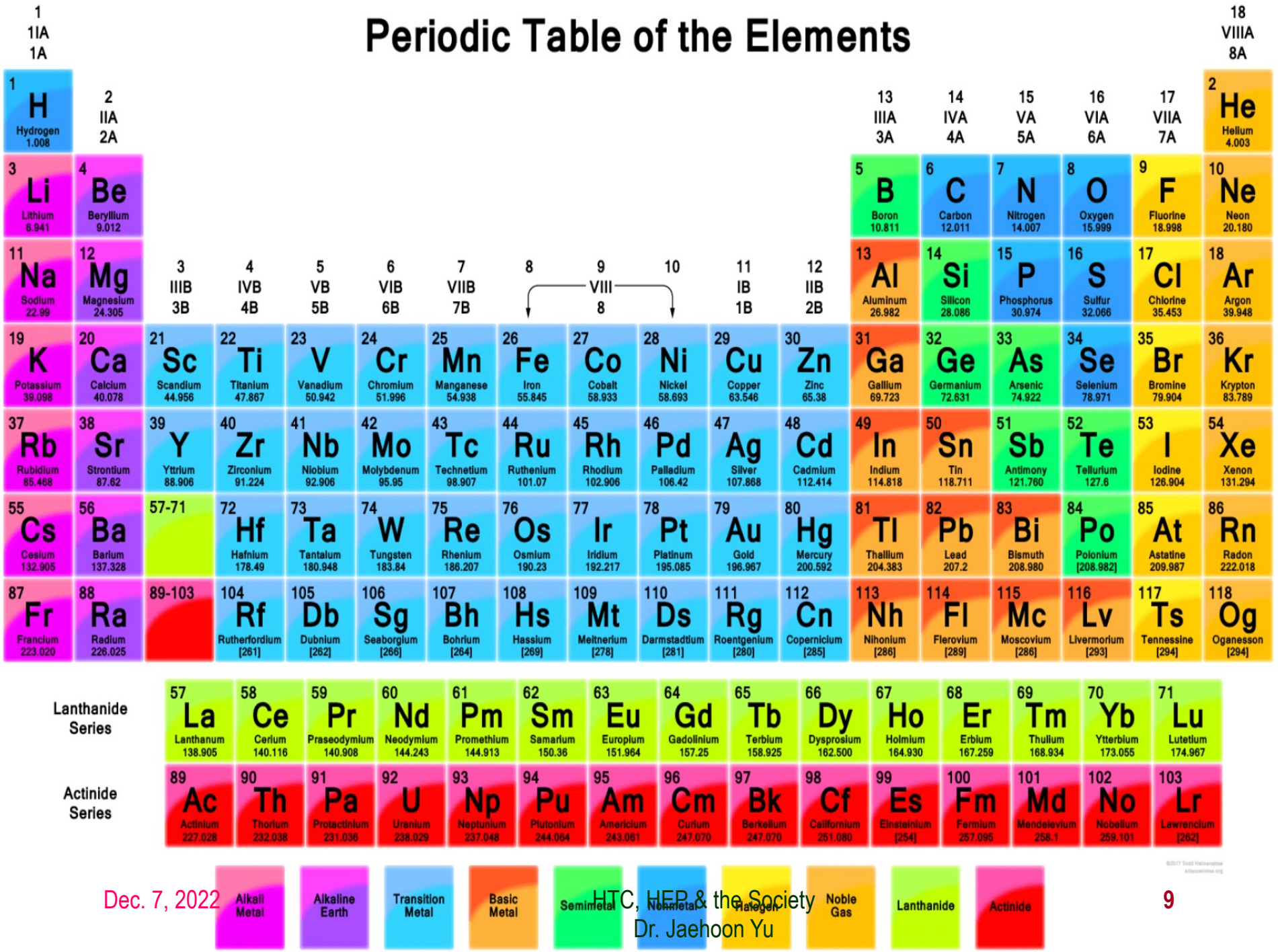


Fermilab Official Poster; photo used in many mass media world-wide

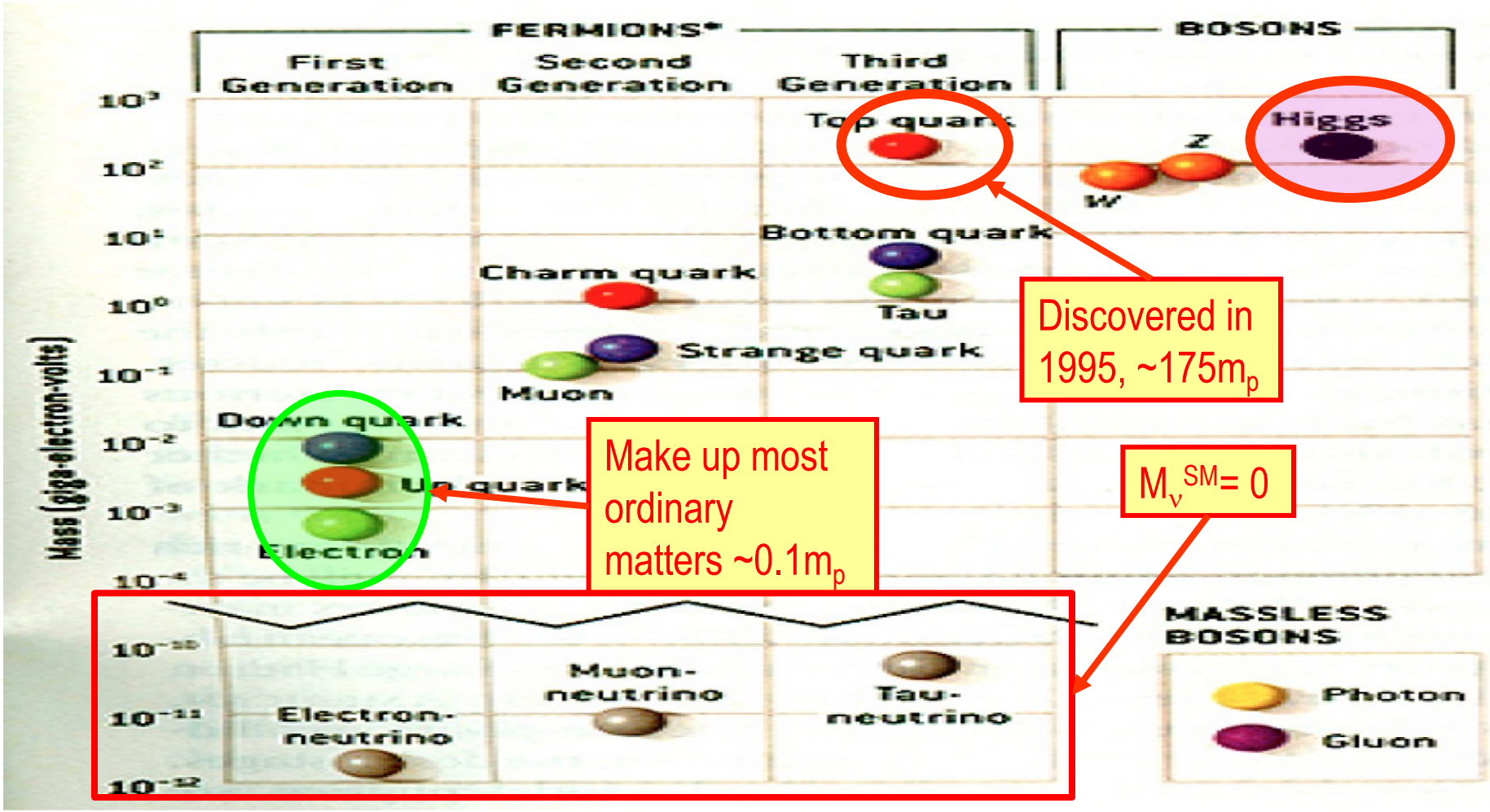
- Organized t… ram in ASP201… at …
  - Secur… In…
- Serving … er
  - Contin… s w/ US NSF
  - Arrang… ool
- Working …
  - Dr. La… ATLAS …
    - Dir…
  - Bright…
    - Has be…
  - Mohamme…

What is High Energy Physics (HEP)?

- The elevator talk: A subfield of physics that seeks to understand what makes up the universe and what the fundamental forces between them are
- Known forces (interactions):
  - Gravitational Force
  - Electromagnetic Force
  - Weak Nuclear Force
  - Strong Nuclear Force
- Current theory: The Standard Model of Particle Physics (SU3xSU2XU1)
- Most important: **Ask yourselves why, what and how**?


Periodic Table of the Elements
1 1IA 1A
18 VIIIA 8A
2 IIA 2A
13 IIIA 3A
14 IVA 4A
15 VA 5A
16 VIA 6A
17 VIIA 7A
3 IIIB 3B
4 IVB 4B
5 VB 5B
6 VIB 6B
7 VIIB 7B
8 9 VIII 10 8
11 IB 1B
12 IIB 2B
1 H Hydrogen 1.008
2 He Helium 4.003
3 Li Lithium 6.941
4 Be Beryllium 9.012
5 B Boron 10.811
6 C Carbon 12.011
7 N Nitrogen 14.007
8 O Oxygen 15.999
9 F Fluorine 18.998
10 Ne Neon 20.180
11 Na Sodium 22.99
12 Mg Magnesium 24.305
13 Al Aluminum 26.982
14 Si Silicon 28.086
15 P Phosphorus 30.974
16 S Sulfur 32.066
17 Cl Chlorine 35.453
18 Ar Argon 39.948
19 K Potassium 39.098
20 Ca Calcium 40.078
21 Sc Scandium 44.956
22 Ti Titanium 47.867
23 V Vanadium 50.942
24 Cr Chromium 51.996
25 Mn Manganese 54.938
26 Fe Iron 55.845
27 Co Cobalt 58.933
28 Ni Nickel 58.693
29 Cu Copper 63.546
30 Zn Zinc 65.38
31 Ga Gallium 69.723
32 Ge Germanium 72.631
33 As Arsenic 74.922
34 Se Selenium 78.971
35 Br Bromine 79.904
36 Kr Krypton 83.789
37 Rb Rubidium 85.468
38 Sr Strontium 87.62
39 Y Yttrium 88.906
40 Zr Zirconium 91.224
41 Nb Niobium 92.906
42 Mo Molybdenum 95.95
43 Tc Technetium 98.907
44 Ru Ruthenium 101.07
45 Rh Rhodium 102.906
46 Pd Palladium 106.42
47 Ag Silver 107.868
48 Cd Cadmium 112.414
49 In Indium 114.818
50 Sn Tin 118.711
51 Sb Antimony 121.760
52 Te Tellurium 127.6
53 I Iodine 126.904
54 Xe Xenon 131.294
55 Cs Cesium 132.905
56 Ba Barium 137.328
57-71
72 Hf Hafnium 178.49
73 Ta Tantalum 180.948
74 W Tungsten 183.84
75 Re Rhenium 186.207
76 Os Osmium 190.23
77 Ir Iridium 192.217
78 Pt Platinum 195.085
79 Au Gold 196.967
80 Hg Mercury 200.592
81 Tl Thallium 204.383
82 Pb Lead 207.2
83 Bi Bismuth 208.980
84 Po Polonium [208.982]
85 At Astatine 209.987
86 Rn Radon 222.018
87 Fr Francium 223.020
88 Ra Radium 226.025
89-103
104 Rf Rutherfordium [261]
105 Db Dubnium [262]
106 Sg Seaborgium [266]
107 Bh Bohrium [264]
108 Hs Hassium [269]
109 Mt Meitnerium [278]
110 Ds Darmstadtium [281]
111 Rg Roentgenium [280]
112 Cn Copernicium [285]
113 Nh Nihonium [286]
114 Fl Flerovium [289]
115 Mc Moscovium [286]
116 Lv Livermorium [293]
117 Ts Tennessine [294]
118 Og Oganesson [294]
Lanthanide Series
57 La Lanthanum 138.905
58 Ce Cerium 140.116
59 Pr Praseodymium 140.908
60 Nd Neodymium 144.243
61 Pm Promethium 144.913
62 Sm Samarium 150.36
63 Eu Europium 151.964
64 Gd Gadolinium 157.25
65 Tb Terbium 158.925
66 Dy Dysprosium 162.500
67 Ho Holmium 164.930
68 Er Erbium 167.259
69 Tm Thulium 168.934
70 Yb Ytterbium 173.055
71 Lu Lutetium 174.967
Actinide Series
89 Ac Actinium 227.028
90 Th Thorium 232.038
91 Pa Protactinium 231.036
92 U Uranium 238.029
93 Np Neptunium 237.048
94 Pu Plutonium 244.064
95 Am Americium 243.061
96 Cm Curium 247.070
97 Bk Berkelium 247.070
98 Cf Californium 251.080
99 Es Einsteinium [254]
100 Fm Fermium 257.095
101 Md Mendelevium 258.1
102 No Nobelium 259.101
103 Lr Lawrencium [262]
Alkali Metal
Alkaline Earth
Transition Metal
Basic Metal
Semimetal
Nonmetal
Halogen
Noble Gas
Lanthanide
Actinide

