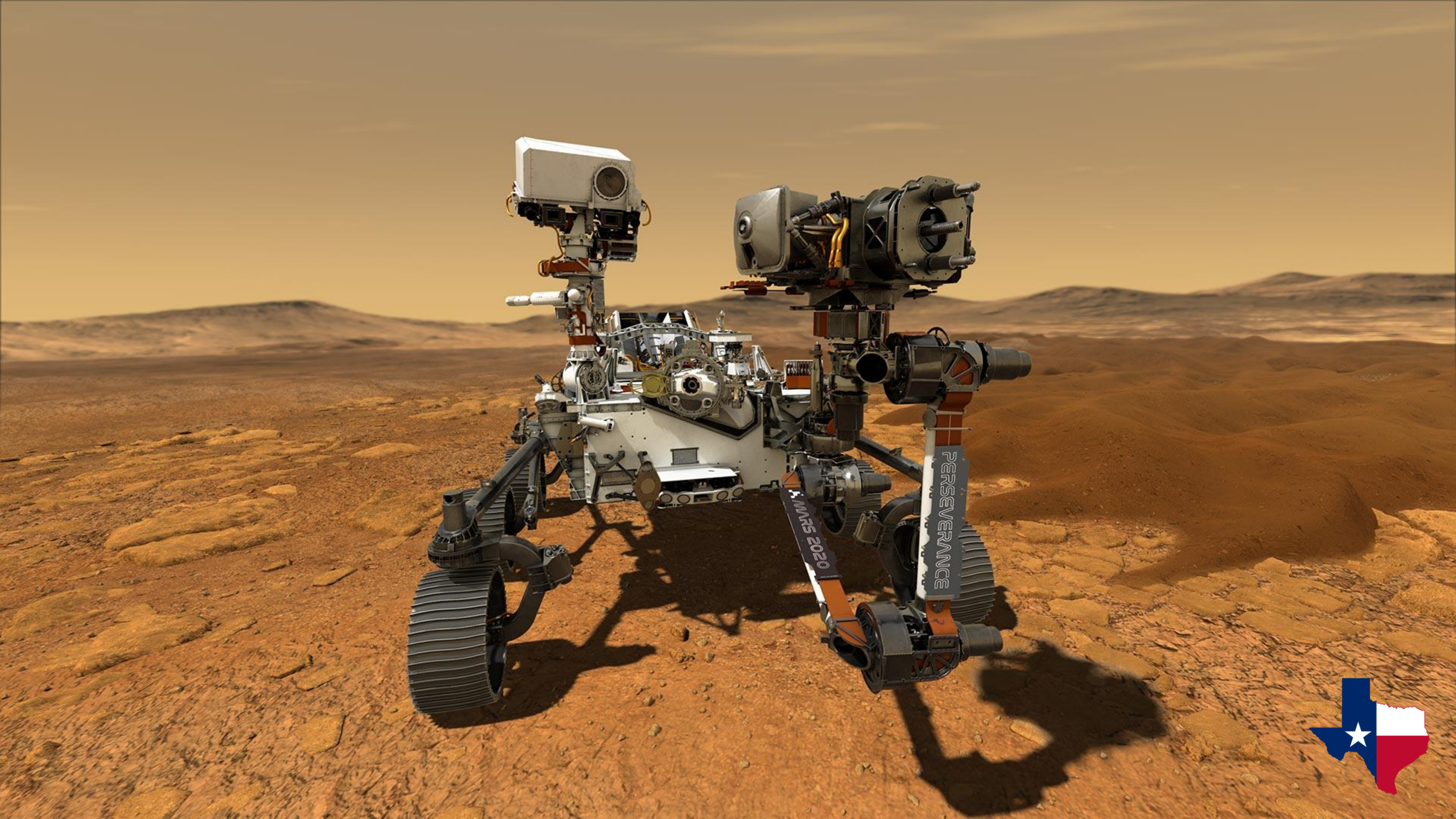
Computing And Particle Physics

G. Watts

2021-02-18

EPE Seminar


MARS 2020
PERSEVERANCE

How Did I End Up Here?

My physics (current) passion is searches for Long Lived Particles

I got into (particle) physics because I've the intersection of physics, computers, and hardware fascinated me.

As I became a professor… I had less and less time… How could I keep doing physics (e.g. making plots)!?

Got involved with others looking to make analysis easy and ended up being deputy executive director of IRIS-HEP.

Software And Particle Physics

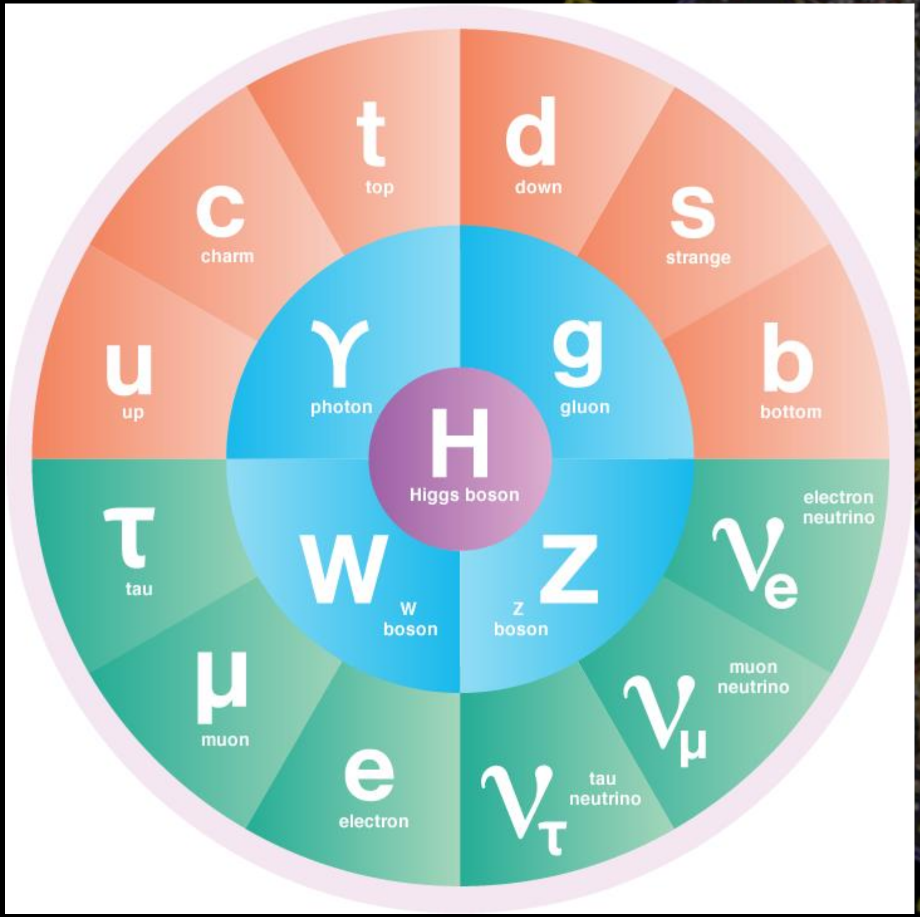
IRIS-HEP: 17 institute, ~30 FTE NSF software institute.
- Data Management
- Algorithms
- Analysis
- Facilities

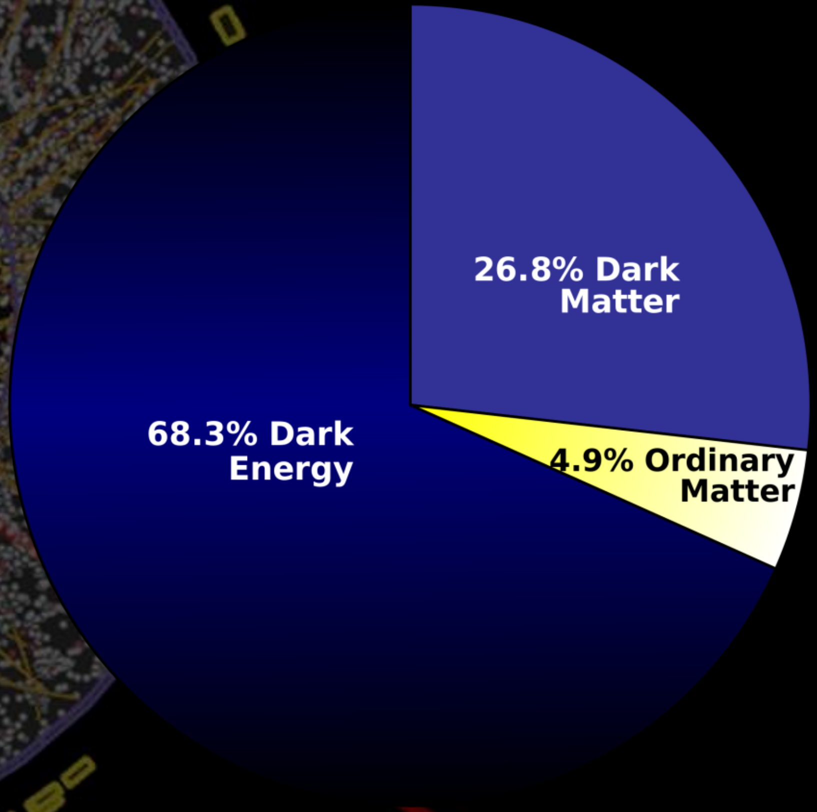
The NSF
- Physics Division
- Office of Advanced Cyberinfrastructure

HEP-CCE: DOE software institute
- GPU Porting of Algorithms and Generators
- I/O and data formats and disk usage

Why Now?

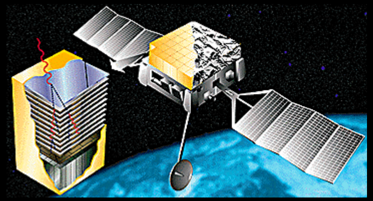
Completing The Story





Challenging the Theory

Astro



(Fermi/GLAST)

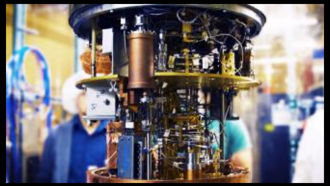
Nuclear Physics

(DUNE)

Tabletop Experiments

($\gamma\gamma$ scattering @ XFEL)

Nuclear Physics

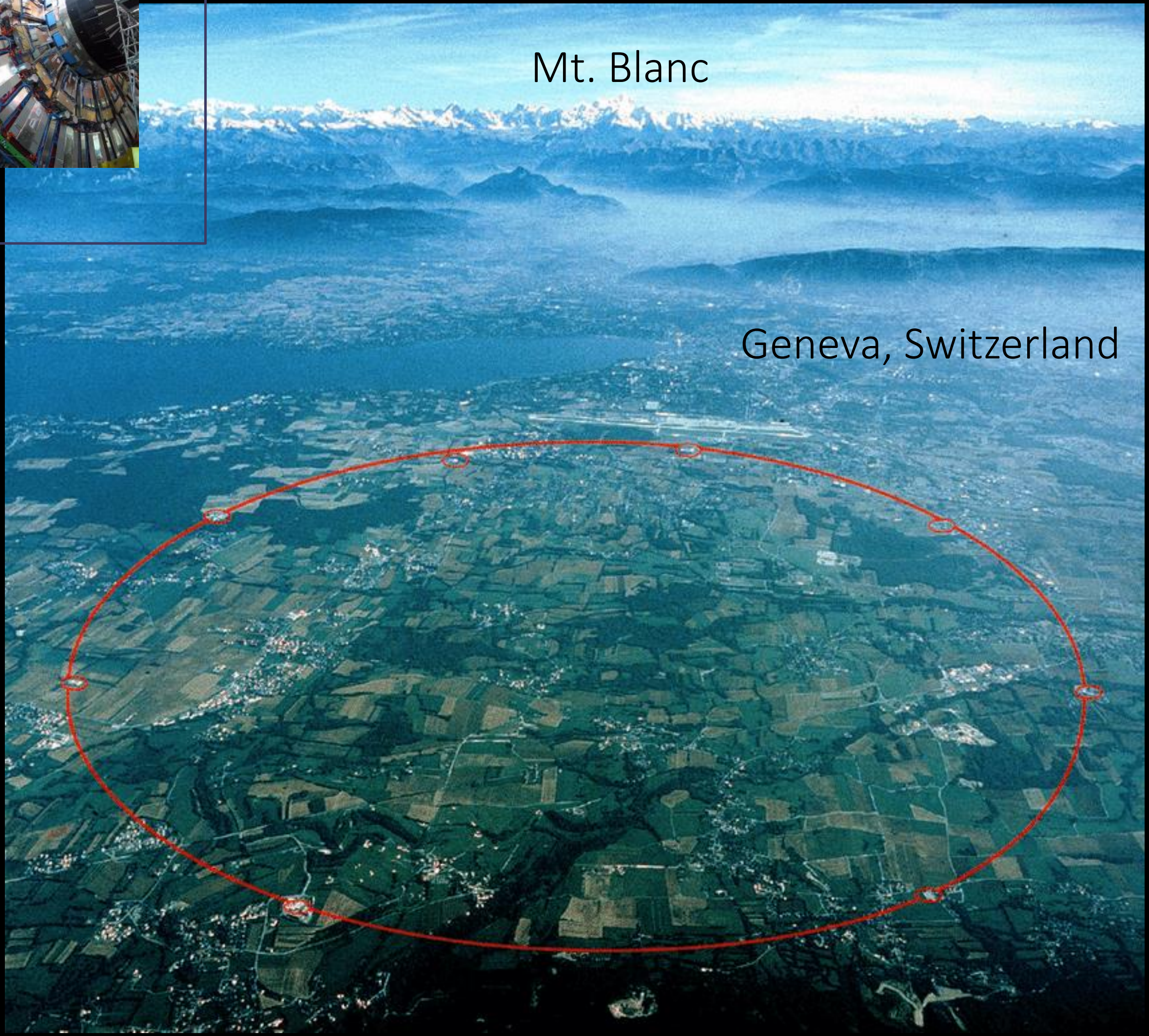


(ADMX)

Particle Physics

(CMS @ LHC)

Particle Physics



$\sqrt{s} = 13$ TeV

$$\mathcal{L} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{8}tr(\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu}) - \frac{1}{2}tr(\mathbf{G}_{\mu\nu}\mathbf{G}^{\mu\nu})$$

