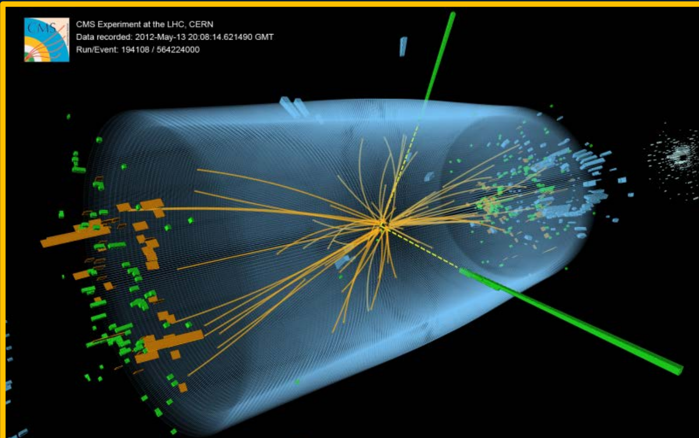
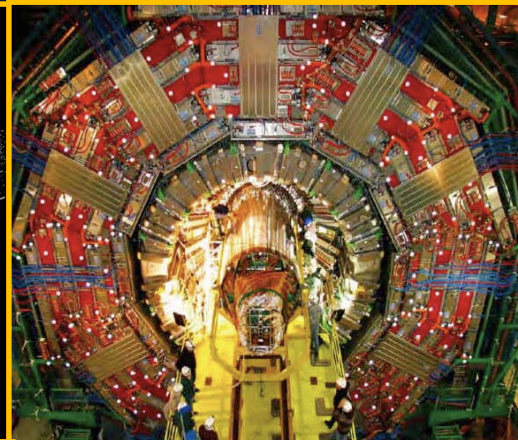
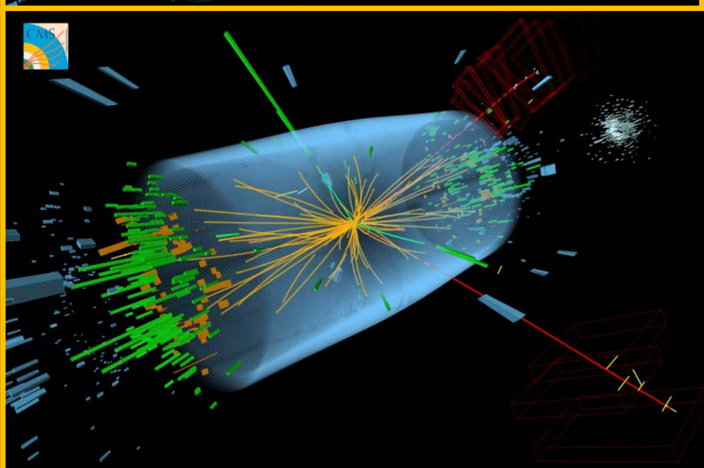
Physics at the LHC: *A New Window on Matter, Spacetime and the Universe*

Meeting the Global Challenges of Exascale Data

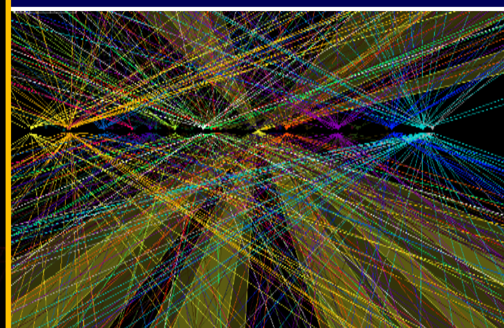




- *A Half-Century Search*
- *The LHC and CMS*
- *Discovery of a New Boson*
- *Beyond the Standard Model: The Future*



50 Vertices
14 Jets, 2 TeV



Gateway to a New Era

Harvey B Newman, Caltech

Simons Institute Workshop on Real-Time Decision Planning

Berkeley, June 27, 2016

On behalf of the Caltech Team + Partners

Discovery of a Higgs Boson
July 4, 2012; Nobel Prize 2013-

Physicists Find Elusive Particle Seen as Key to Universe

The New York Times

Englert | Higgs

2013

Theory : 1964

LHC + Experiments Concept: 1984

Construction: 2001

Operation and Discovery: 2009-12

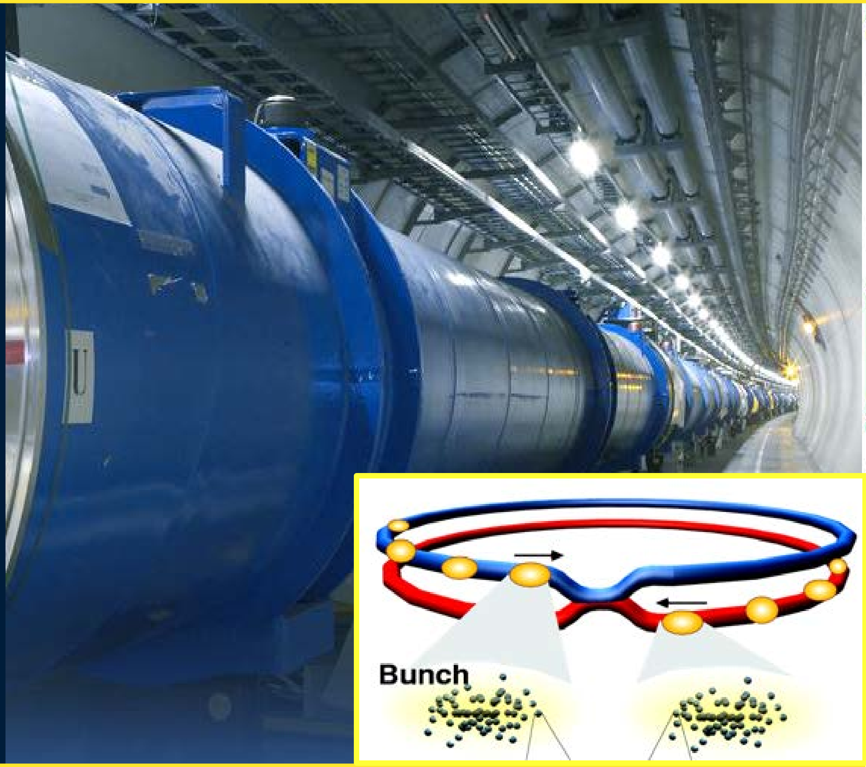
Advanced Computational Methods and Networks Were Essential to the Higgs Discovery and Every Ph.D Thesis of the last 20+ Years

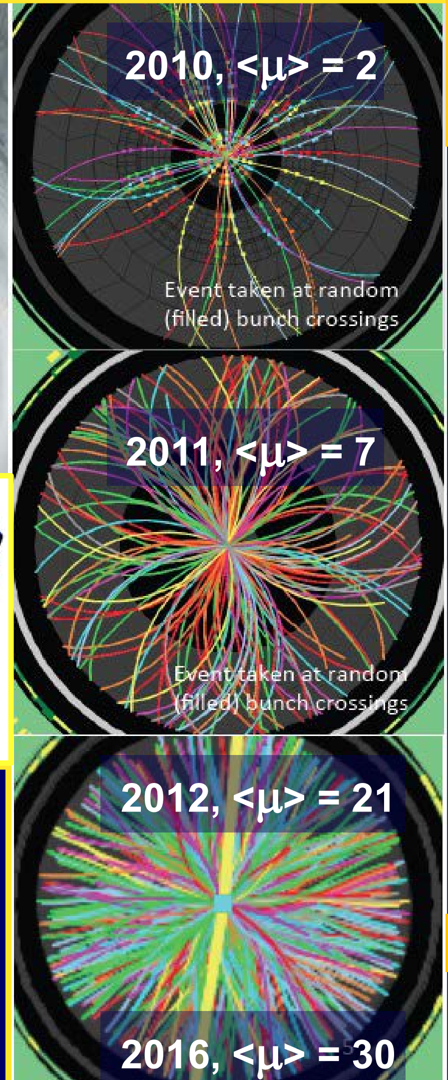
New Innovative Methods will be Essential to Future Discoveries, the Ph. D Theses to Come

The LHC: Deep into the Multi-TeV Scale

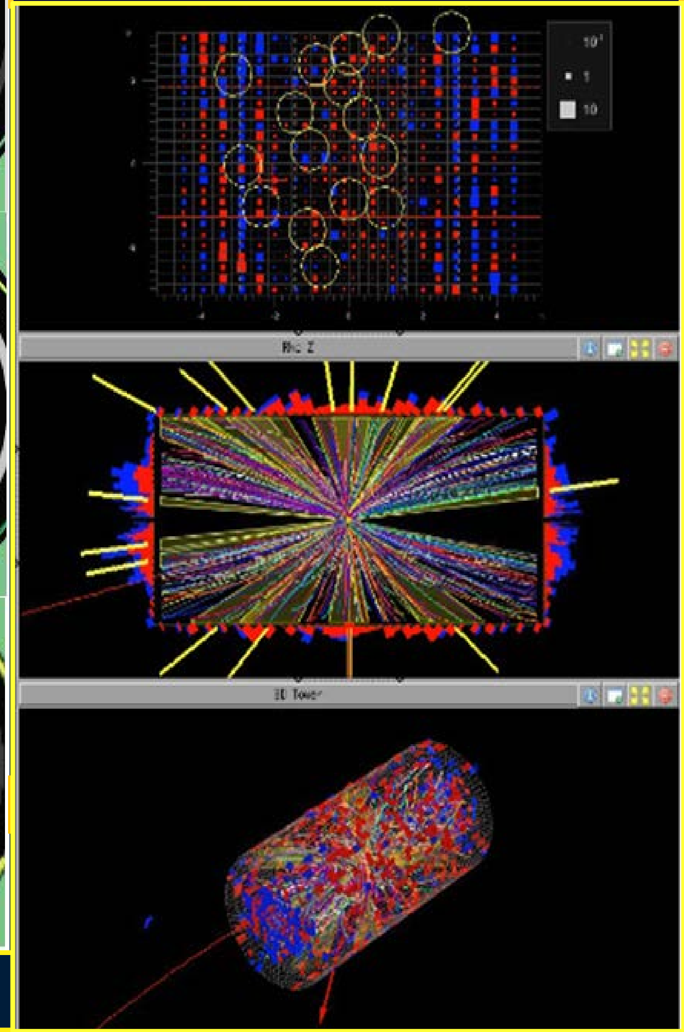
HEP: Complex Data. Challenge of Pileup

~4 X 10^{15} pp Collisions
~2M Higgs Bosons
created So Far





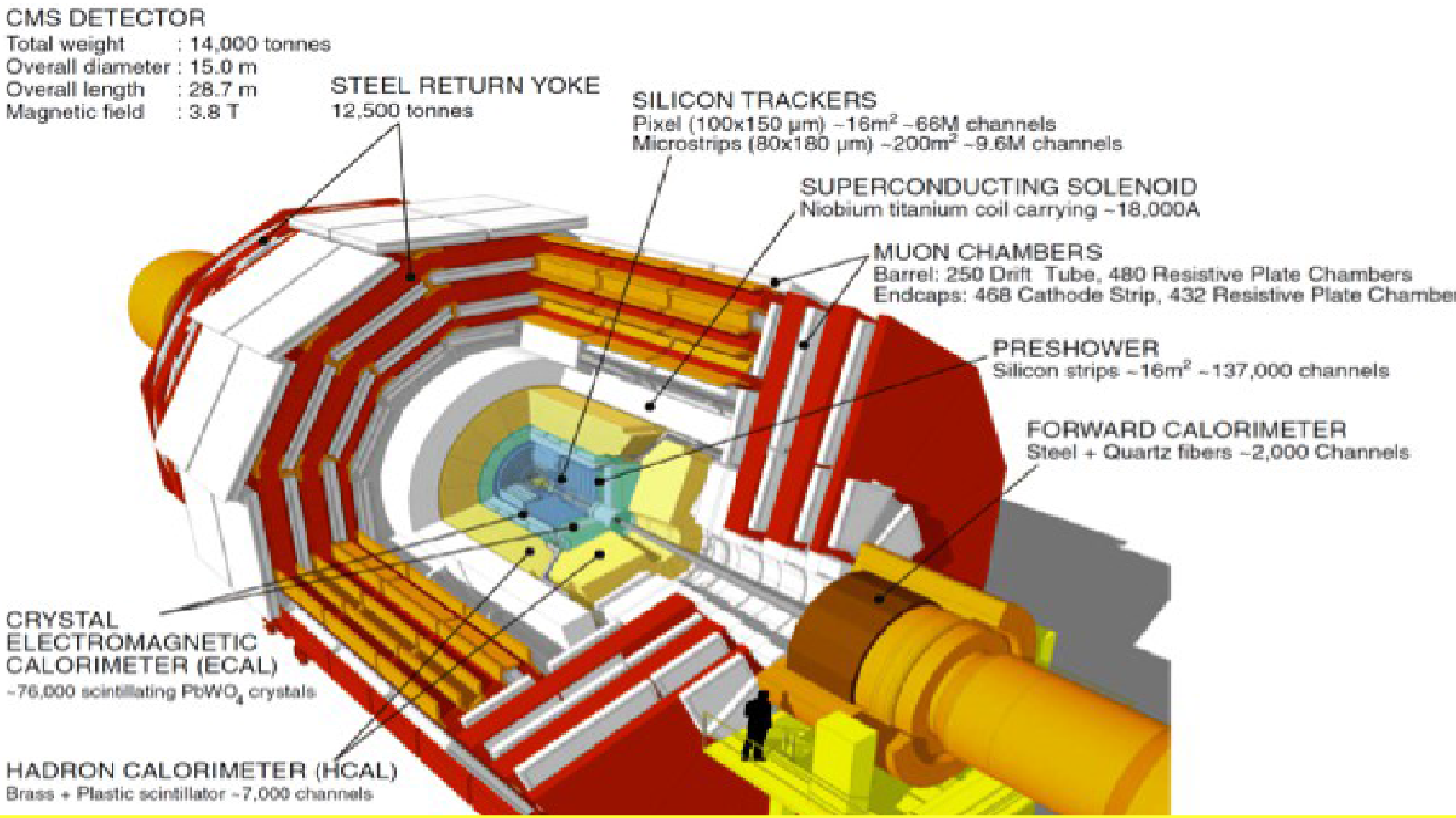
HL LHC 140-200



Run2 and Beyond will bring:

- Higher energy and intensity
- Greater science opportunity
- Greater data volume & complexity
- A new Realm of Challenges

the Compact Muon Solenoid

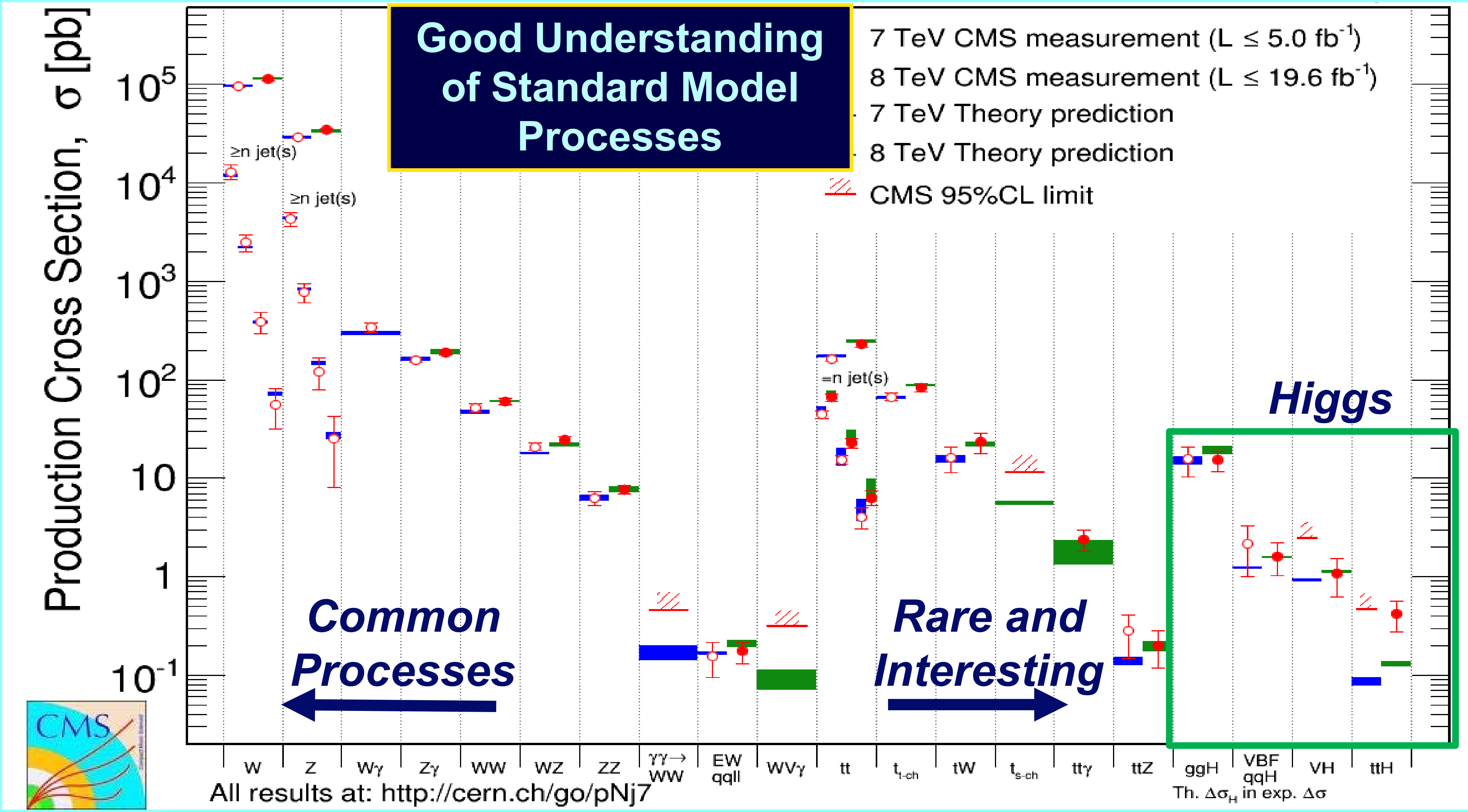


CMS is a Highly Heterogeneous System

Raw data is 100M channels sampling every 25 nsec: 1 petabit/sec

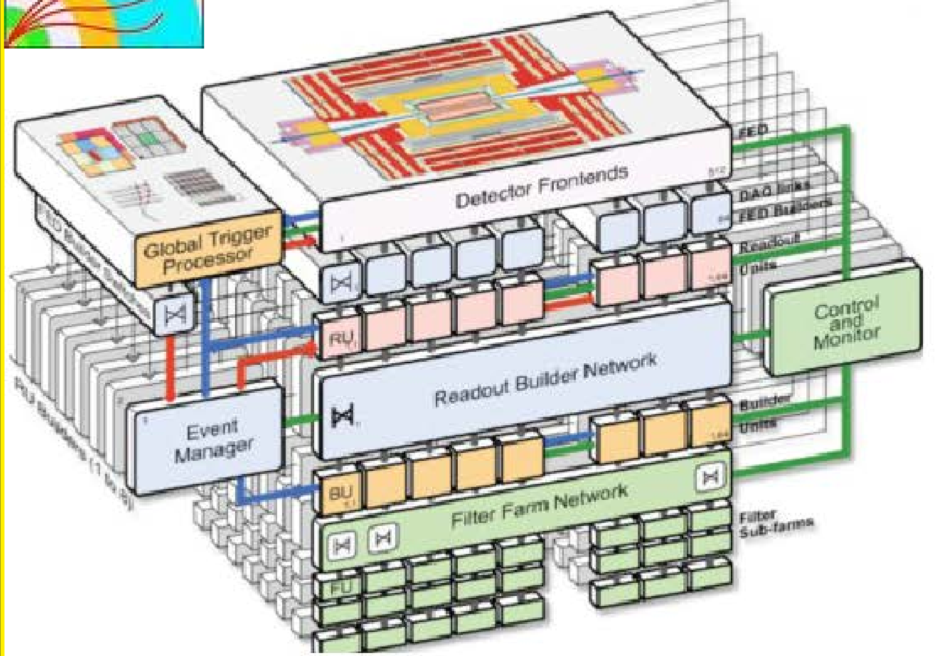
50 Exabytes/Day in Readout and Online Processing

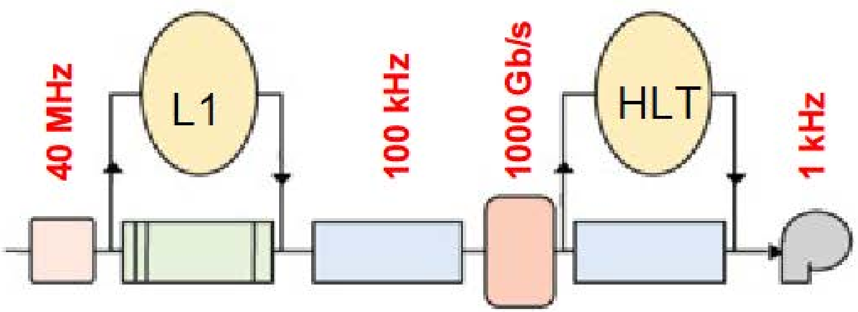
Finding Rare Signals: Down to 1 in 10^{13}

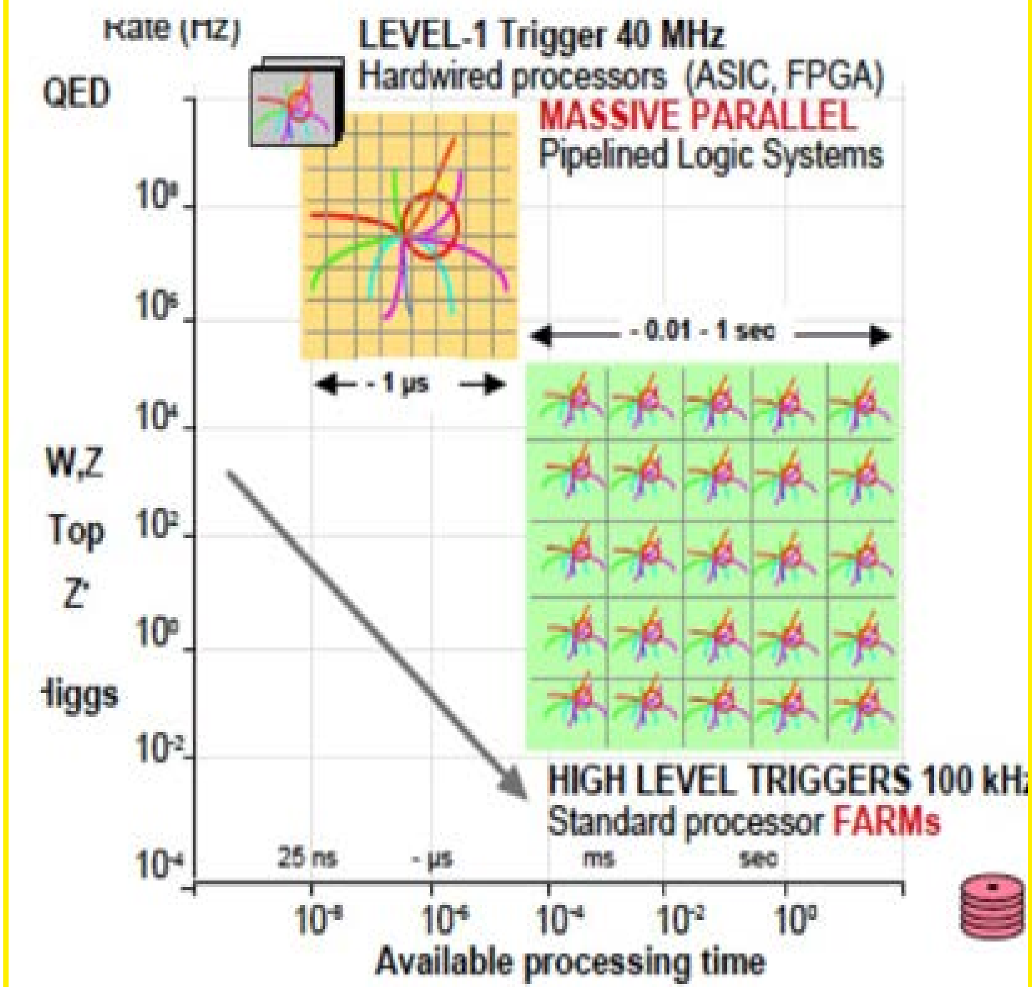


Many Orders of Magnitude Rejection Required
in Order to Extract Interesting Events