HEP and the Standard Model

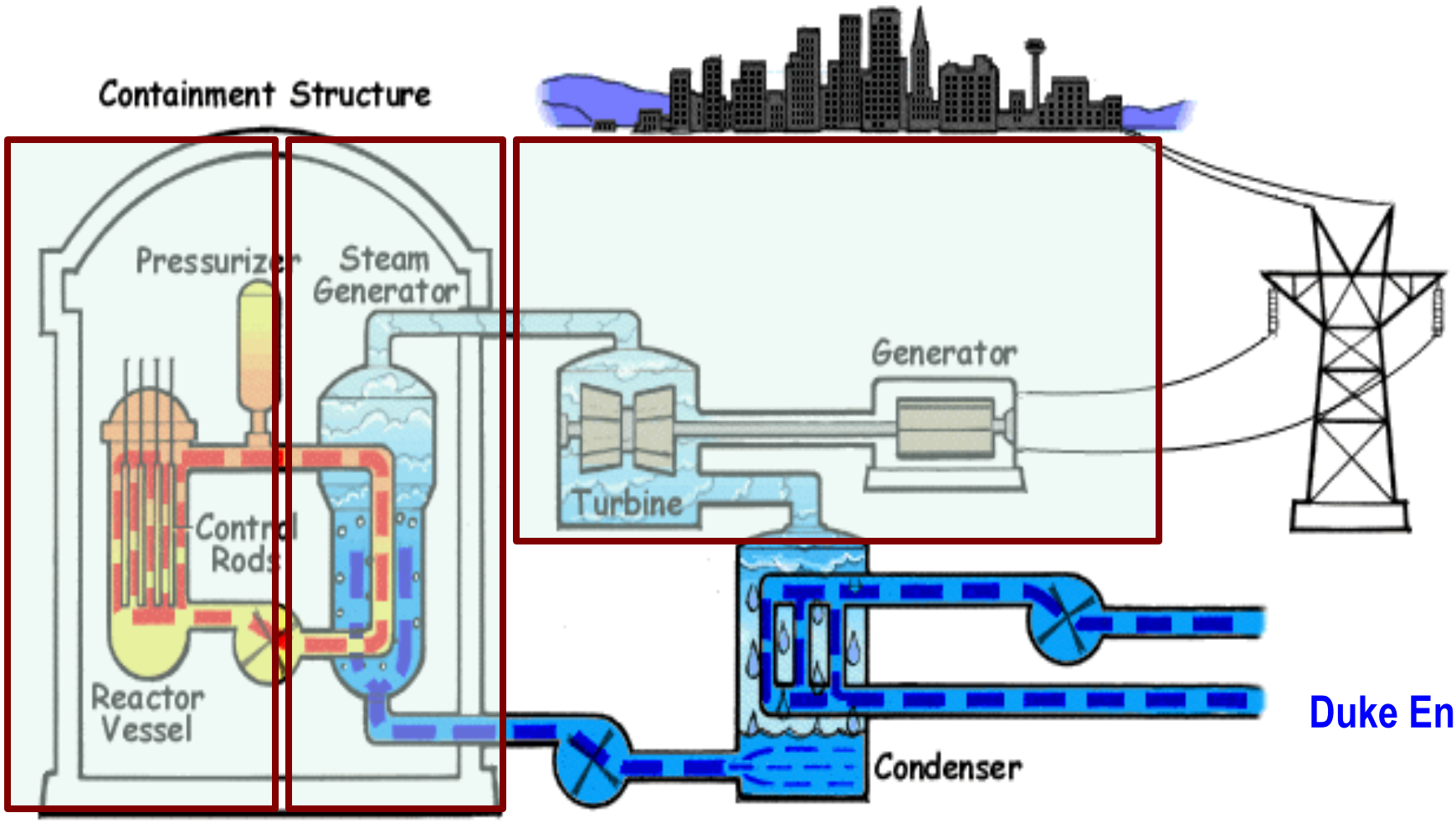


- Total of 16 particles (12+4 force mediators) make up all the visible matter in the universe! ➔ Simple and elegant!!!
- Tested to a precision better than 1 part per million!

Some remaining issues in HEP..

- Why is the mass range so large ($0.1m_p$ – $175\ m_p$)?
- Is the particle discovered at the LHC really the Higgs particle?
- Why is the matter in the universe made only of particles?
- Neutrinos have mass!! (**OMG!! The SM is broken!!!**)
  - What are the mixing parameters, particle-anti particle asymmetry and the neutrino mass ordering?
- Why are there only four apparent forces?
  - Were they all unified at the Big Bang?

How does a nuclear power plant work?



Duke Energy

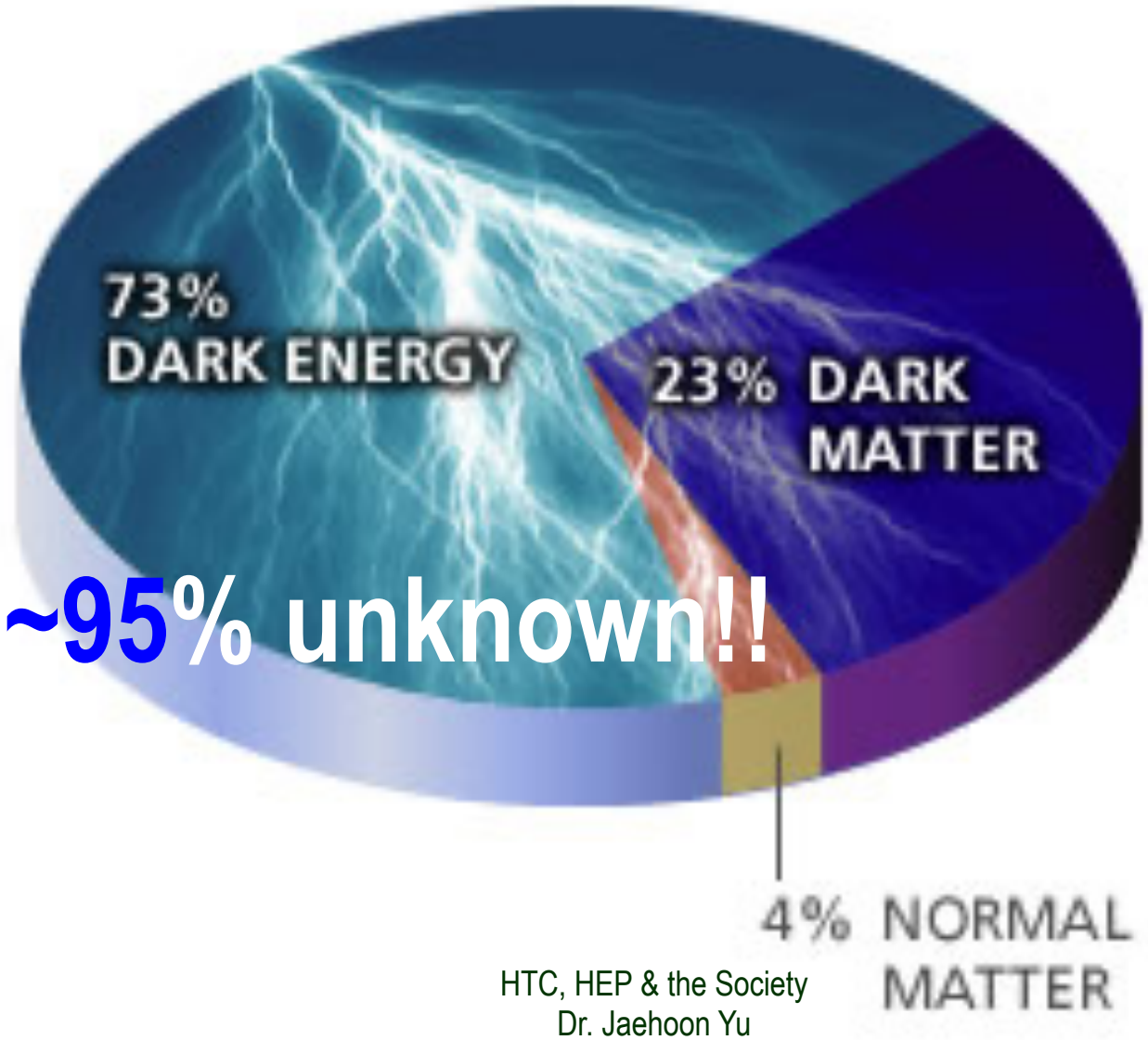
My 1000 year dream: Skip the whole thing!

Make electricity directly from nuclear forces!

So what's the problem?

- Why is the mass range so large ($0.1m_p$ – $175\ m_p$)?
- Is the particle we discovered really the Higgs particle?
- Why is the matter in the universe made only of particles?
- Neutrinos have mass!! What are the mixing parameters, particle-anti particle asymmetry and mass ordering?
- Why are there only four apparent forces?
  - Were they all unified at the Big Bang?
- Is the picture of the universe we present the real thing?

What makes up the universe?



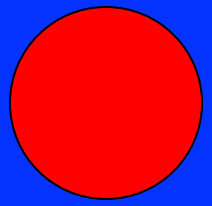
~95% unknown!!

So what's the problem?

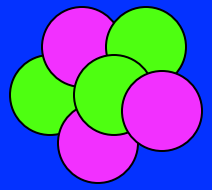
- Why is the mass range so large ($0.1m_p$ – $175\ m_p$)?
- Is the particle we discovered really the Higgs particle?
- Why is the matter in the universe made only of particles?
- Neutrinos have mass!! What are the mixing parameters, particle-anti particle asymmetry and mass ordering?
- Why are there only four apparent forces?
  - Were they all unified at the Big Bang?
- Is the picture of the universe we present the real thing?
- Are there any other particles we don't know of?
  - Big deal for the new LHC Run that started now and in the new experiments starting up in the US!
- Where do we all come from?
- **Can we live well in the universe as an integral partner?**

Accelerators are Powerful Microscopes.

They make high energy particle beams
that allow us to see small things.



seen by
low energy beam
(poorer resolution)



seen by
high energy beam
(better resolution)

Accelerators are also Time Machines.

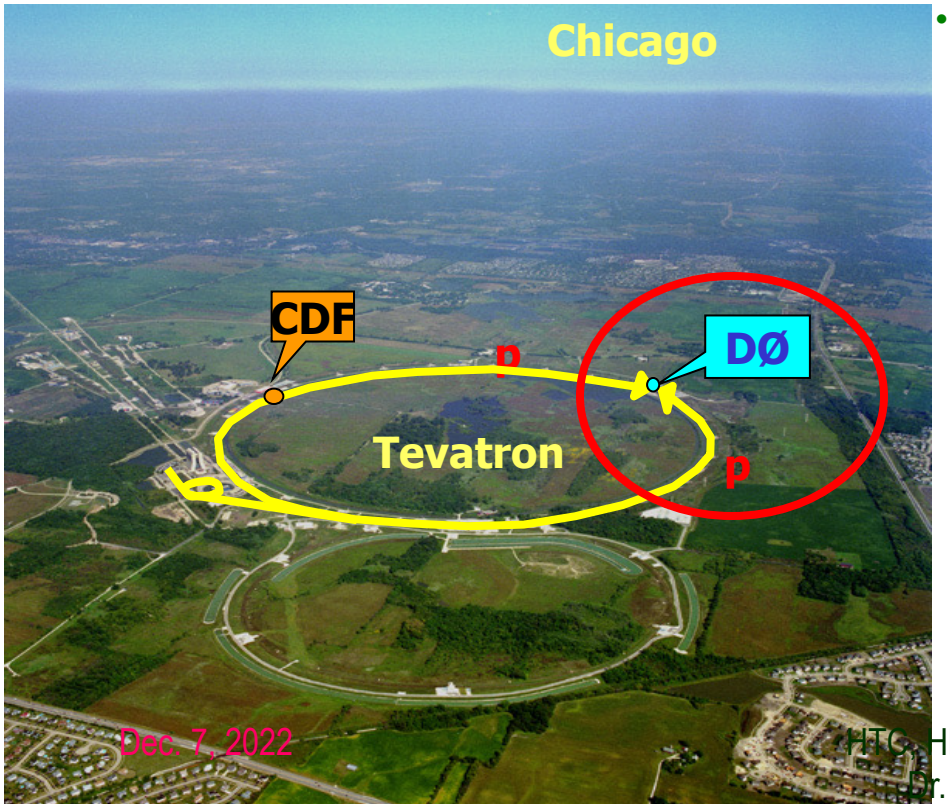
They make particles last seen
in the earliest moments of the universe.

Particle and anti-particle annihilate.

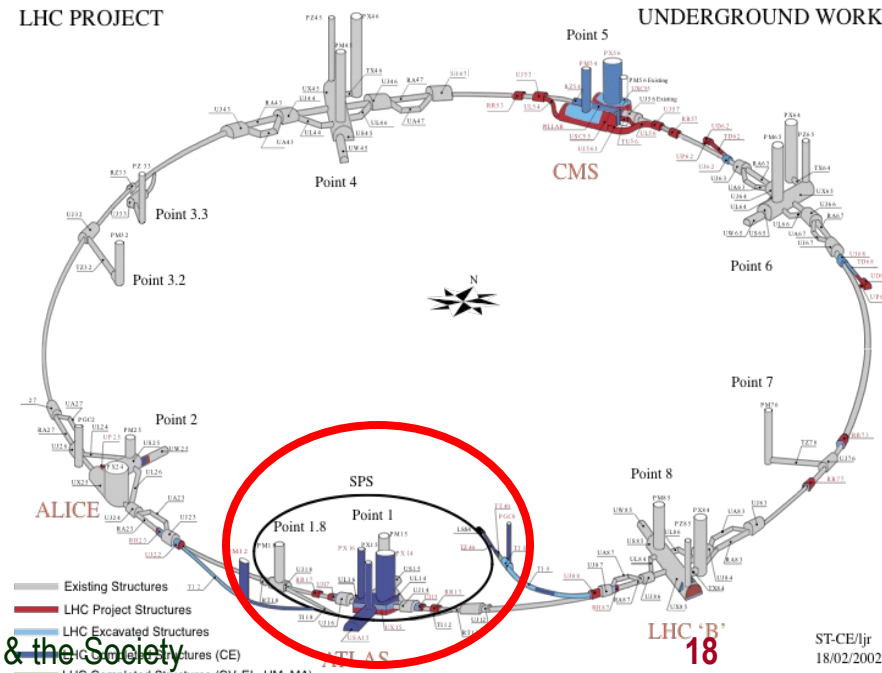
$$E = mc^2$$

Fermilab Tevatron and LHC at CERN

- World's Highest Energy proton-anti-proton collider
  - 4km (2.5mi) circumference
  - E_{cm}=2 TeV (=6.3x10^{-7}J/p➔ 13M Joules on the area smaller than 10^{-4}m^2)
  - Same as the KE of a 20t truck w/ speed 130km/hr
    - ~100,000 times the energy density at the ground 0 of the Hiroshima atom bomb
  - **Tevatron was shut down in 2011**
  - **New frontiers with high intensity proton beams including the search for dark matter with beams!!**