U(1), SU(2) and SU(3) gauge terms

$$+(\bar{\nu}_L, \bar{e}_L)\,\tilde{\sigma}^{\mu} iD_{\mu}\begin{pmatrix}\nu_L\\ e_L\end{pmatrix} + \bar{e}_R\sigma^{\mu} iD_{\mu}e_R + \bar{\nu}_R\sigma^{\mu} iD_{\mu}\nu_R + \text{Hermitian conjugate}$$

lepton dynamical term

$$-\frac{\sqrt{2}}{v}\left[(\bar{\nu}_L, \bar{e}_L)\,\phi M^{e} e_R + \bar{e}_R\bar{M}^{e}\bar{\phi}\begin{pmatrix}\nu_L\\ e_L\end{pmatrix}\right]$$

electron, muon, tauon mass term

$$-\frac{\sqrt{2}}{v}\left[(-\bar{e}_L, \bar{\nu}_L)\,\phi^{*} M^{\nu}\nu_R + \bar{\nu}_R\bar{M}^{\nu}\phi^{T}\begin{pmatrix}-e_L\\ \nu_L\end{pmatrix}\right]$$

neutrino mass term

$$+(\bar{u}_L, \bar{d}_L)\,\tilde{\sigma}^{\mu} iD_{\mu}\begin{pmatrix}u_L\\ d_L\end{pmatrix} + \bar{u}_R\sigma^{\mu} iD_{\mu}u_R + \bar{d}_R\sigma^{\mu} iD_{\mu}d_R + \text{Hermitian conjugate}$$

quark dynamical term

$$-\frac{\sqrt{2}}{v}\left[(\bar{u}_L, \bar{d}_L)\,\phi M^{d} d_R + \bar{d}_R\bar{M}^{d}\bar{\phi}\begin{pmatrix}u_L\\ d_L\end{pmatrix}\right]$$

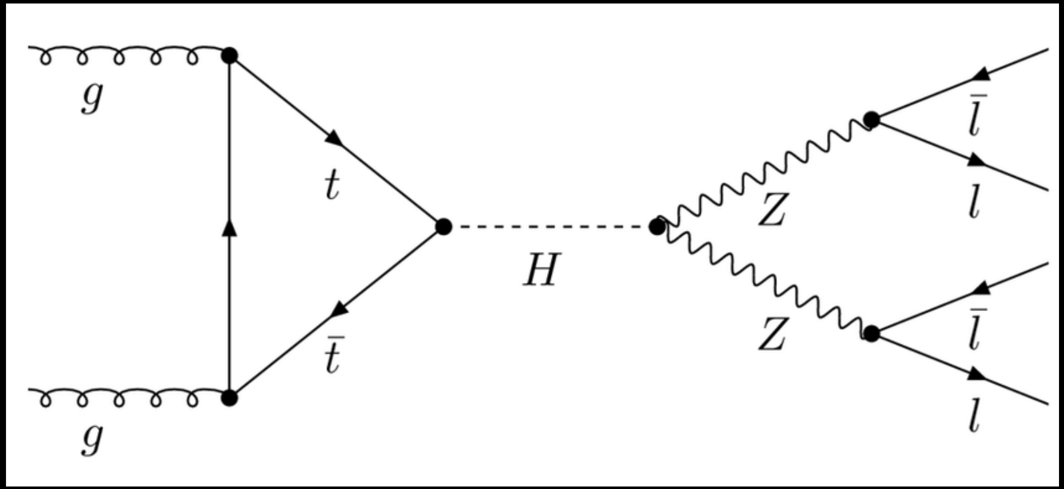
down, strange, bottom mass term

$$-\frac{\sqrt{2}}{v}\left[(-\bar{d}_L, \bar{u}_L)\,\phi^{*} M^{u} u_R + \bar{u}_R\bar{M}^{u}\phi^{T}\begin{pmatrix}-d_L\\ u_L\end{pmatrix}\right]$$

up, charmed, top mass term

$$+\overline{(D_{\mu}\phi)}D^{\mu}\phi - m_h^2[\bar{\phi}\phi - v^2/2]^2/2v^2.$$

Higgs dynamical and mass term



$$\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{BSM}$$

Carefully measure every constant

Look for something we have missed

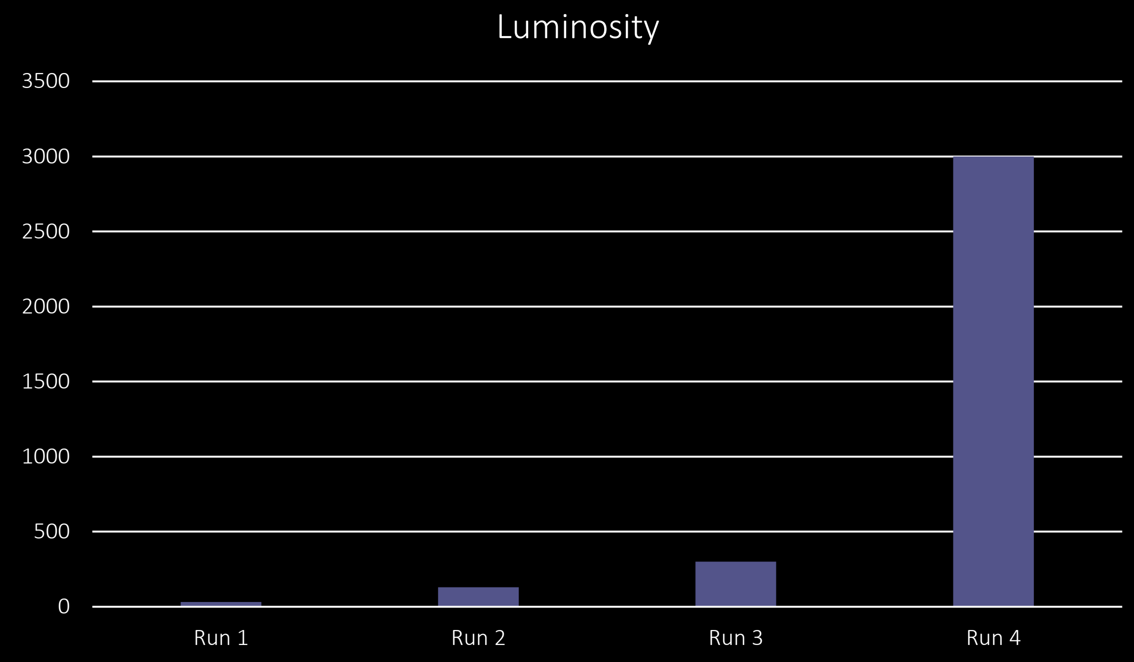
Search for new particles

$\sqrt{s}$ Energy, Mass Scale Probed

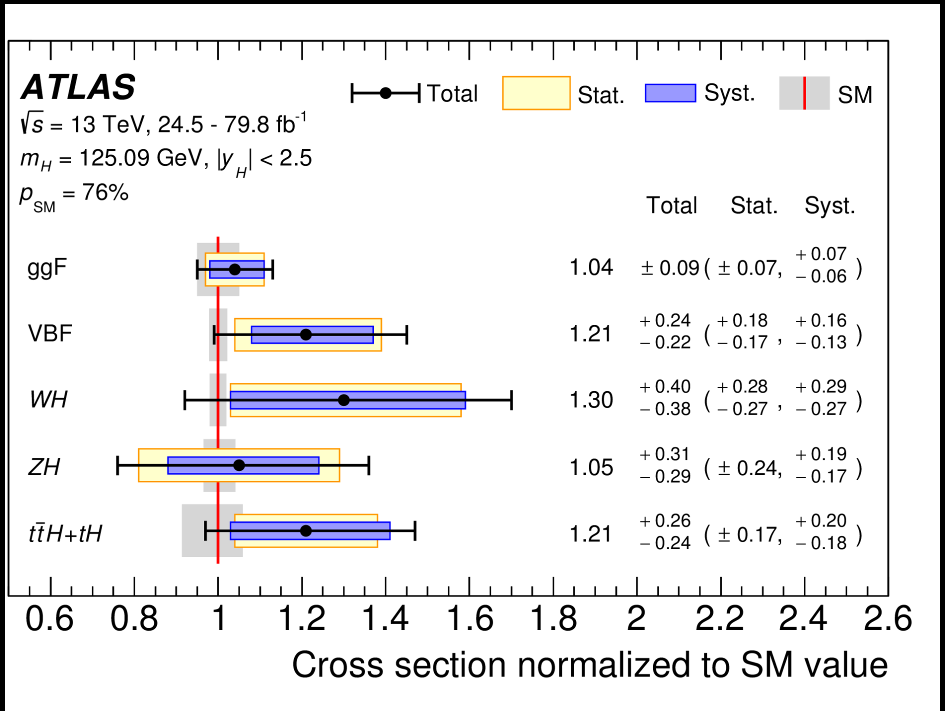
$$N = \sigma_{H \to ZZ} \times L$$

- L: How often the accelerator collides protons
- $\sigma_{H \to ZZ}$: How often a particular process happens
- N: How many events we can expect in our detector