Event Triggering





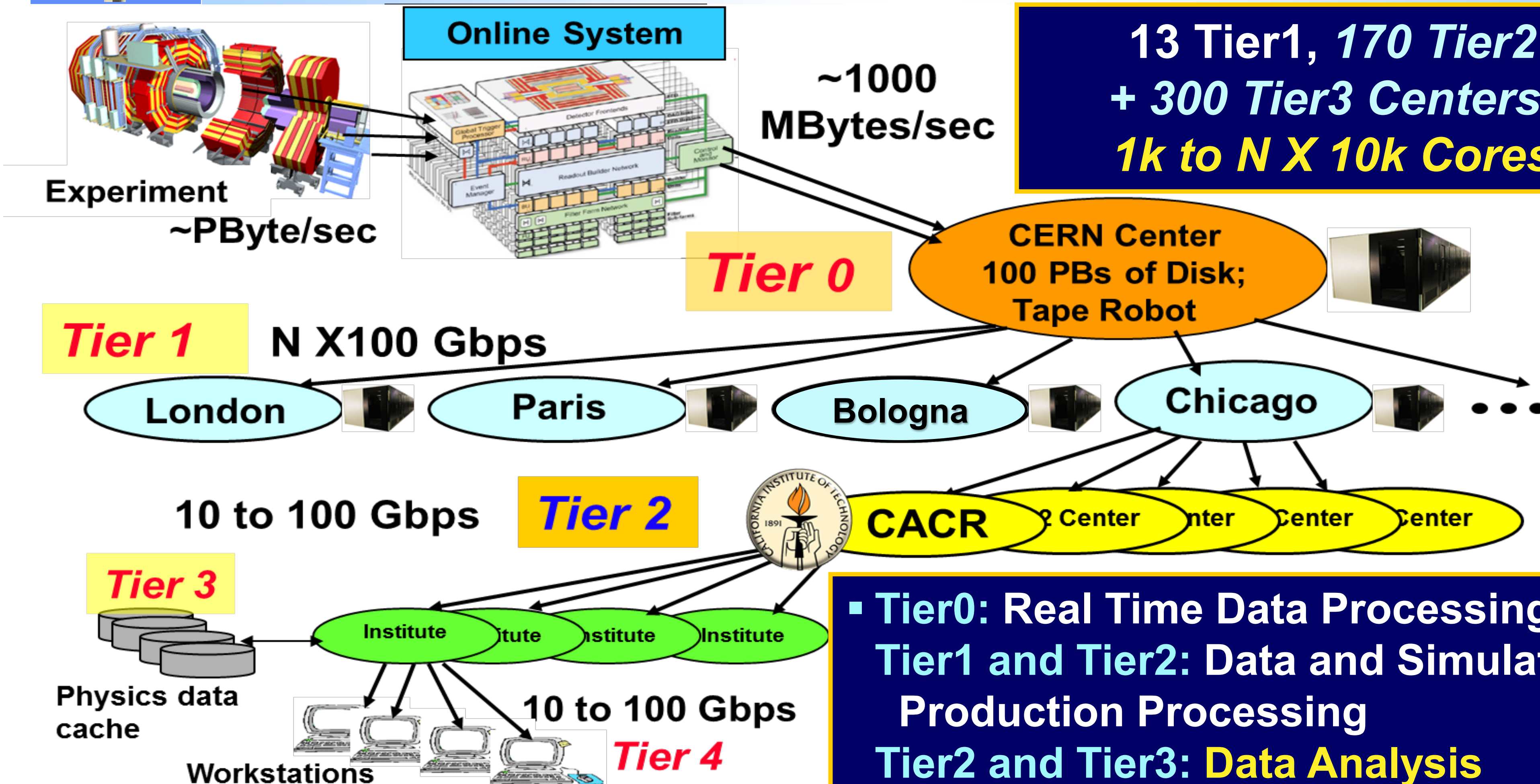


Massively parallel electronic infrastructure makes a prime selection

Refined decision in a software defined trigger: N X 10k cores

Little processing time for selection: ML for a faster algorithms

Global Data Flow: LHC Grid Hierarchy
A Worldwide System Invented by Caltech (1999)



13 Tier1, *170 Tier2*
+ 300 Tier3 Centers:
1k to N X 10k Cores

- Tier0: Real Time Data Processing
  Tier1 and Tier2: Data and Simulation Production Processing
  Tier2 and Tier3: Data Analysis

Increased Use as a Cloud Resource (Any Job Anywhere)
Increasing Use of Additional HPC and Cloud Resources
A Global Dynamic System: Fertile Ground for Control with ML

Global Networks Today and Tomorrow
Science Program Data Flow Challenges

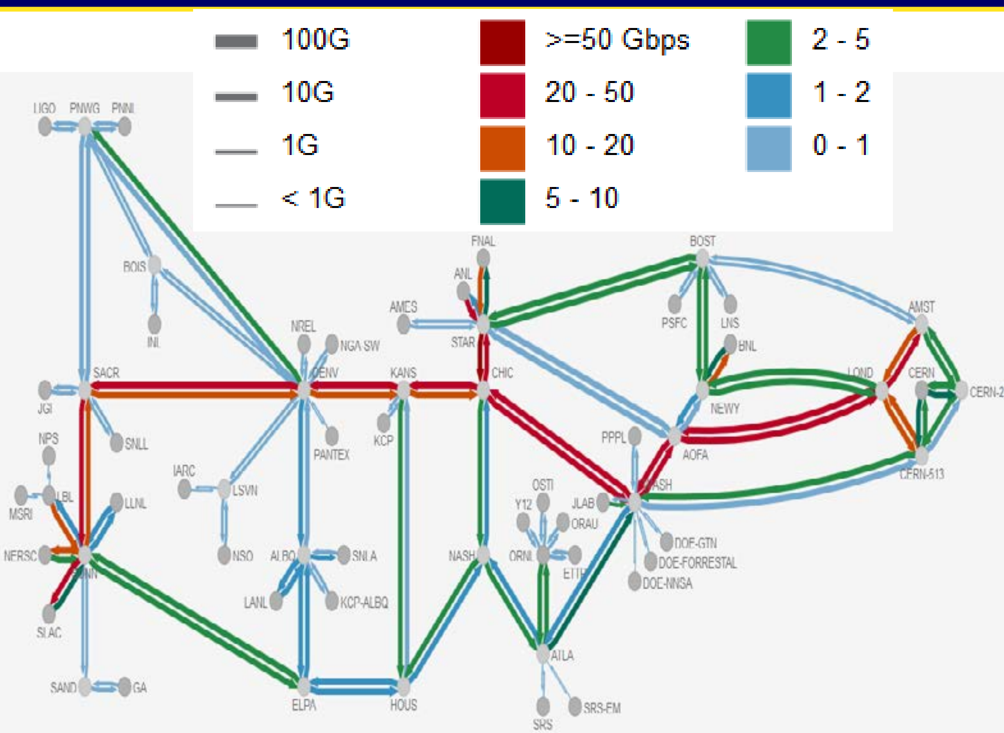
- Volume of Global Data Flow + Expansion Rate: Challenging the World's Research and Education Networks
- Complex Data and Workflow
- Worldwide Inter-Facility Connections: A Complex System Over Networks of Varying Capacity and Reliability
- Plan to Meet the Challenges: Integrating Worldwide Operations in an Intelligent, SDN-Driven System
  - Optimized Using Deep Learning
  - Coupled to Modeling and Simulation, Pervasive Monitoring and State Tracking
  - Game Theory to Find Effective Metrics and Stable Solutions

Energy Sciences Network

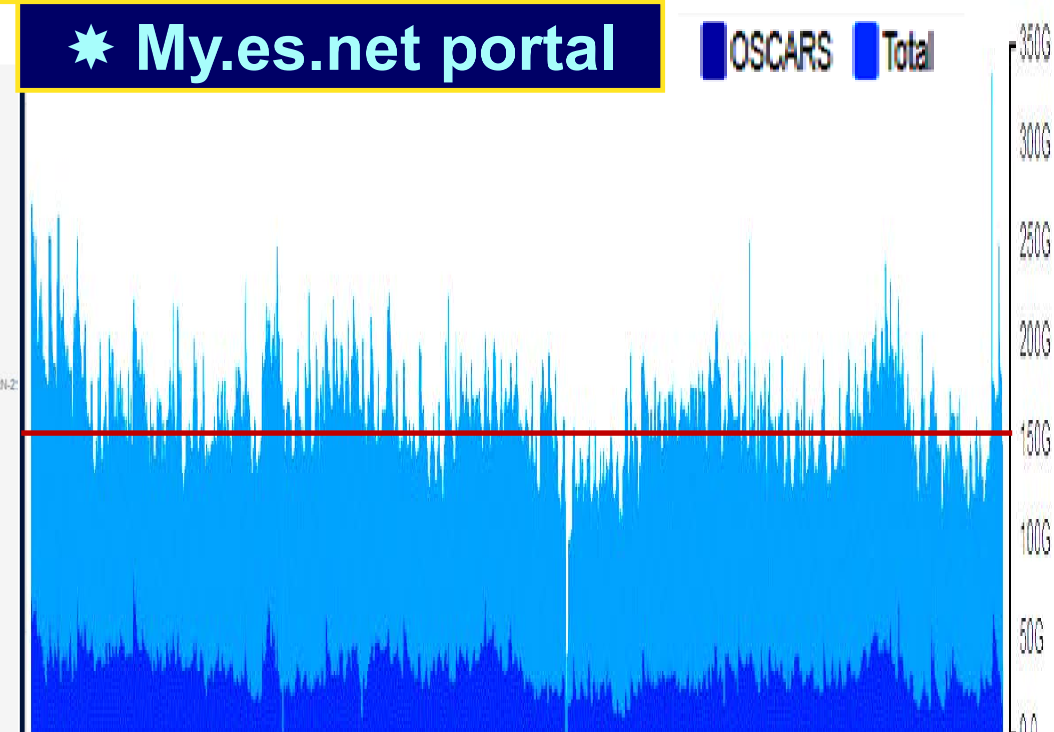
ESnet

- 150-250 Gbps Typical; Peaks to 300+ Gbps
- 45.7 PB input data volume in May 2016

- Long term traffic growth trend is 72%/year (*10X per 4 Yrs*)
- But 2015-16 growth is above this trend: *+104% in 12 Months*
- LHCONE growth in 2015-16: *+254%, to 16.4 Pbytes/month*



✸ My.es.net portal

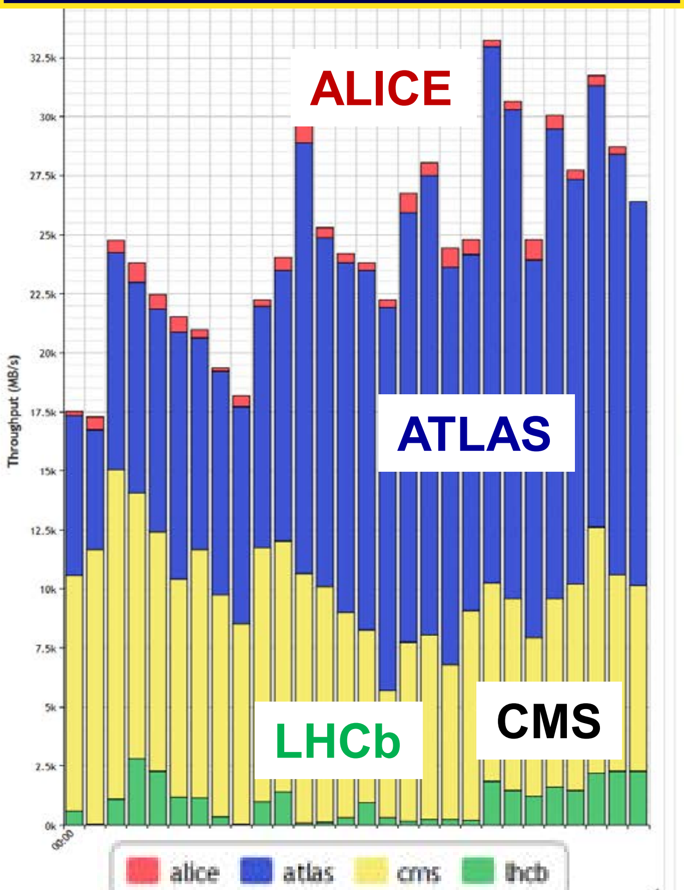


✸ ESnet6: the next SDN-enabled generation, is planned by ~2019

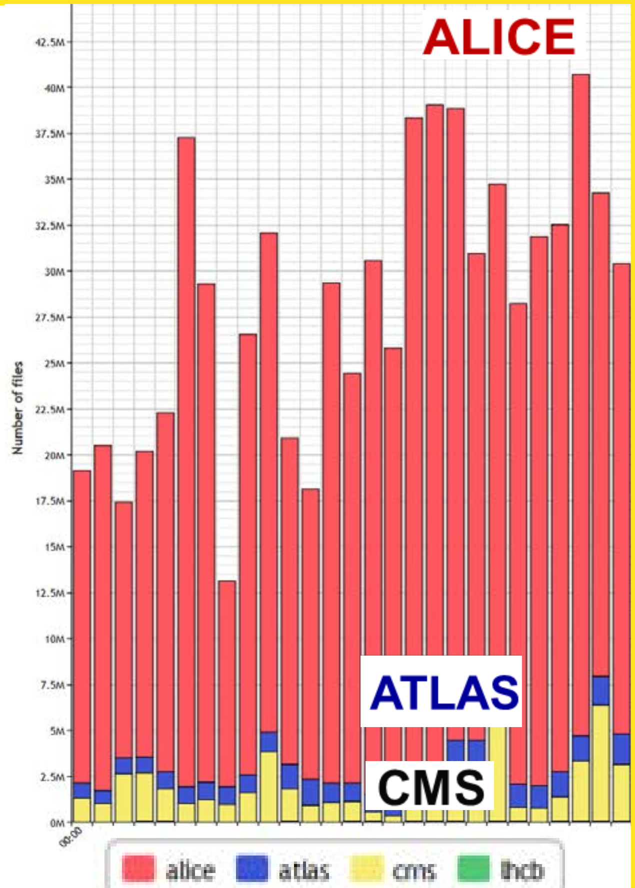
Complex Workflow: the Flow Patterns Have Increased in Scale and Complexity, even at the start of LHC Run2

WLCG Dashboard Snapshot April-May: Patterns Vary by Experiment

Transfer Throughput



Transfers Done/Day



28 GBytes/s Typical

To 40 GBytes/s Peak Transfer Rates

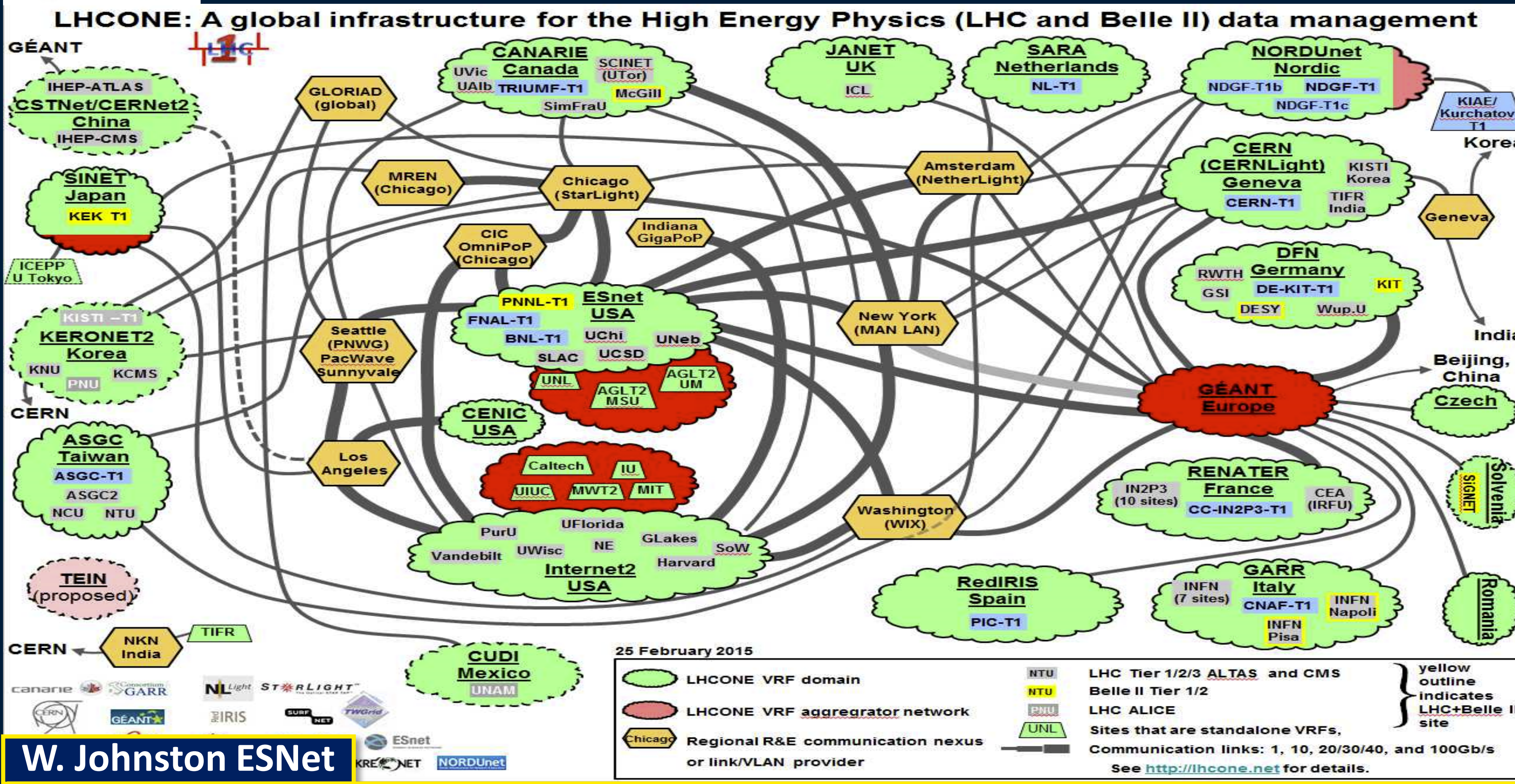
Complex Workflow

- **Multi-TByte Dataset Transfers**
- **Transfers of 13-41 Million Files Daily**
- **Access to Tens of Millions of Object Collections/Day**
- **>100k of remote connections (e.g. AAA) simultaneously**

2.7X Traffic Growth (+166%) in Last 12 Months; +60% in April

LHCONE: a Virtual Routing and Forwarding (VRF) Fabric

A global infrastructure for HEP (LHC and Belle II) data management



W. Johnston ESNet

Good News: The Major R&E Networks Have Mobilized on behalf of HEP

Issue: A complex system with limited scaling properties.

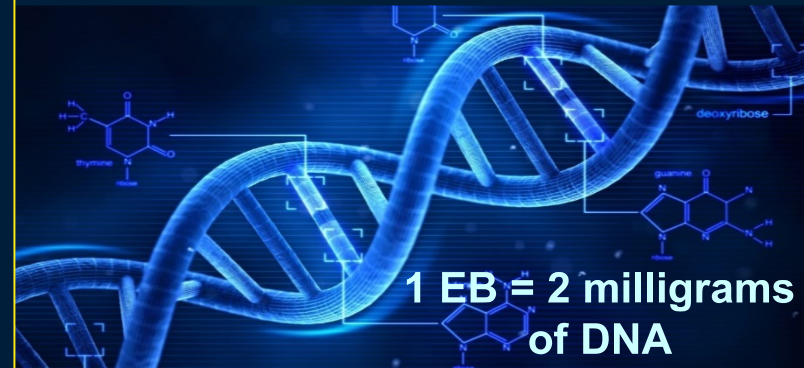
LHCONE traffic grew by *3.5X* in last 12 months: *a challenge during Run2*

High Luminosity LHC Era 2026-2037

A New Era of Exascale Network Challenges

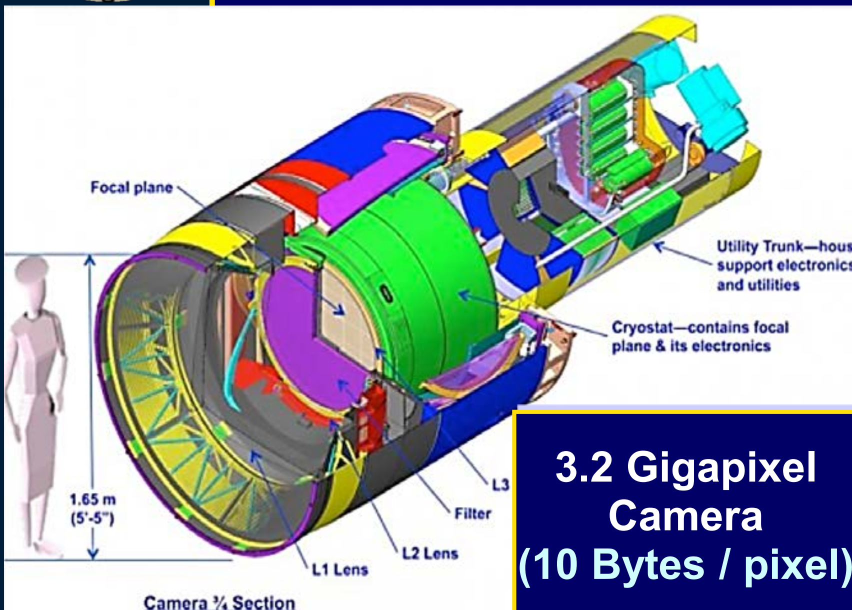
- **Networks**
- Projected needs are growing at an exponential rate: beyond affordable budgets, Moore's Law.

  100-1000X by HL LHC in 2026
- Needs of other fields continue to grow, HEP will face stiff competition for network resources.
- Need for innovation: a new generation of Intelligent software driven global systems coordinating computing, storage and network resources



Earth Observation

LSST + SKA Data Movement
Upcoming *Real-time* Challenges for Astronomy



3.2 Gigapixel Camera (10 Bytes / pixel)

SKA

- ❒ **Networks**: Dedicated 2 X 100G for image data, Additional 100Gs for other traffic, and diverse paths
- ❒ Lossless compressed Image size = 2.7GB (~5 images transferred in parallel over a 100 Gbps link)
  - ❒ Custom transfer protocols for images (UDP Based)
- ❒ Real-time Challenge: delivery in seconds **to catch cosmic "events"**
- ❒ **+ SKA in Future: 3000 Antennae covering > 1 Million km2;** *15,000 Terabits/sec to the correlators ➡ 1.5 Exabytes/yr Stored*

Higgs Boson Discovery and Properties

Example Analysis with BDTs

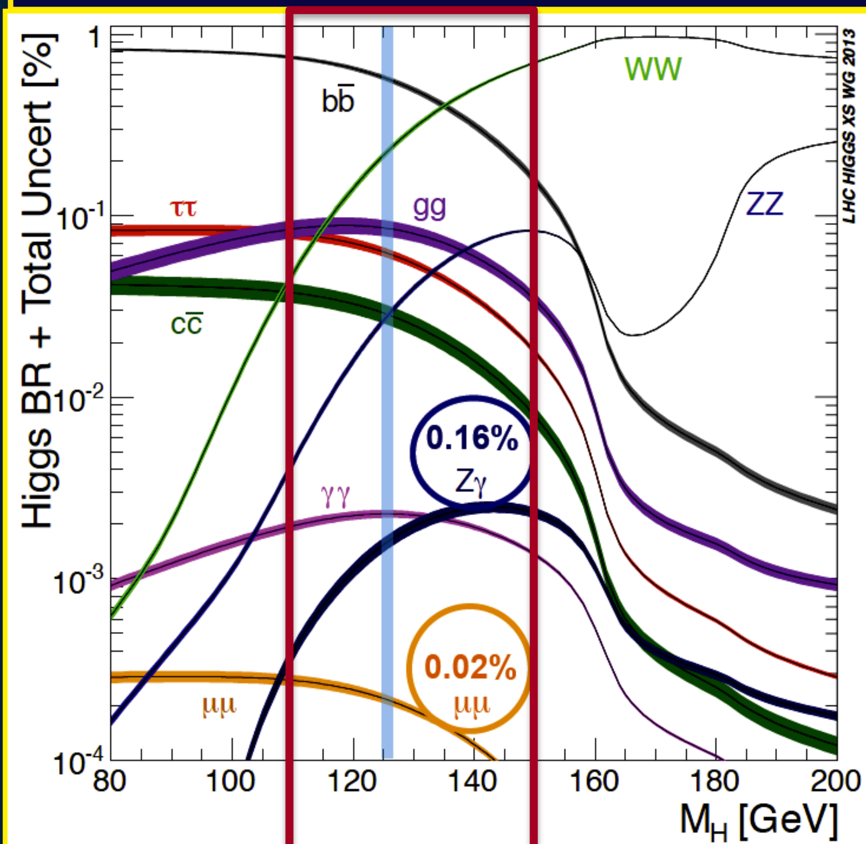
Higgs Boson Decays

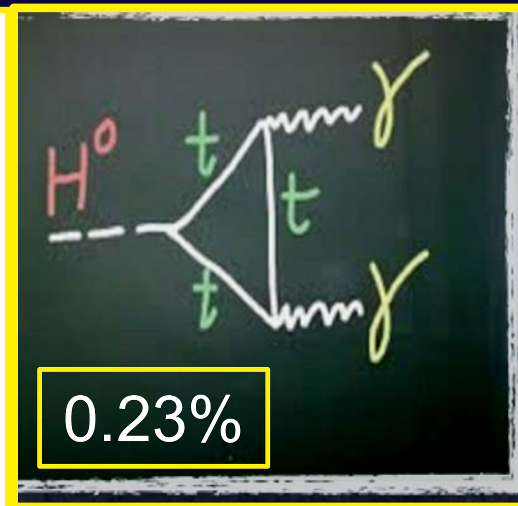
Many Modes Contribute

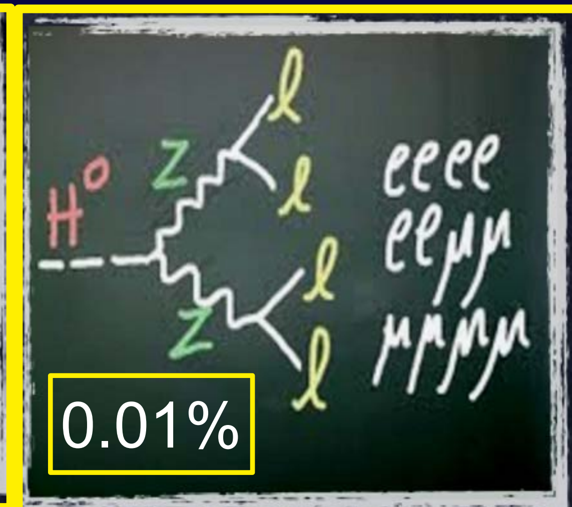
125 GeV Region: Rich and Challenging: ZZ, $\gamma\gamma$, WW, $\tau\tau$, bb