- World's Highest Energy p-p collider
  - 27km (17mi) circumference, 100m (300ft) underground
  - Design E_{cm}=14 TeV (=44x10^{-7}J/p➔ 362M Joules on the area smaller than 10^{-4}m^2)
  - ➢ KE of a B727 (80t) w/ speed 310km/hr
    - ➢ ~3M times the energy density at the ground 0 of the Hiroshima atom bomb
- **Discovered a new heavy particle that looks Higgs in 2012**
- **Search for new particles has been ongoing!!**
- **The LHC started back up in 2021 at high intensity**



LHC @ CERN Aerial View

CMS

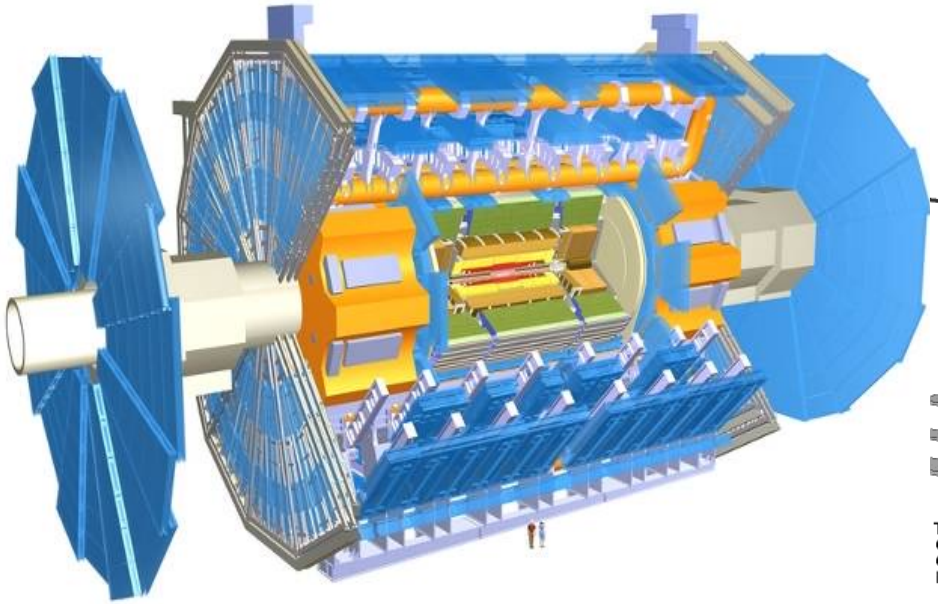
France

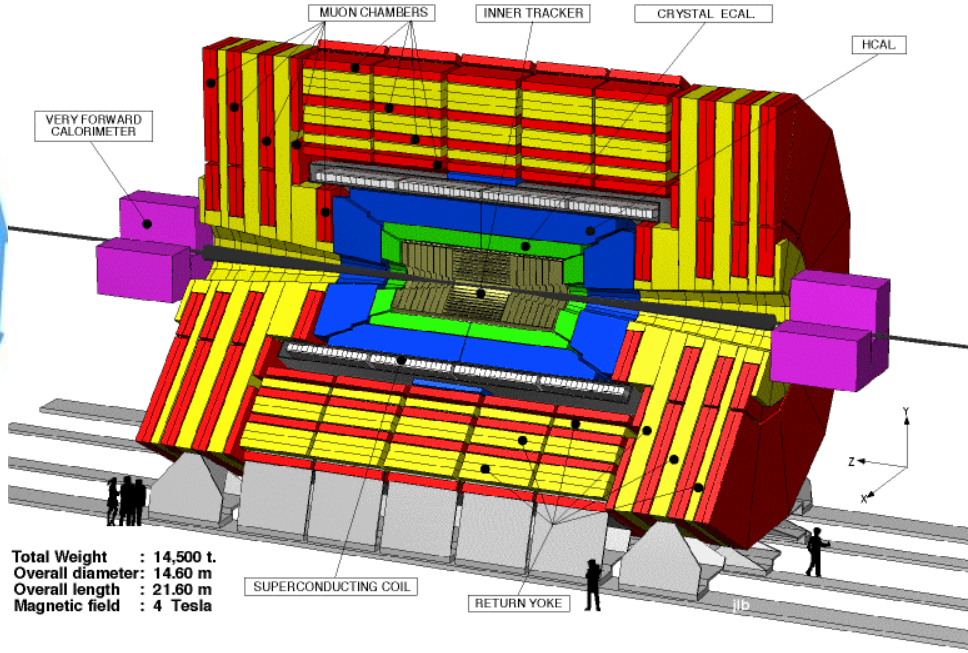
Geneva Airport

ATLAS

Swizerland

The ATLAS and CMS Detectors





- **Weighs 7000 tons and ~10 story tall**
- Records **200 – 400 collisions/second (out of 50million)**
- Records approximately **350** MB/second
- Records >2 PB per year ➔ **200*Printed material of the US Lib. of Congress**

200x

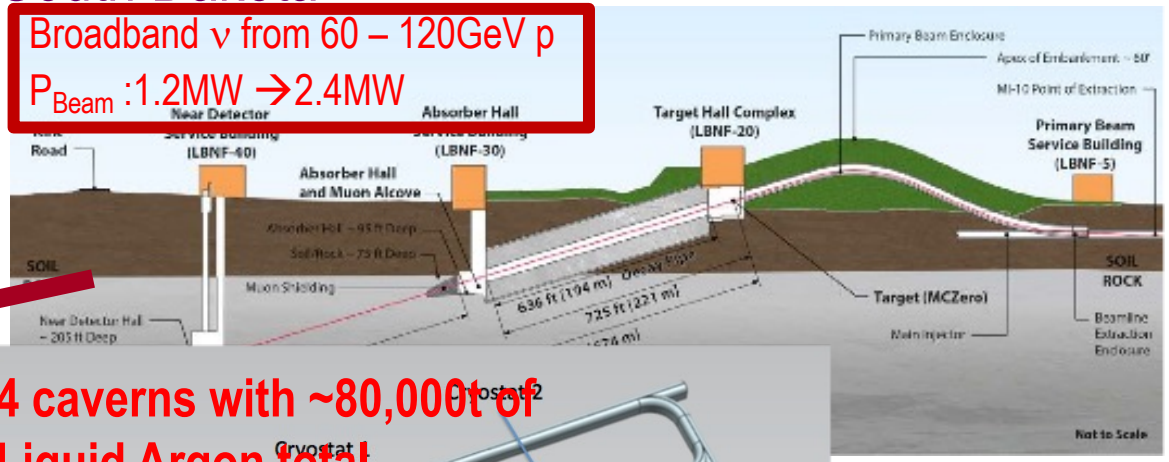
The Next Big Thing - DUNE

DEEP UNDERGROUND NEUTRINO EXPERIMENT

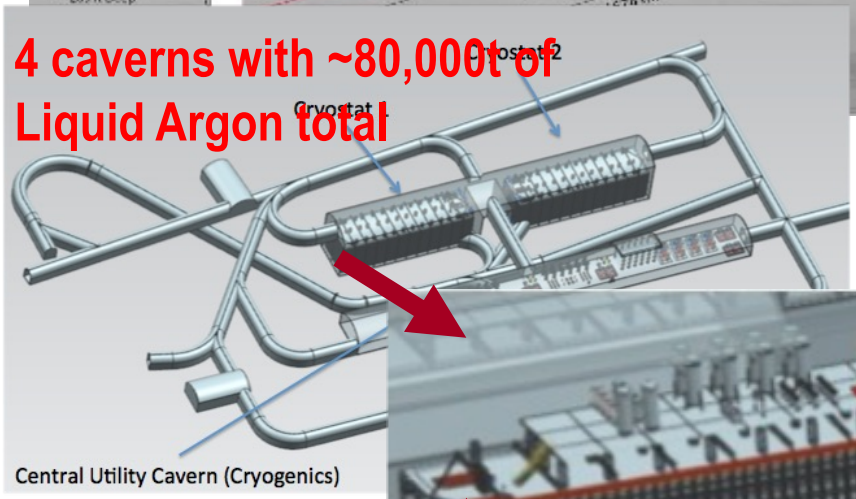
- Stands for Deep Under Ground Neutrino Experiment
- The $2.5B US flagship long baseline (1300km) ν experiment
  - 1500m underground in South Dakota

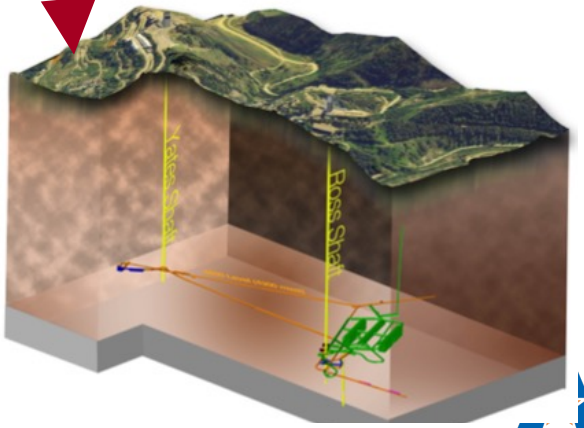
Broadband ν from 60 – 120GeV p

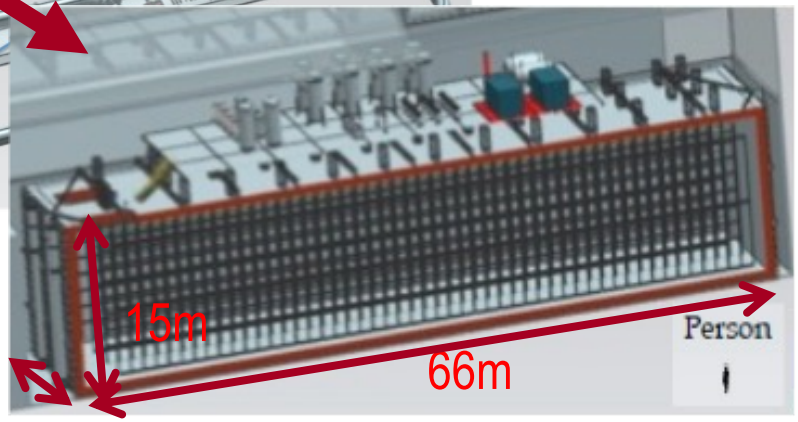
P_{Beam} :1.2MW →2.4MW



4 caverns with ~80,000t of Liquid Argon total



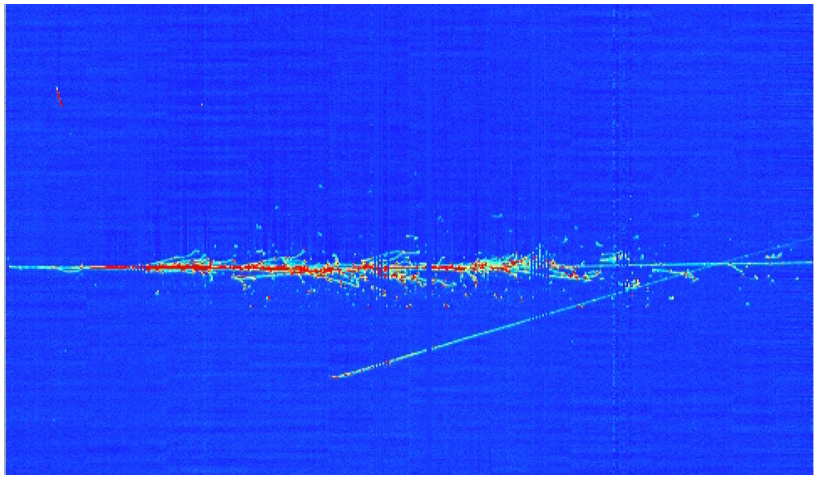




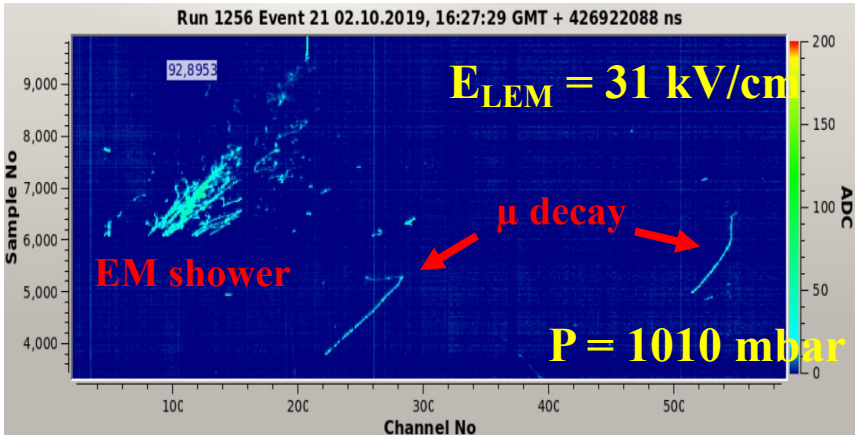
15m

Images in DUNE LAr-TPC Prototypes

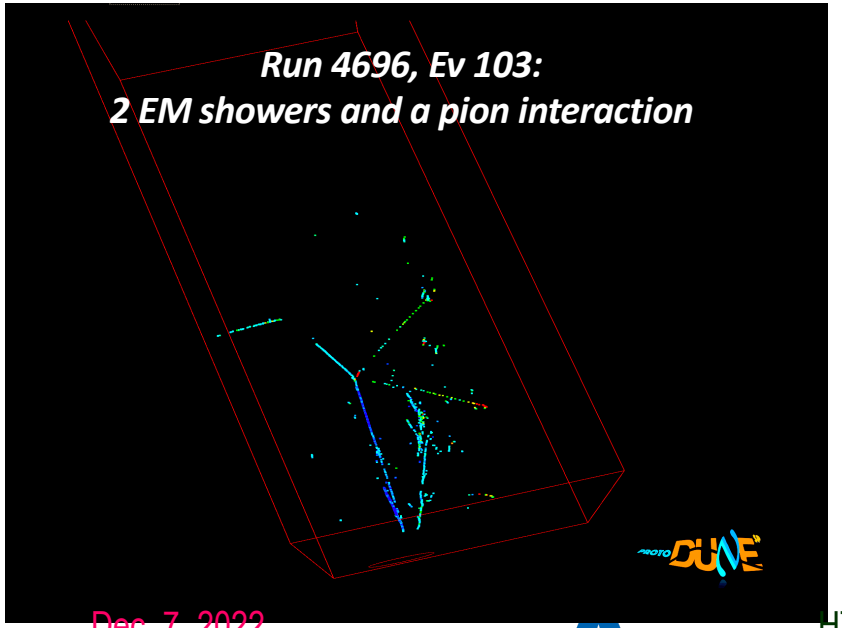
Throughgoing μ



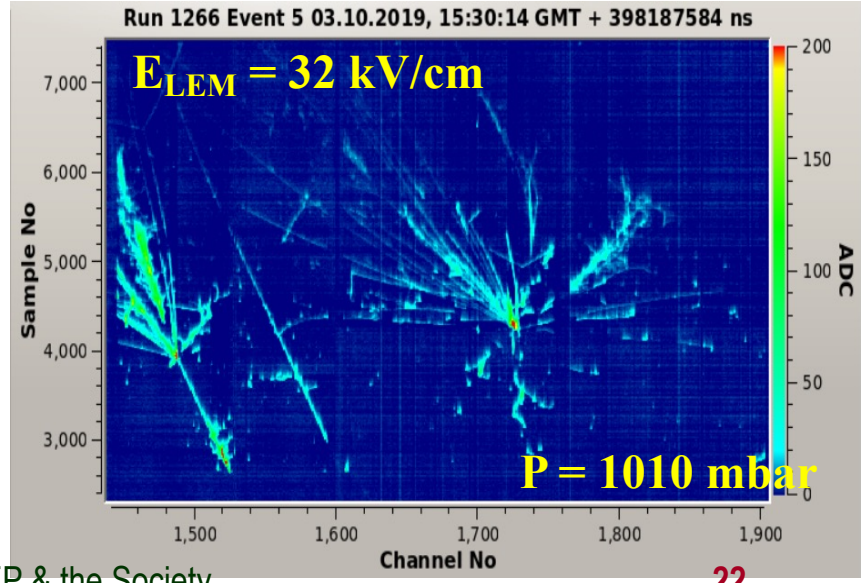
Electromagnetic shower + two muon decays