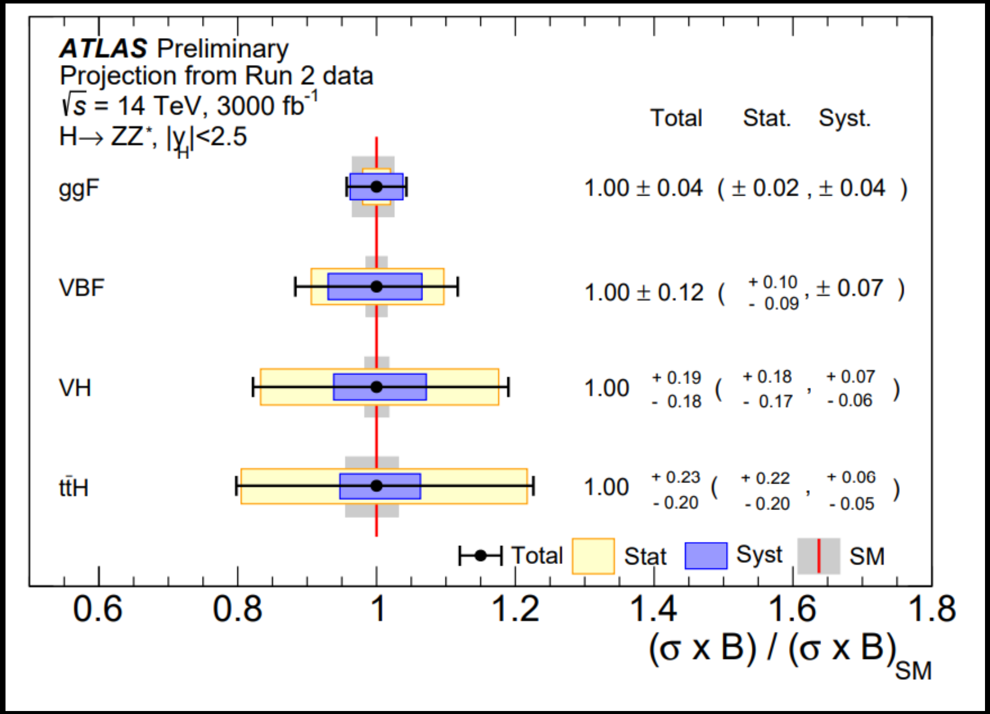
Run 4


Luminosity
3500
3000
2500
2000
1500
1000
500
0
Run 1
Run 2
Run 3
Run 4

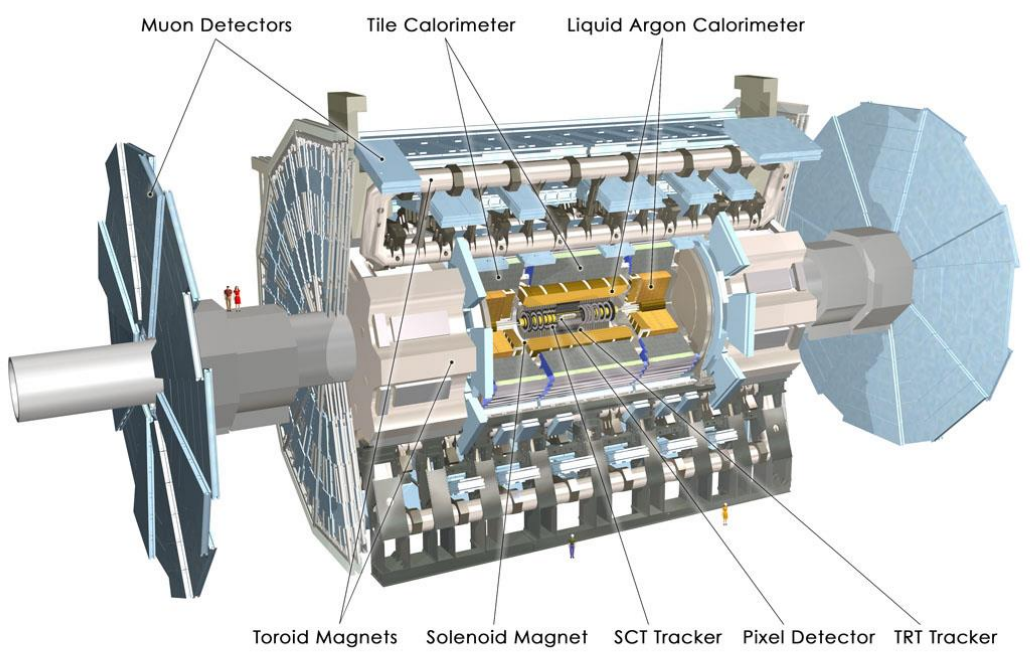
Now

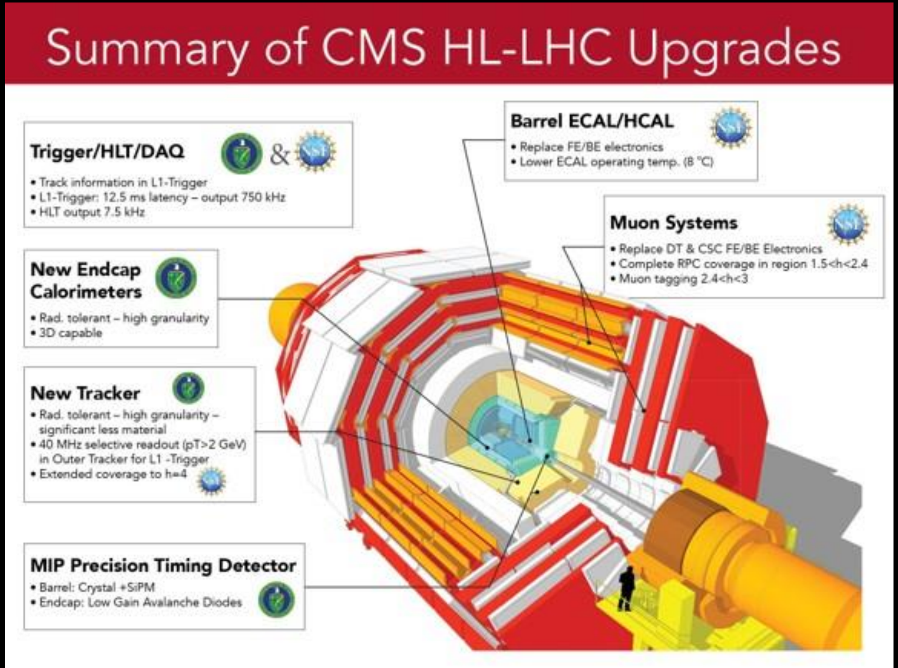


HL-LHC Projections

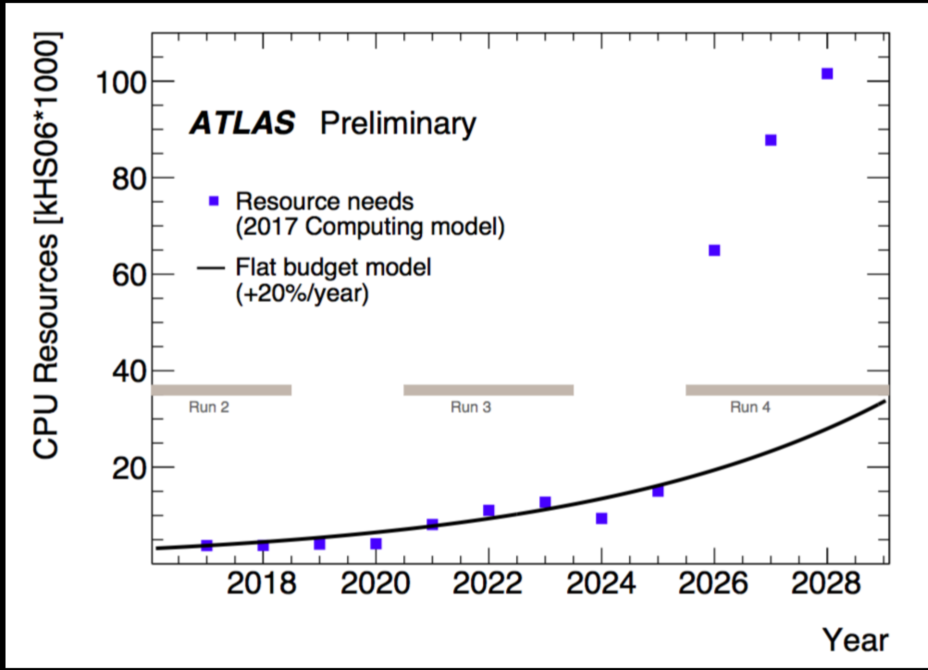


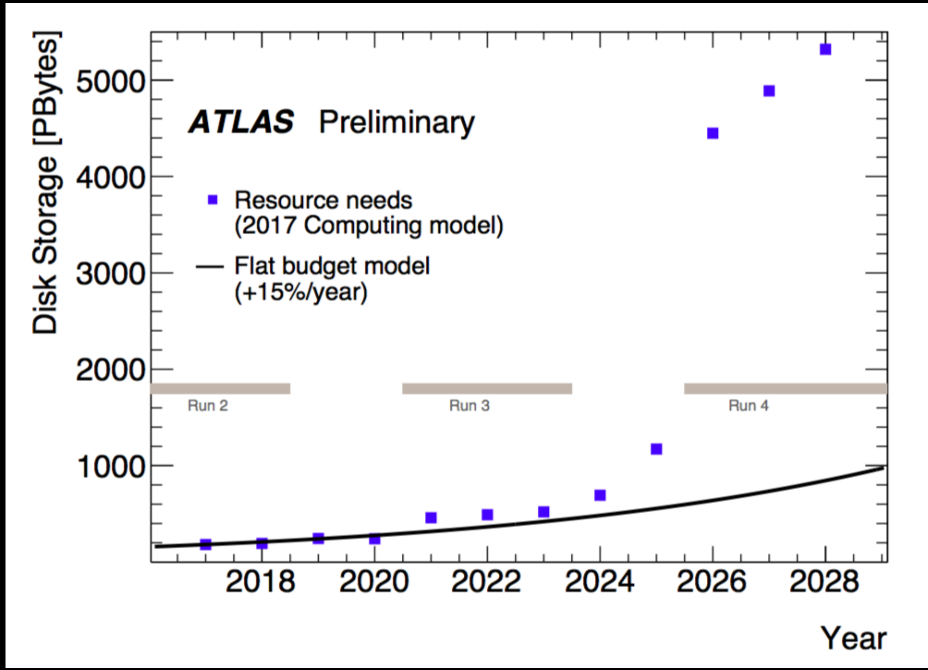
Plans for Run 4 – Upgrade the Detectors



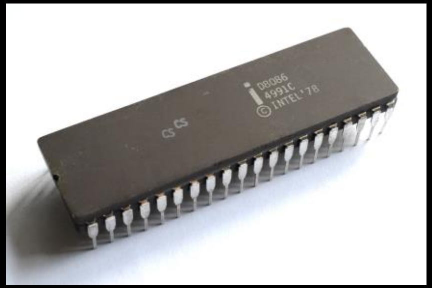


Plans for Run 4 – Computing





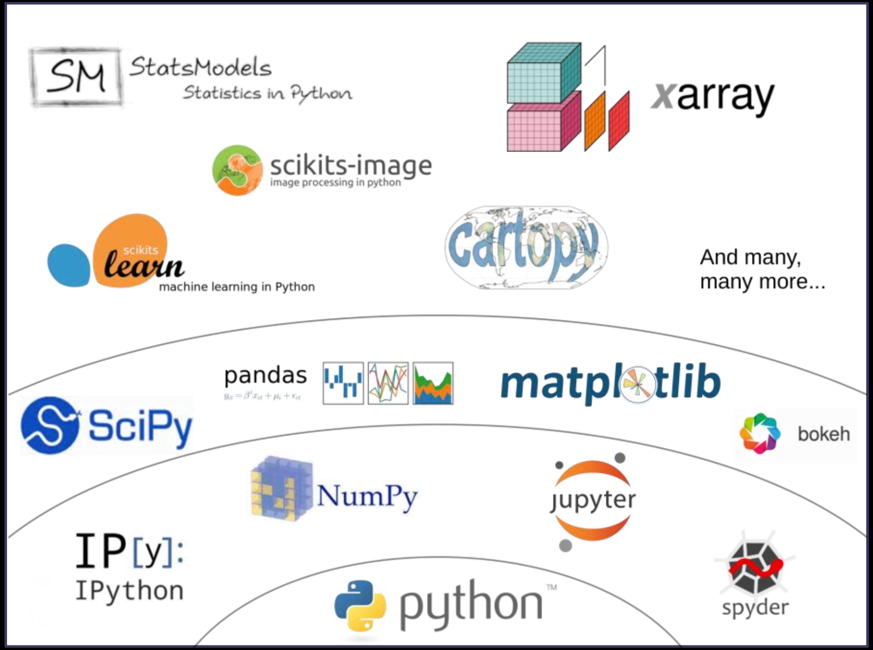
The World Is Changing I



≠

The World Is Changing II

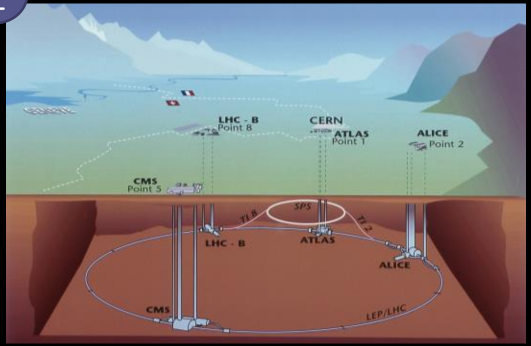
ROOT
Data Analysis Framework

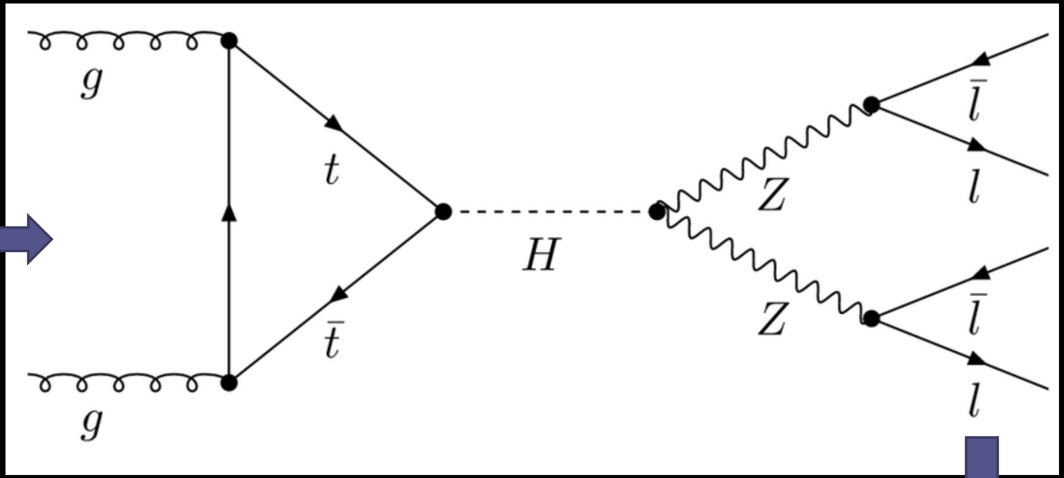


Computing @ LHC

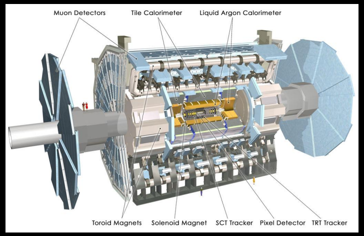
4 Analysis
Reconstructs the underlying physics

1





2 Measures

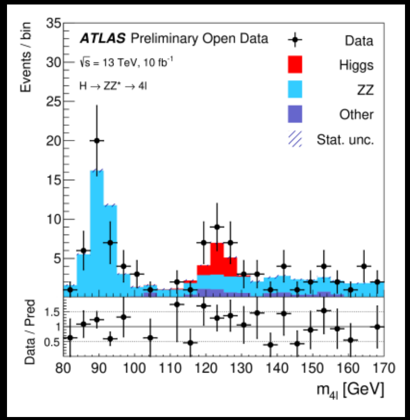


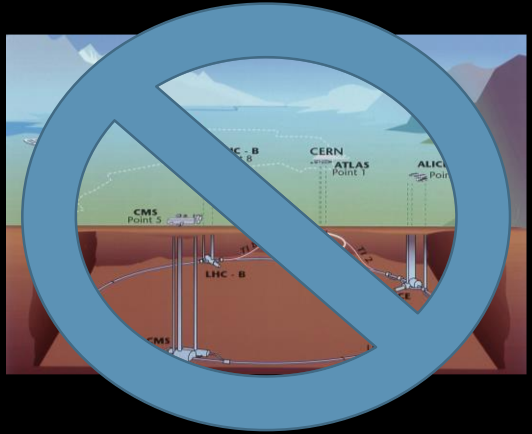
3 Reconstructs hits into objects (electrons, etc.)

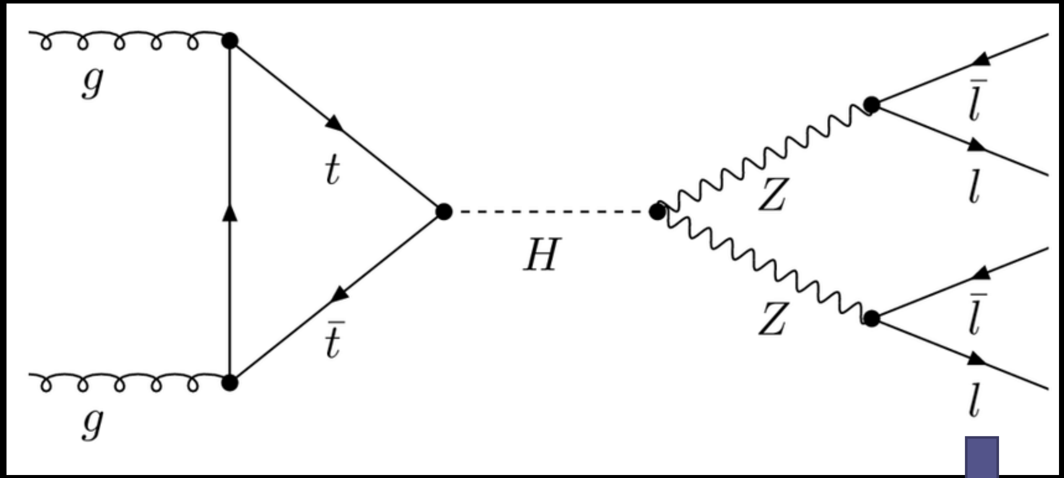
Computing @ LHC

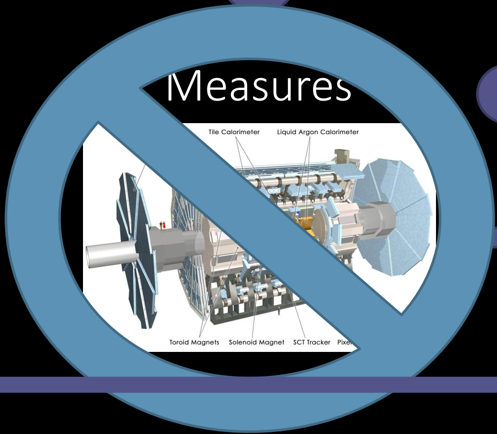
4 Analysis
Reconstructs the underlying physics