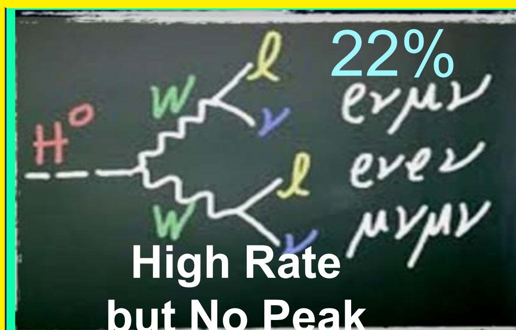
Rare High Mass Resolution Channels Have a Special Role:

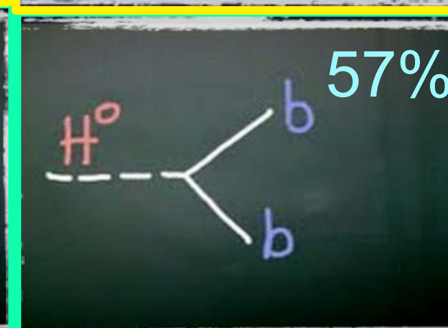
$H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow$ 4 Leptons





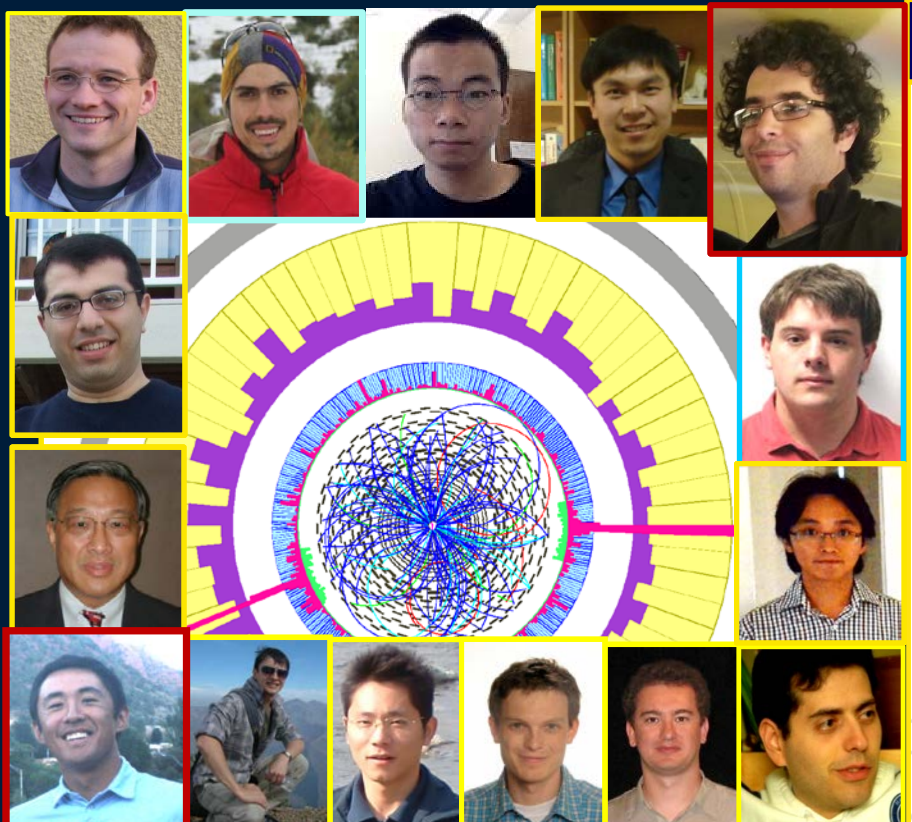




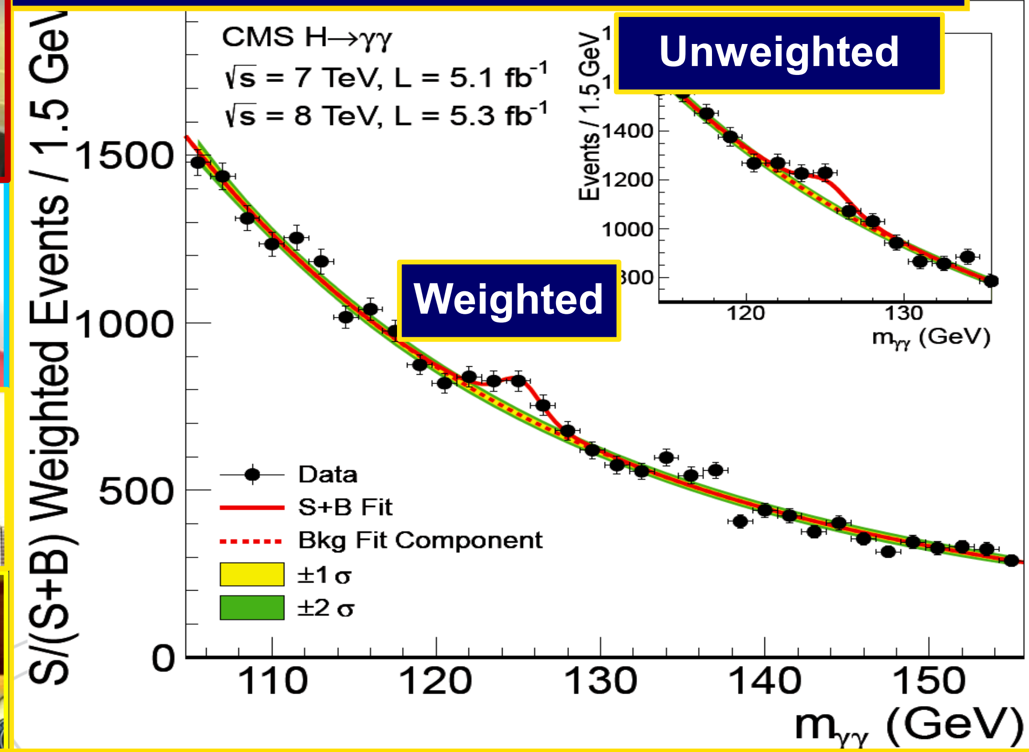


Higgs Discovery at the LHC: H→ γγ

Narrow diphoton mass peak over smooth background



Diphoton Mass Spectrum



Keys: 1) Precise Calibration 2) Optimized Photon Identification 3) Precise Energy Scale 4) Innovative Analysis Methods

\+ *Many* Caltech postdocs and students over last 20+ years

A Stream of Innovations; from the first BDT Analysis

"Razor Variables" for New Physics Searches

Next: Deep Learning Approaches

CMS Experiment at the LHC, CERN
Data recorded: 2012-May-13 20:08:14.621490 GMT
Run/Event: 194108 / 564224000

$H \rightarrow \gamma\gamma$

Search for a narrow mass peak with two isolated high E_T photons on a smoothly falling background

- **High Resolution: ~1% in barrel**

Photon

Thesis Defense 11/7/12

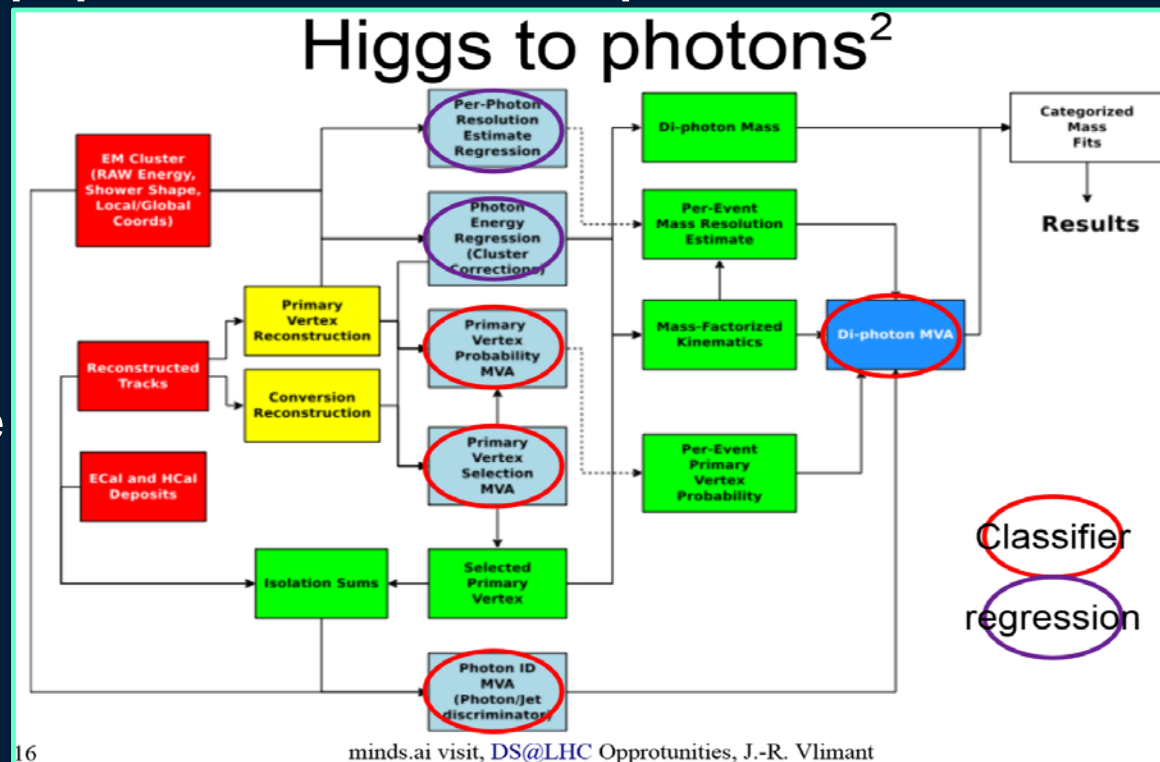
- Analysis optimized categorizing events by γ ID and vertex efficiency; purity & mass resolution.
- Specific di-jet tag categories targeting VBF production mode (Higher S/B)
- Exclusive categories (e, μ, E_T^{Miss}) targeting WH, ZH Associated Production

Photon

$M_{\gamma\gamma}$=125.9 GeV
σ_M/M=0.9%

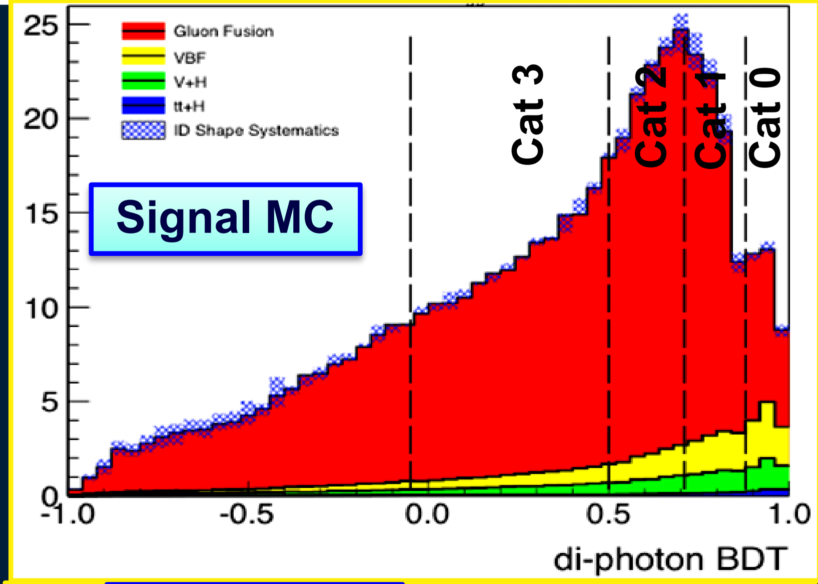
H $\rightarrow \gamma\gamma$ Analysis Overview

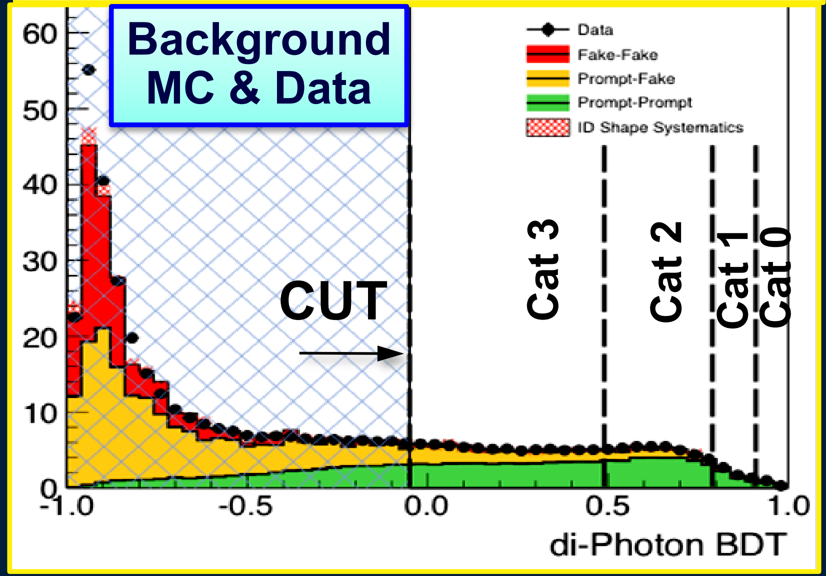
- BDT Analysis: **Fit to Diphoton mass $m_{\gamma\gamma}$ in event categories**
  - 4 event classes based on a diphoton BDT output , 2 di-jet categories (VBF) + 3 Exclusive categories (VH): Electron, Muon, E_T^{Miss}
  - Score according to Probability (correct vertex), per-event $m_{\gamma\gamma}$ resolution estimate, prompt photon ID score, + diphoton kinematics
- **Cross-checked with traditional cut based analysis**
  - photon ID & mass fit in categories
  - 2 angular x 2 shower shape categories with different Signal/Background ratios; + 2 di-jet + 3 Exclusive Categories



Diphoton MVA

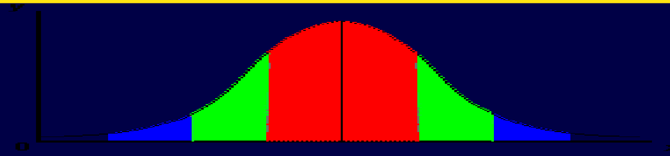
- **Encode all relevant information on signal vs background (aside from $m_{\gamma\gamma}$ itself) into a single MVA diphoton discriminant,** with input variables largely independent of $m_{\gamma\gamma}$
  - **Photon ID MVA for each photon:** based on isolation, shower shape, energy density per event
  - **Kinematics and Topology:** p_T and η of each photon, and cos $\Delta\phi$ between the two photons
  - **Per-event mass resolution** and **correct-vertex probability**
- **Trained on MC signal and background**
- **Validation of the inputs (photon ID, energy resolution):** uses $Z\to ee$, $\mu\mu\gamma$
- **Validation of the output with $Z\to ee$**





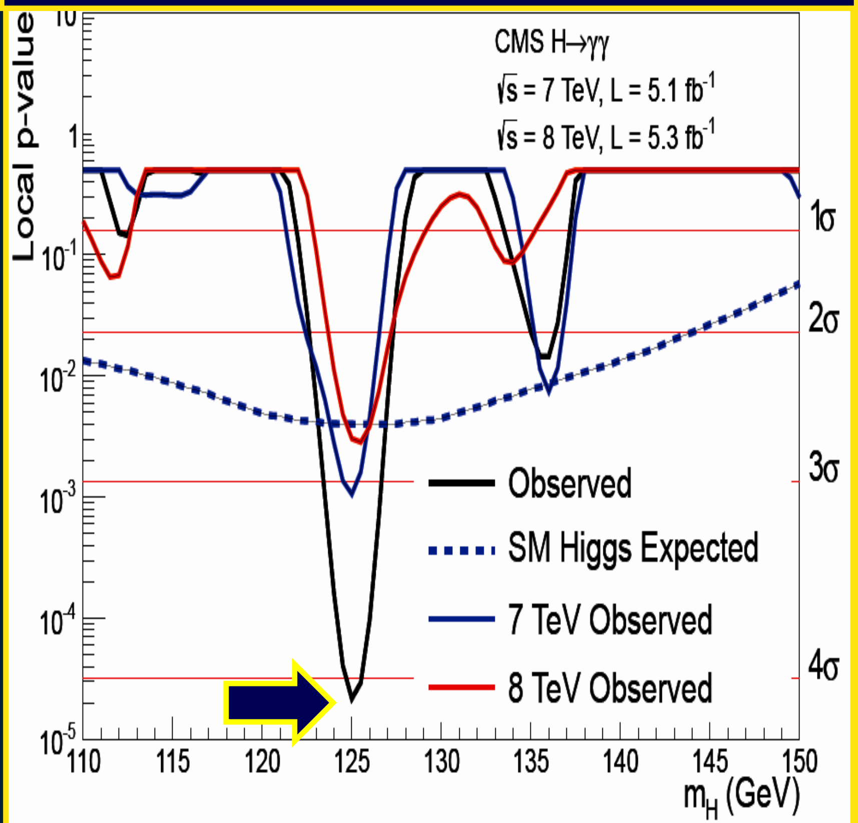
H $\rightarrow\gamma\gamma$: Extracing the Signal
Leading to the Discovery

- Measure the **probability** that an excess in the data can be explained by an **upward background fluctuation,** ***without a SM Higgs boson***



By Convention, you need "5 Sigma" for a Discovery

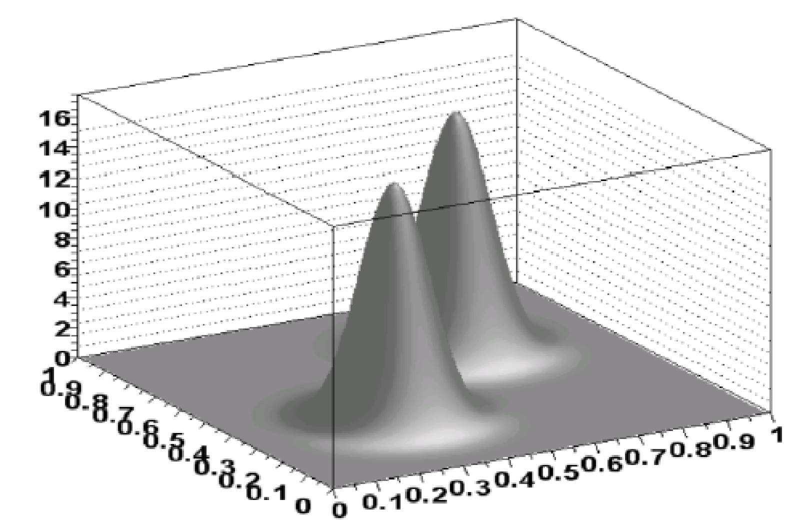
- *Had 4 Sigma significance in this channel alone at the discovery (June 2012)*
- *In final 2012 analysis (with 4X Data), we had 6 Sigma in this channel alone*

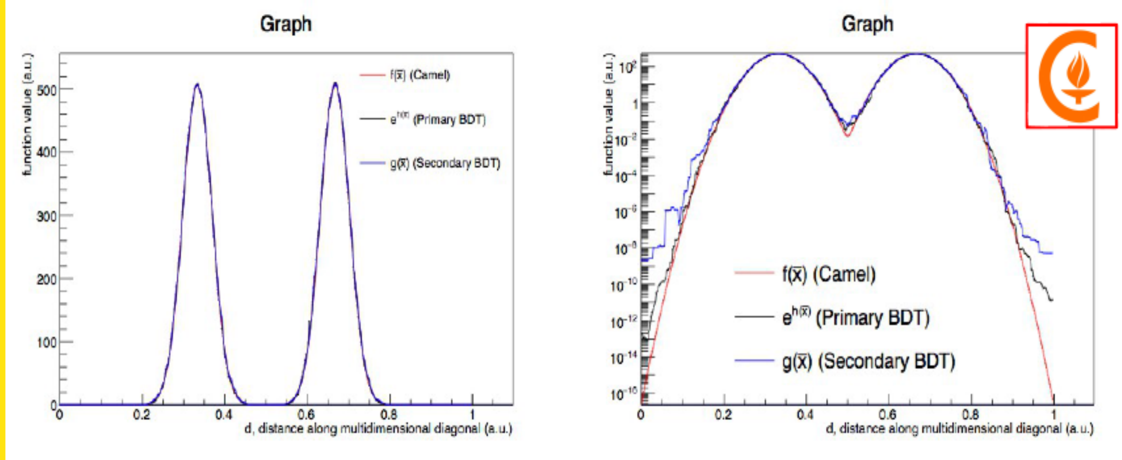


- ***Use of MVAs (BDTs) Boosts Sensitivity by a Factor of 1.8 relative to traditional "cut and count" method***
- ***Without MVA we would have had run 80% Longer for the discovery, and***
- ***Many years longer to reach the same final sensitivity to rare and/or new processes***

Function Sampling and Event Generators

- **Using Boosted Decision Trees** to build a multidimensional function approximation that can be directly sampled
- **For high efficiency** numerical Integration or phase space sampling
- **Application to** Complex (Higher Order) Event Generation





(a) linear (b) log

- **Example of 4D** Camel Function Integration
- 10X Acceleration with an order of magnitude improved accuracy **over the best previous methods in HEP**

J. Bendavid, Caltech

The LHC Program Areas
Great Potential for Machine Learning

- Science at the LHC
  - Triggering on rare signals
  - Data processing and simulation
  - Data movement and + computation
  - Search strategy

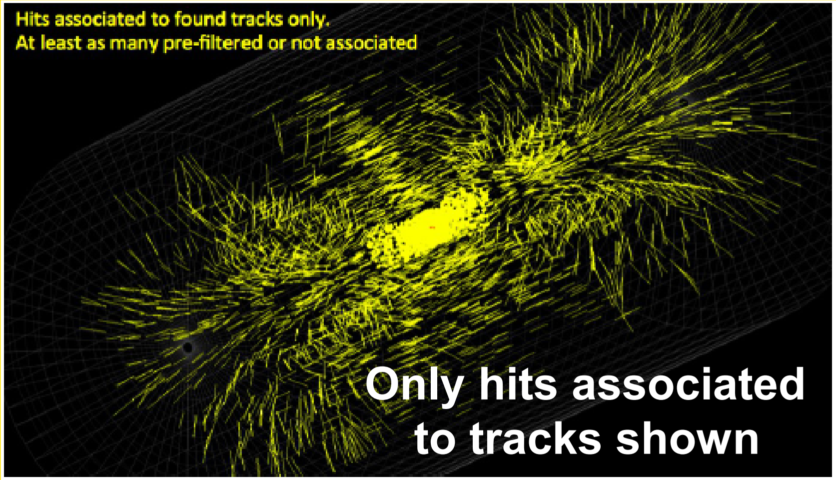
- High Luminosity LHC
  - Ever Increasing Event Complexity
  - Global Computing and Network Challenges of with Exascale Data

- Data Science at the LHC: *Deep Learning Approaches to Solutions*
  - Advanced Tracking Algorithms
  - Object Identification
  - Faster Simulation
  - Low Energy Computation

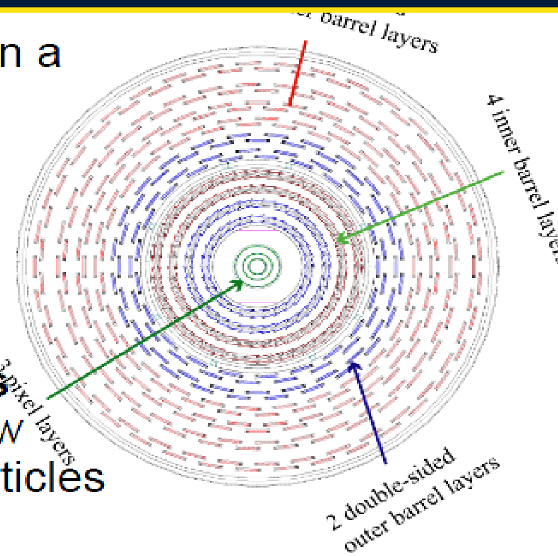
- Beyond Computing Needs Alone
  - Optimized Computing; Global Workflow
  - Model Independent Searches
  - Faster Time to Discovery
  - New Opportunities for Science Discoveries Not Otherwise Realized

Reconstruction: From raw measurements in subdetectors to kinematics and properties of the particles created in Collisions