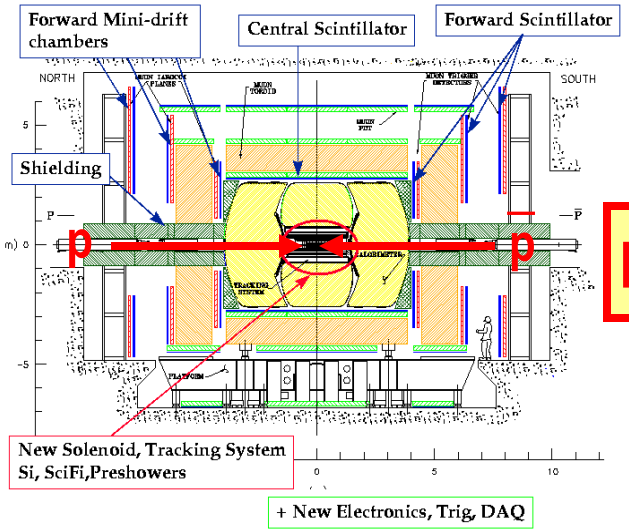


Run 4696, Ev 103:
2 EM showers and a pion interaction



Multiple hadronic interactions in a shower





Digital data

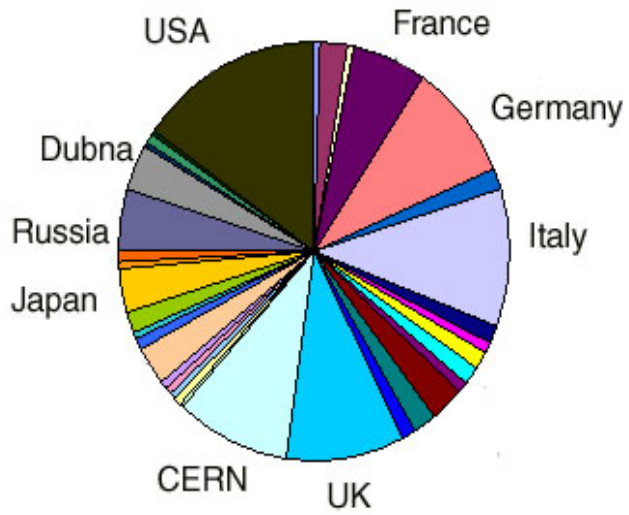
Data Reconstruction

The Problem

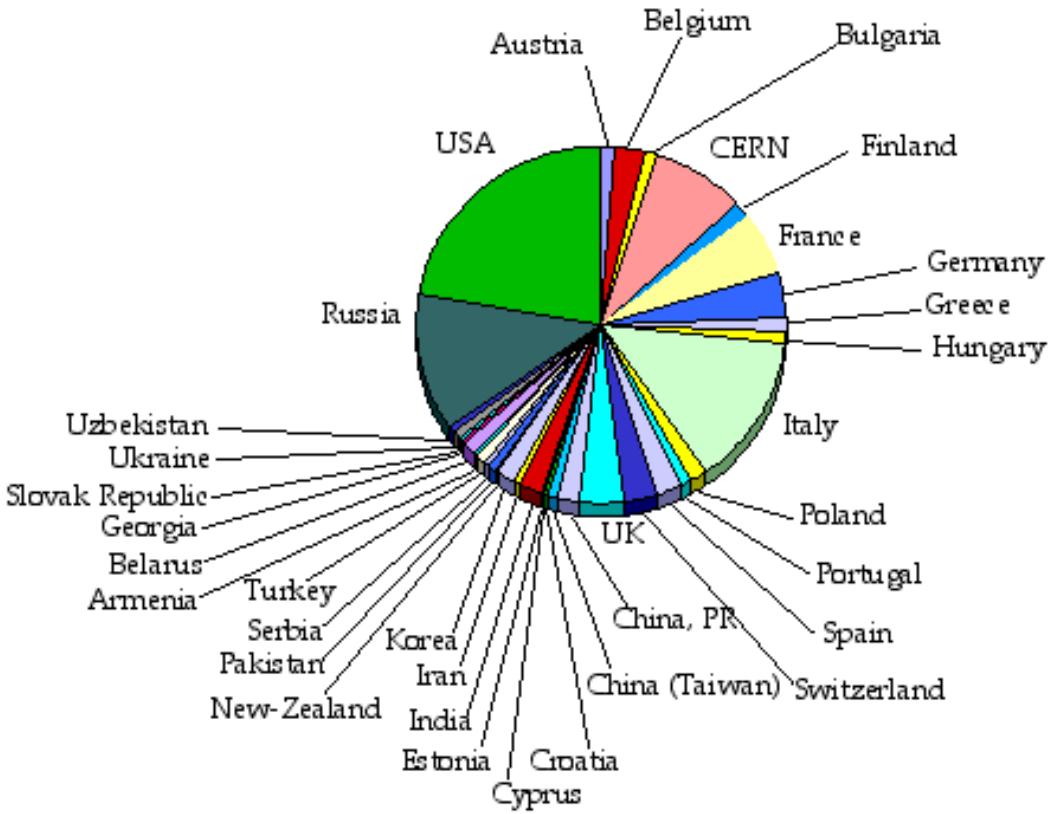
- Detectors are complicated and large ➔ Need large number of collaborators
  - They are scattered all over the world!

LHC Collaborations

ATLAS



CMS



ATLAS+CMS over 6000 Physicists and Engineers
Over 60 Countries, 250 Institutions

The Map of the DUNE Experiment

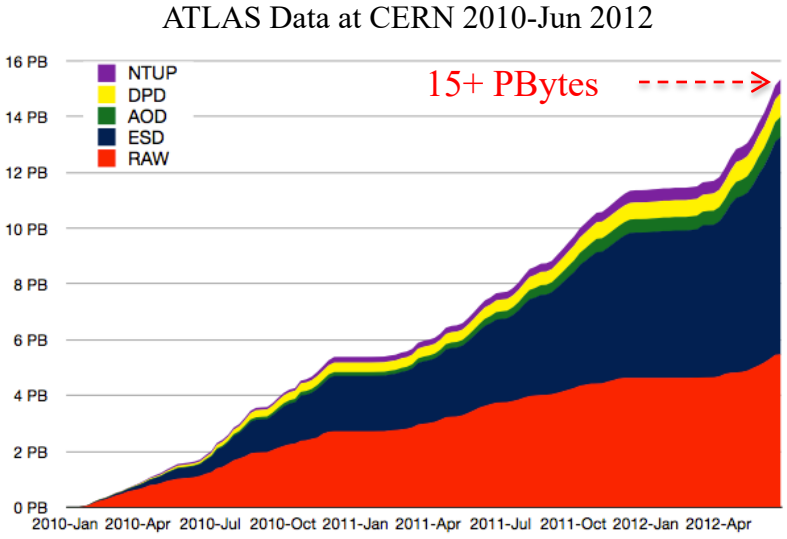
DEEP UNDERGROUND NEUTRINO EXPERIMENT

>1400 collaborators
>200 institutions
>30 countries + CERN

https://www.dunescience.org/about-the-collaboration/

The Problem

- Detectors are complicated and large ➔ Need a large number of collaborators
  - They are scattered all over the world!
  - How do we get them communicate quickly and efficiently?
  - How do we leverage collaborators' capabilities?
  - How do we efficiently utilize all the computing resources?
- Data size is large >>10 PB per year for raw data only
  - Entire data set 15+PB on disc
  - Where and how to store the la
  - How do we allow collaborators to access data in an efficient fa

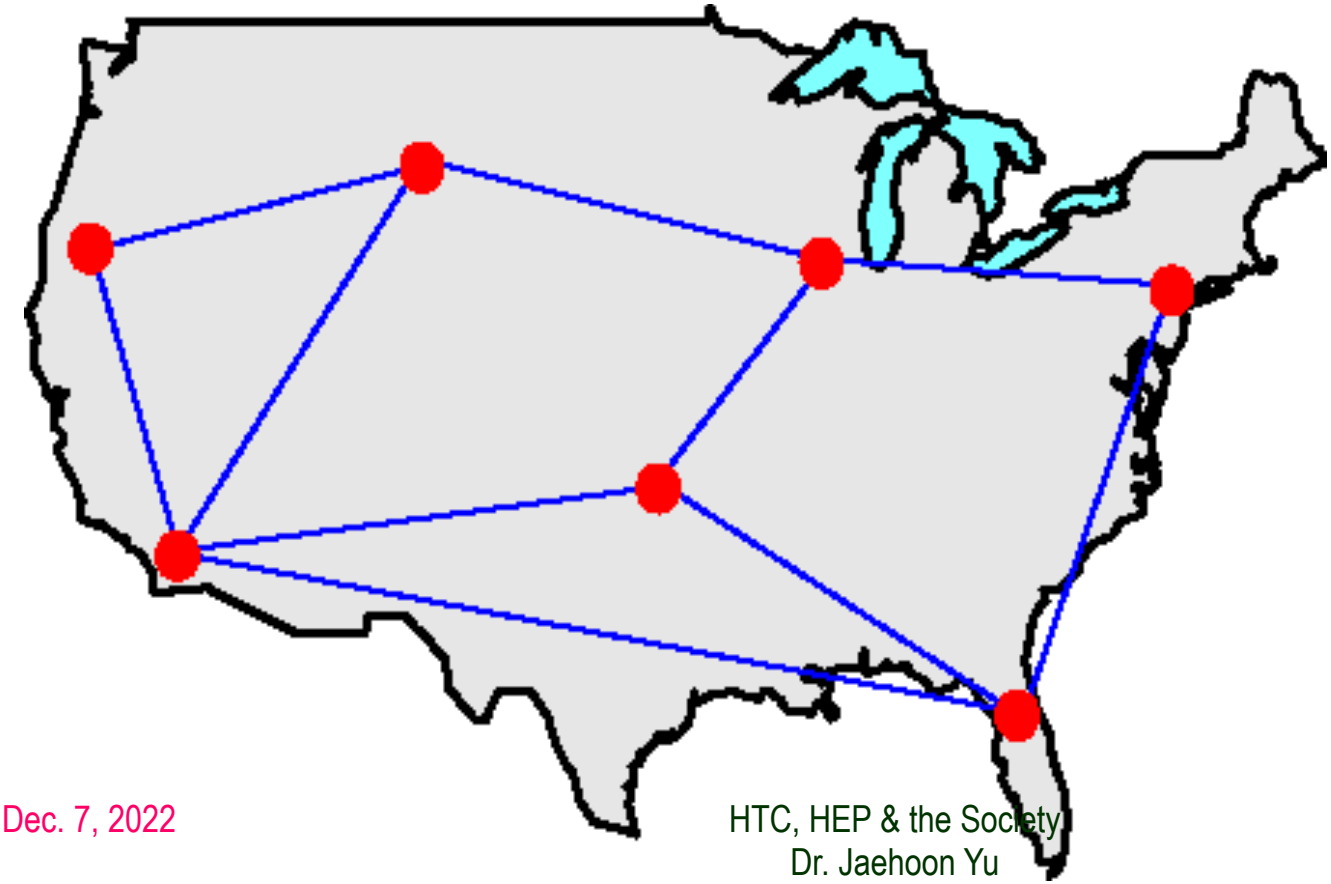


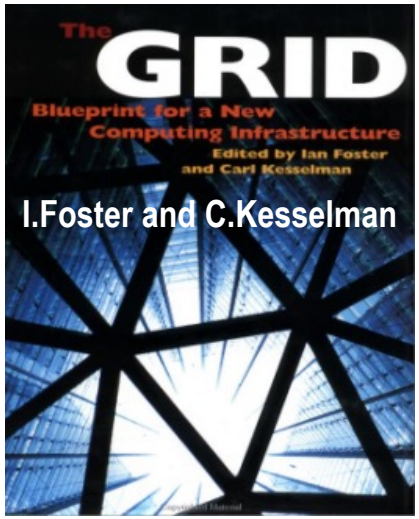
The Problem, cont'd

- How do we allow people's analysis jobs to access data and make progress rapidly and securely?
  - What is the most efficient way to get jobs' requirements matched with resources?
  - Should jobs go to data or data go to jobs?
  - What level of security should there be?
- How do we allow experiments to reconstruct data and generate the large amount of simulated events quickly?
  - How do we garner the necessary compute and storage resources effectively and efficiently?
  - What network capabilities do we need in the world?
- How do we get people to analyze at their desktops?