MonteCarlo & Data









3 Reconstructs hits into objects (electrons, etc.)

1 Madgraph5
(The Standard Model)

2 Detector Simulation

The Frontiers of Computing Evolution

Machine Learning

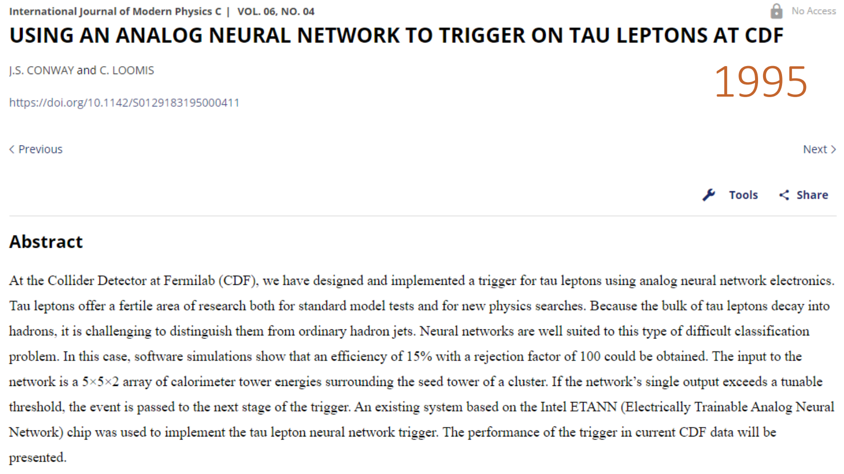
Coprocessors (GPU)

Analysis

OpenData/Reuse

Machine Learning

ML Has Been Around…


International Journal of Modern Physics C | VOL. 06, NO. 04

No Access

USING AN ANALOG NEURAL NETWORK TO TRIGGER ON TAU LEPTONS AT CDF

J.S. CONWAY and C. LOOMIS

1995

https://doi.org/10.1142/S0129183195000411

< Previous

Next >

Tools Share

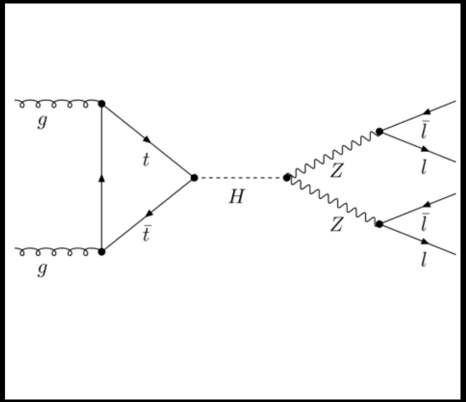
Abstract

At the Collider Detector at Fermilab (CDF), we have designed and implemented a trigger for tau leptons using analog neural network electronics. Tau leptons offer a fertile area of research both for standard model tests and for new physics searches. Because the bulk of tau leptons decay into hadrons, it is challenging to distinguish them from ordinary hadron jets. Neural networks are well suited to this type of difficult classification problem. In this case, software simulations show that an efficiency of 15% with a rejection factor of 100 could be obtained. The input to the network is a 5×5×2 array of calorimeter tower energies surrounding the seed tower of a cluster. If the network's single output exceeds a tunable threshold, the event is passed to the next stage of the trigger. An existing system based on the Intel ETANN (Electrically Trainable Analog Neural Network) chip was used to implement the tau lepton neural network trigger. The performance of the trigger in current CDF data will be presented.

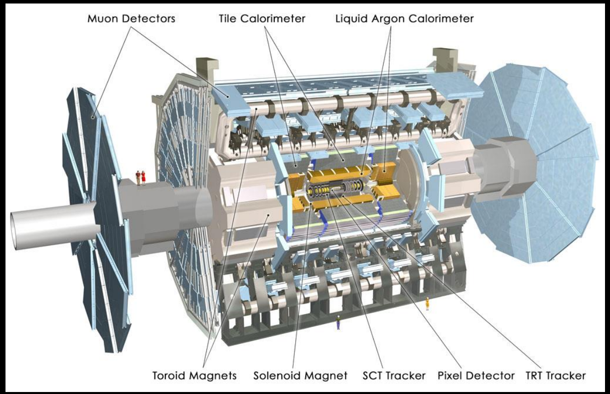
Machine Learning is a function fit, or minimization problem

Deep Learning: function has millions of parameters

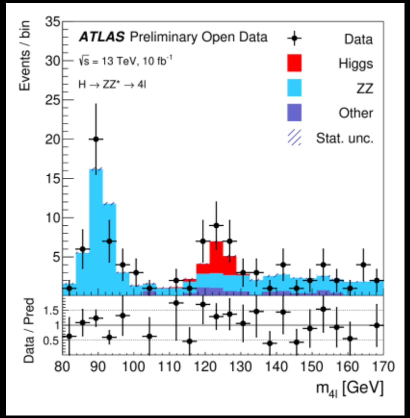
We learned how to calculate the gradient exactly…



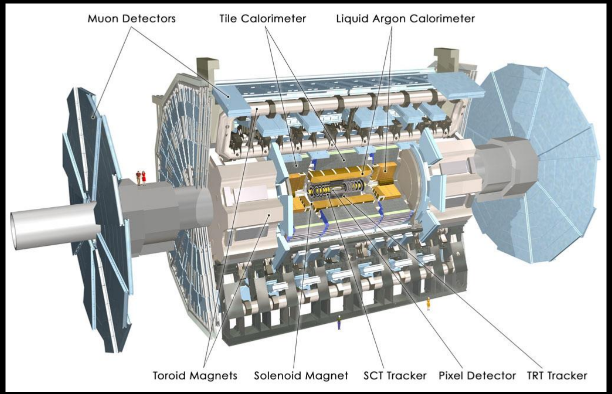
We can “calculate” this



$f(x)$



What you want



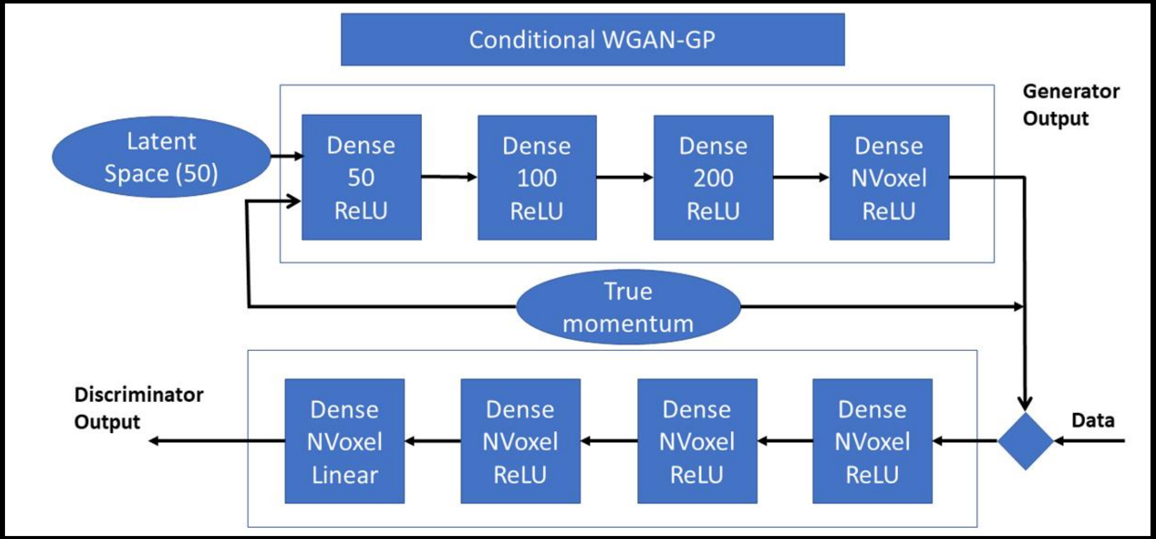
$f(x)$

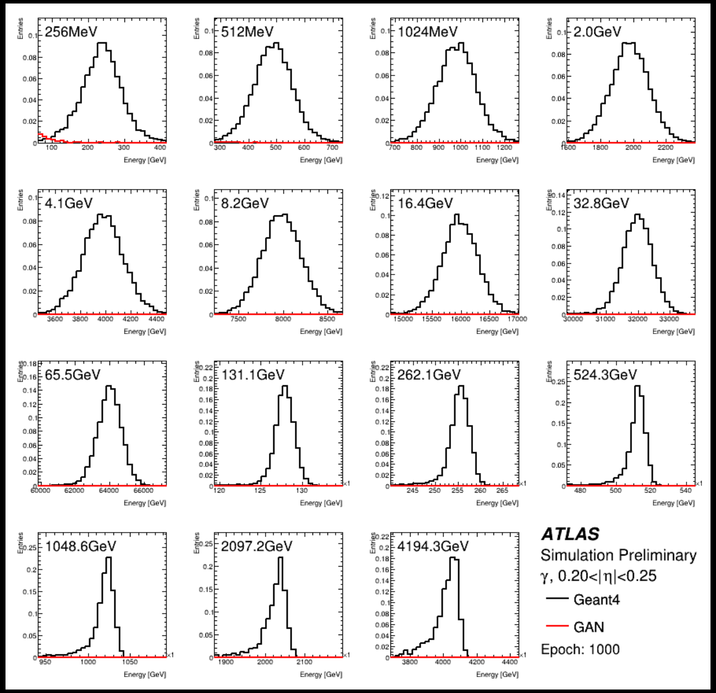
Simulation of complex detector elements

Calibration

Separating signal and background

In Simulation



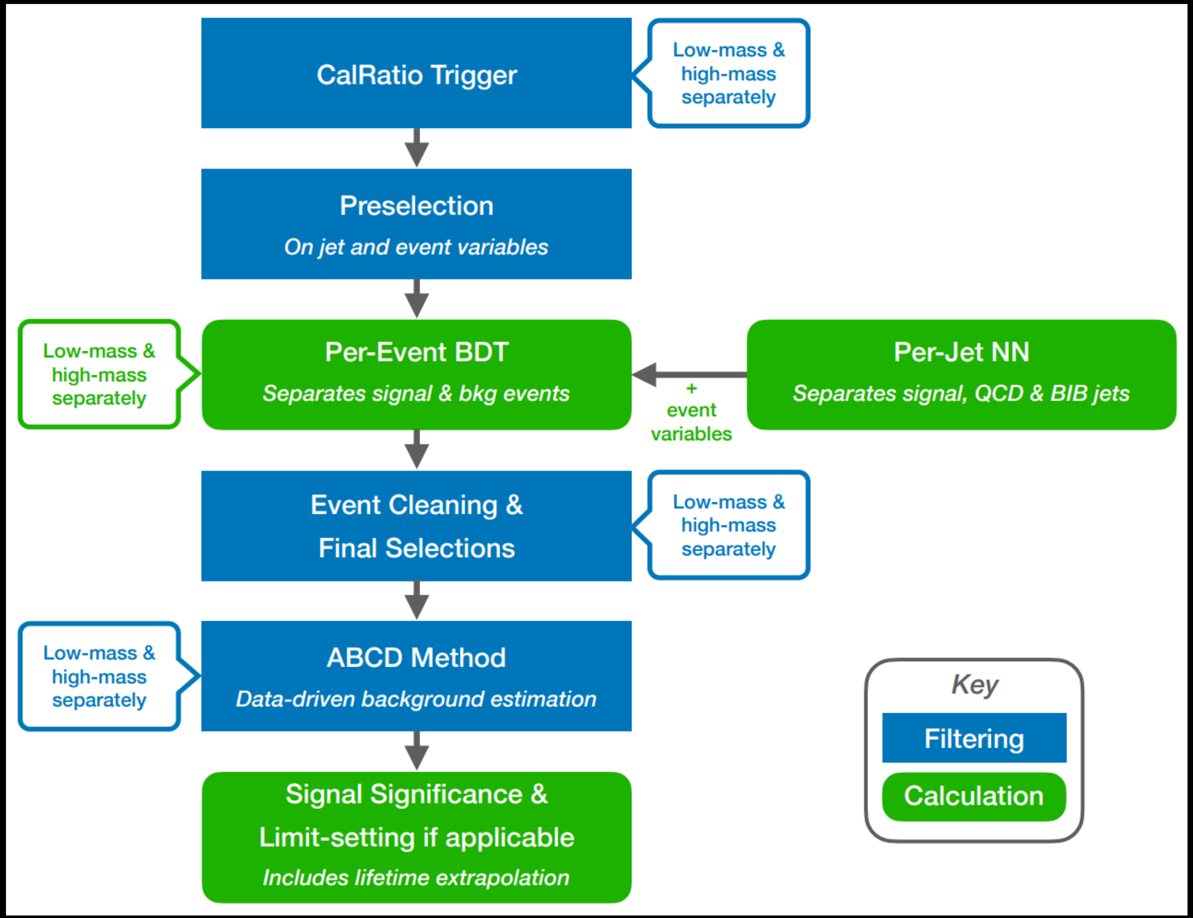


Photons – Geant4 vs GAN

In Analysis

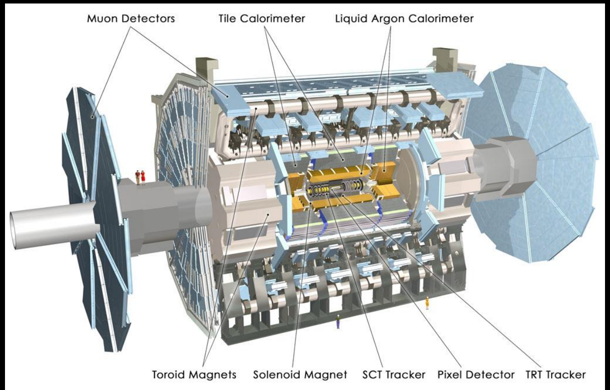
Grab decay from CMS or ATLAS detector picture

Make sure to save to LLP PR plots

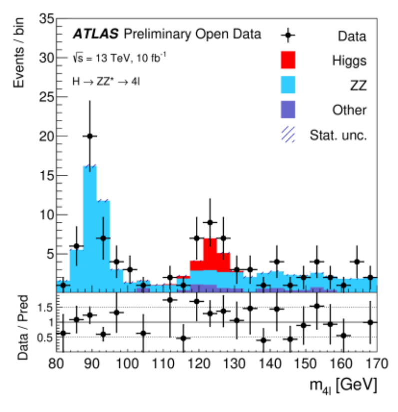


20-30% Gain in Sensitivity

Differentiable Programming



→ Black Box (NN) →



Traditional NN Approach

Lepton Information
Environment Information
→ Black Box (NN) → m_{4l}

Traditional Approach

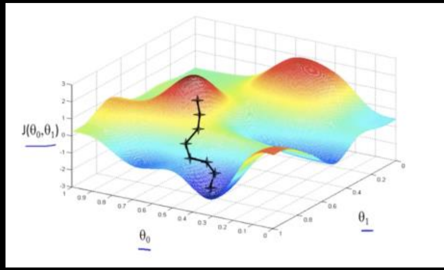
1. Select 4 muons
2. Selection cuts on muons
3. Combine into Z bosons
4. Selection cuts on Z's
5. Combine into $4l$ object
6. Plot m_{4l}

Differentiable Programming

Combine the Approaches

1. Select 4 muons
2. Selection cuts on muons
3. Combine into Z bosons
4. Selection cuts on Z's
5. Combine into $4l$ object
6. Plot m_{4l}

Use ML techniques to optimize!
Based on figure of merit: expected sensitivity



Differentiable Programming allows you to do gradient decent through loops, if statements, etc.