Example 1: Charged Particle Reconstruction

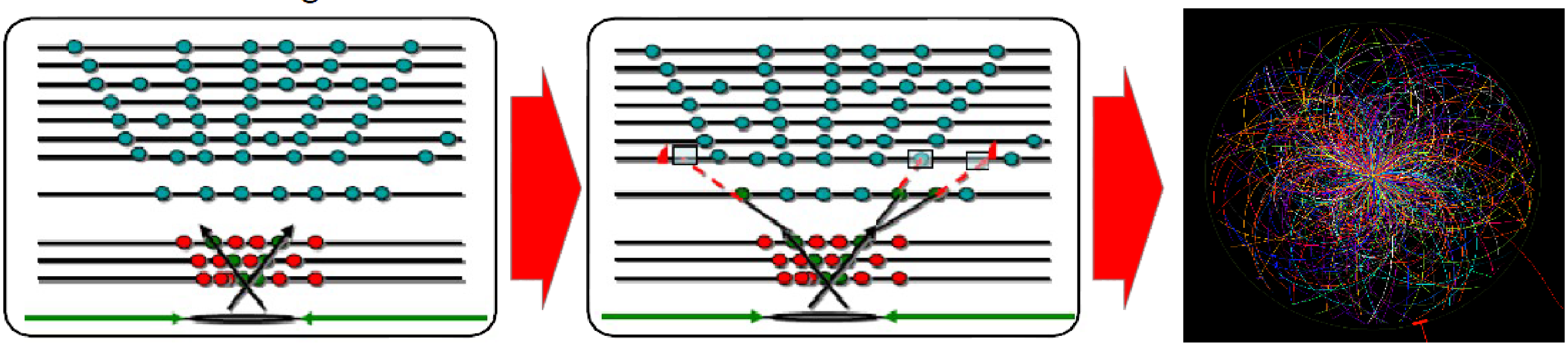


- Particle trajectory bended in a solenoid magnetic field
- Curvature is a proxy to momentum
- Particle ionize silicon pixel and strip throughout several concentric layers
- **Thousands of sparse hits**
- Lots of hit pollution from low momentum, secondary particles



Seeding

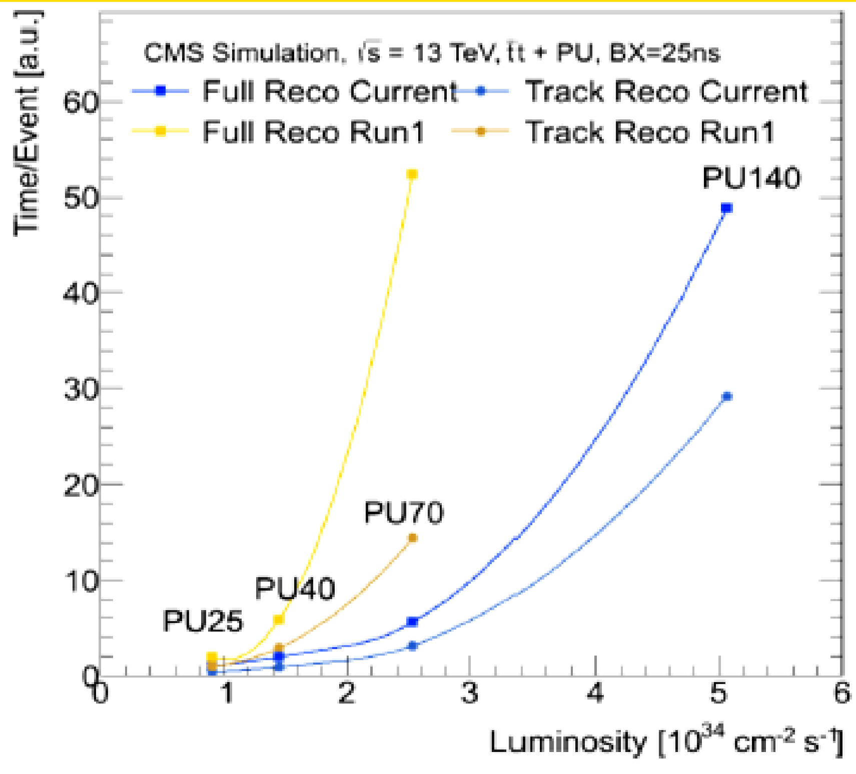
Kalman Filter



- **Explosion in hit combinatorics** in both seeding and stepping pattern recognition
- **Highly time consuming task** in extracting physics content from LHC data

Cost of Charged Particle Tracking

- ***65-200X Greater overall CPU need in the HL LHC Era (Est.)***
- **In spite of more optimization, Moore's Law, *needs will surpass the computing budget by 4-12X***

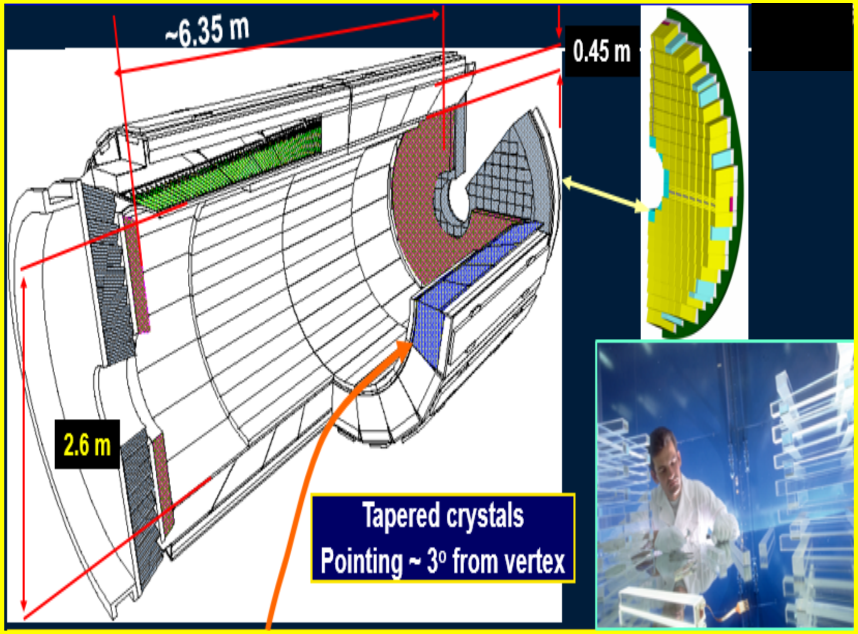


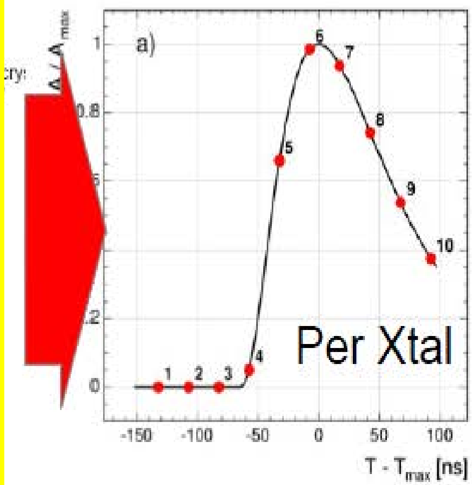
- **Charged particle tracking is one of the most CPU consuming tasks**
- **Code Optimizations (to fit in computing budgets) are saturated**
- **Large fraction of available CPU is required in the HLT**
  - **So can only perform tracking for pre-filtered subsamples**

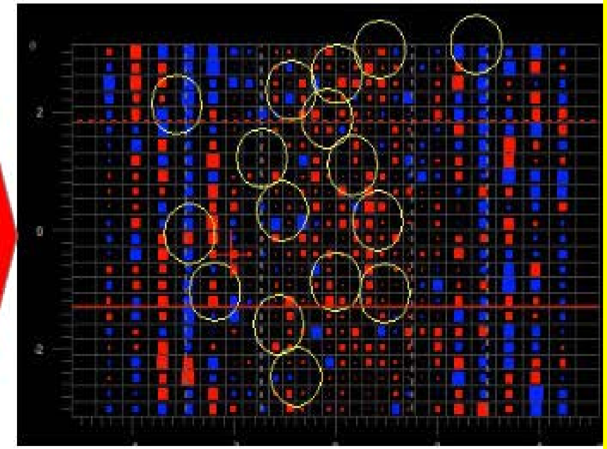
✸ ***Need for much faster algorithms***

✸ ***Apply machine learning to the challenge***

Calorimeter Reconstruction





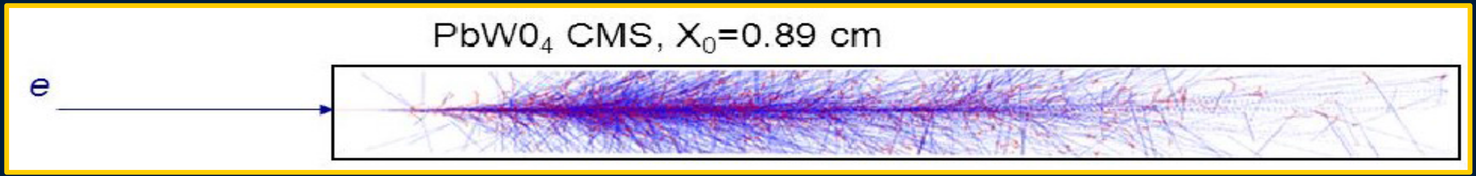


- Energy deposit per Crystal computed from time-samples
- Crystal energies collected in clusters then super-clusters
- Photon, electron, jets identified; then energy calibrated

- Multi-step reconstruction process
- More and more challenging with higher granularity, pileup
- Pattern recognition, identification and regression: become all-in-one with machine learning

ML and HEP: Simulation of Collisions

- **Most analyses have data driven background estimations**
- **Cross checks and analysis tailoring nevertheless require *large* sets (to billions) of simulated events for the main backgrounds**
- **Simulating events is a costly process**
  - **Particle vectors are generated randomly according to physics processes computed from theoretical matrix elements and amplitude calculations**
  - **Particles propagate, bend, slow down, interact, and deposit energy in a complete representation of the CMS experimental apparatus**
  - **Energy in sensitive elements is digitized, emulating the real CMS readout**
- **Billions of CPU hours are spent in Monte Carlo Simulation**



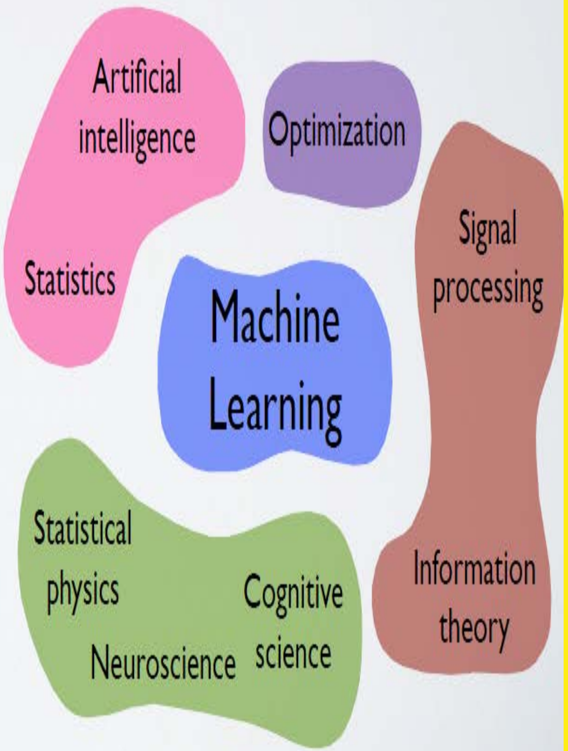
- **Complex showering process in One Crystal of 76000 in CMS**

✸ **An opportunity for generative models with machine learning**

✸ **Generate the energies and topology of the resulting pattern seen in the crytals (and its fluctuations) directly from raw data**

Machine Learning: Learn to Discover in HEP

- "The science of getting computers to act without being explicitly programmed" - Andrew Ng (Stanford/Coursera)
- part of standard computer science curriculum since the 90s
- inferring knowledge from data
- generalizing to unseen data
- usually no parametric model assumptions
- emphasizing the computational challenges



▪ Taxonomy

- Supervised learning: non-parametric (model-free) input - output functions
  - classification (Trees, BDT, SVM, NN) - what you call MVA
  - regression (Trees, NN, Gaussian Processes)
- Unsupervised learning: non-parametric data representation
  - clustering (k-means, spectral clustering, Dirichlet processes)
  - dimensionality reduction (PCA, ISOMAP, LLE, auto-associative NN)
  - density estimation (kernel density, Gaussian mixtures, the Boltzmann machine)
- Reinforcement learning:
  - learning + dynamic control: learn to behave in an environment to maximize cumulative reward

Balazs Kegl at CERN 2014: https://indico.cern.ch/event/316800/

Machine Learning: Scene Labeling

Approach to Calorimeter and Track Reconstruction

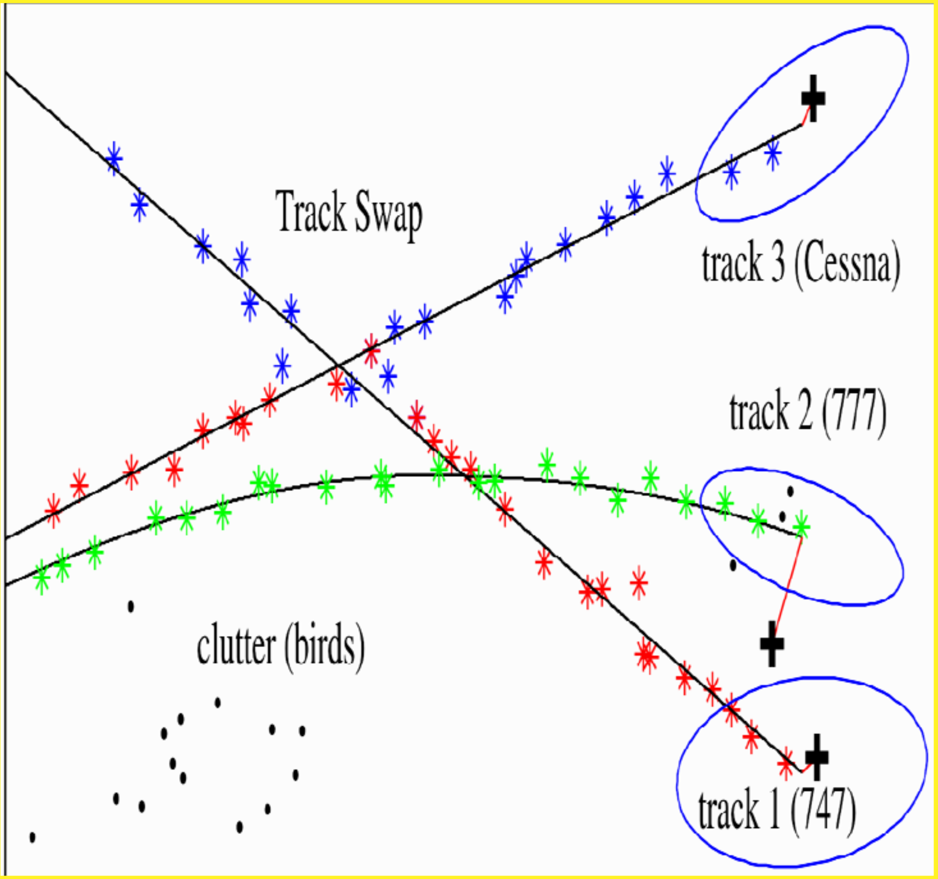
Farabet et al. ICML 2012, PAMI 2013

- Group and classify what each pixel belongs to:
- Real-time video processing with deep learning

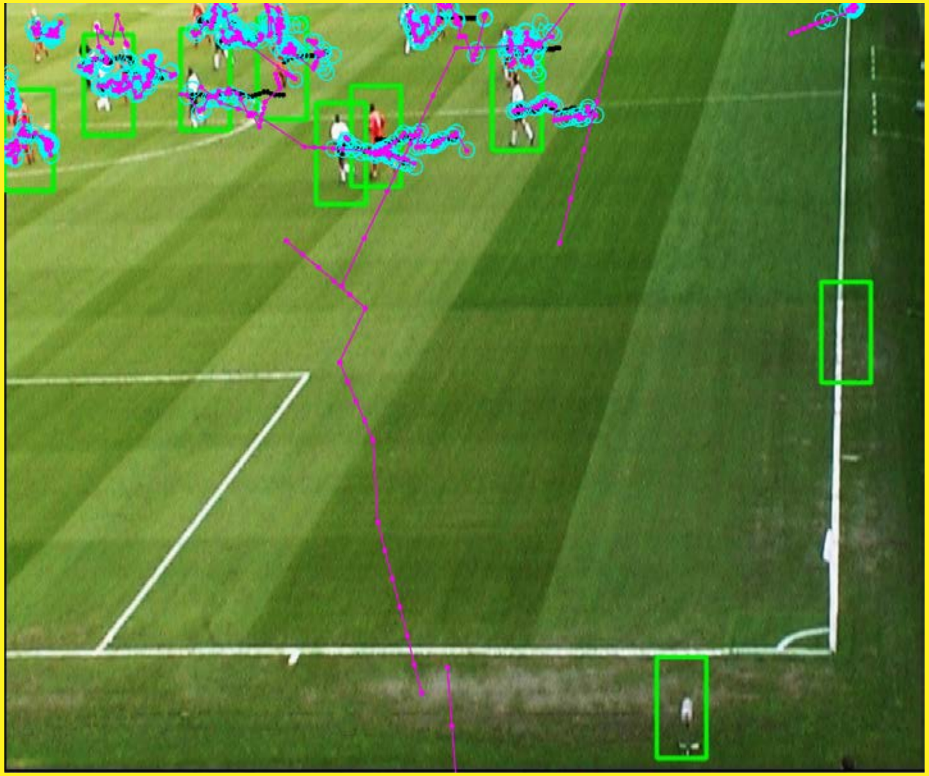
- Associate each Crystal energy to a cluster with DL
- Associate each tracker hit to a charged particle with DL

Application: "Something Like Tracking"

One example among many from NIPS 2014 :
http://papers.nips.cc/paper/5572-a-complete-variational-tracker.pdf

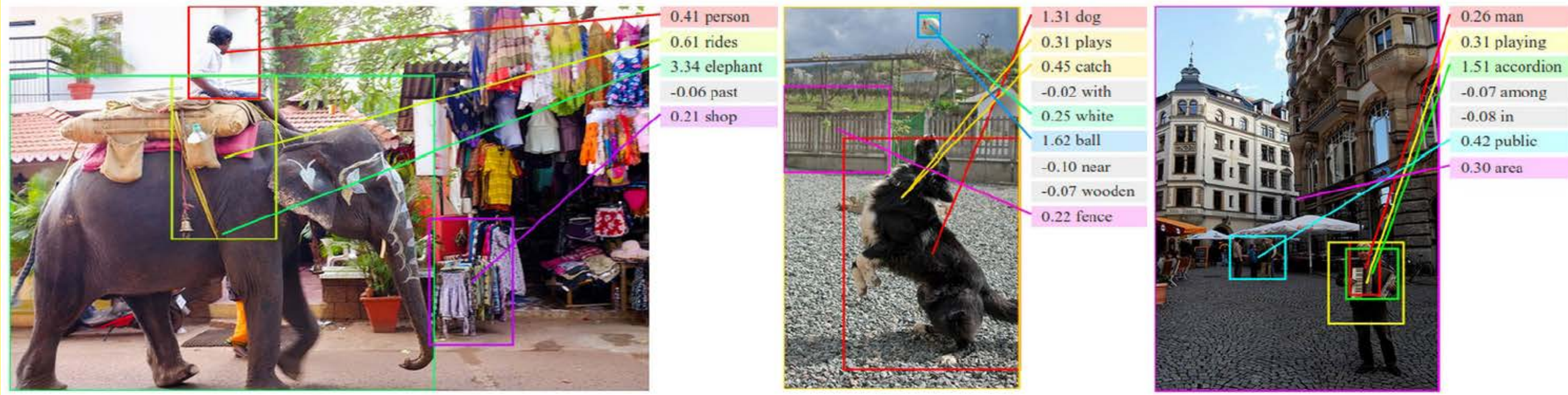


Also



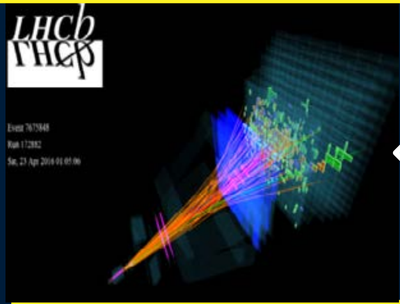
- Note that these are real-time applications, with CPU constraints
- Worry about efficiency, "track swap"

Object ID to Sentence Generation from “Raw” Images

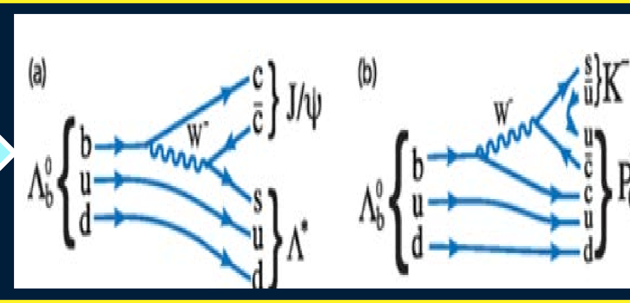


Region-Level ID and Annotations with a Multimodal RNN
Karpathy and Fei-fei Li, CVPR 2015

Create Description of a (Scene) Image ⟷	**Generate a Decay Process Description from a Collision representation**



Raw hits Not even Tracks !



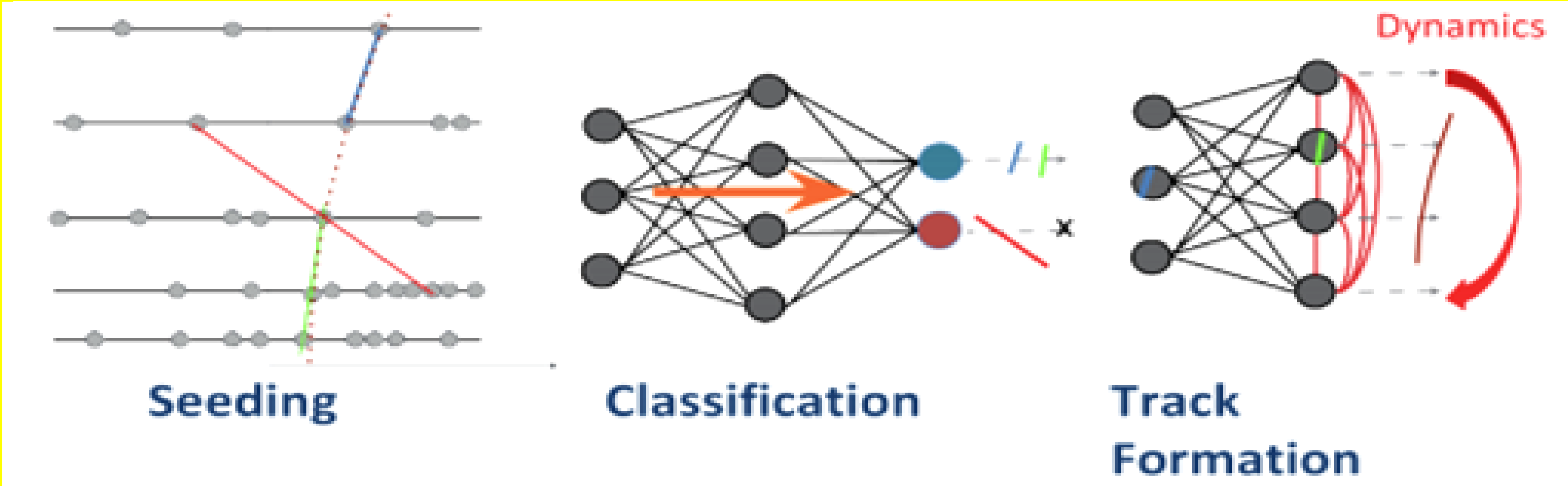
Penta quark !

Deep Visual Image Alignments for Generating Image Descriptions
http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf

Deep Learning Applications
Pilot Projects in CMS and Other HEP Experiments

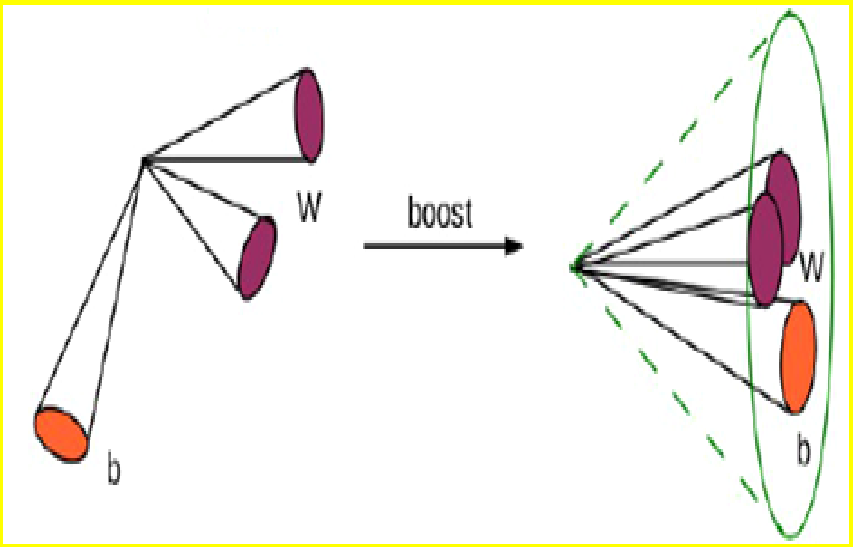
J-R Vlimant,
M. Spiropulu et al.