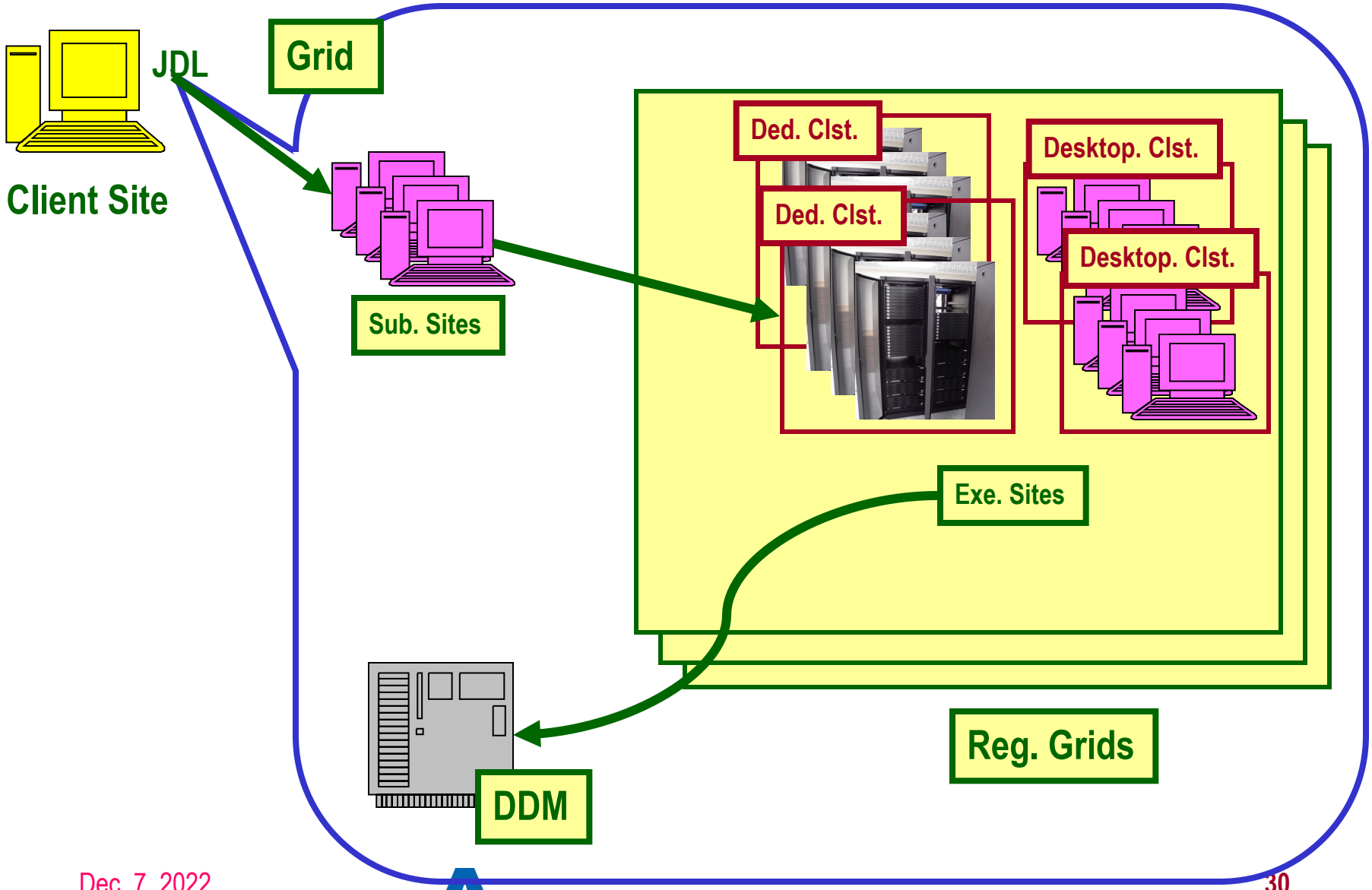
What is a Computing Grid?

- Grid, the definition: Geographically distributed computing resources configured for a coordinated use
- Physical resources & good network provide hardware capability
- The "Middleware" software ties them together

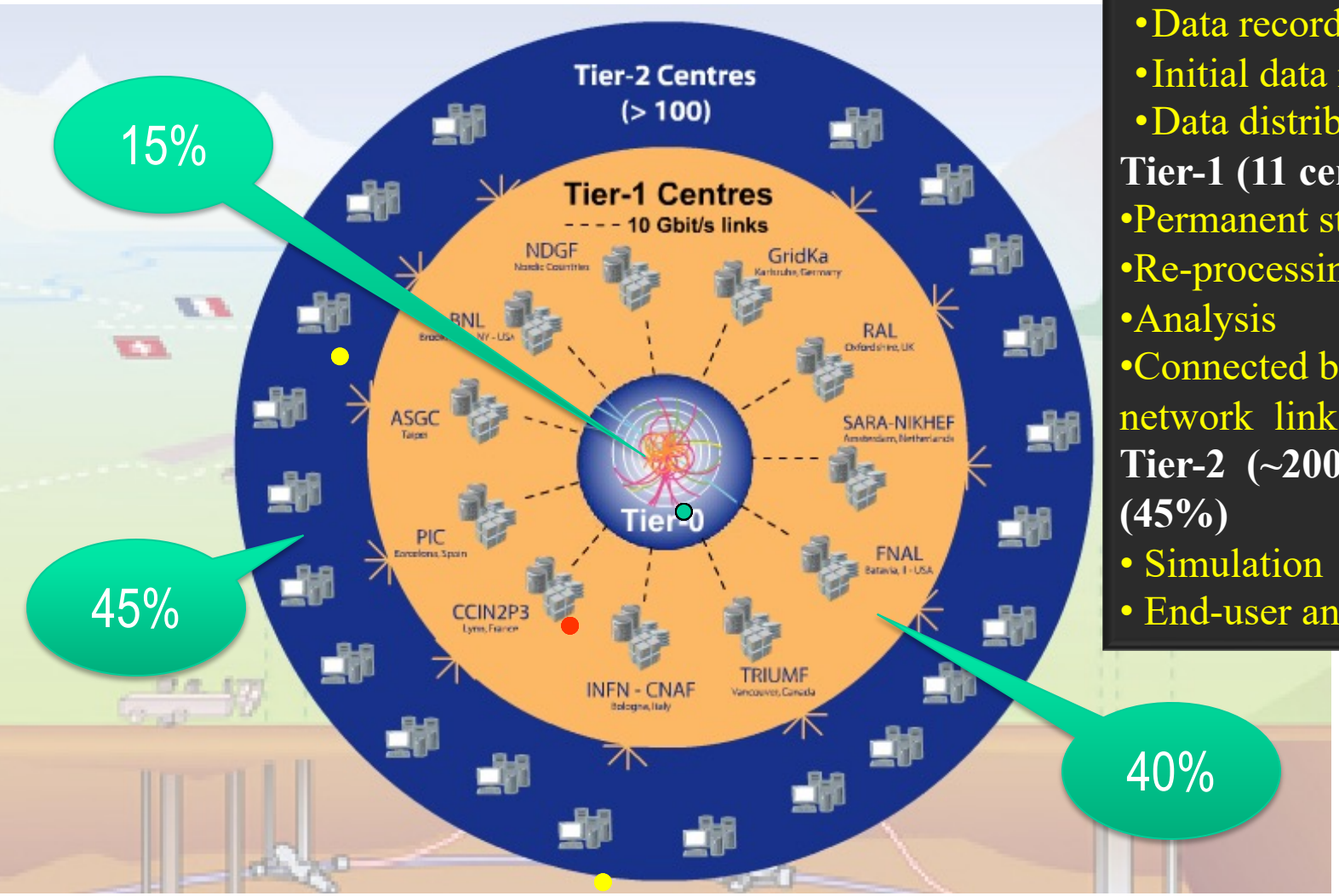




How does a computing Grid work?



Implemented ATLAS Grid Structure



Tier-0 (CERN): (15%)
- Data recording
- Initial data reconstruction
- Data distribution

Tier-1 (11 centres): (40%)
- Permanent storage
- Re-processing
- Analysis
- Connected by direct 10 Gb/s network links

Tier-2 (~200 centres): (45%)
- Simulation
- End-user analysis

NORDUGRID
Grid Solution for Wide Area Computing and Data Handling

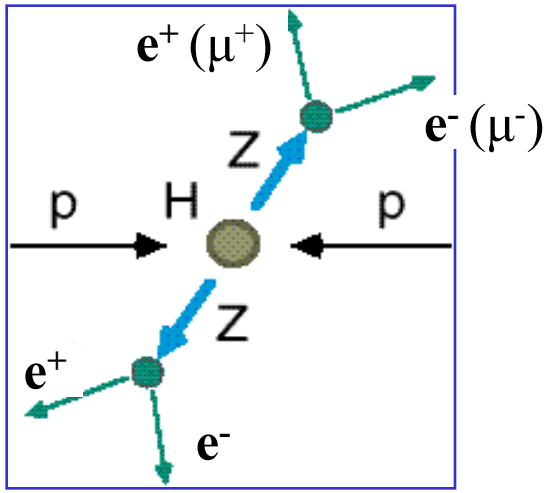
Open Science Grid

How to look for rare particles?

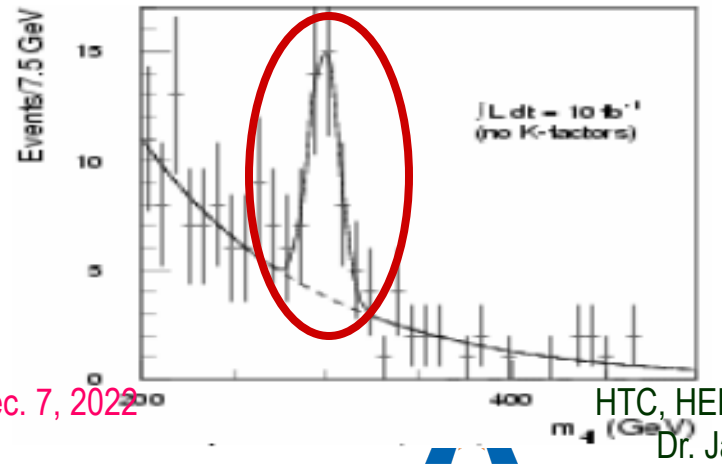
- Many of these rare particle are so heavy they decay into other lighter particles instantaneously
- When one searches for a new particle, one looks for the easiest way to get at them
- Of many signatures of the rare particle final states, some are much easier to find ➔e.g. for the Standard Model Higgs particle
  - H➔ $\gamma\gamma$
  - H ➔ ZZ* ➔ 4e, 4μ, 2e2μ, 2e2ν and 2μ2ν
  - H➔ WW*➔2e2ν and 2μ2ν
  - And many more complicated signatures

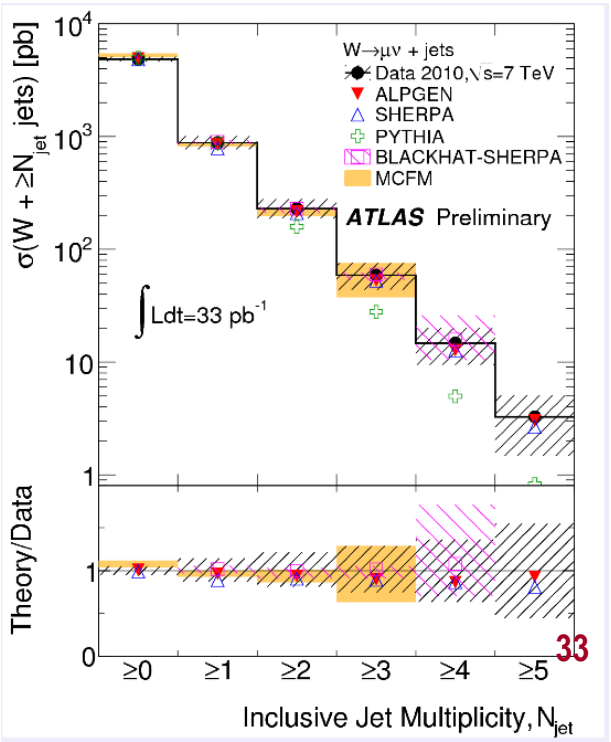
How do we look for a rare particle?