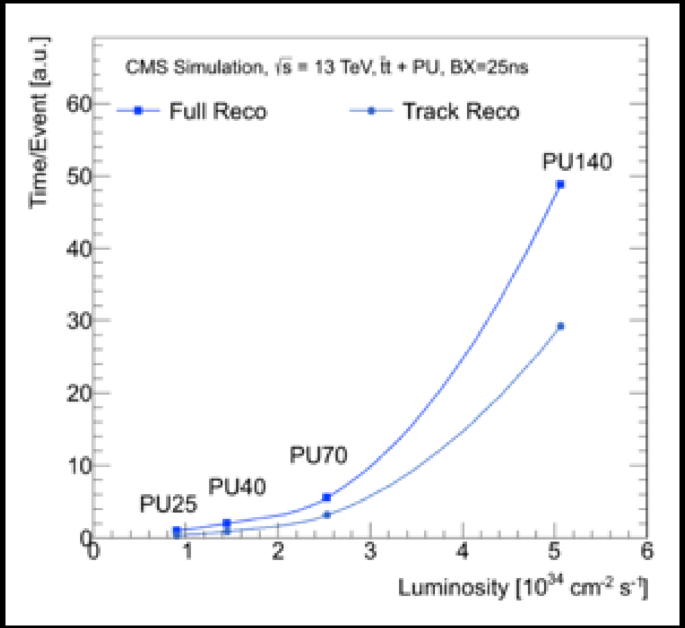
Differentiable Programming allows you to do gradient decent through loops, if statements, etc.

Take into account systematic errors!

Coprocessors

Why Are Coprocessors Interesting?

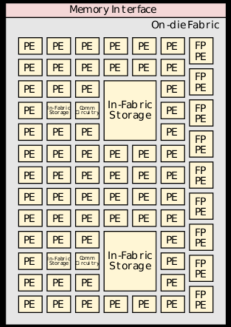
A Co-Processor does a specific calculation much more efficiently than a general-purpose CPU



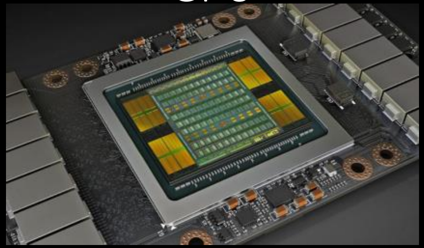
ASIC

FPGA's

Configurable Spatial Accelerator (CSA)



GPU



The Lay Of The Land…

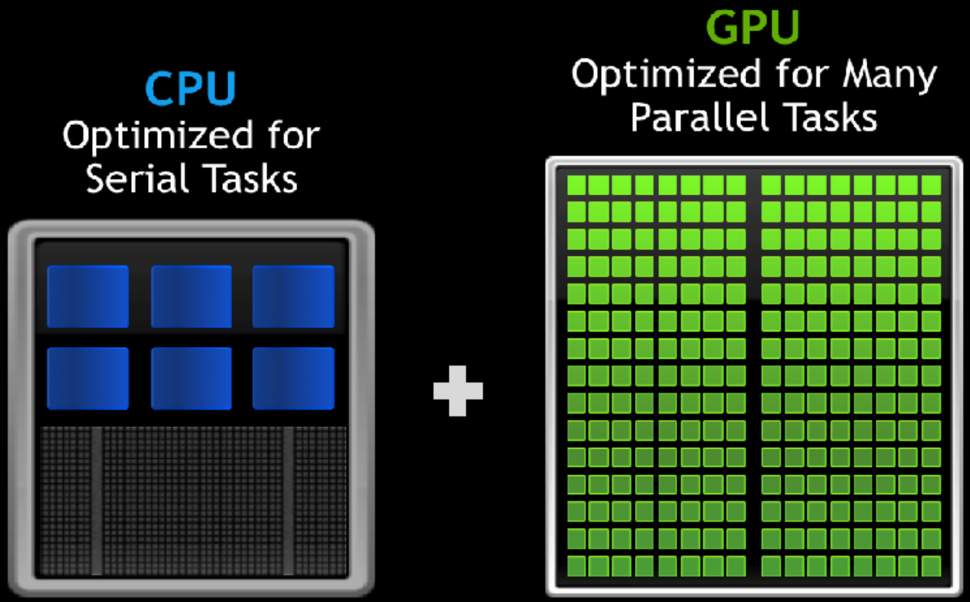
Next Generation HPC Architectures

BERKELEY LAB

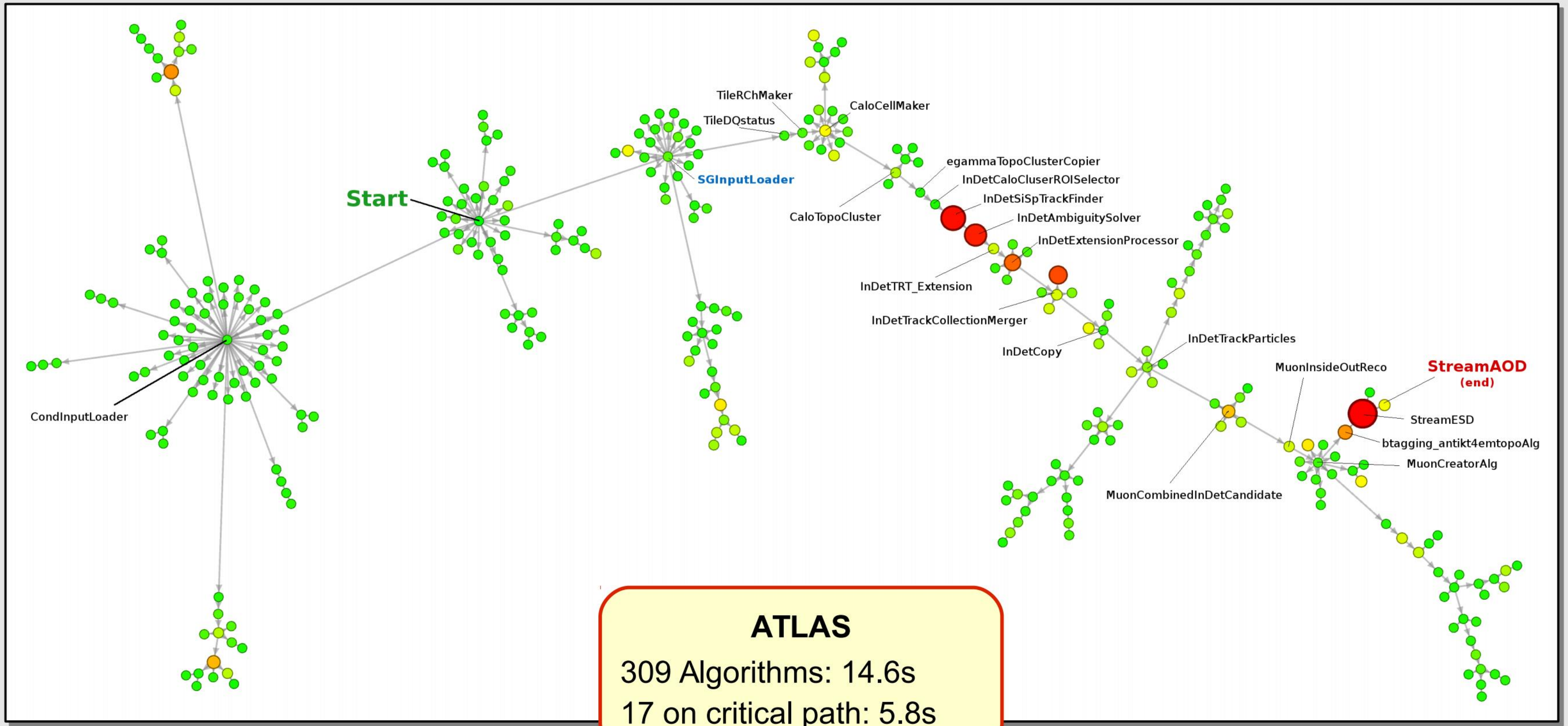
- In the next generation of supercomputers we see extensive use of accelerator technologies
  - Oak Ridge: **Summit** (2018)
    - 4608 IBM AC922 nodes *w/* 2x Power9 CPU
    - 3x NVIDIA Volta V100 + NVLink / CPU
  - LBL: NERSC-9 **"Perlmutter"** (2020)
    - AMD EPYC "Milan" x86 only nodes + mixed CPU / "next gen" NVidia GPU
  - Oak Ridge: **Frontier** (2021)
    - 1.5 exaflop
    - AMD EPYC CPU + 4x AMD "Instinct" GPU
  - Commercial clouds:
    - Brainwave / Azure FPGA
    - Google Cloud TPU
    - Amazon EC2 P3
  - LLNL: **Sierra** (2018)
    - 4320 IBM AC922 nodes *w/* 2x Power9 CPU
    - 2x NVIDIA Volta V100 + NVLink / CPU
  - Argonne: **Aurora A21** (2021)
    - possibly first exascale HPC
    - Intel Xeon CPU + Intel X^e/gen12 GPU + Optane
  - Tsukuba: **Cygnus** (2020)
    - 2x Intel Xeon 6162 + 4x NVidia V100 GPU
    - 2x CPU + 4x GPU + 2x Intel Stratix FPGA
  - Japan: **Fugaku** (2021)
    - manycore ARM A64fx (48+2)
    - integrated "SVE" 512 bit GPU-like accelerator
  - Spain: **MareNostrum**
    - Xeon 8268 + Power9 + V100 GPU
  - Switzerland: **Piz Daint**
    - Xeon E5 2690 + NVidia P100 GPU

What is viable?

- GPU's – for general purpose applications
- FPGA – for tailor-made applications



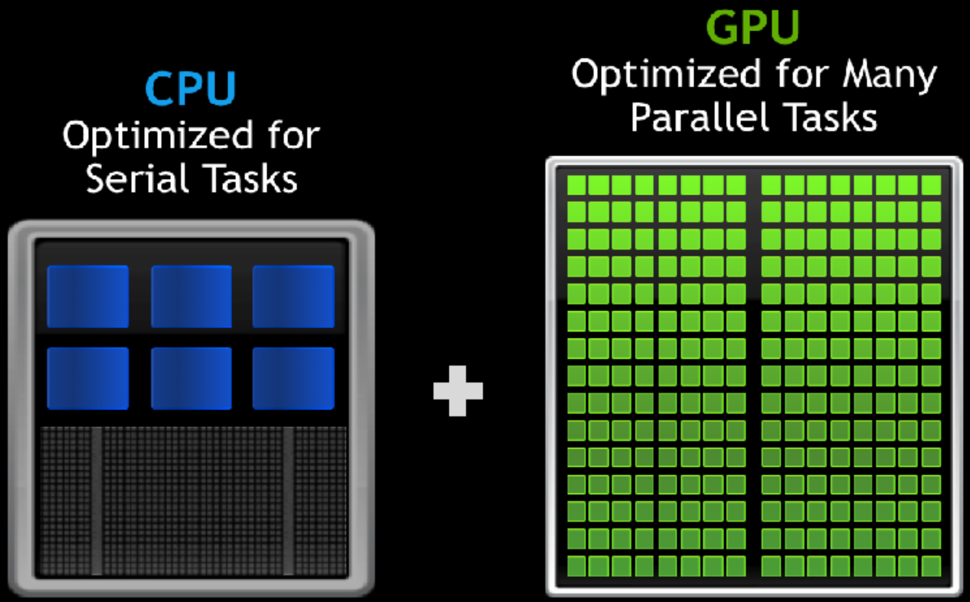
The GPU is optimized for the same parallel task



ATLAS

309 Algorithms: 14.6s

17 on critical path: 5.8s

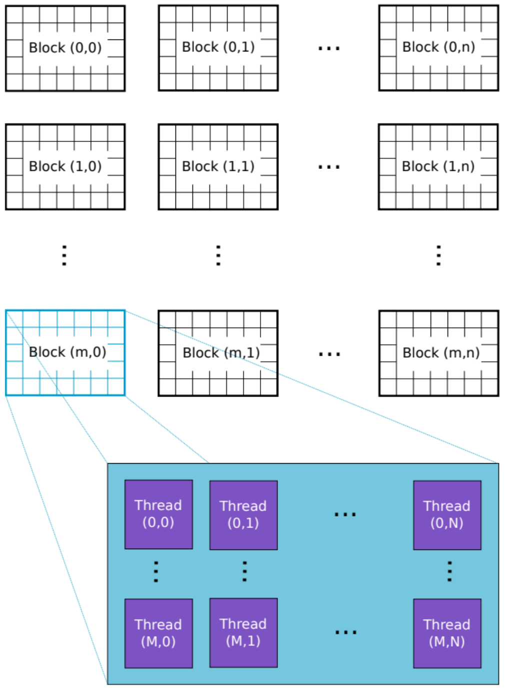


- GPU’s are very good at Linear Algebra
- GPU’s are very bad at if statements
- Many threads operate in lock-step
- There is no i86 instruction set

Porting algorithms is difficult or nearly impossible

Those who say it can't be done are usually interrupted by others doing it.