Advanced Tracking Algorithms



- Pilot project: **recurrent neural nets for Kalman filtering**
- Further investigation may involve
  - **Application of scene labeling to seed formation**
  - **Application of object detection to track assembling**
- Medium/High risk, very high reward problem
  - **Exploratory phase on the model definition**

Jet Tagging

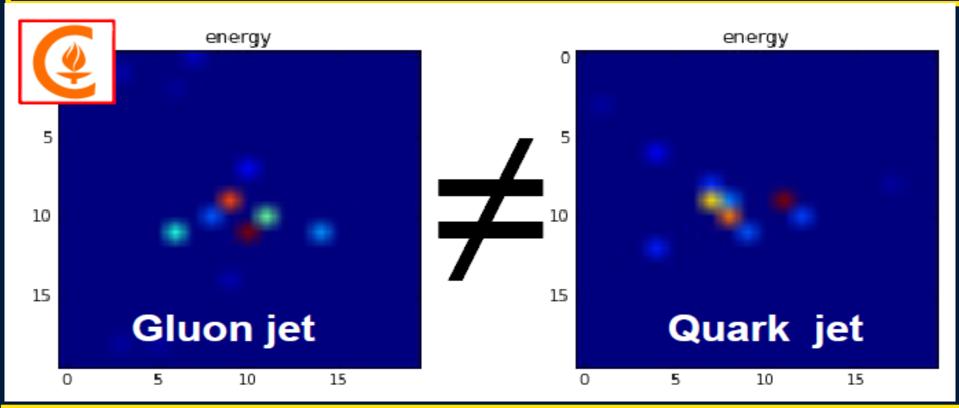


- Hadronic activity results in a bundle of collimated particles
- The more energetic, the more collimated : W-jet
- With even higher energy, even mother particles are collimated: top-jet, Higgs-jet

- Available discriminators are performing well
- Not yet taking advantage of the full substructure of the jets

- Image processing methods are natural candidates to perform the classification

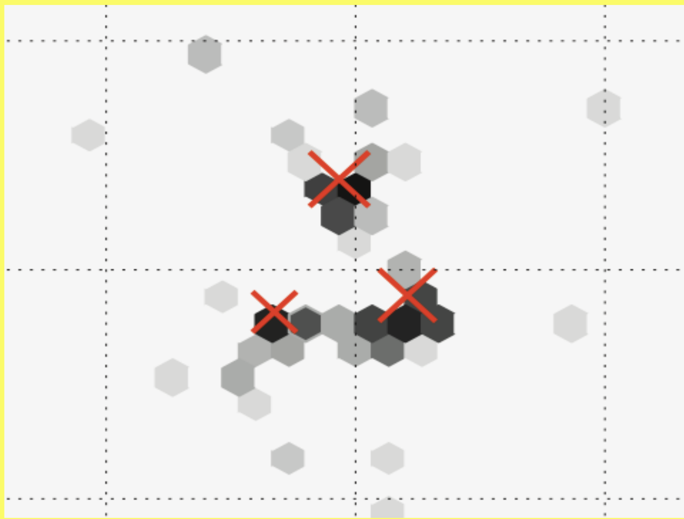
Energy Topology



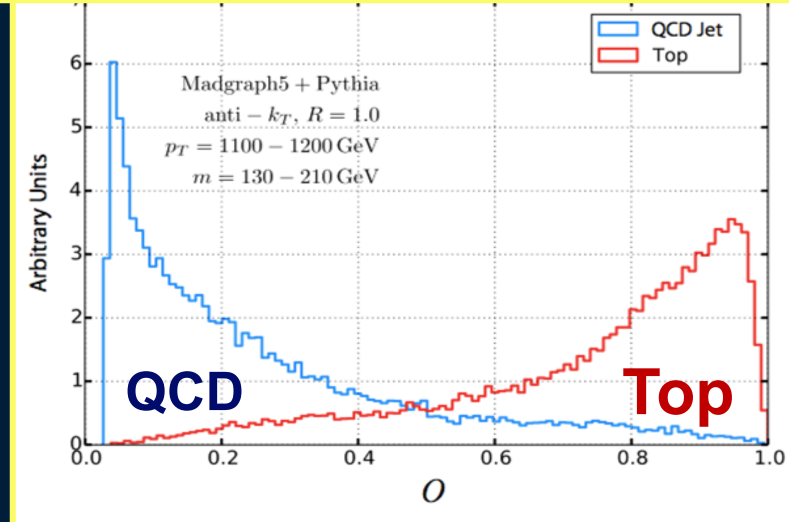
Small dataset, 11 categories, 60% accuracy on gluon jet versus any quark jet. Pre-preliminary

Tagging Boosted Objects: W and Top

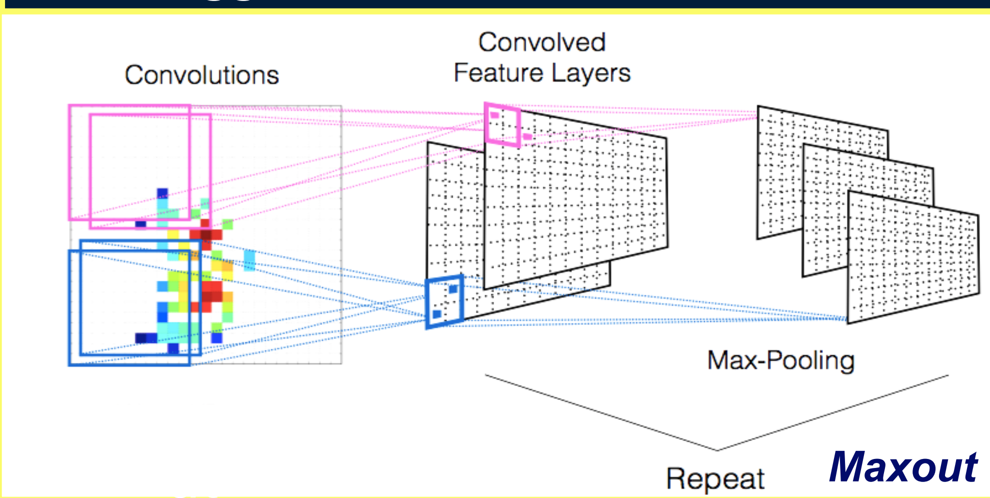
Top Tagger arXiv: 1501.05968 Almeida, Backovic, Cliche, Lee, Perelstein



Neural net

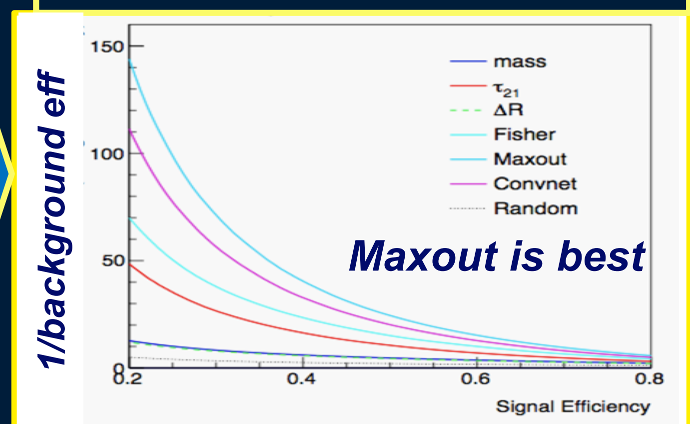


W tagger arXiv: 1511.05190, Oliveira, Kagan, Mackey, Nachman, Schwartzman



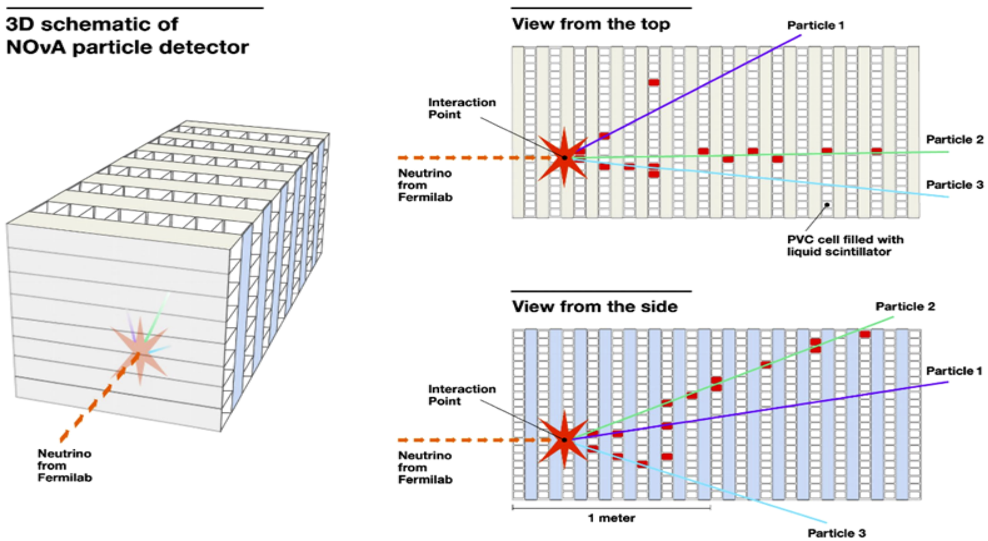
Train

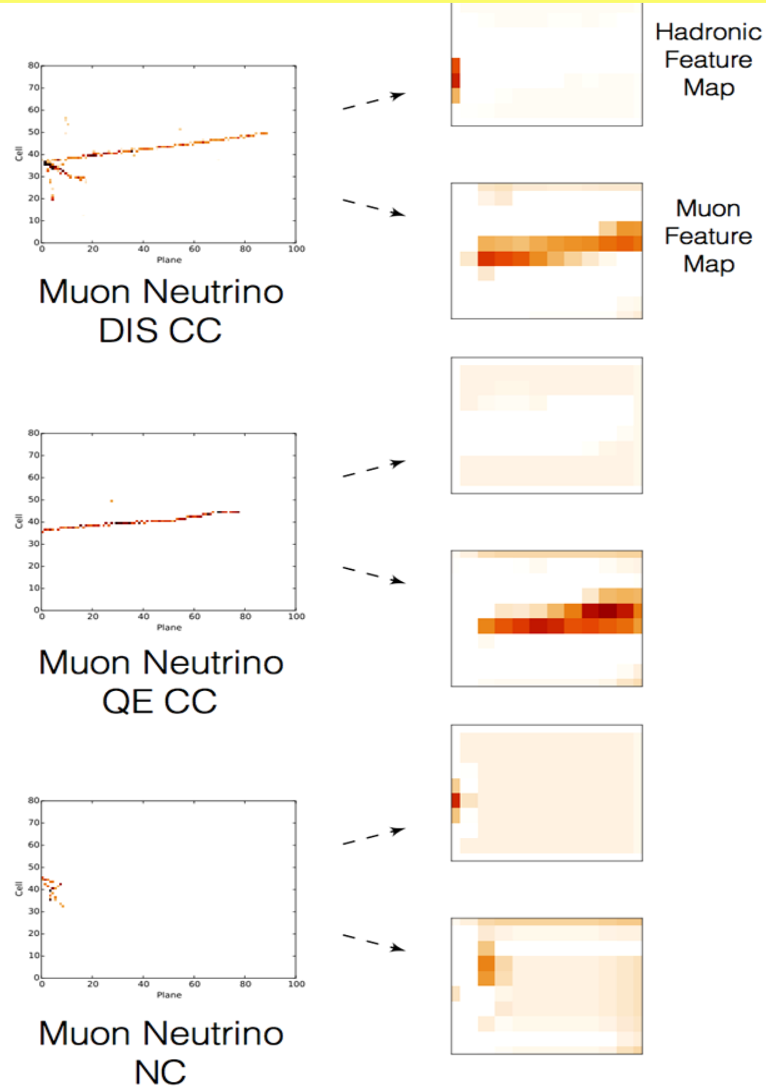
W Vs QCD Jet Discrimination



NOνA: Long Baseline (Fermilab – Minn.) Muon to Electron Neutrino Oscillations

CNN to Convolutional Visual Net Neutrino Event Classifier





	CVN Selection Value	ν_e sig	Tot bkg	NC	ν_μ CC	Beam ν_e	Signal Efficiency	Purity
Contained Events	–	88.4	509.0	344.8	132.1	32.1	–	14.8%
$s/\sqrt{b}$ opt	0.94	43.4	6.7	2.1	0.4	4.3	49.1%	86.6%
$s/\sqrt{s+b}$ opt	0.72	58.8	18.6	10.3	2.1	6.1	66.4%	76.0%

	CVN Selection Value	ν_μ sig	Tot bkg	NC	Appeared ν_e	Beam ν_e	Signal Efficiency	Purity
Contained Events	–	355.5	1269.8	1099.7	135.7	34.4	–	21.9%
$s/\sqrt{b}$ opt	0.99	61.8	0.1	0.1	0.0	0.0	17.4%	99.9%
$s/\sqrt{s+b}$ opt	0.45	206.8	7.6	6.8	0.7	0.1	58.2%	96.4%

40% Better Electron Efficiency for same background; *Faster*

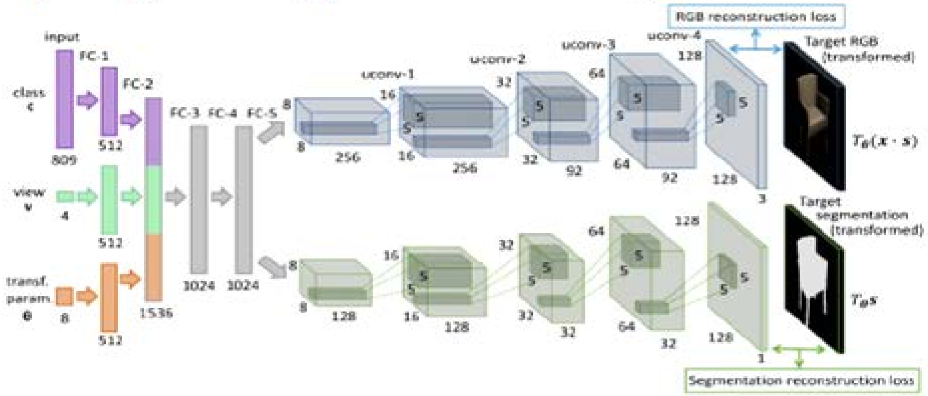
•

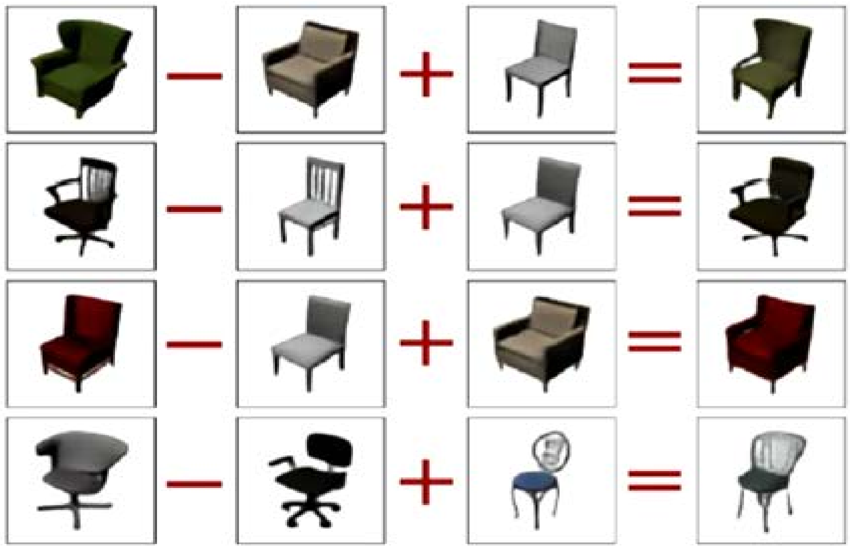
http://arxiv.org/pdf/1604.01444.pdf

Generative Models: Rendered 3D Models to 3D Object Recognition, Correspondence, *Invention*

Learning to Generate Chairs, Tables and Cars with CNNs

Arxiv:1411:5928, Dosovitskiy, Springenberg, Tatarchenko, Brox

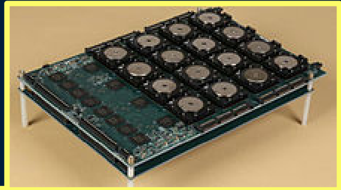


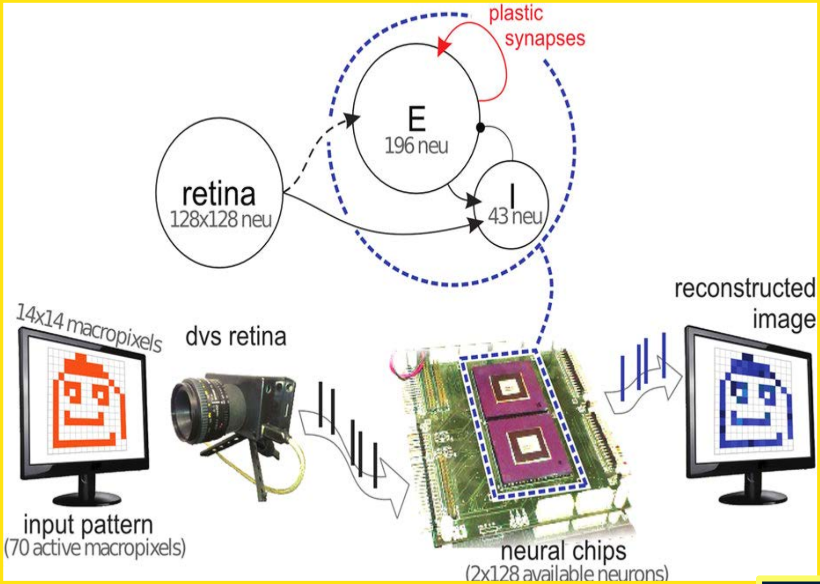


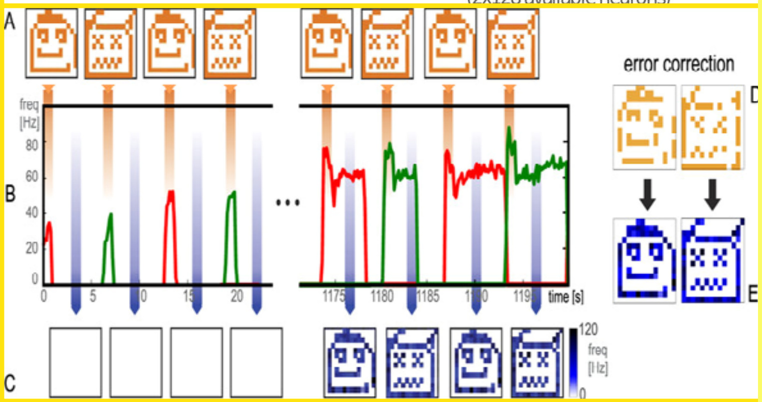
- Generate, 3D objects from rendered images
- Derive a 3D object's key attributes and component parts
- Evaluate similarity, perform Image Arithmetic
- Generate New Views and/or New Objects from known ones

- HEP Application: Replace complex multi-step simulation with Generative Models
- Address Computing Bottleneck
- Enable science program:
- ✷ *Increased speed, agility and/or scope of investigations*

Neuromorphic Computing
Chip and Systems Hardware







Real time unsupervised learning of visual stimuli in neuromorphic VLSI systems
http://www.nature.com/articles/srep14730

J-R Vlimant, Caltech

- Brain emulating low power systems of silicon neural chips
- Spiking neurons and spike-driven synapses for general computation
- Demonstrated to perform well on pattern recognition problems
- Unsupervised learning capabilities shown on some chip types

- ❒ Ongoing collaboration with iniLab & INI Zurich
- ❒ Aimed at calorimeter pattern recognition using NM chips in CMS Trigger level 1
- ❒ Potential application: NM accelerator cards for tracking, patt. rec. etc
- ❒ Involvement with IBM TrueNorth team
  - ❒ Application: pattern recognition in HEP
  - ❒ Possible synergy with LBNL

Machine Learning So Far

- Faster algorithms
- Highly relevant for triggers
- Reduce software maintenance
- Faster event simulation
- Facilitate detector design
- Mitigate operational cost
- ✸ Offer New Science Opportunities

Plus: Data Analytics

Beyond the Need for Computation [a la Watson]

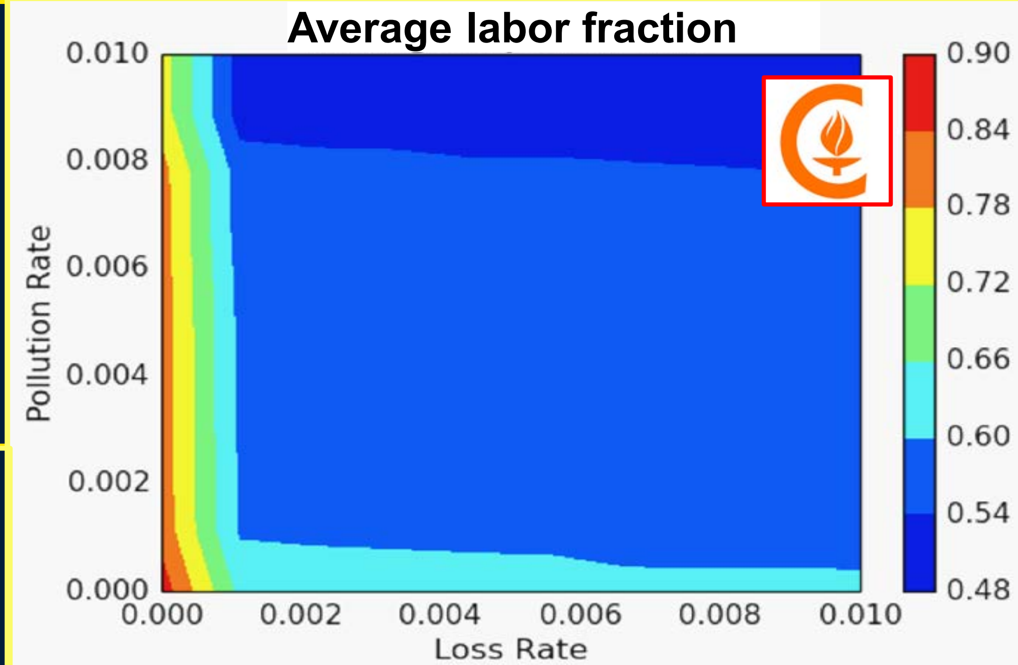
J-R Vlimant,
M. Spiropulu et al.