1. Identify Higgs candidate events

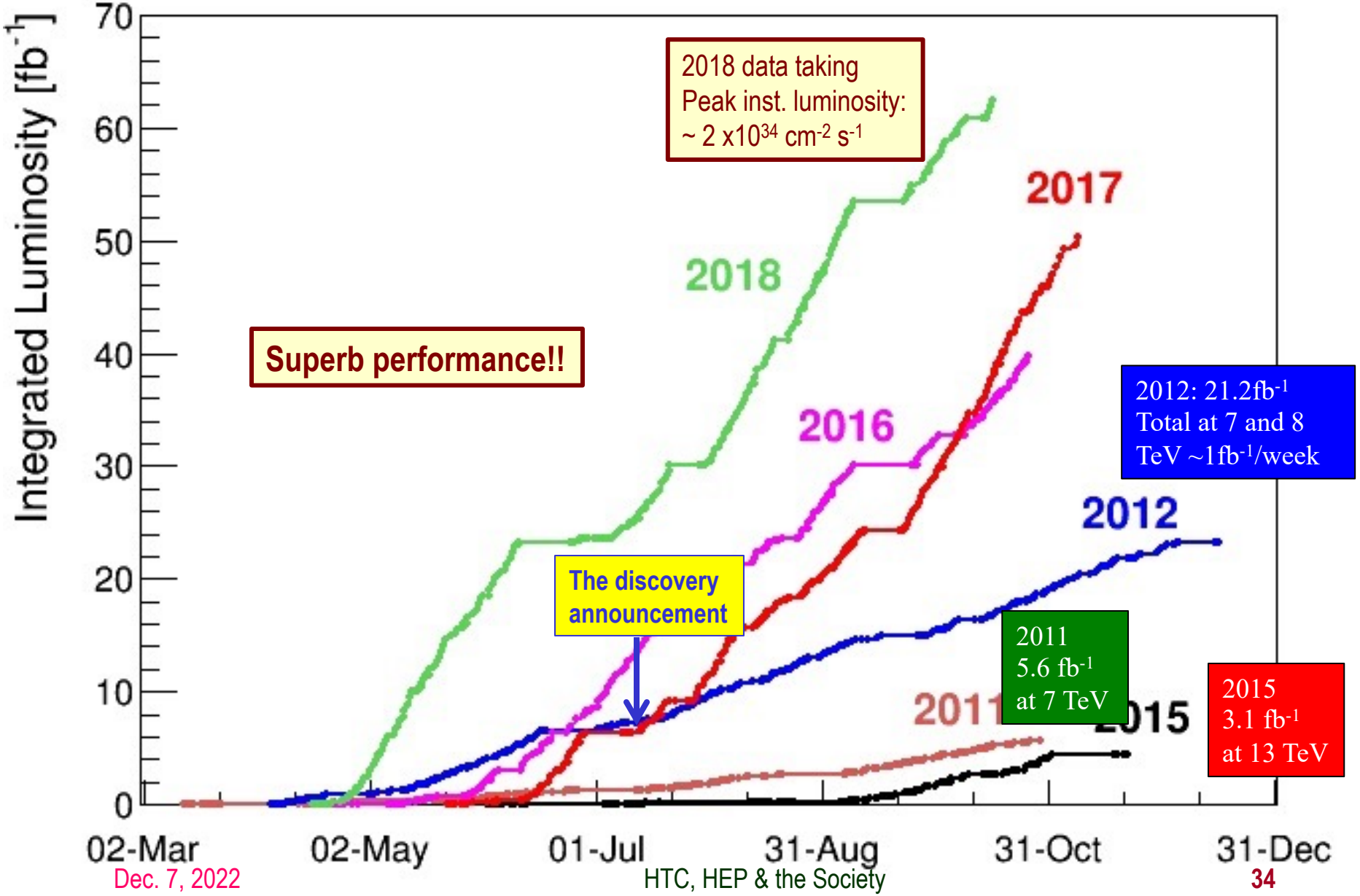


2. Understand fakes (backgrounds)
3. Look for a bump!!
   - Large amount of data absolutely critical

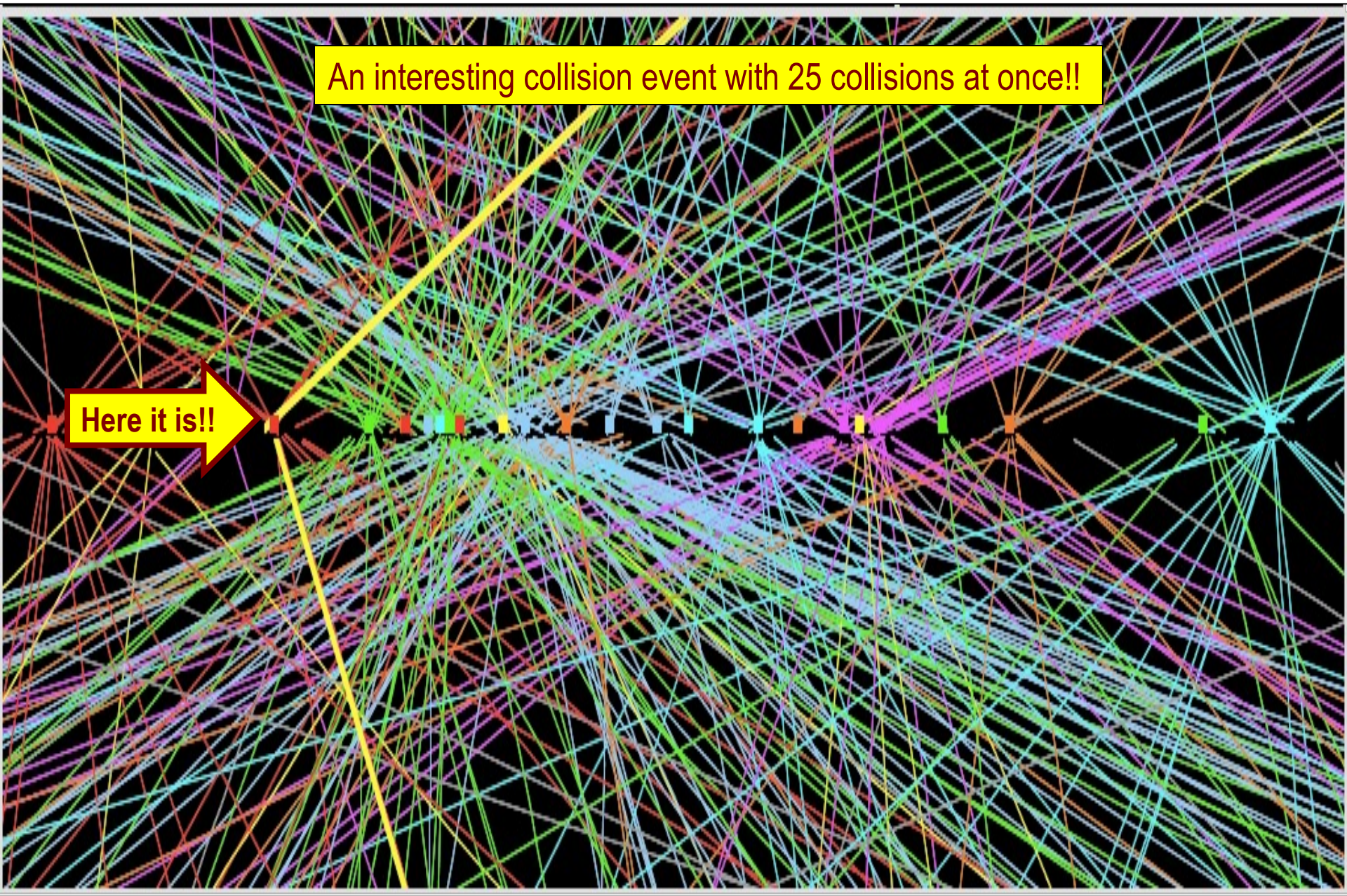




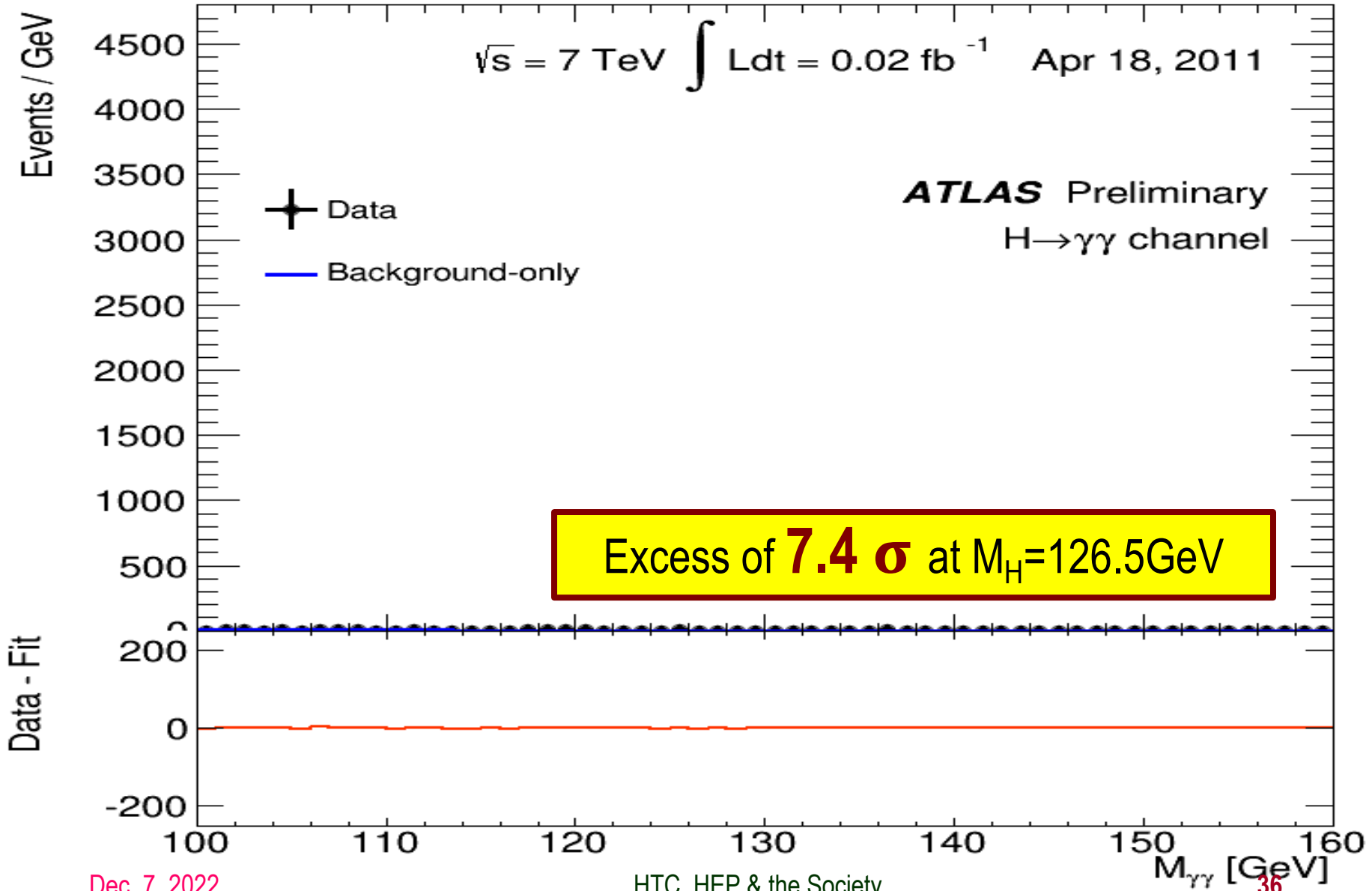
Amount of the LHC Data



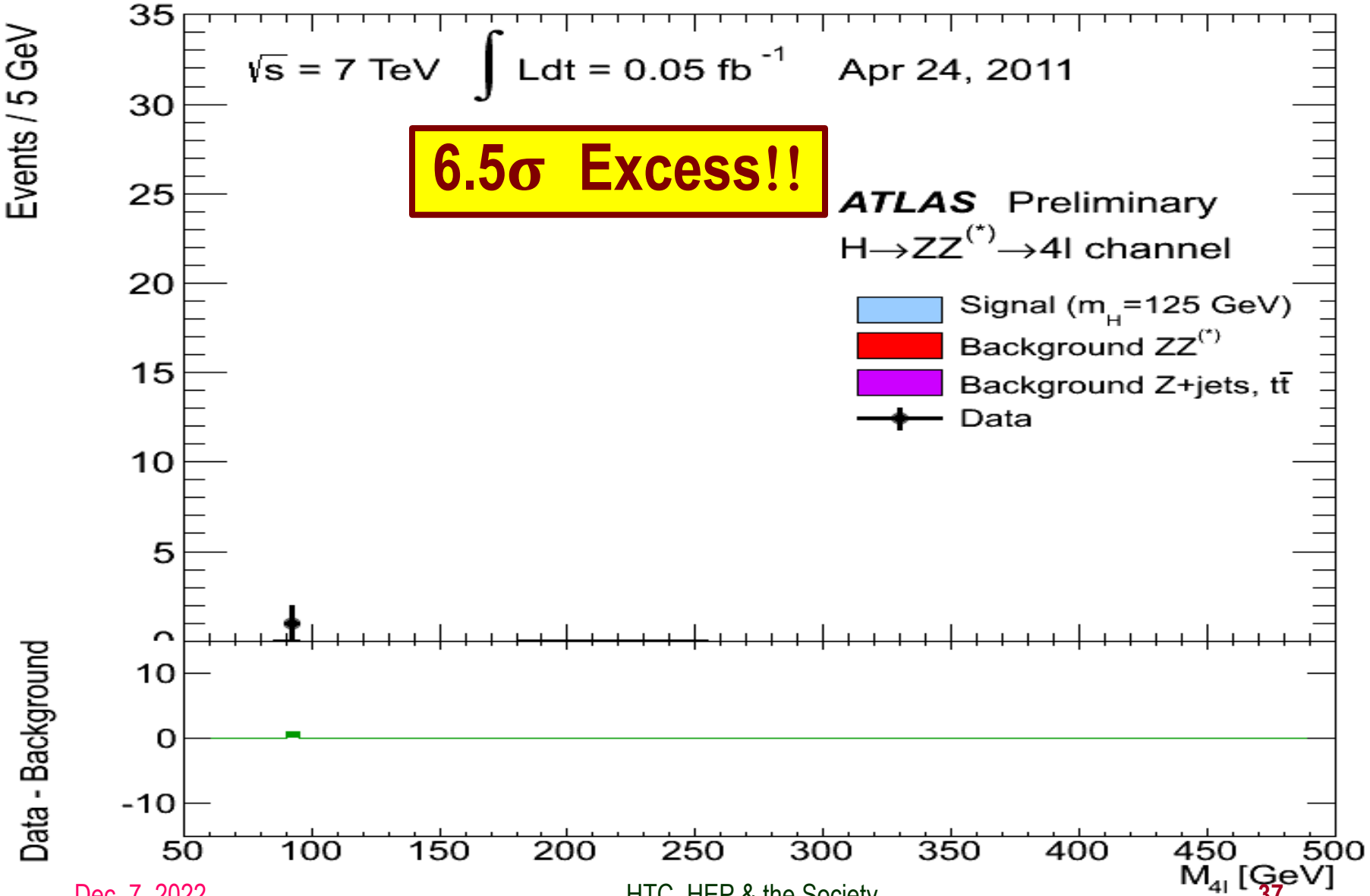
Challenges? No problem!



What did statistics do for Higgs→ γγ?