(James Baldwin)



Run each event in one block

Decoding → parallelise by readout unit

Clustering → parallelise in (overlapping) detector regions

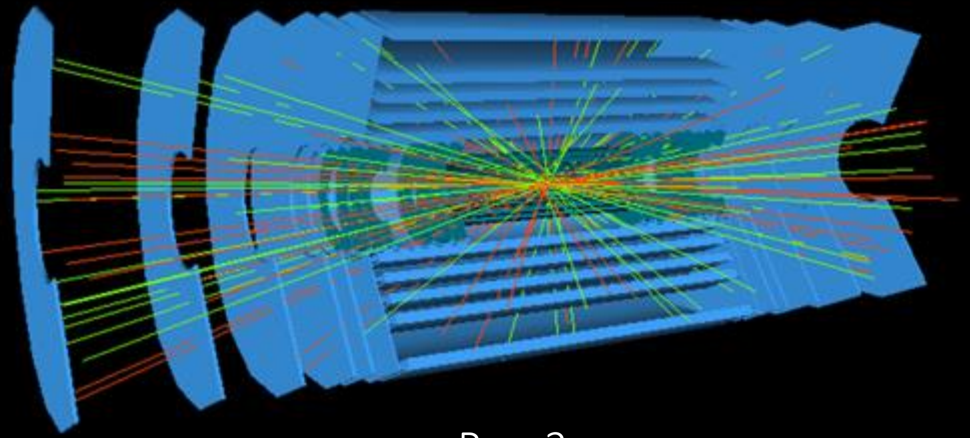
Tracking → parallelise by track

Vertexing → parallelise by combination

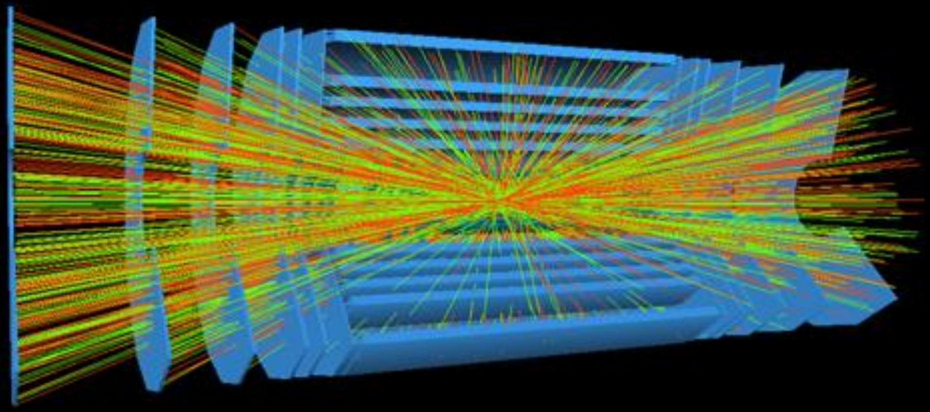
Is LHCb the exception? Or the Rule?

Two of the most expensive operations:

- Tracking
- Jet reconstruction



Run 2

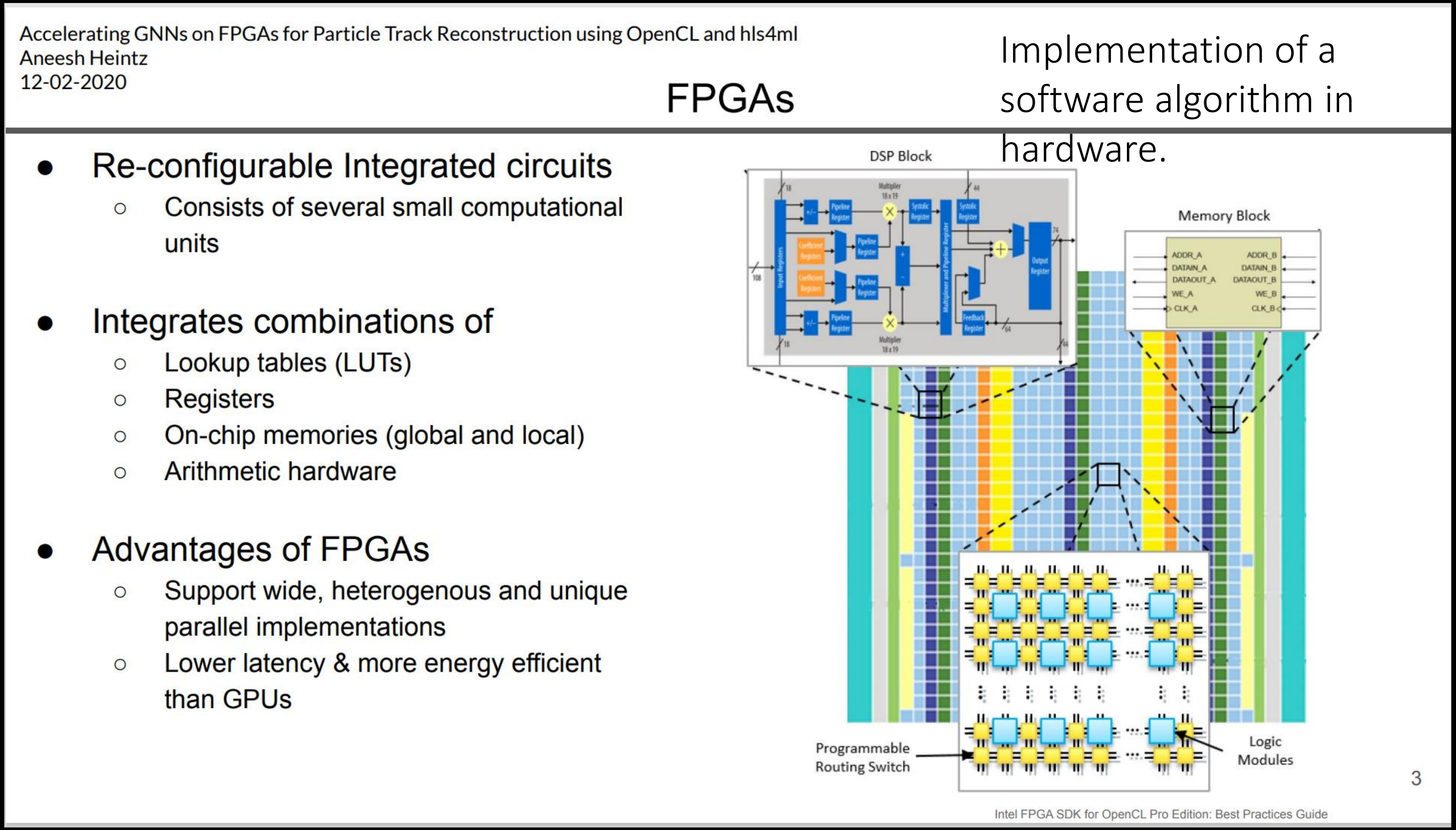


Run 4

Let's hope it is the rule…

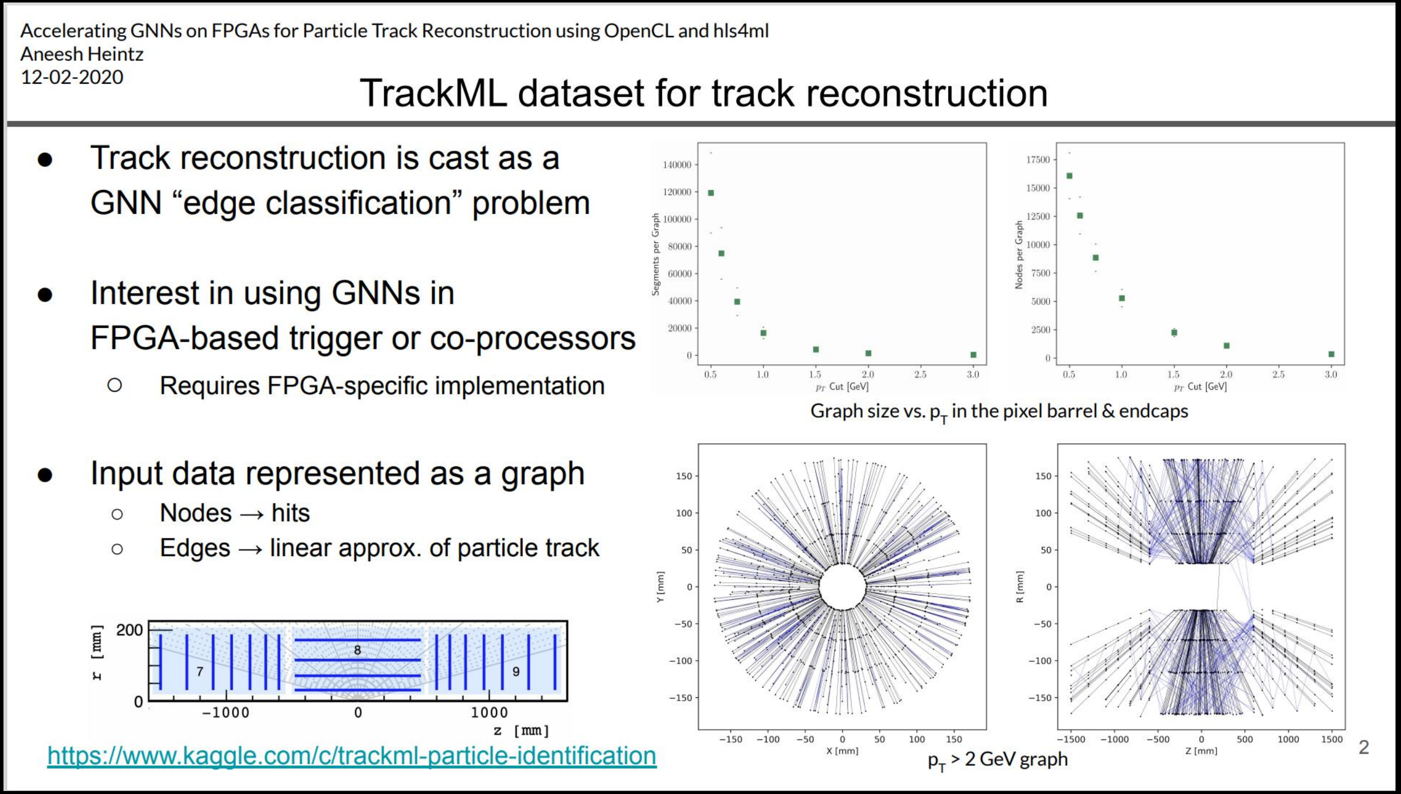
Custom Programming of FPGA's often used in the trigger

Field Programmable Gate Arrays



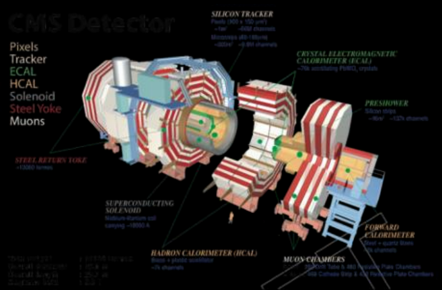
Implementation of a software algorithm in hardware.

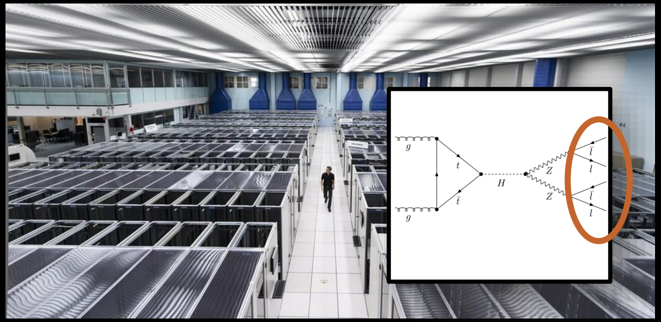
Mostly Arbitrary NN's → FPGA Code



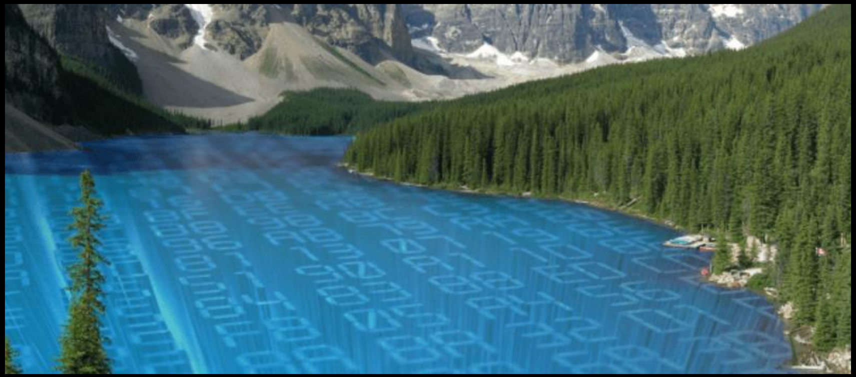
Interesting because of the trigger

Analysis





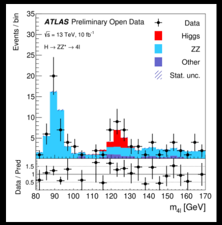
Reconstruction/Simulation



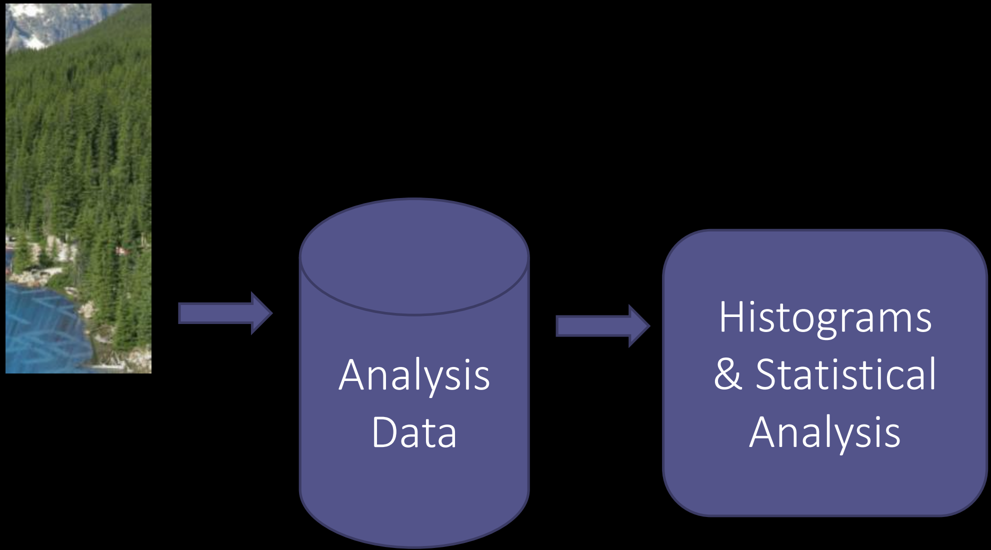
Data Lake

Analysis Data

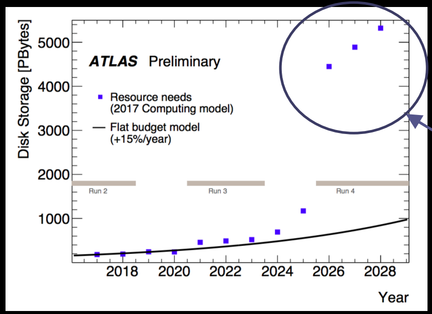
Histograms & Statistical Analysis



“Analysis”