Data Certification Robot

- ❐ Not all of the data taken at the experiments is good for analysis: (detector channel or readout effects, software defect, calibration)
- ❐ Histograms made per luminosity block (23s of beam time) are scrutinized by experts to decide on good/bad data
- ❐ Huge number of histograms, several layers of scrutiny: *Labor Intensive*

- ❐ The machine learning approach identifies relevant features
- ❐ Calculates good data percentage per lumiblock
- ❐ Trains rolling classifiers

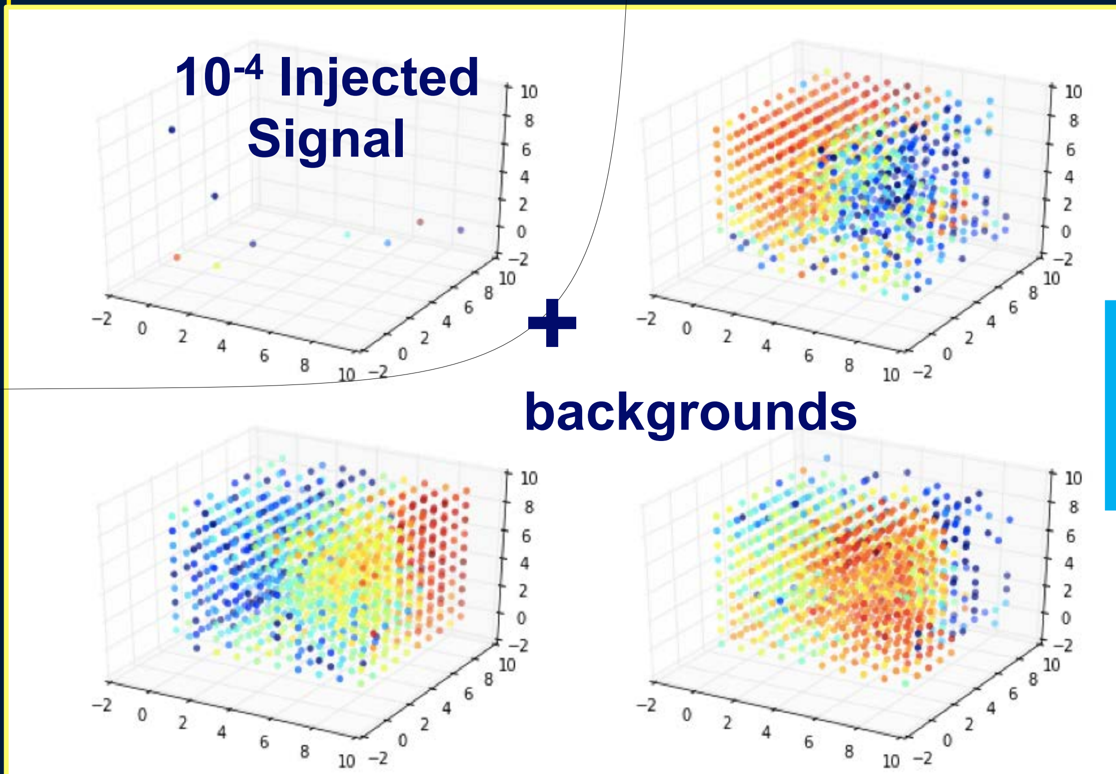
✸ By accepting 1% data loss we could save 40% of the certification team's workload

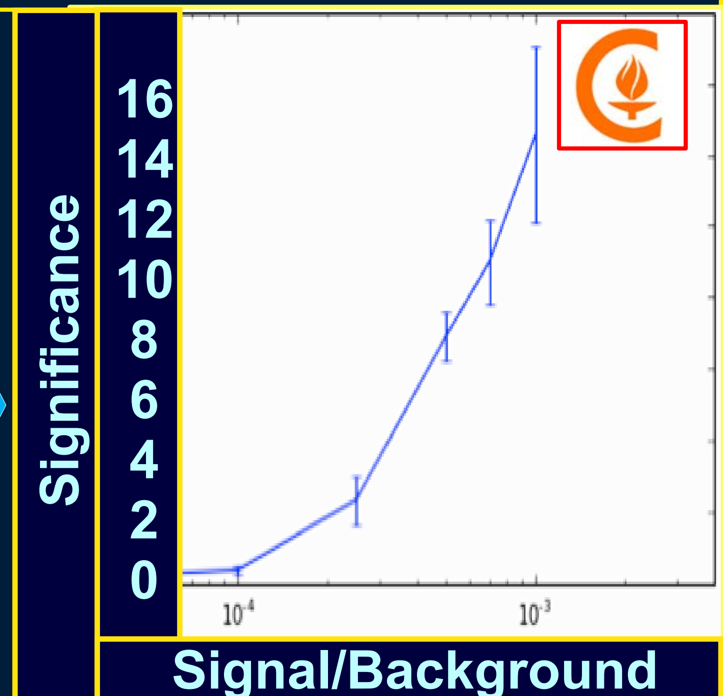


J-R Vlimant, Caltech with Yandex

Self Organizing Analysis

- Train a 4D self organizing map (SOM) on synthetic data composed of one signal and 3 backgrounds
- Injection performed at varying signal/background ratio
- Interpretation using only backgrounds allows one to single out the events from signal: *deviation*
- Significance of deviation estimated as function of signal injection

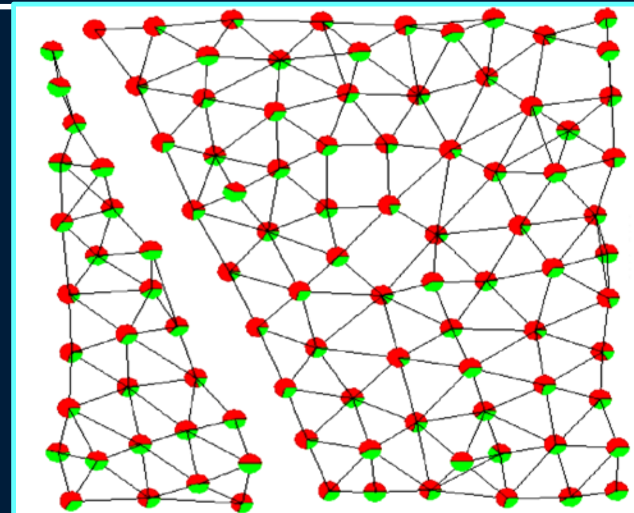


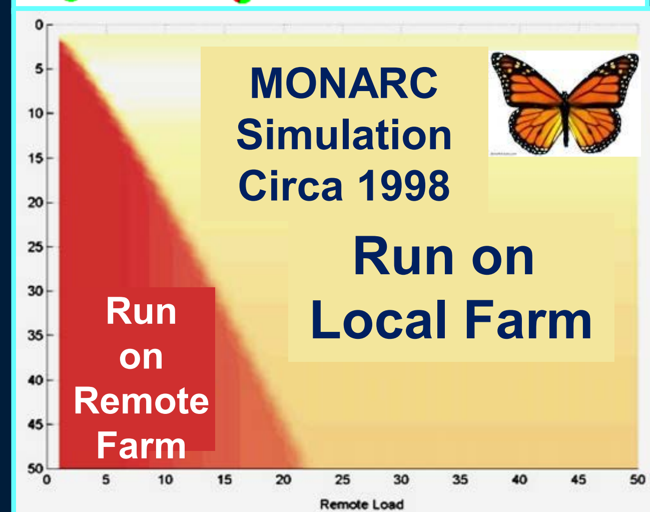


J-R Vlimant, Caltech

Key Developments from the HEP Side
Enabling the Vision: Machine Learning

- **Applying** Deep Learning + Self-Organizing systems methods **to optimize LHC workflow**
  - Unsupervised: **extract key variables/functions**
  - Supervised: **to derive optima**
  - Iterative and model based: **to find** effective metrics **and stable solutions** [*]
- **Complemented by** game theory methods, modeling and simulation
- Shown to be effective to solve traffic, communications and workflow problems
- **Starting with** logged monitoring information
- **Progressing to** real-time agent-based pervasive monitoring





Self-organizing neural network for job scheduling in distributed systems

[*] T. Roughgarden (2005). *Selfish routing and the price of anarchy*

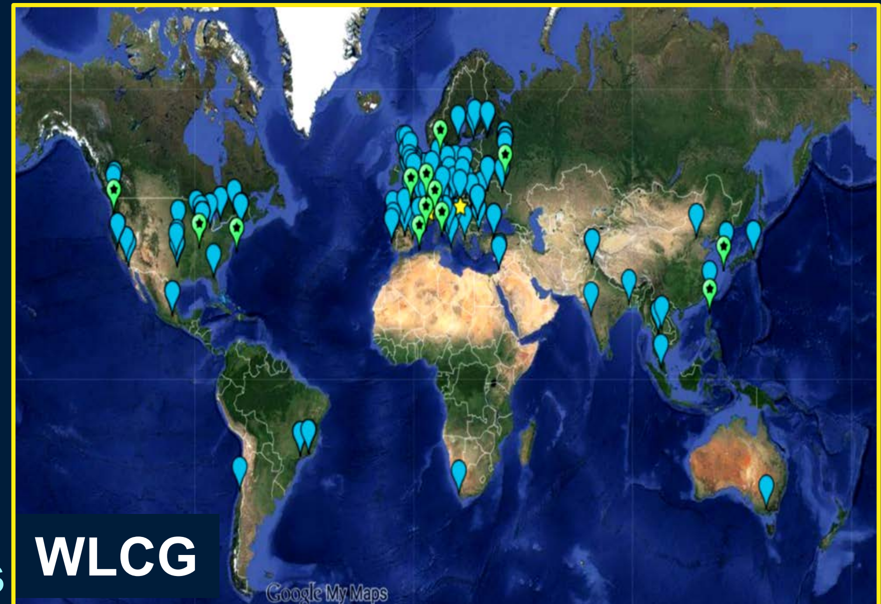
Towards a Next Generation
Network-Integrated System

Facing the Challenges
of Exascale Global Data
with Deep Learning

Vision: Next Gen Integrated Systems for Exascale Science: Synergy ➡ a Major Opportunity

Exploit the Synergy among:

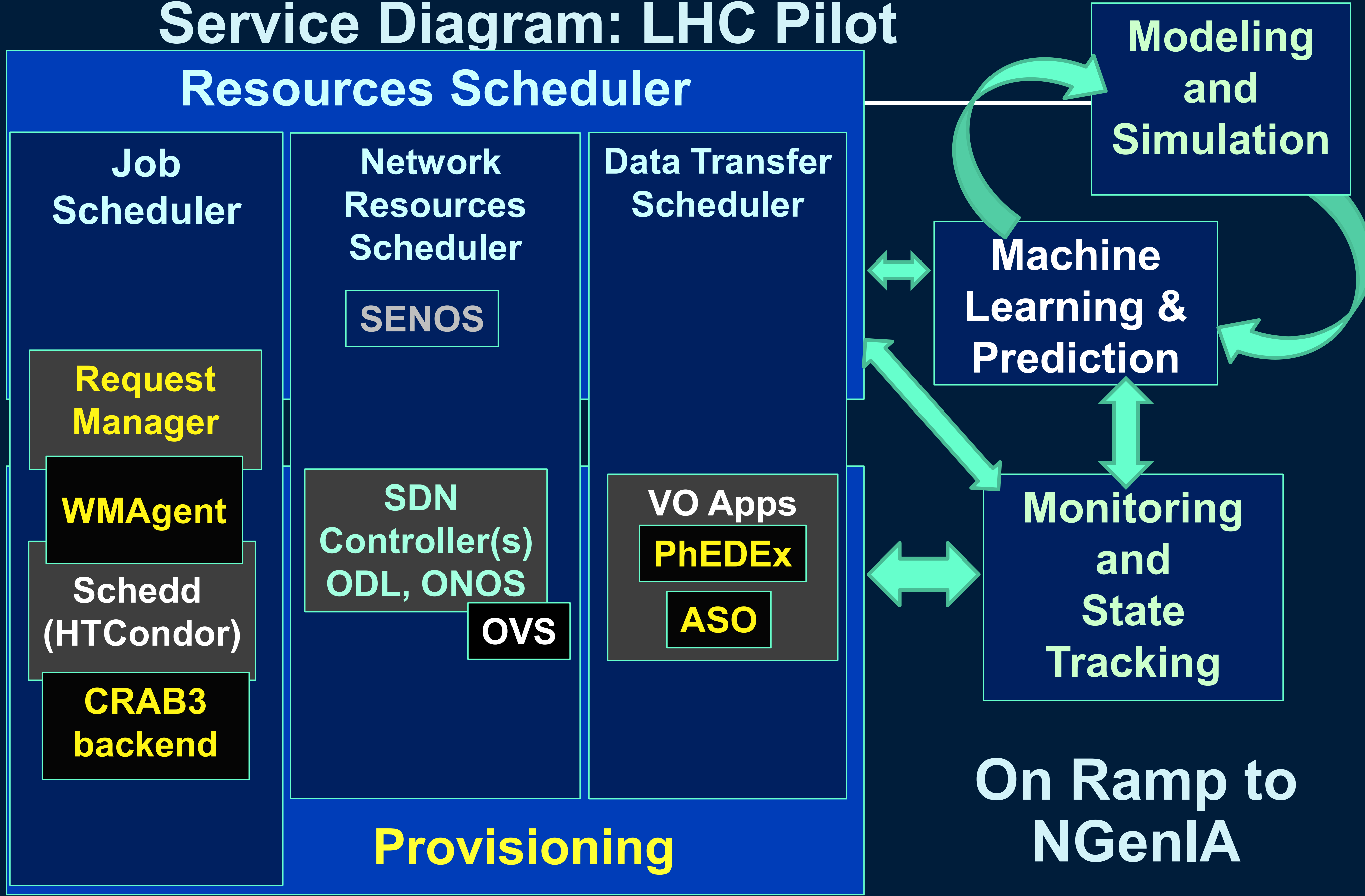
1. Global operations data and workflow management systems developed by HEP programs
   - Enabled by distributed operations and security infrastructures
   - Riding on high capacity (but mostly still-passive) networks
   - Being geared to more diverse resources



2. Deeply programmable, agile software-defined networks (SDN) Emerging as multi-domain network "operating systems"
   - With Proactive and reactive site-network interactions
3. ✸ *Machine Learning, modeling and simulation, and game theory methods; Extract key variables; Optimize; move to real time self-optimizing workflows*

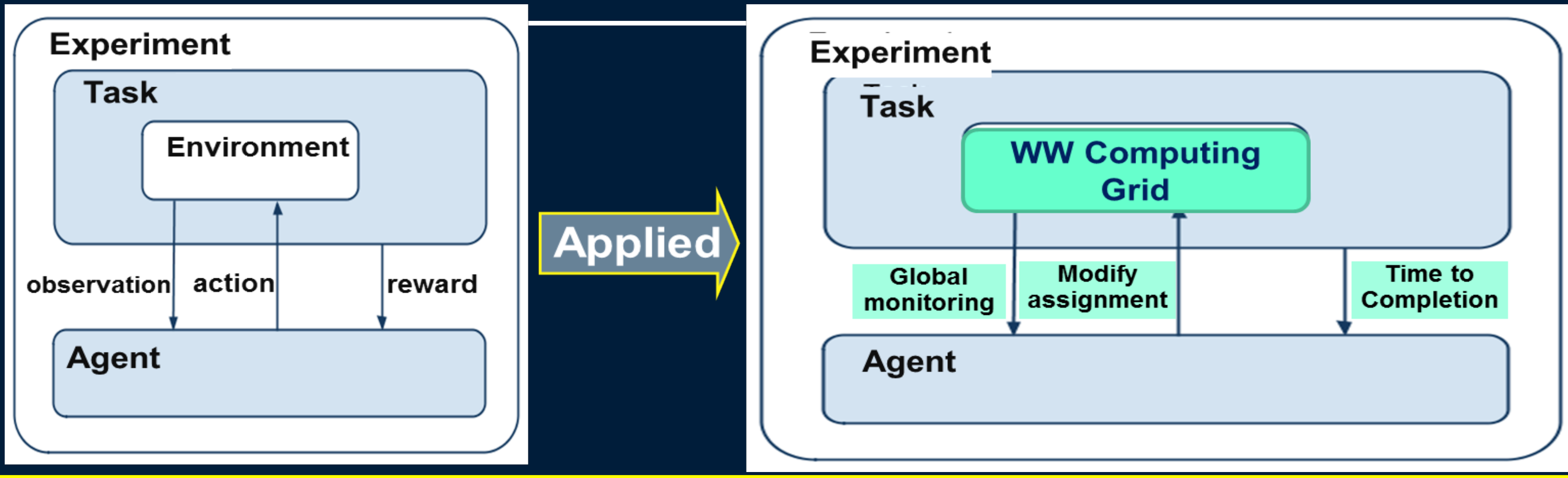
✸ Watershed: A new ecosystem with ECFs as focal points in the global workflow

Service Diagram: LHC Pilot



Computing Optimization R&D

Machine Learning Coupled to Modeling and Simulation



- ❐ Learn complex models using deep learning with monitoring data and the chosen metric(s)
- ❐ Use simulations together with game theory techniques or a reinforcement learning method to find optima
  - ❐ Balancing among max throughput, balanced resource use, predicability of time to completion (predictable workflow) etc.
  - ❐ Variations: evolve towards the metrics yielding stable solutions with good throughput
- ❐ Steering computing, storage and network elements like robot arms

Game Theory and the Future of Networking

http://blog.eai.eu/game-theory-and-the-future-of-networking/

- Game theory: Mathematical models of conflict and cooperation among intelligent rational decision-makers
  - Studies participants' behavior in strategic situations.
- Motive and the need for Increased Reach induce selfish entities to cooperate in pursuit of a common goal
- Application Pull: the Internet calls for analysis and design of systems that span multiple entities with diverging information and interests
- Technology Push: math and science mindset of game theory is similar to that of many (computer) scientists
- Diverse Fields of Use: economics, political science, psychology, logic, computer science, biology, poker and now HEP

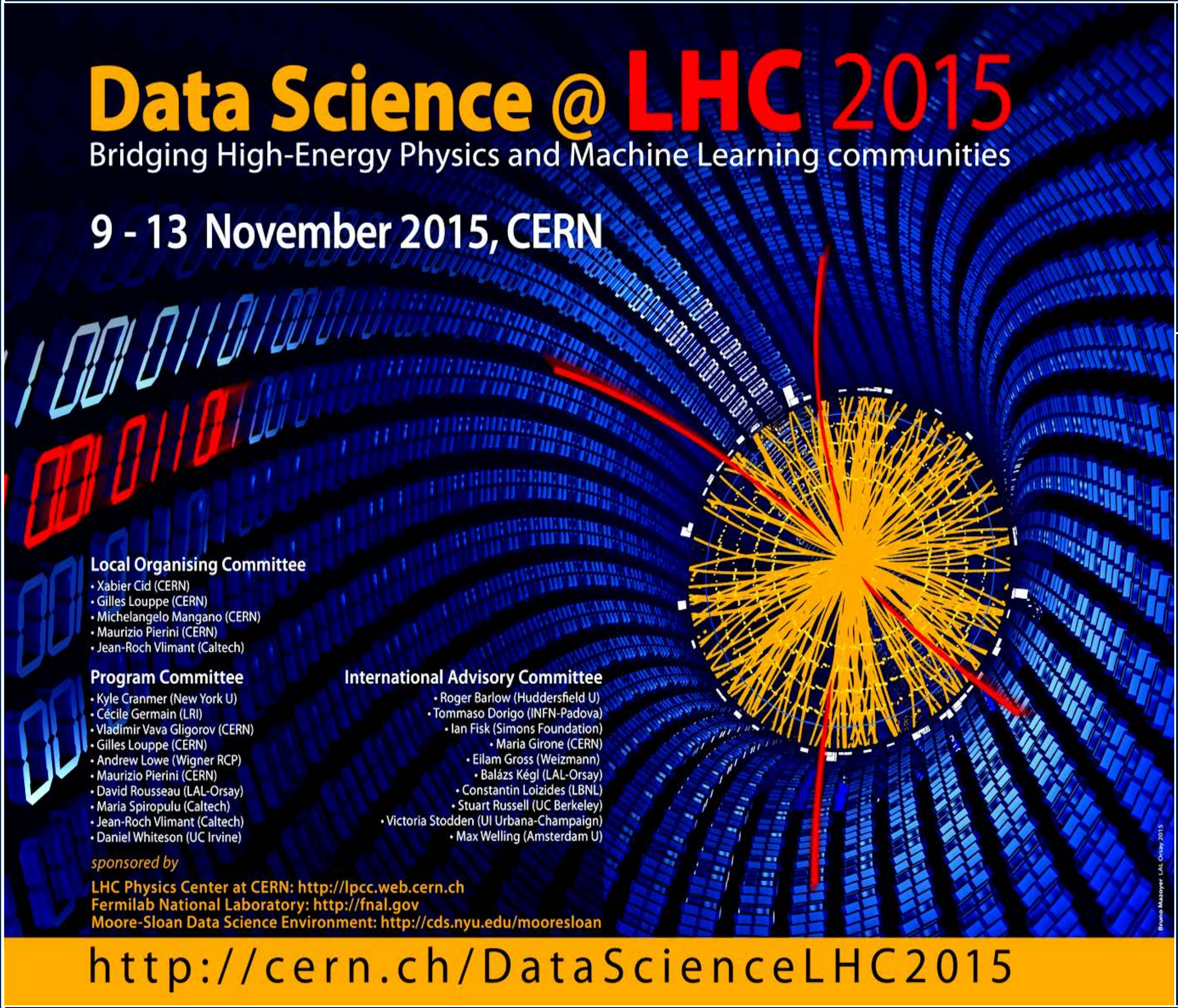
- Emergence of the internet has motivated development of GT algorithms for finding equilibrium in games, markets, auctions, peer-to-peer systems, security and information markets

- GT is now applied to a wide range of behaviors
- It has become an umbrella term for the science of logical decision making
- In and among humans and computers

- *Coherent Interactions* among the experiments' workflow management systems, the end sites, the network and the user groups *as a System*


LHC Run2 and Beyond
We have launched on a River of Discovery
Amazon Sunrise

Organized by M. Spiropulu, J-R. Vlimant, et al.



Hands-on workshops on contemporary machine learning techniques

To foster HEP ↔ ML Community Collaboration

Follow On: DS@HEP 2016 Workshop

July 5-7 at Simons Foundation, NYC

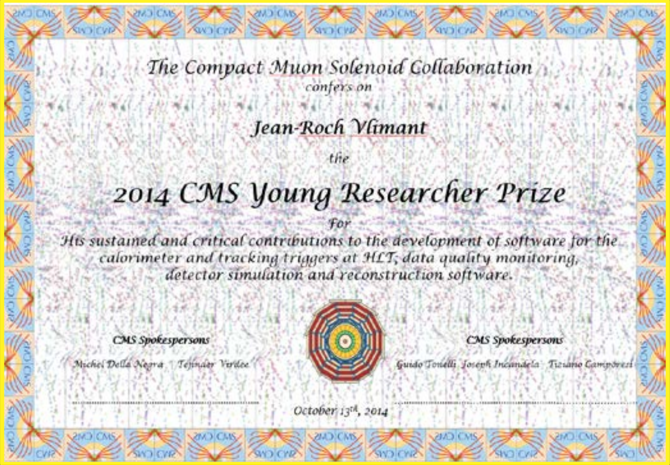
Focus on new ideas + solutions in tracking, calorimetry, anomaly detection, and

New paradigms in machine learning; close to the raw data

+ Contacts with NVIDIA, Orange Labs Silicon Valley, INI Zurich, IBM True North, Yandex, Minds.ai, etc.

With Thanks to J-R Vlimant, J. Bendavid and Prof. Maria Spiropulu

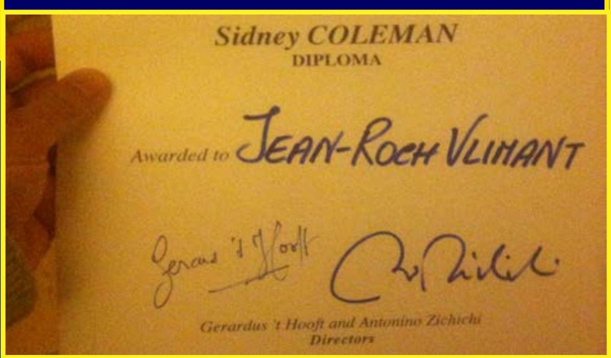
2014+2015 CMS Young Researcher Awards to J-R Vlimant, J. Bendavid



The Compact Muon Solenoid Collaboration
confers on
Jean-Roch Vlimant
the
2014 CMS Young Researcher Prize
For
His sustained and critical contributions to the development of software for the calorimeter and tracking triggers at HLT; data quality monitoring, detector simulation and reconstruction software.
CMS Spokespersons
Michel Della Negra Tejinder Virdee
CMS Spokespersons
Guido Tonelli Joseph Incandela Tiziano Camporesi
October 13th, 2014

For "His sustained and critical contributions to the development of software for the calorimeter and tracking triggers at HLT; data quality monitoring, detector simulation and reconstruction software."

J-R Vlimant: Sidney Coleman Diploma at the 54th Erice Subnuclear Physics School for his Talk on Machine Learning for HEP June 2016



Sidney COLEMAN
DIPLOMA
Awarded to JEAN-ROCH VLIMANT
Gerardus 't Hooft and Antonino Zichichi
Directors

The Compact Muon Solenoid Collaboration
confers on
Joshua Bendavid
the
2015 CMS Young Researcher Prize
For
His sustained and critical contributions to the development of photon and electron energy reconstruction, the discovery of the Higgs boson via its two photon decay mode, and the Tier-0 operation at LHC startup.
CMS Spokespersons
Michel Della Negra Tejinder Virdee
CMS Spokespersons
Guido Tonelli Joseph Incandela Tiziano Camporesi
December 7th, 2015

For "His sustained and critical contributions to the development of sphoton and electron energy Reconstruction, the discovery of the Higgs boson in its two photon decay mode, and the Tiero operation at LHC startup

THANK YOU!

Harvey Newman

newman@hep.caltech.edu

Machine Learning: Take Away Messages

- ❒ Machine Learning solves problems
- ✸ Which are very hard to model ab initio
- ✸ Working from the ground up
- ✸ *Then* extracting relationships and even deriving models
  - ✸ Without Domain Knowledge
- ➡ Coding vision: scene labeling, face and other object recognition; understanding properties
- ❒ Physicists can solve complex HEP problems:
  - ❒ Real-time event filtering
  - ❒ Object composition + identification
  - ❒ Needle in the hay-stack analysis
- ❒ **But within severe limits**

- ❒ But: We spend a great deal of effort, time and money
  - ❒ Operating our experiments
  - ❒ Handling worldwide data
  - ❒ Dealing with hardware and software complexity, faults and human error
- ❒ All of this narrows or blocks our path to science discovery
- ✸ We are looking to Machine Learning for New Paths with
  - ✸ Greater speed + simplicity
  - ✸ Lower cost and ultimately
  - ✸ ***Greater insight***

Machine Learning: Our Approach

- Deep Learning represents an ongoing *leap forward* in computer science and industry to solve complex problems
- Discipline scientists including HEP are beginning to follow; are already reaping benefits, and are contributing
- Practical advantages are *Compelling*
  - GPU Computing power per $, and Joules/flop are increasingly, very favorable
  - Training is complex, but execution can be very fast +cheap with the right processor (neuromorphic, FPGA, M4-type)
- We will ride and support the deep learning trends towards
  - Affordable computation
  - Faster algorithms for trigger, pattern rec. and analysis
  - Optimized workflow for globally distributed exascale data
  - Enhanced science-industry interface
- Focusing physicists' efforts on science rather than software
  - To meet the challenges in computing and science
  - Now and through the next Generation

Extra Slides Follow

US LUA

US LHC Users Association

A Special Time in Particle Physics

The List of Outstanding Questions Grows

- 2012 Higgs Discovery; 2013 Nobel Prize
- 2011 Nobel: Accelerated Expansion of the Universe
- 2014 Nobel: Neutrino Oscillations Large neutrino mixing: θ_{13}

AND New Physics Hints

- Dark Matter in cosmic positrons and photons ?
- BSM Effects in the Flavor Sector ?
- Gravitational Waves !