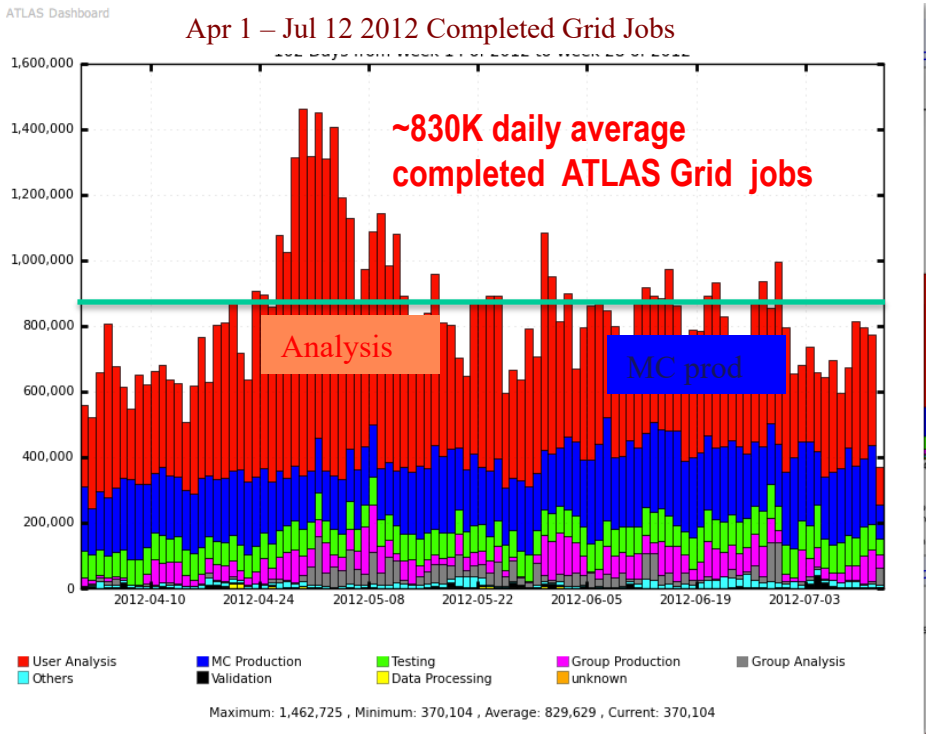


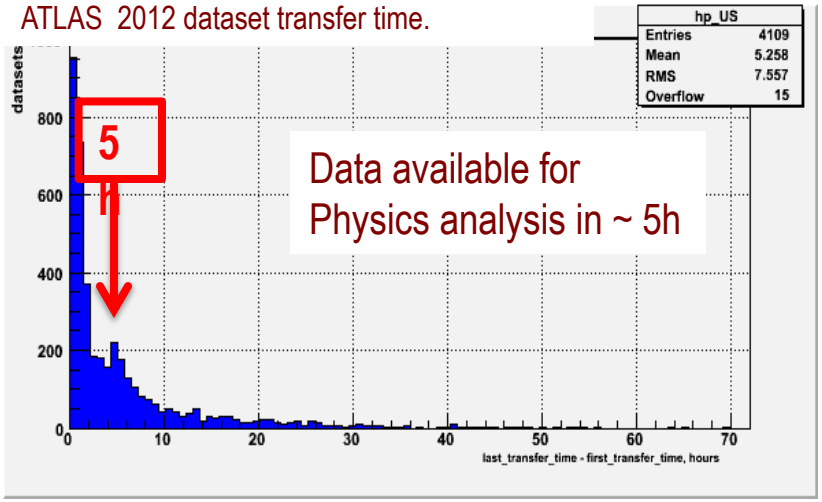
ATLAS Mass Bump Plot ($H \rightarrow 4\ell$)?

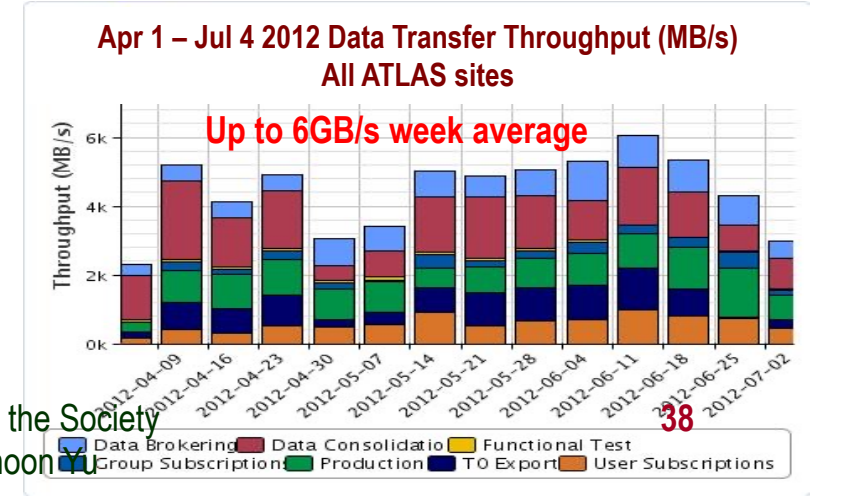


Performance of the Grid for LHC

- ATLAS Distributed Computing on the Grid : 10 Tier-1s + CERN + ~70 Tier-2s +…(more than 80 Production sites)
- High volume, high throughput process through fast network!!

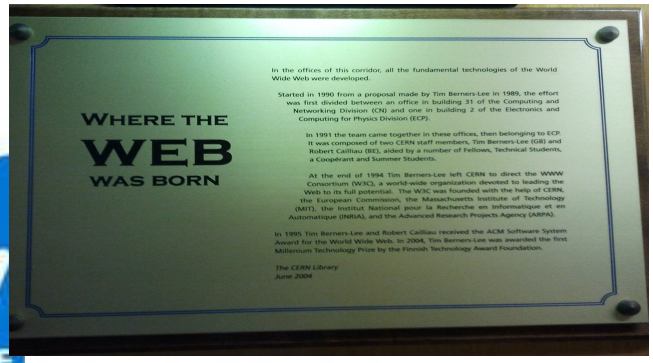






The commercial world picked up..

Early 90's



2004

facebook

1996

amazon

1998

NETFLIX

2006

zoom

amazon web services EC2

Many private entities fully utilized the internet communication we've developed to multi-trillion dollar venture!!

The concept of cloud and the HTC turned into a new area of study, the Data Science!!

Dropbox

Google Cloud Platform

So why is HEP relevant to me?

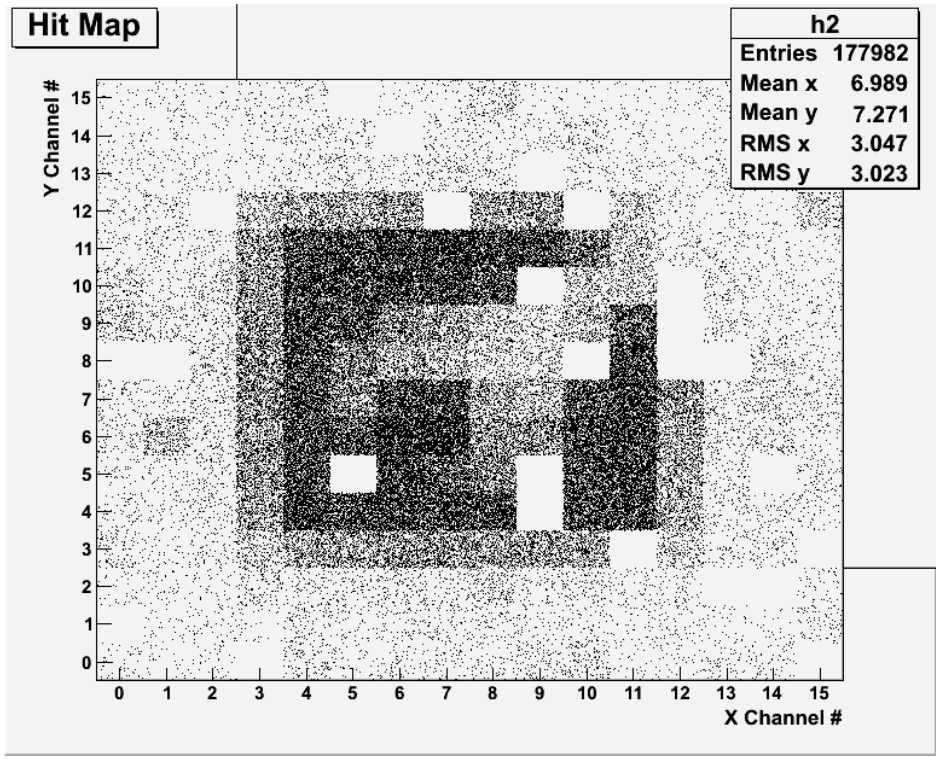
- HEP explores the most fundamental nature of the universe!
- The discovery of the dark matter and making of dark matter beams will take us to the next quantum level
- Discoveries will realize our 1000 year dreams
- Outcome and bi-products of HEP research improves our daily lives directly and indirectly
  - WWW came from HEP

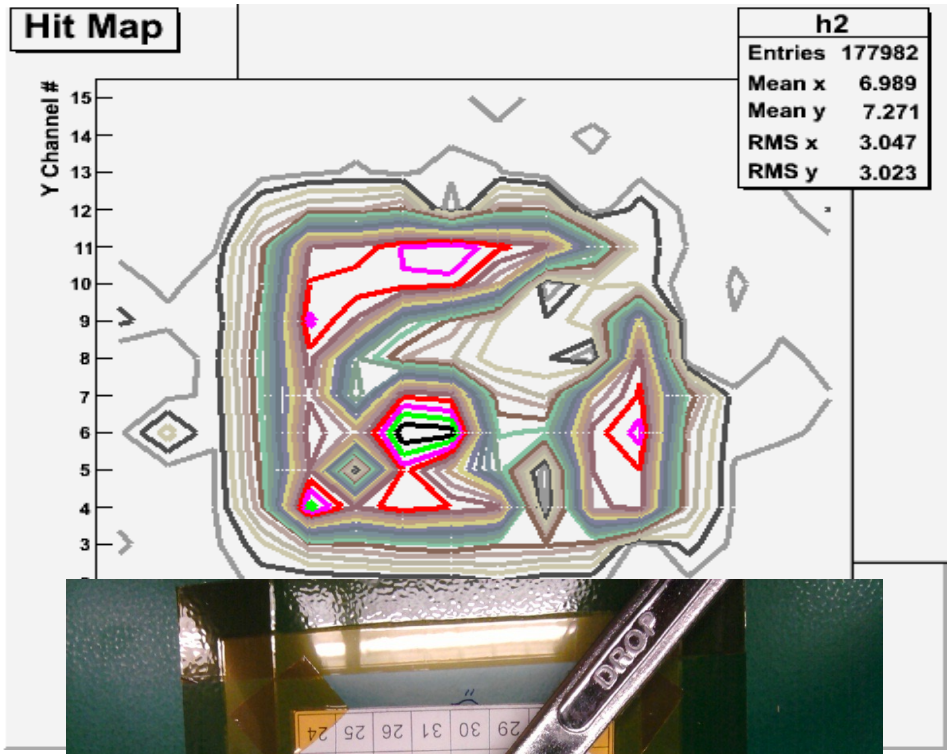

WHERE THE
WEB
WAS BORN

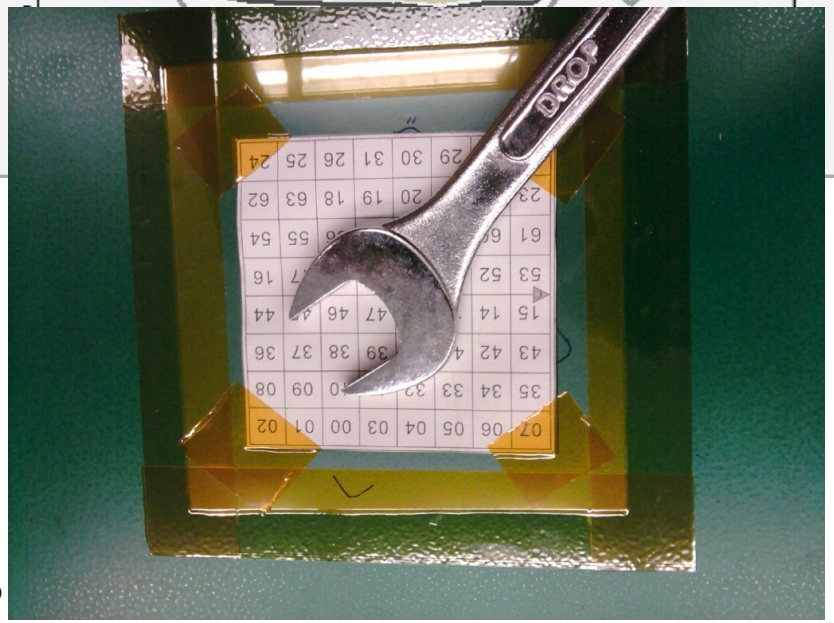
So why is HEP relevant to me?

- HEP explores the most fundamentals of the Universe!
- Discoveries will realize our 1000 year dreams
- The discovery of the dark matter and making of dark matter beams will take us to the next Quantum level
- Outcome and bi-products of HEP research improves our daily lives directly and indirectly
  - WWW came from HEP
  - Advanced detector technologies like GEM will make a large screen low dosage X-ray imaging possible

Bi-product of High Energy Physics Research







Can you tell what the object is?
(GEM Detector X-ray Image)

So are we done with the HTC?

- LHC has performed extremely well!
- The data size will increase by over 10 fold in HLLHC
  - Computing will be stressed even further!!
- Grid computing infrastructure has served well thus far
  - LHC users process PBs of data & billions of jobs
- High Intensity, large scale experiments, such as DUNE will record even larger amount of data than the LHC
- Identified limits in databases scalability, CPU resources, storage utilization, etc, are being addressed
- Utilization of quantum computing and machine learning technologies actively sought

Conclusions

- HEP is an exciting endeavor in understanding the universe
- In the quest for the origin of the universe, High Energy Physics
  - Uses accelerators to "look" into extremely small distances
  - Uses large detectors to explore nature and uncover secrets of universe
  - Uses large number of computers to process data in a timely fashion
  - Large amount of accumulated data ➔ computing grid performed marvelously for expeditious data analyses
- Physics analyses at one's own desktop using computing grid sitting behind has happened and is improving fast!!
- Computing grid used in other disciplines with large data sets
- Computing grid fully integrated into everyday lives
  - The pandemic accelerated this process
- A true computing grid is revolutionizing everyday lives ….