1. Write C++ code to access experiment Data Format
2. Locate Data in Data Lake
3. Submit jobs to run on world-wide GRID of computers on many files
4. Download & combine data locally
5. Make plots/look at data



Not just too long, but also leads to bad habits

A cycle takes ~weeks

Particle Physics Software Eco-System

ROOT

ROOT Data Analysis Framework

The computational backbone of particle physics for the last 25 years

- Maintained by the community and a core of ~10 developers at CERN and Fermilab.
- C++ (even in a command line)
- A framework
- I/O, data analysis, histogram filling, fitting, etc., all carefully put together
- Designed by and for particle physics
- Limited use outside

Python

python™

Used by industry to analyze data similar to HEP data

- Ecosystem of many packages maintained by industry and open-source developers
- Uses C/C++ for speed, python for productive interface
- A collection of libraries that communicate via standards
- I/O, data analysis, histogram filling, fitting, etc., all availible
- Needs additional libraries to satisfy all of science's needs
- Growing use by particle physicists

ROOT
Data Analysis Framework

python™

Declarative "Analysis"

SQL Structured Query Language

```
SELECT CustomerName, City FROM Customers
```

- What parts of the data you want (CustomerName, City)
- The data you want to start from (Customers)

1. Write C++ code to access experiment Data Format
2. Locate Data in Data Lake
3. Submit jobs to run on world-wide GRID of computers on many files
4. Download & combine data locally
5. Make plots/look at data

→

1. Get Name of Data in Data Lake
2. Write "simple" query
3. Make plots/look at data

Industry Has Tackled This

e.g.

Apache Spark

Given a query against a registered dataset, automates all aspects and returns your requested data. Tested against PB of data.

Many tools → Kafka, Hive, Fink, Storm, etc.

Many tools are mature… why not take advantage of them?

- Independent collider collisions: extreme parallelism
- Rectilinear Data Model
- Data Lake Interfaces

There are many working prototypes

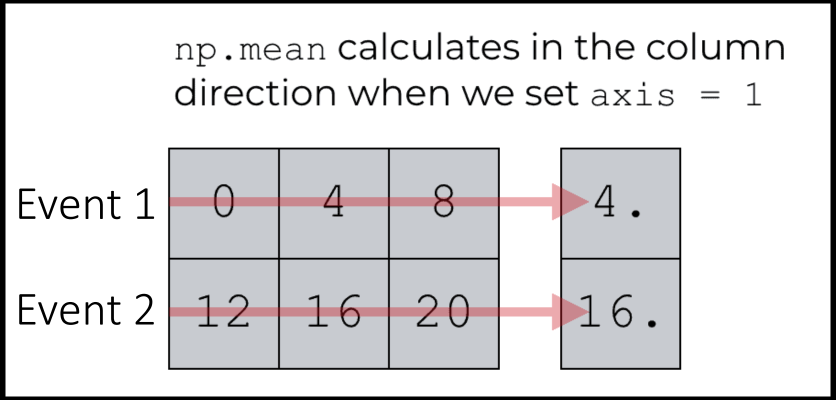
Extreme Parallelism

Industry Solution: Columnar Calculations

- Don’t loop over events
- Loop over objects/numbers/arrays

Calculate the mean energy of all jets in an event

1. Load electron energies into memory for all events
2. Scan all at the same time, calculating the mean



No loops, no if statements: very GPU friendly!

- What if your data is too large to fit in memory?
- Complex operations generate unneeded temporaries

Rectilinear Data

Particle Physics Data is Jagged

Energy						
Column1	Electron 1	Electron 2	Electron 3	Electron 4	Electron 5	Electron 6
Event 1	180	91.2	22.1			
Event 2	121	100.6	77.3	55.5	33.4	21.1
Event 3	115.4	56	45	25.1		

Gaps!

Each Event does not look the same!

Outlander Spices

Quarterly Sales 2004 - 2010

	2004					2005					2006		
Product	**Qtr1**	**Qtr2**	**Qtr3**	**Qtr4**	**Total**	**Qtr1**	**Qtr2**	**Qtr3**	**Qtr4**	**Total**	**Qtr1**	**Qtr2**	**Qtr3**
Angelica Root	$838	$1,117	$906	$739	$3,600	$1,338	$1,450	$1,157	$1,239	$5,184	$1,160	$1,340	$1,339
Anise	$541	$820	$609	$442	$2,412	$1,041	$1,153	$860	$942	$3,996	$863	$1,043	$1,042
Anise Seed Powder	$497	$776	$565	$398	$2,236	$997	$1,109	$816	$898	$3,820	$819	$999	$998
Anise Seeds	$508	$787	$576	$409	$2,280	$1,008	$1,120	$827	$909	$3,864	$830	$1,010	$1,009
Annatto Seed (Ground)	$519	$798	$587	$420	$2,324	$1,019	$1,131	$838	$920	$3,908	$841	$1,021	$1,020
Annatto Seed (Whole)	$530	$809	$598	$431	$2,368	$1,030	$1,142	$849	$931	$3,952	$852	$1,032	$1,031
Asafoetida Powder	$662	$941	$730	$563	$2,896	$1,162	$1,274	$981	$1,063	$4,480	$984	$1,164	$1,163
Asafoetida Seed Powder	$673	$952	$741	$574	$2,940	$1,173	$1,285	$992	$1,074	$4,524	$995	$1,175	$1,174
Basil Leaf (Ground)	$783	$1,062	$851	$684	$3,380	$1,283	$1,395	$1,102	$1,184	$4,964	$1,105	$1,285	$1,284
Basil Leaf (Whole)	$816	$1,095	$884	$717	$3,512	$1,316	$1,428	$1,135	$1,217	$5,096	$1,138	$1,318	$1,317
Basil Leaf Powder	$827	$1,106	$895	$728	$3,556	$1,327	$1,439	$1,146	$1,228	$5,140	$1,149	$1,329	$1,328
Bay Leaf (Ground)	$442	$721	$510	$343	$2,016	$942	$1,054	$761	$843	$3,600	$764	$944	$943
Bay Leaf (Whole)	$563	$842	$631	$464	$2,500	$1,063	$1,175	$882	$964	$4,084	$885	$1,065	$1,064
Bay Leaf Powder	$574	$853	$642	$475	$2,544	$1,074	$1,186	$893	$975	$4,128	$896	$1,076	$1,075
Caraway Seed (Ground)	$552	$831	$620	$453	$2,456	$1,052	$1,164	$871	$953	$4,040	$874	$1,054	$1,053
Caraway Seed (Whole)	$354	$633	$422	$255	$1,664	$854	$966	$673	$755	$3,248	$676	$856	$855
Caraway Seed Powder	$365	$644	$433	$266	$1,708	$865	$977	$684	$766	$3,292	$687	$867	$866
Cardamon Seed Powder	$464	$743	$532	$365	$2,104	$964	$1,076	$783	$865	$3,688	$786	$966	$965
Cardamon Seed (Ground)	$475	$754	$543	$376	$2,148	$975	$1,087	$794	$876	$3,732	$797	$977	$976
Cardamon Seed (Whole)	$486	$765	$554	$387	$2,192	$986	$1,098	$805	$887	$3,776	$808	$988	$987
Carob Pods (Ribbled)	$409	$688	$477	$310	$1,884	$909	$1,021	$728	$810	$3,468	$731	$911	$910
Carob Pods (Ribbled) Powder	$420	$699	$488	$321	$1,928	$920	$1,032	$739	$821	$3,512	$742	$922	$921
Carob Powder (Raw)	$596	$875	$664	$497	$2,632	$1,096	$1,208	$915	$997	$4,216	$918	$1,098	$1,097

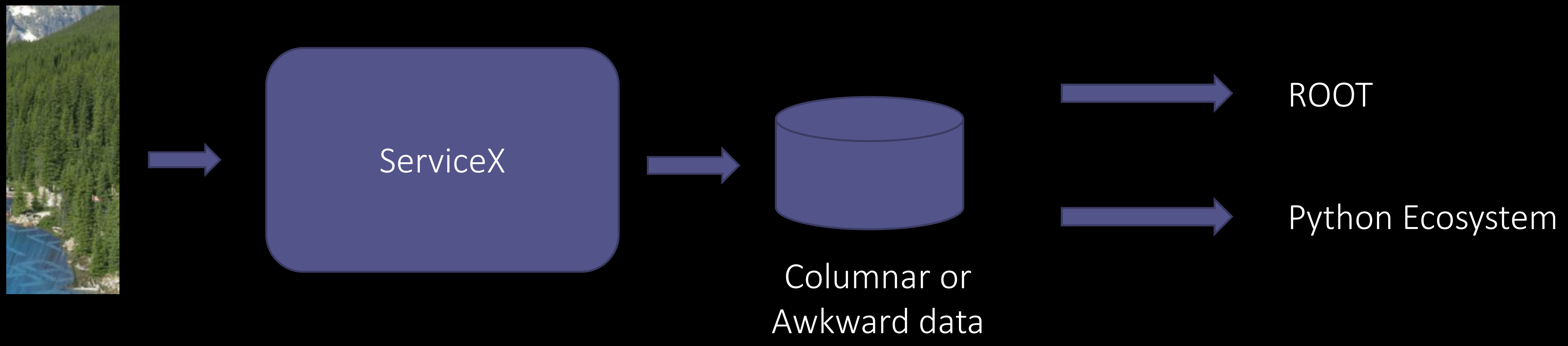
Custom Views / Sales Data / Bulk Order Discounts / Filters and Tables / Grouping Sales / Subtotal / Worksh

(numpy ecosystem in Python)

Build out and Extend the numpy ecosystem

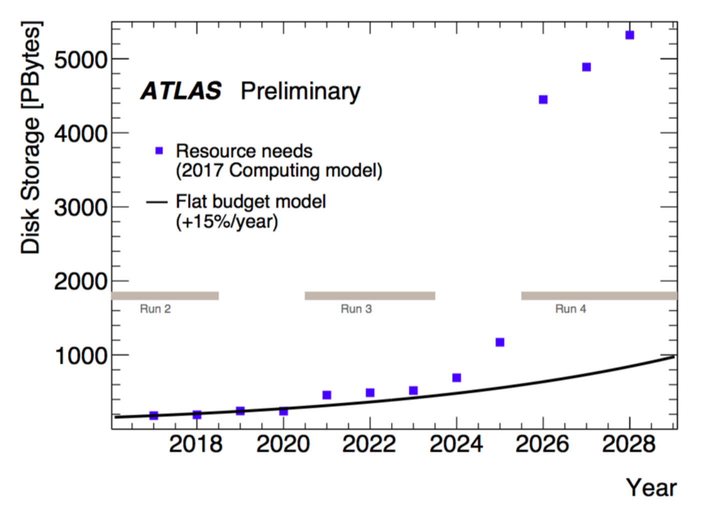
Awkward Array

Data Access



- Give it a simple data query: grab only the data you want
- Transmits the data in a compact format to the world

ROOT Tools



Particle physics lives and dies by the size of its on-disk data

- PB’s of data
- SSD’s are very expensive – but fastest way to access data

RNTuple

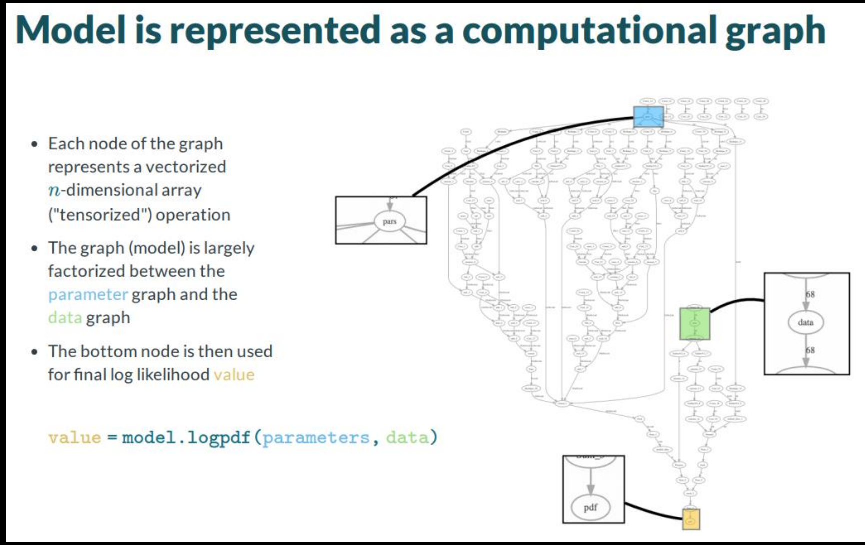
- New data format tied to ROOT
    - ~25% smaller than current file formats
    - 2-5 times faster read times
- Should be favorable compared with industry formats!