AND Mystery: Higgs and SUSY

Nature is More Subtle

Exciting times just ahead

Higgs boson and EWSB
- ❑ m_H natural or fine-tuned ?
- → if natural: what new physics/symmetry?
- ❑ does it regularize the divergent $V_L V_L$ cross-section at high $M(V_L V_L)$? Or is there a new dynamics ?
- ❑ elementary or composite Higgs ?
- ❑ is it alone or are there other Higgs bosons ?
- ❑ origin of couplings to fermions
- ❑ coupling to dark matter ?
- ❑ does it violate CP ?
- ❑ cosmological EW phase transition

Quarks and leptons:
- ❑ why 3 families ?
- ❑ masses and mixing
- ❑ CP violation in the lepton sector
- ❑ matter and antimatter asymmetry
- ❑ baryon and charged lepton number violation

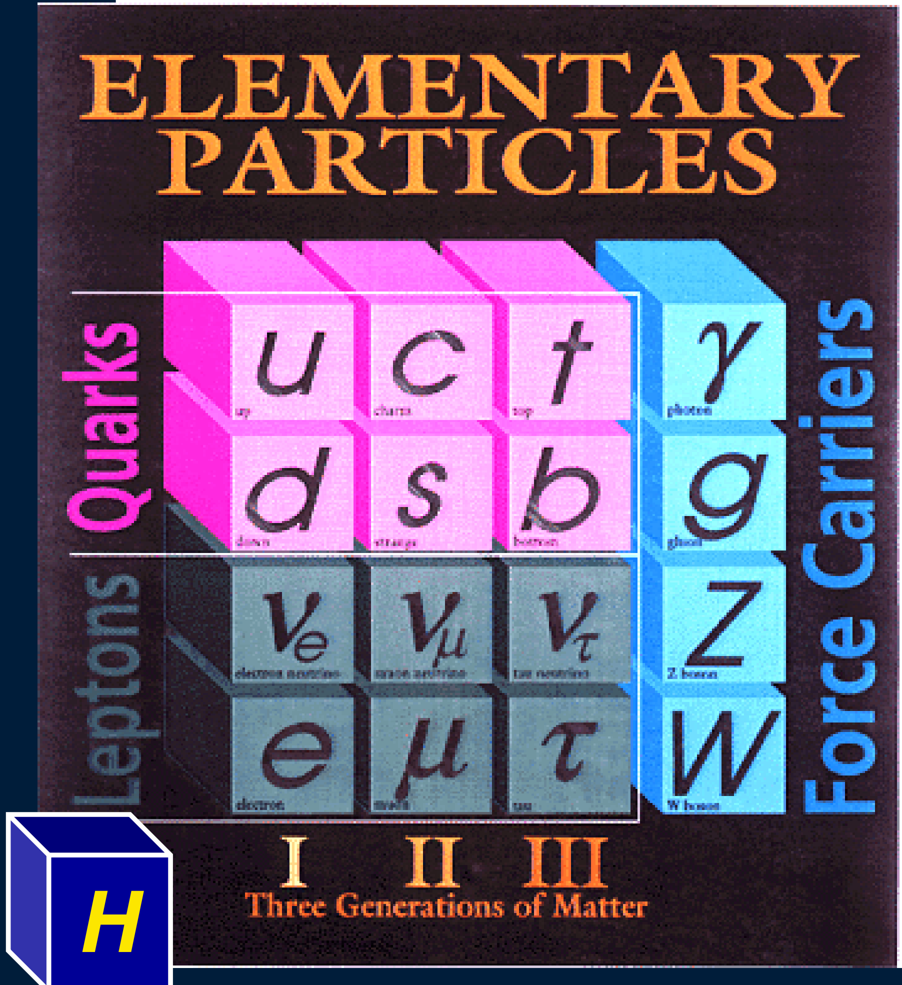
Neutrinos:
- ❑ ν masses and and their origin
- ❑ what is the role of H(125) ?
- ❑ Majorana or Dirac ?
- ❑ CP violation
- ❑ additional species → sterile ν ?

The two epochs of Universe's accelerated expansion:
- ❑ primordial: is inflation correct ? which (scalar) fields? role of quantum gravity?
- ❑ today: dark energy (why is Λ so small?) or is GR wrong on large scales?

Dark matter:
- ❑ composition: WIMP, sterile neutrinos, axions, other hidden sector particles, ..
- ❑ one type or more ?
- ❑ only gravitational or other interactions ?

Physics at the highest E-scales:
- ❑ how is gravity connected with the other forces ?
- ❑ do forces unify at high energy ?

The Standard Model of Particle Physics:
3 Quark, 3 Lepton Families, 3 of 4 Forces



35 Nobel prizes have been awarded for the experimental discoveries & theoretical breakthroughs

[Higgs Boson Generates Masses]

The SM describes the known forces and particles, with one important exception:

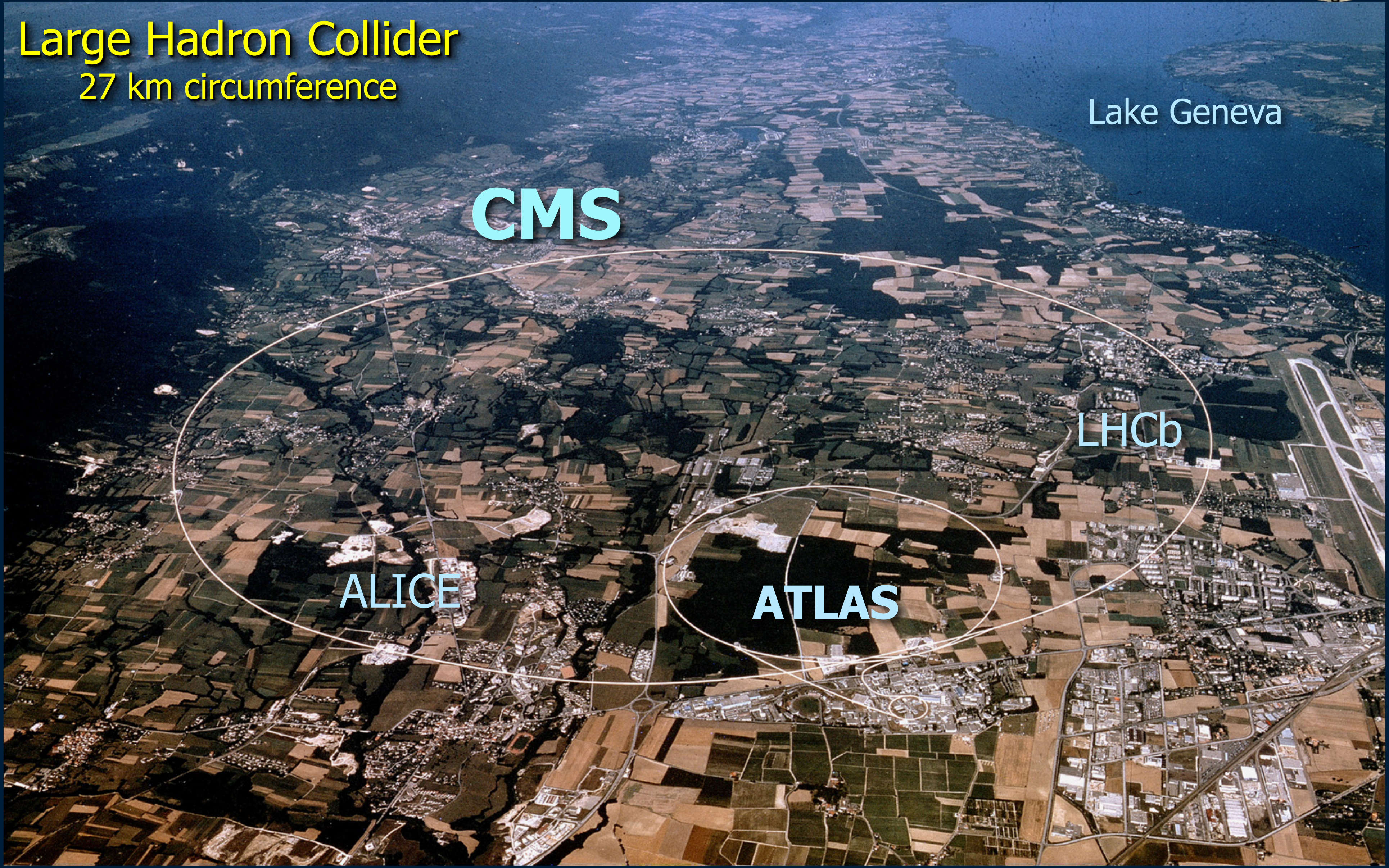
Gravity

And it does not explain:

- The existence of dark matter
- The pattern of particle masses
- The unification of all forces
- The matter-antimatter asymmetry
- Dark energy

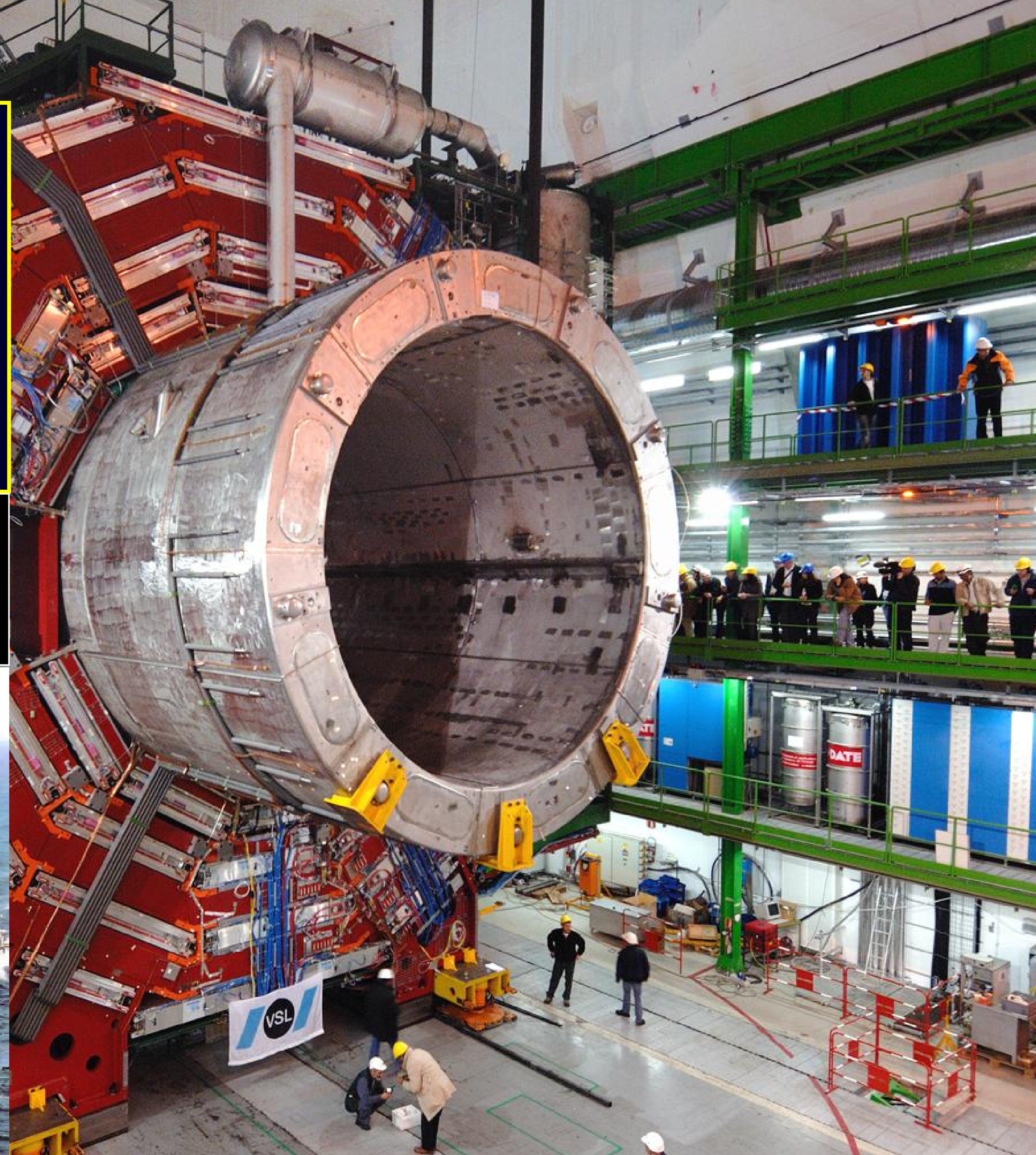
A beautifully simple picture with great predictive power.
Leaving many questions unanswered

ENTER the LHC and the LHC Experiments

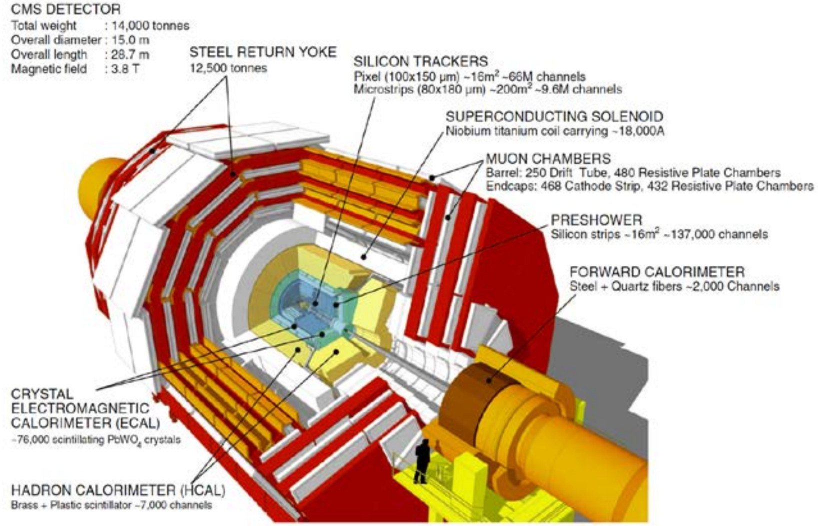


Magnetic length	12.5 m
Free bore diameter	6 m
Central B Field	3.8 Tesla
Temperature	4.2^{o} K
Nominal current	19 kA
Radial Pressure	64 Atm.
Stored energy	***2.7 GJ***

CMS: KE of a Nimitz Class 117,000 Ton Carrier Moving at 20 mph



CMS Design



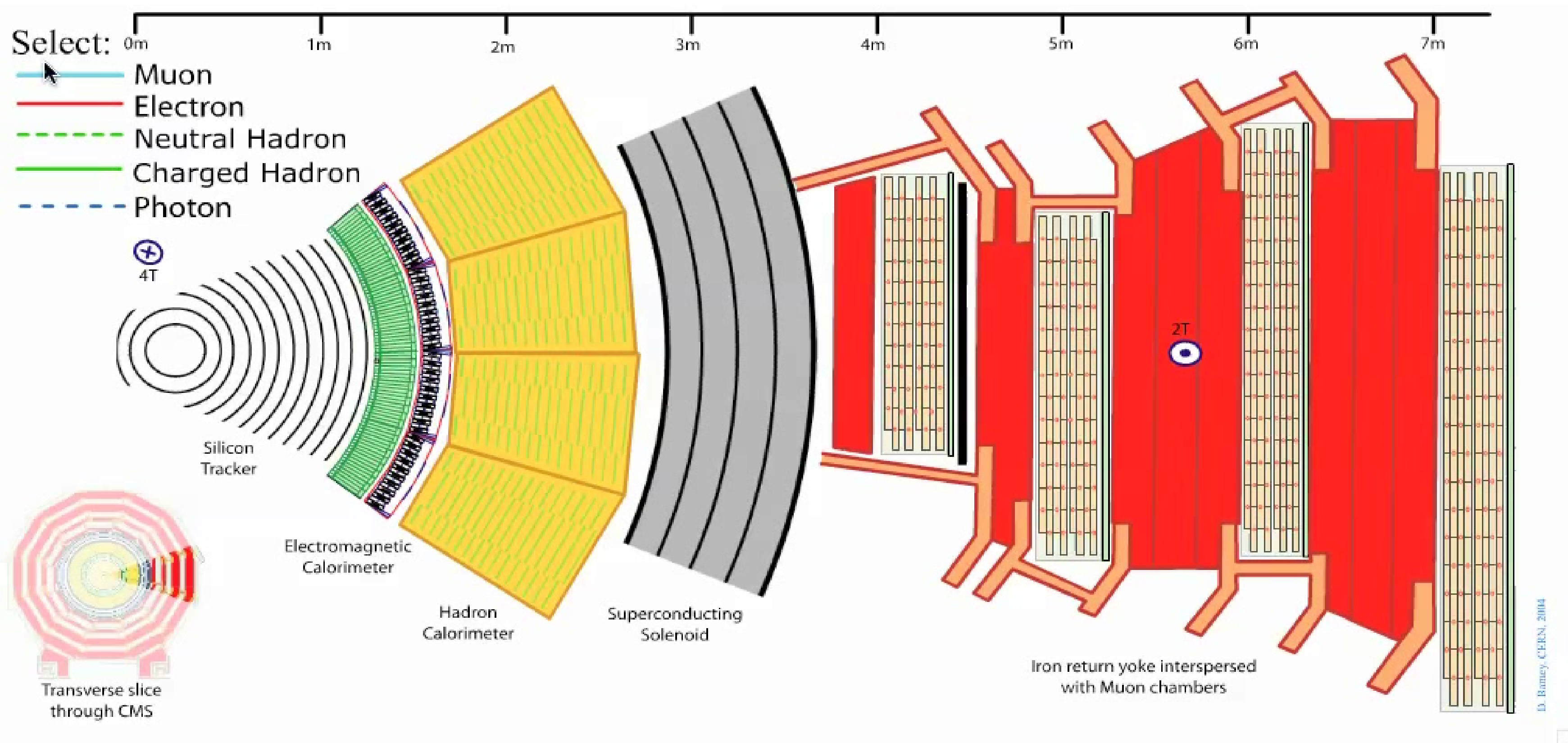
High Field (3.8T)

Modular

Compact Tracker

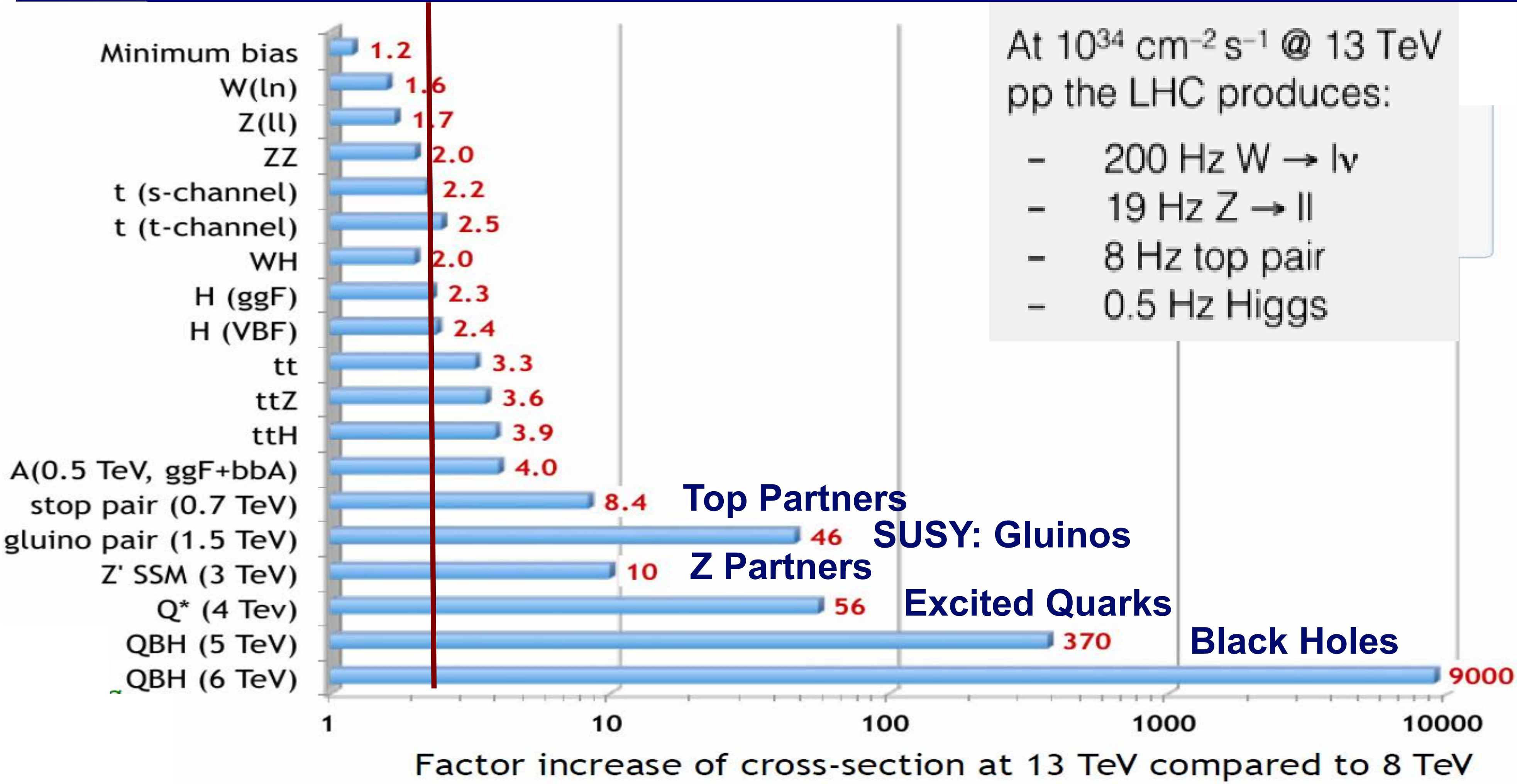
Precise ECAL: inside Coil

Muons in & Out



LHC Run2 Production Rates: 13 Vs 8 TeV

Ratio 2 to 9k Times: Entering a New Era of Discovery



Greater Sensivity to New Physics and Higgs Properties Across the Board; *Especially for High Masses*

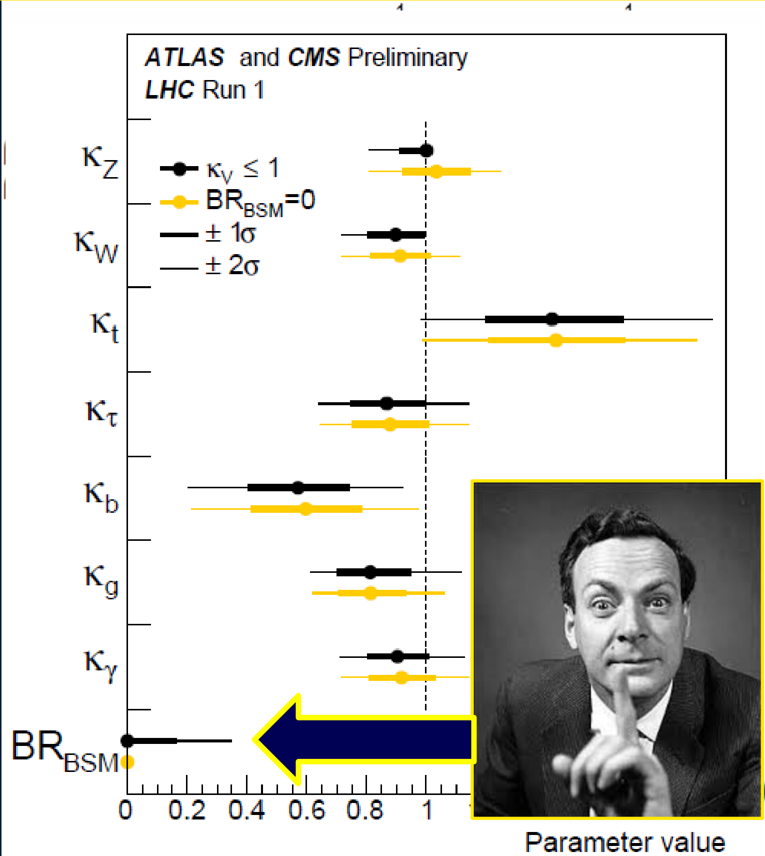
Prospects for Run2 and Beyond: 2016-37

"There's Plenty of Room at the Bottom"

An Invitation to Enter a New Field of Physics

(Feynman Lecture at Caltech, December 29, 1959)

There is So Much Room



CMS

L (fb^{-1})	κ_γ	κ_W	κ_Z	κ_g	κ_b	κ_t	κ_τ	$\kappa_{Z\gamma}$	κ_μ	BR_{invis}
300	7%	6%	6%	8%	13%	15%	8%	41%	23%	28%
3000	5%	5%	4%	5%	7%	10%	5%	12%	8%	17%

ATLAS

L (fb^{-1})	κ_γ	κ_W	κ_Z	κ_g	κ_b	κ_t	κ_τ	$\kappa_{Z\gamma}$	κ_μ	BR_{invis}
300	9%	9%	8%	14%	23%	22%	14%	24%	21%	22%
3000	5%	5%	4%	9%	12%	11%	10%	14%	8%	14%

And if We Improve

→ Reduce Theory Systematics by 50% | → Reduce Exp Syst by √Lumi

	κ_γ	κ_W	κ_Z	κ_g	κ_b	κ_t	κ_τ	$\kappa_{Z\gamma}$	κ_μ	BR_{invis}
ATLAS	5→4	5→5	4→4	9→7	12→11	11→9	10→9	14→14	8→7	14→11
CMS	5→2	5→2	4→2	5→3	7→4	10→7	5→2	12→10	8→8	17→6