HEP's Impacts to the Society?

- WWW and other advanced computing technologies from HEP greatly reduced the physical distances between us
  - Help freeing oppressed people and protecting their freedom
  - Keeping people from being imprisoned by their physical limitations or even by a pandemic
- HTC generates petrillions (=1000 trillions) dollars of economy
- Data science becomes a major area of education
  - Helps recording and analyzing enormous data in the COVID-19 fight
- All these technologies that can do good things, however, are instead harmful if used by those lack humanity and fundamental human decency
  - See how spreading misinformation hurts the very humanity we care!!
- Be a good person first with a heart toward the good of humanity

Let's all dream,

not just for tomorrow,

not even just for the next year,

but for 1000 years into the future for the whole humanity!!

Additional Materials

FFT: Number of beam particles per sec?

- What is the number of particles per second for an accelerator facility that can provide:
  - **P** MW of total beam power
  - of charged particles of energy **E** GeV?

$$N_p\left(\ /\sec; E\ GeV; P\ MW\right) = P/E \cdot 6.3\times10^{15}\left(particles/\sec\right)$$

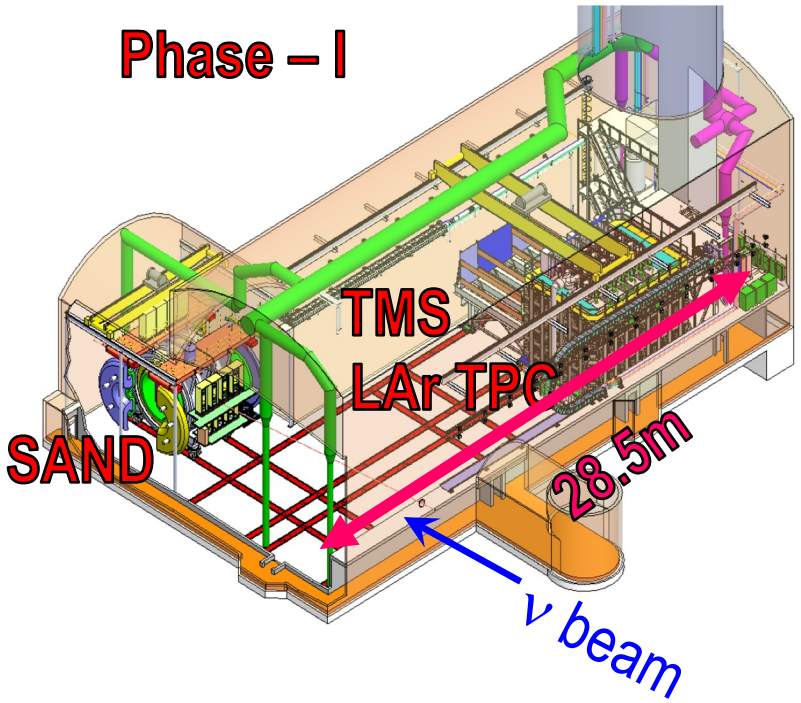
  - What is the number of protons per second for 120GeV beams at 1.2MW?

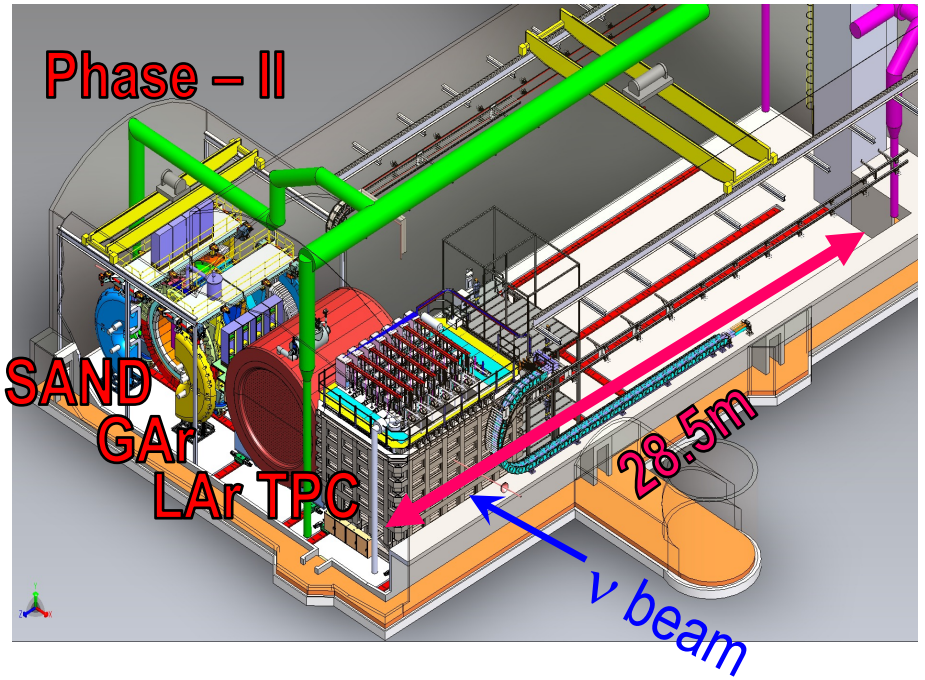
$$N_p\left(\ /\sec; 120\ GeV; 1.2\ MW\right) = \frac{1.2}{120}\cdot 6.3\times10^{15}\left(particles/\sec\right)$$
$$= 6.3\times10^{13}\left(particles/\sec\right)$$

  - What is the beam current? $I = N_p \cdot 1.6\times10^{-19}$

$$= 1.2\times10^{-5}\left(C/\sec\right) = 12\mu A$$

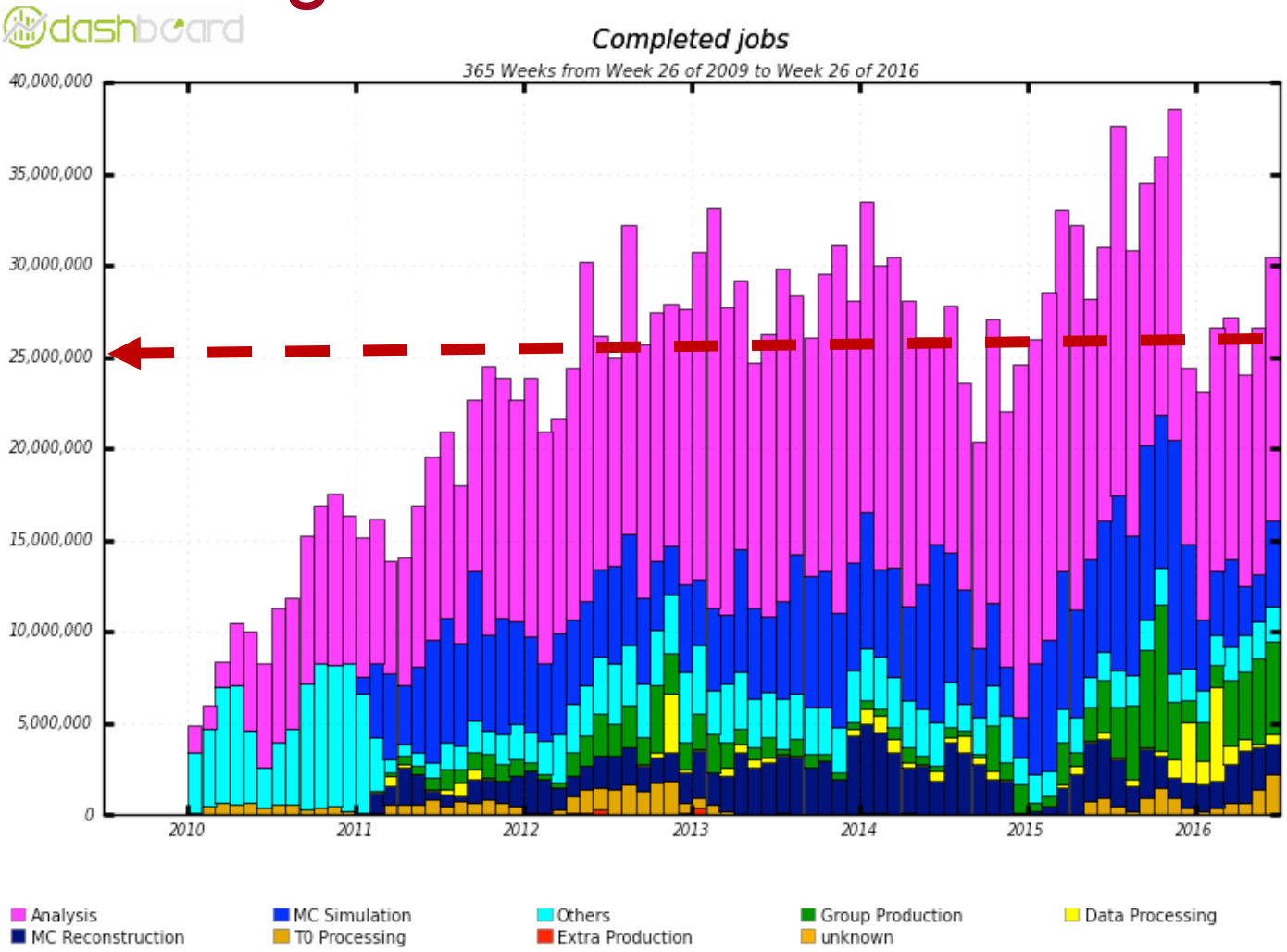
DUNE Near Detector Complex





- **Phase I ND** consists of [LAr TPC (M_A=150t, V_A=105m^3) – TMS, making up the PRISM] – SAND
- **Phase II Full Suite ND consists** of [LArTPC - Magnetized (0.5T) large volume HPGAr TPC (10atm – M_A=1t, V_A=108m^3) w/ ECAL, making up the PRISM] – SAND

Data Management Software Performance



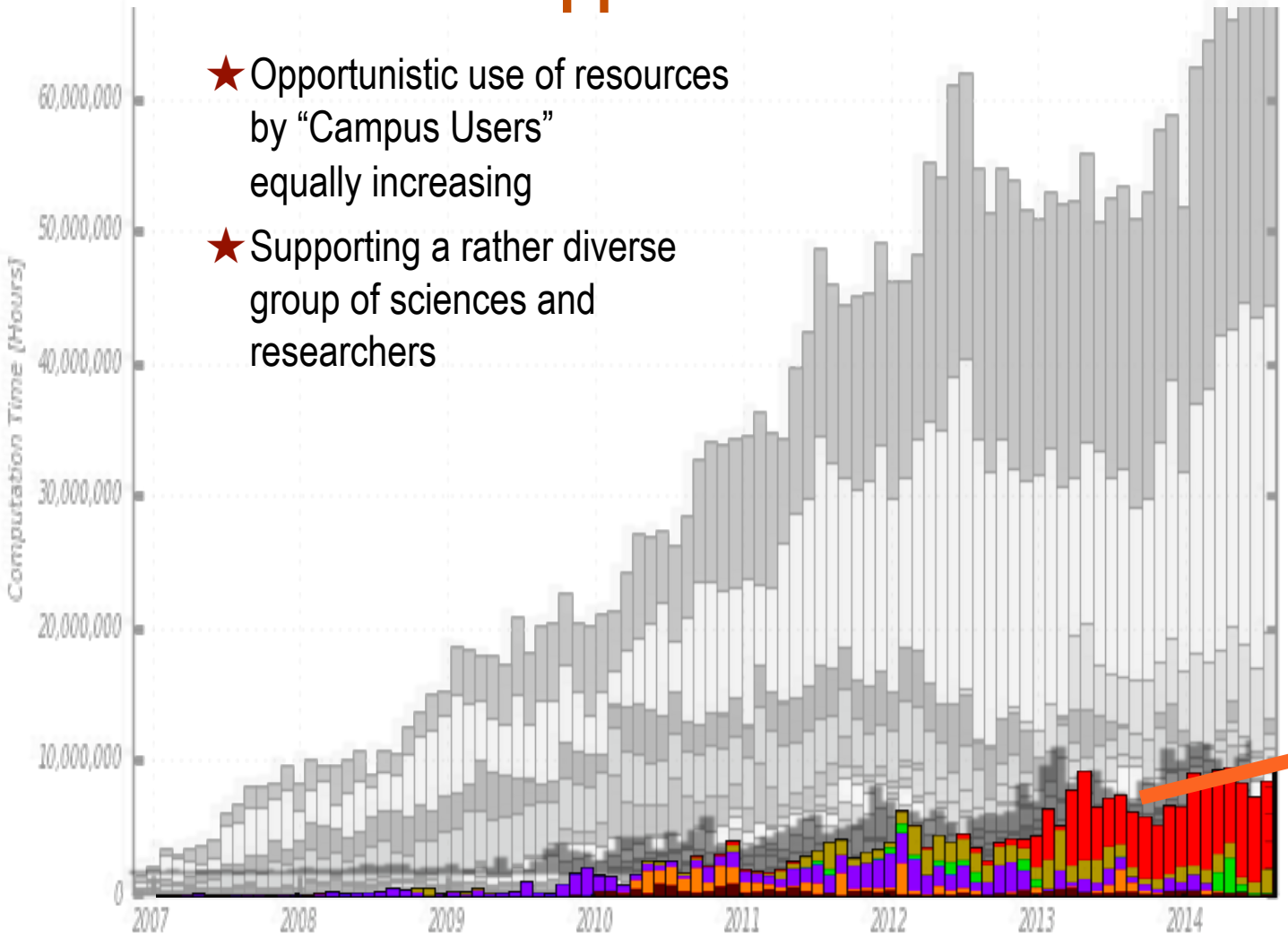
Current scale – 25M jobs completed every month at >hundred sites

Kaushik De

First exascale system in HEP – 1.2 Exabytes processed early in the LHC run

Growing Use of "Owned" and of "Opportunistic" Resources

- ★ Opportunistic use of resources by "Campus Users" equally increasing
- ★ Supporting a rather diverse group of sciences and researchers



Opportunistic Use?

>90M CPU hours opportunistic use past 12 Months

Lotha