Statistics

Fundamentally, calculating the sensitivity, limit, and discovery of a signal is fitting

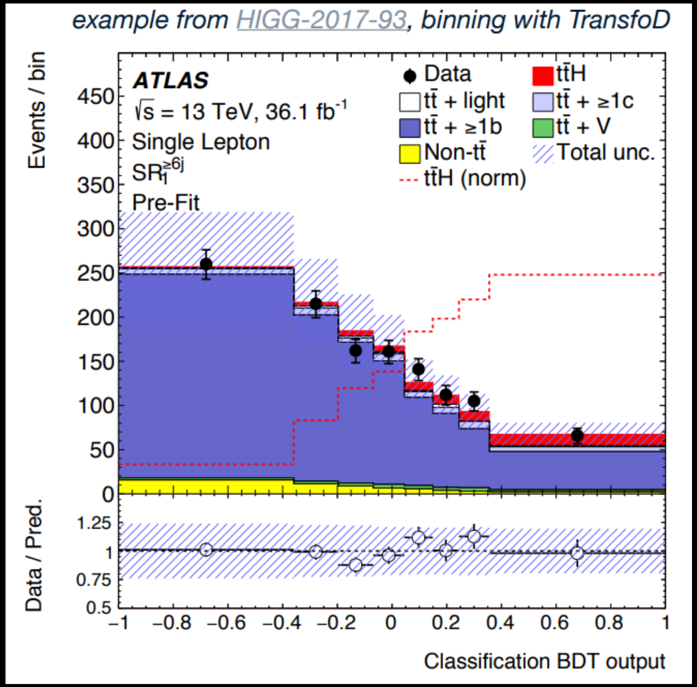
- Expected Background
- Systematic and Statistical Errors on background
- Data

→ How Compatible?



TensorFlow

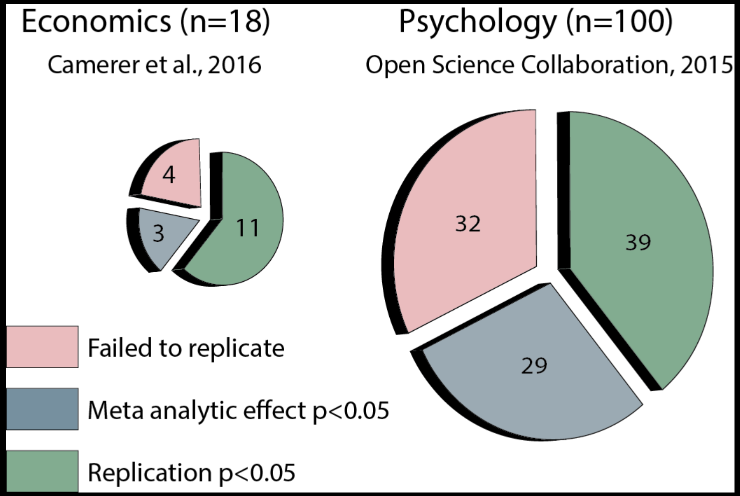
PyTorch



OpenData and Reuse

OpenData and Particle Physics

"The replication crisis"



Publicly Availible Data Has Benefits Too

- Non-collaboration members can produce science
- New models can be tested
- Tool authors have realistic datasets to experiment on
- Easy to link in other communities for tooling

U.S. DEPARTMENT OF ENERGY | Office of Science

NSF

CERN

- Funding agencies
- Laboratories
- The Community

CERN Has an OpenData Policy!

CERN Open Data Policy for the LHC Experiments

2020-12-11 by CERN

News

The CERN Open Data Policy reflects values that have been enshrined in the CERN Convention for more than sixty years that were reaffirmed in the European Strategy for Particle Physics (2020)[1], and aims to empower the LHC experiments to adopt a consistent approach towards the openness and preservation of experimental data. Making data available responsibly (applying FAIR standards[2]), at different levels of abstraction and at different points in time, allows the maximum realisation of their scientific potential and the fulfillment of the collective moral and fiduciary responsibility to member states and the broader global scientific community. CERN understands that in order to optimise reuse opportunities, immediate and continued resources are needed. The level of support that CERN and the experiments will be able to provide to external users will depend on available resources.

This policy relates to the data collected by the LHC experiments, for the main physics programme of the LHC — high-energy proton–proton and heavy-ion collision data. The foreseen use cases of the Open Data include reinterpretation and reanalysis of physics results, education and outreach, data analysis for technical and algorithmic developments and physics research. The Open Data will be released through the CERN Open Data Portal which will be supported by CERN for the lifetime of the data. The data will be tailored to the different uses, and will be made available in formats defined by each experiment that afford a range of opportunities for long-term use, reuse and preservation. In general, four levels of complexity of HEP data have been identified by the Data Preservation and Long Term Analysis in High Energy Physics (DPHEP) Study Group[3], which serve varying audiences and imply a diversity of openness solutions and practices.

What it means...

	ALICE	ATLAS	CMS	LHCb
Run-2	2 PB	0.5 PB	2 PB	10 PB (including Run-1)
Run-3	4 PB	1 PB	4 PB	45 PB
Total	**6 PB**	**1.5 PB**	**6 PB**	**55 PB**

Can you imagine anyone running over 1 PB of data?

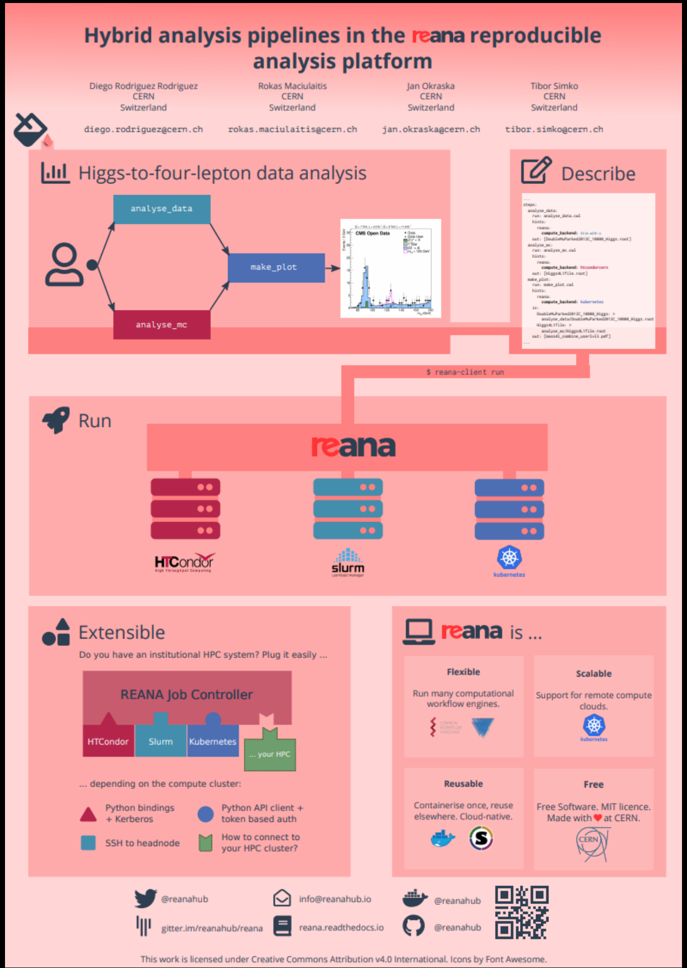
reana

Reproducible research data analysis platform

ATLAS EXPERIMENT

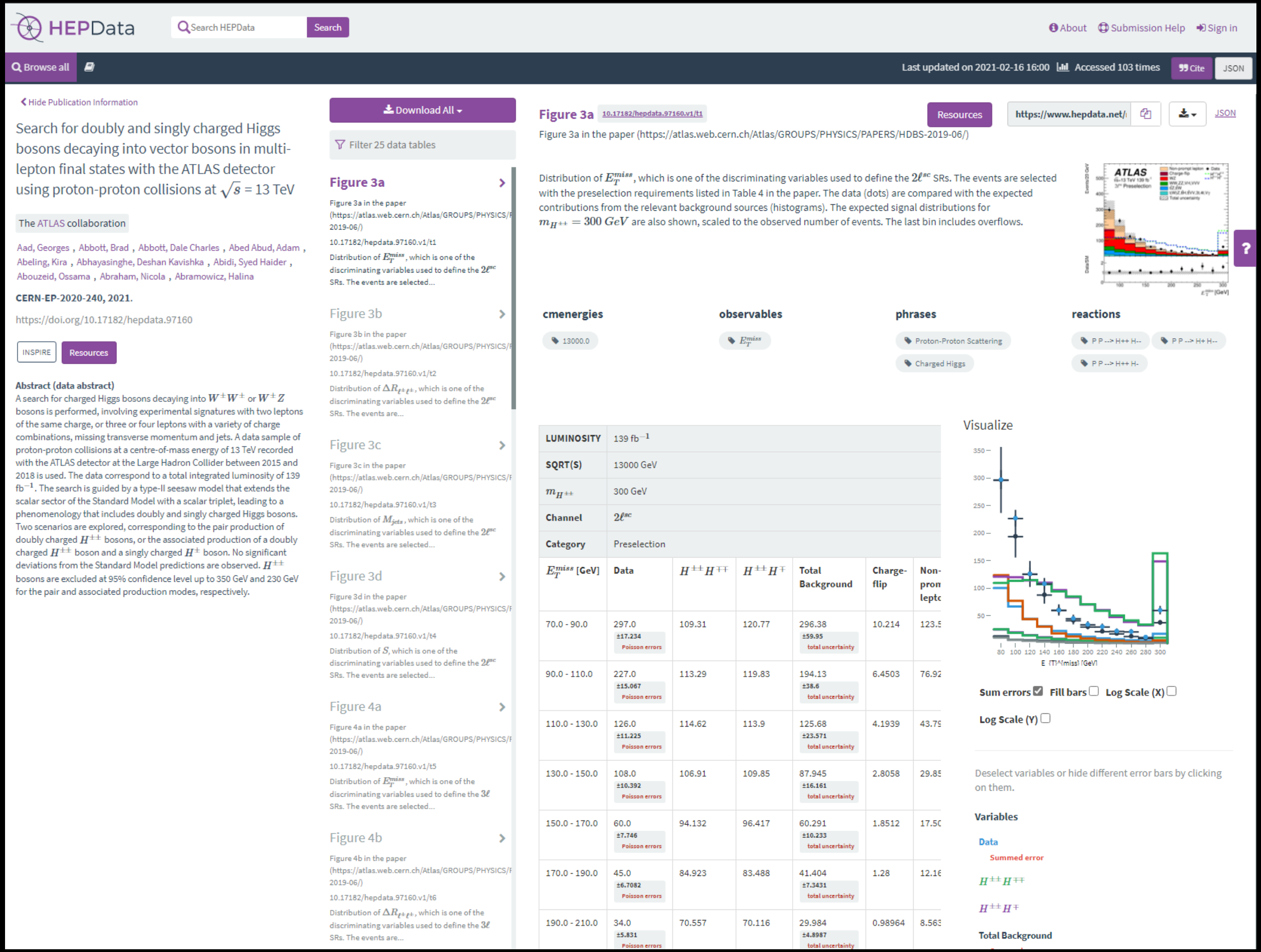
- Defines the analysis
  - Background Model
  - All Errors
- Defines sample set of signals
- Analysis is freeze-dried

- New set of signals
- Uses reana resources



HEPData

Repository for publication-related High-Energy Physics data



- High Level Results for published papers
- Computer readable

No more overlaying transparencies!

Last Thoughts

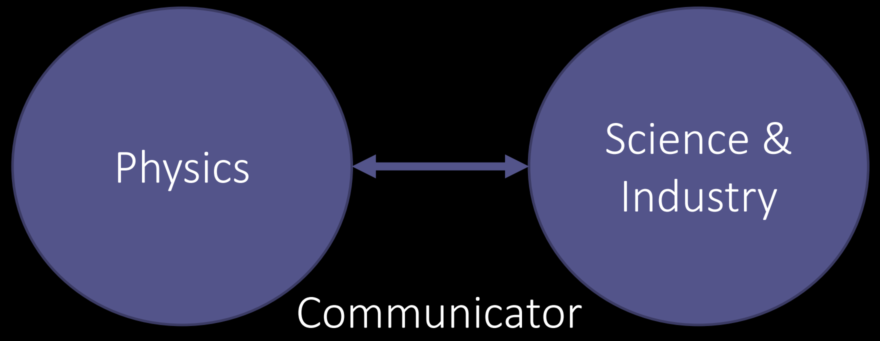
The Community Outside of HEP

All branches of science are taking advantage of the recent advances in computing

"Can a physicist get a Ph.D. without at least understanding Machine Learning?"

- Astrophysics
- Particle Physics
- Nuclear Physics (Jefferson Lab's efforts)

Several examples in this talk
- Pulling techniques and software from industry into the community
- Designing our own software and putting out into industry



Conclusions

Mention Quantum Computing
Facilities, data management

$100M is being spent on LHC upgrades

- We must extract as much physics from this and past investments as possible

Larger dataset, higher luminosity, new detectors

- Software and computing infrastructure will require a major upgrade

Physics is not an island

- Industry and others have many mature and useful tools
- We have years of experience to give back

Software and Computing is a huge subject

- Quantum Computing
- Major advances in computing facilities
- Data management
- Machine Learning Details would require several talks all on its own

The community is vibrant

- And not at all limited to just Particle Physics

Community White Paper

Contents

A Roadmap for HEP Software and Computing R&D for the 2020s

HEP Software Foundation[1]

Abstract: Particle physics has an ambitious and broad experimental programme for the coming decades. This programme requires large investments in detector hardware, either to build new facilities and experiments, or to upgrade existing ones. Similarly, it requires commensurate investment in the R&D of software to acquire, manage, process, and analyse the shear amounts of data to be recorded. In planning for the HL-LHC in particular, it is critical that all of the collaborating stakeholders agree on the software goals and priorities, and that the efforts complement each other. In this spirit, this white paper describes the R&D activities required to prepare for this software upgrade.

Mathusla

We can still do it
on one computer

Stuff to check

CMS or ATLAS institute?