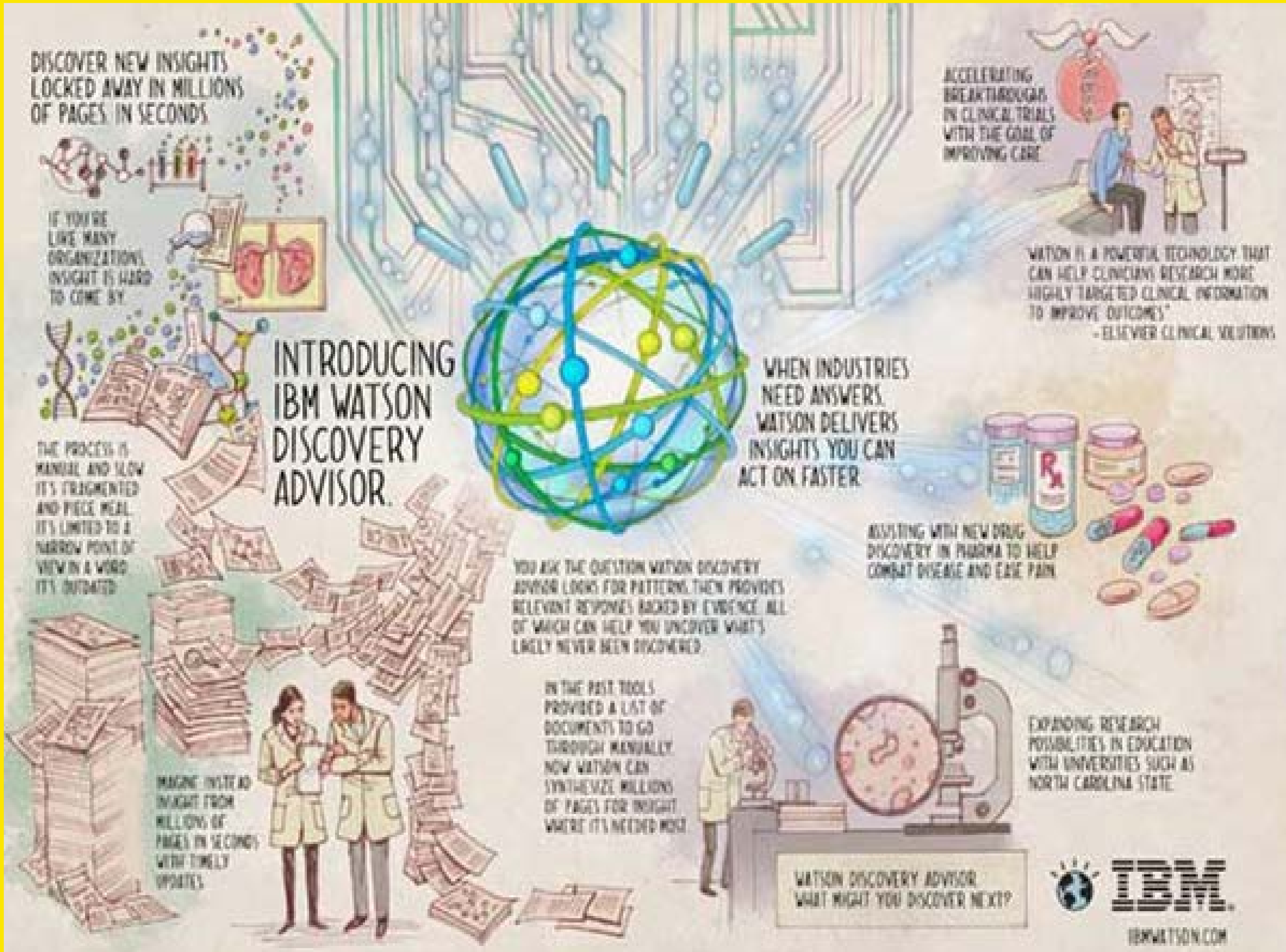
Plus Rare Higgs Decays, DiHiggs and BSM Higgs Production,

We have only just begun: Time for Deep Learning and Innovation

The View in LHC Run1

Mining Documentation

IBM Watson Discovery Advisor

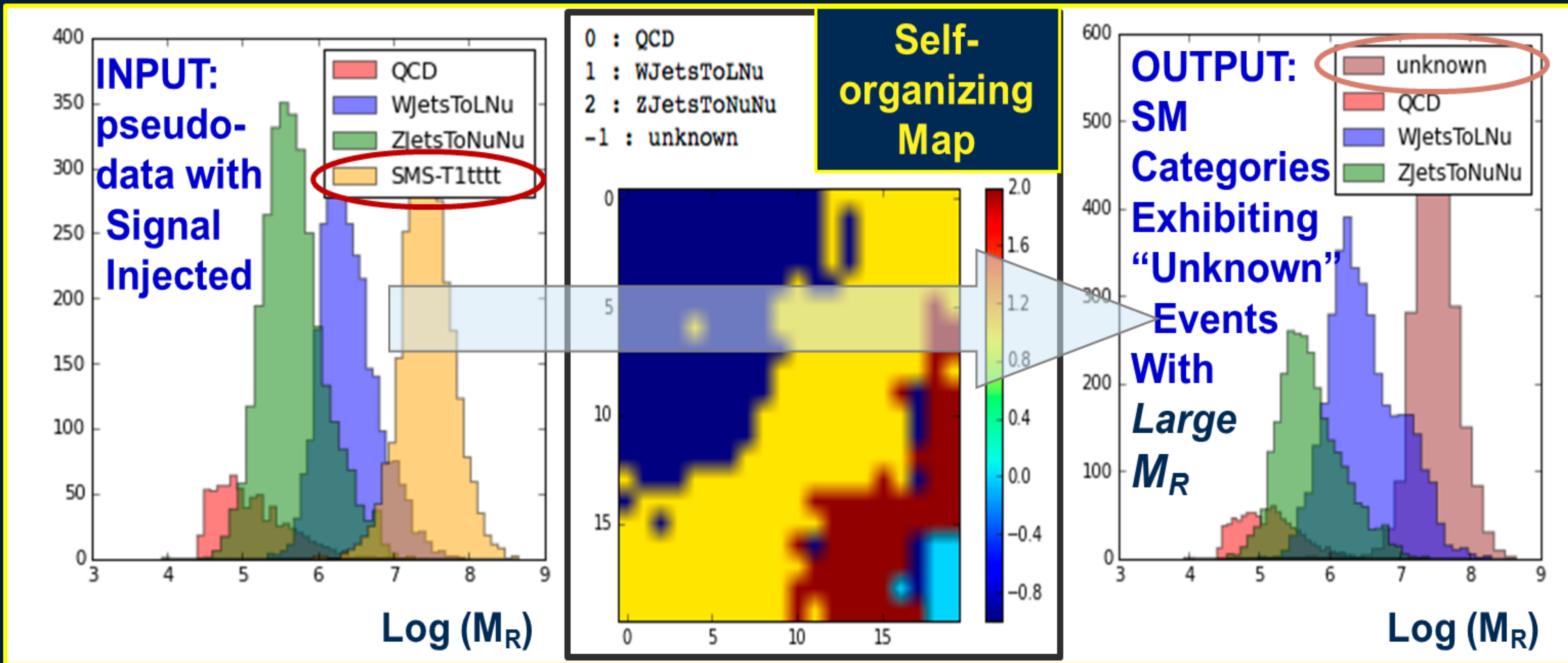


http://www.ibm.com/smarterplanet/us/en/ibmwatson/discovery-advisor.html

Demonstrated the ability to make sense of a large volume of research papers and provide insights

Machine Learning: Exploring New Methods

Aim to extend CMS' (and HEP's) Discovery Reach



Targets: Analysis - Identification/discovery of *unknown BSM signals; Optimization of LHC workflow and distributed system operations*

- Synergy with previous Computing Model work on optimization of global grid and network systems using **Self-organizing Neural Nets in MONARC**

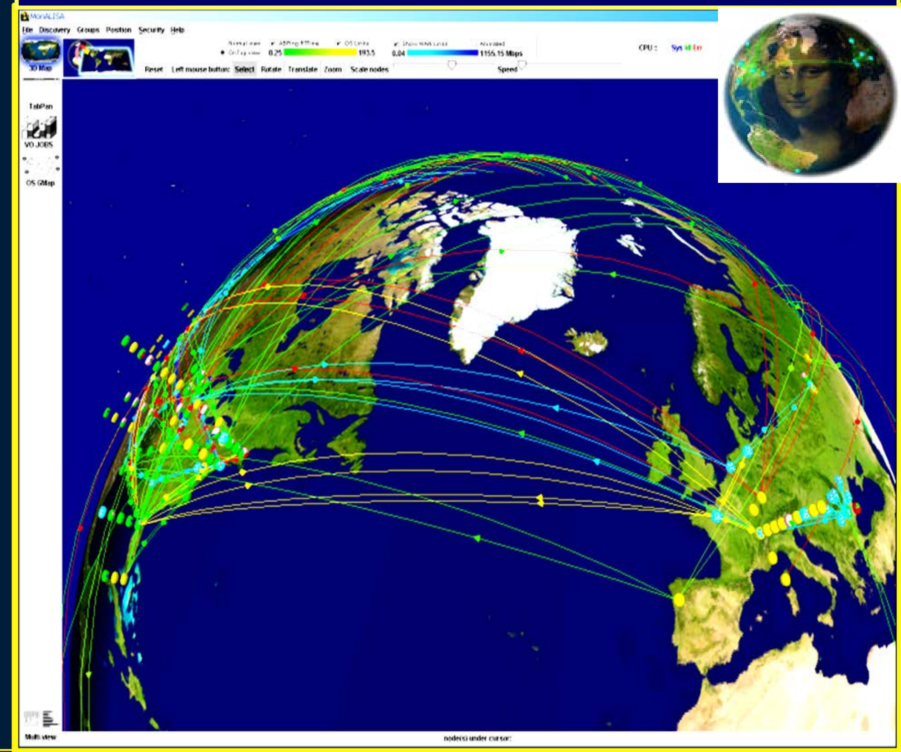
Building Consistent Agile Network Operations
At the Edges and in the Core

A New Era of Technical Challenges as we Move to Exascale Data and Computing

- **Beyond network capacity and reliability alone,** the keys to future success are next generation systems **able to:**
  - Respond agilely **to peak and shifting workloads**
  - Accommodate a more diverse set of computing systems **from the Grid to the Cloud** to HPC
  - Coordinate the use of globally distributed computing and storage, and networks that interlink them
    - **In** a manner compatible across fields **sharing common networks**
- **The complexity of the data, and hence the needs for CPU power, will grow disproportionately:** by a factor of several hundred during the same period

MonALISA: Monitoring Agents in a Large Integrated Services Architecture

A Global Autonomous Real Time System



Next Gen SDN Systems for Exascale Science

Vision: Distributed environments where resources can be deployed flexibly to meet the demands

- SDN is a natural path to this vision:
  - Separating the functions that control the flow of traffic, from the switching infra-structure that forwards the traffic
  - Through open deeply programmable "controllers".

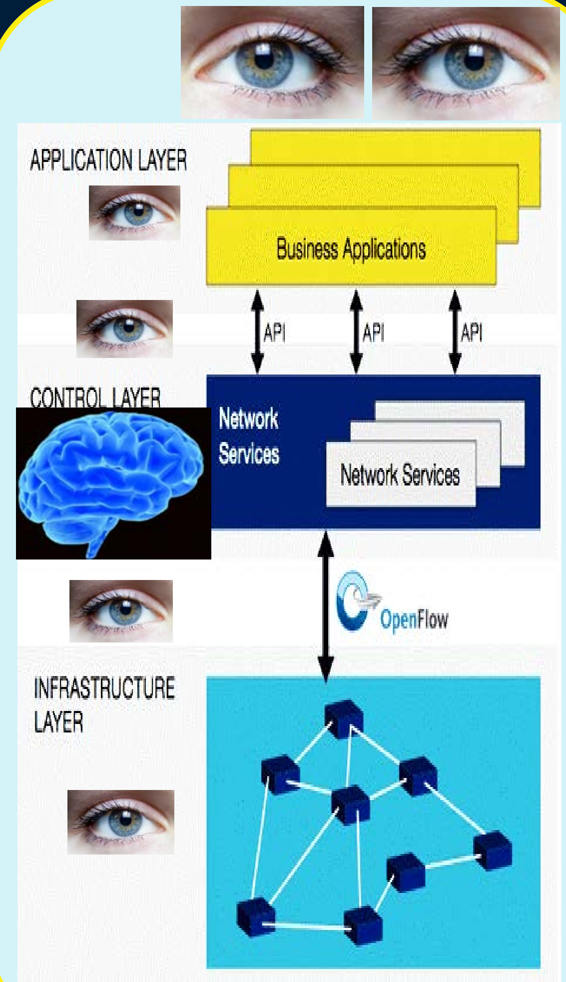
With many benefits:

- ❒ Replacing stovepiped vendor HW/SW solutions by open platform-independent software services
- ❒ Virtualizing services and networks: lowering cost and energy, with greater simplicity
- ❒ Adding intelligent dynamics to system operations

A major direction of Research networks + Industry

- ❒ A Sea Change that is still emerging and maturing

Building on the Caltech/ESnet/Fermilab Pilot Experience



opennetworking.org

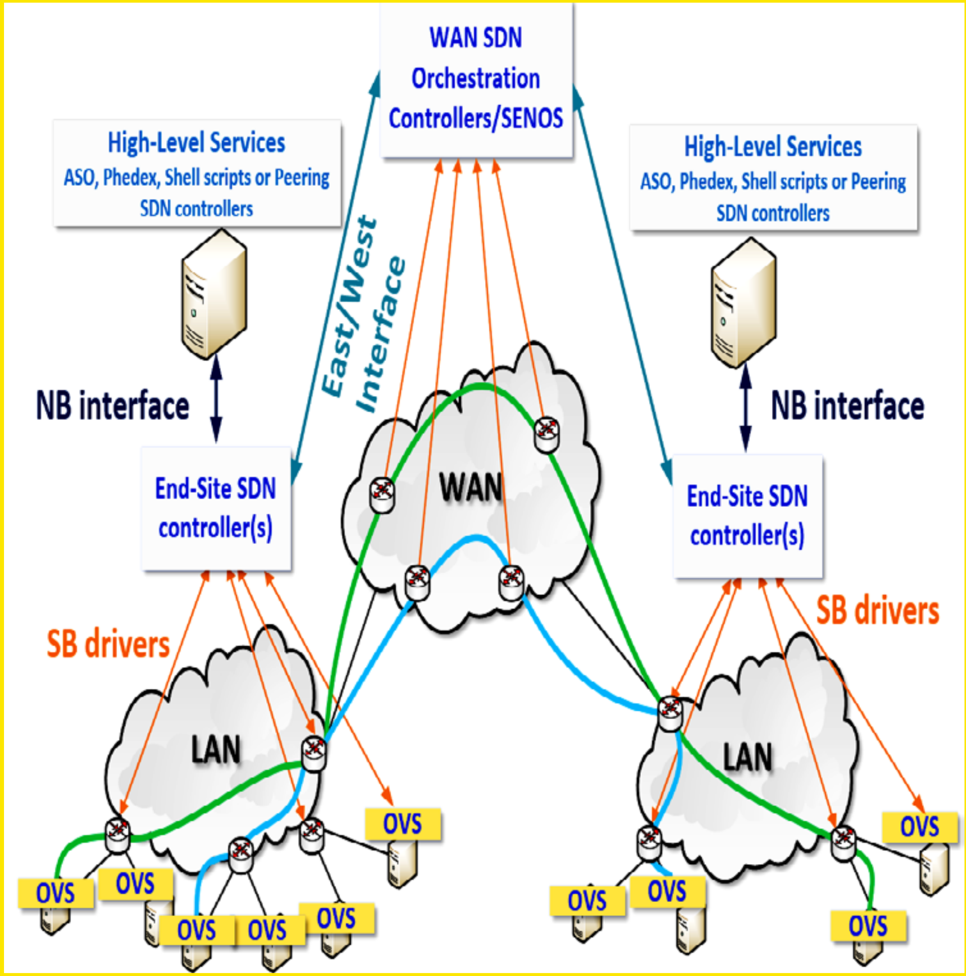
A system with built in intelligence

Requires excellent monitoring at all levels

OVS End- and Inter-Site Orchestration

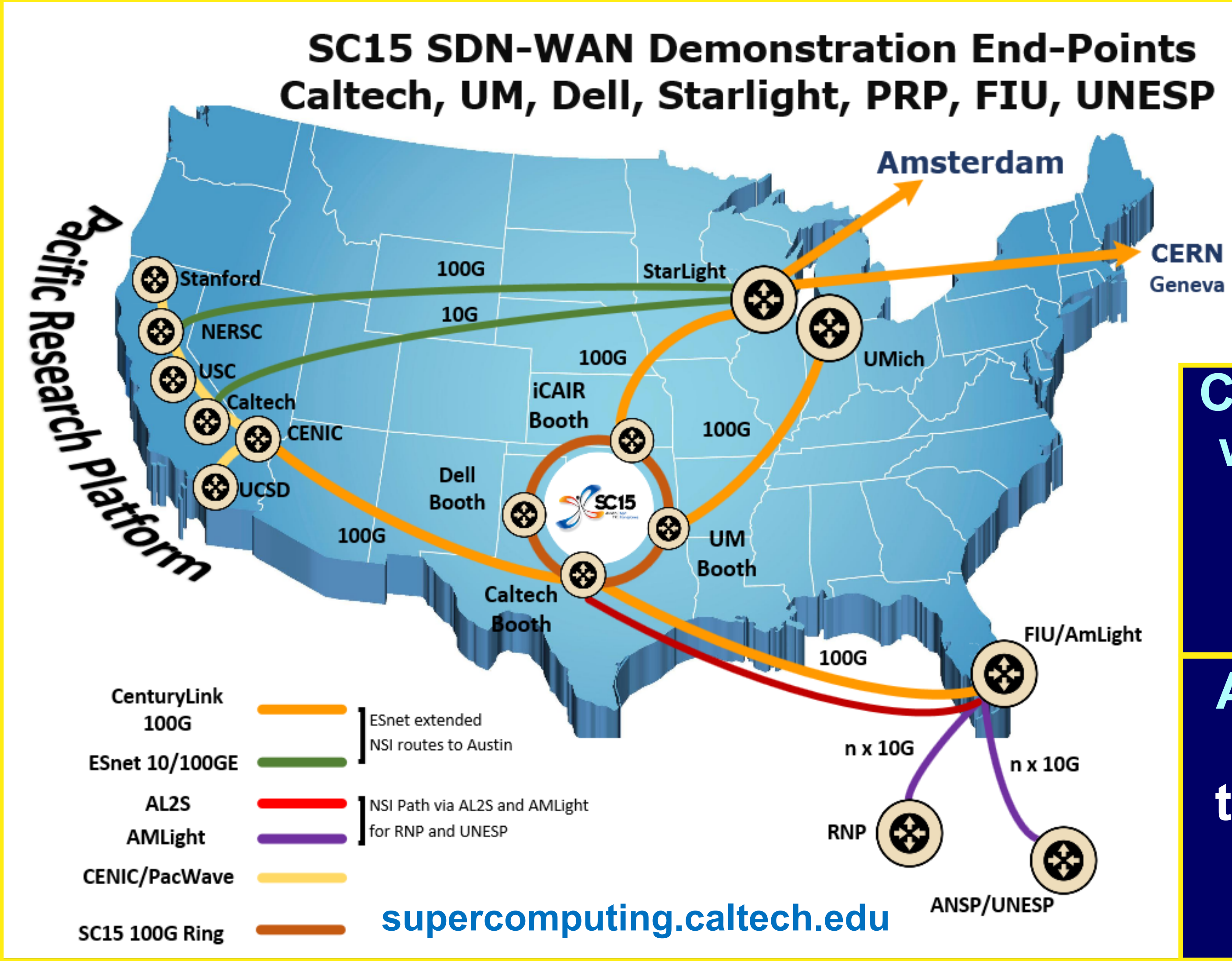
Design + Implementation: Multiple Host Groups, Paths, Policies

- ❑ Diverse network paths to support flows among multiple host groups
- ❑ Diverse policies governing path setup and prioritization of flows
- ❑ Assigned bandwidth individually or in groups in response to users, applications [e.g. PhEDEx, ASO], upstream SDN controllers
- ❑ Real-time adjustment of allocations triggered by: (1) new requests, (2) real-time feedback on progress of transfers, (3) network state changes or error conditions, (4) proactive load-balancing operations, or (5) rate-limiting operations imposed by controllers or emerging network operating systems (e.g. SENOS)



Northbound Interaction with SDN Controller(s)

SC15: SDN Driven Next Generation Terabit/sec Integrated Network for Exascale Science



SDN-driven flow steering, load balancing, site orchestration Over Terabit/sec Global Networks

Consistent Operations with Agile Feedback: Major Science Flow Classes Up to High Water Marks

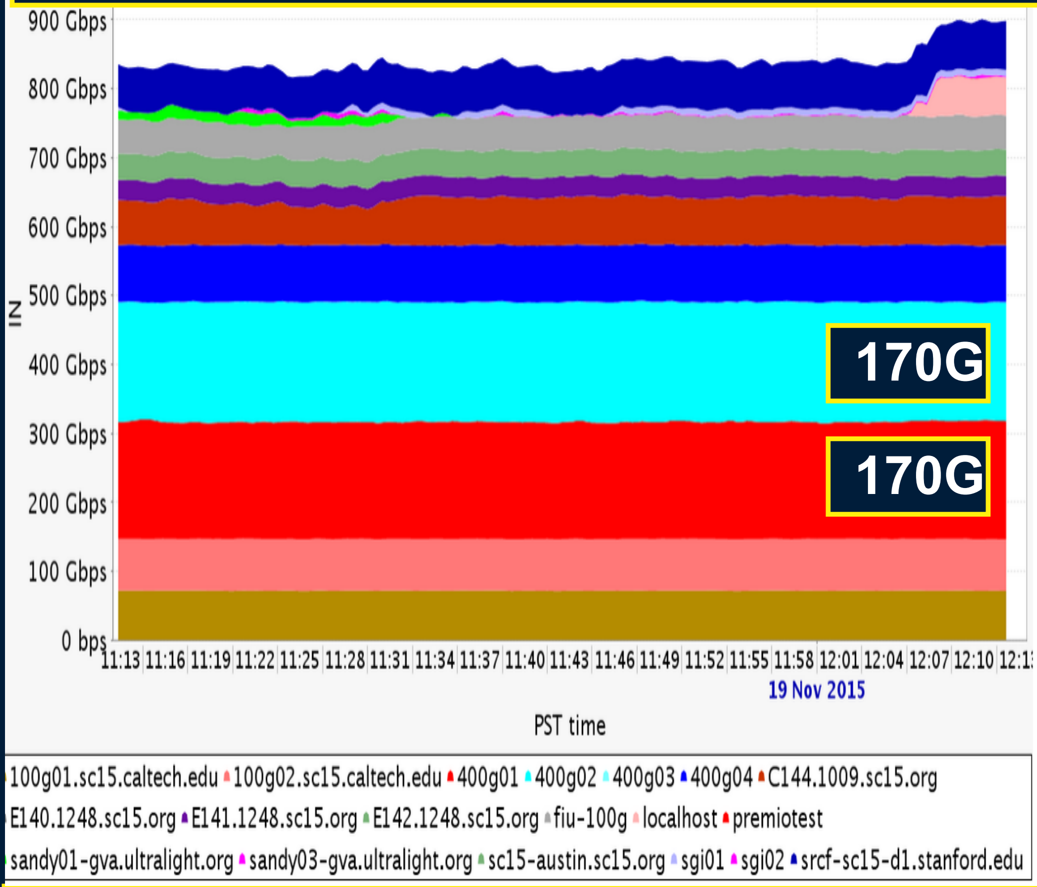
Added Goal: Preview PetaByte Transfers to/from Site Edges of Exascale Facilities With 400G DTNs

Open Daylight SDN Controller

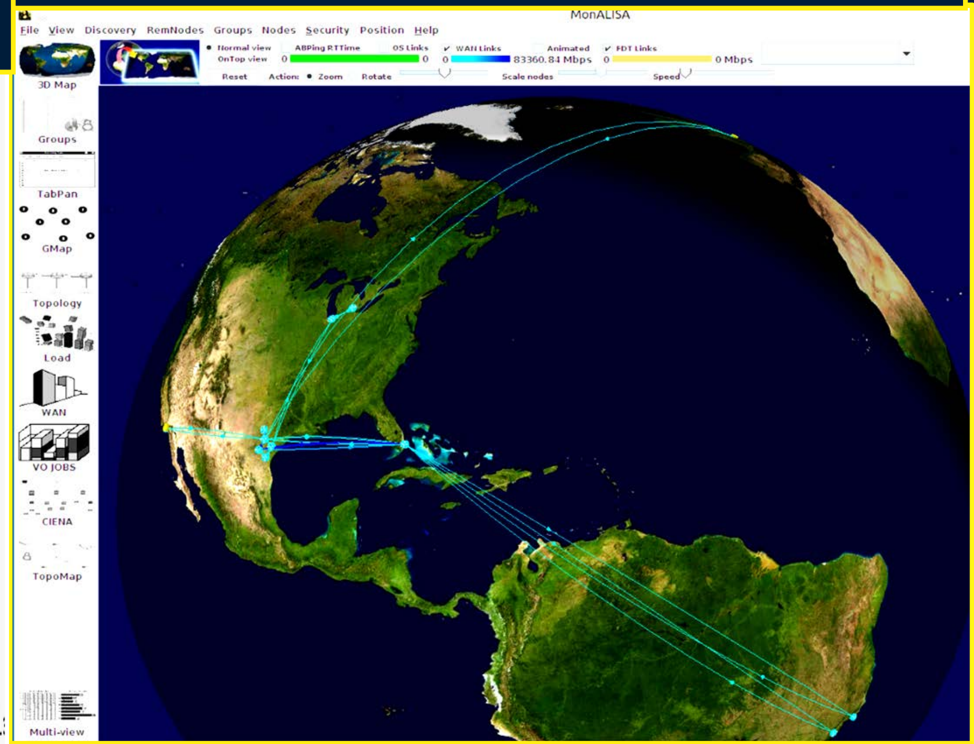
Tbps Ring Planned for SC16

SC15: Terabit/sec SDN Driven Agile Network
Aggregate Results

900 Gbps Total
Peak of 360 Gbps in the WAN



Global Topology



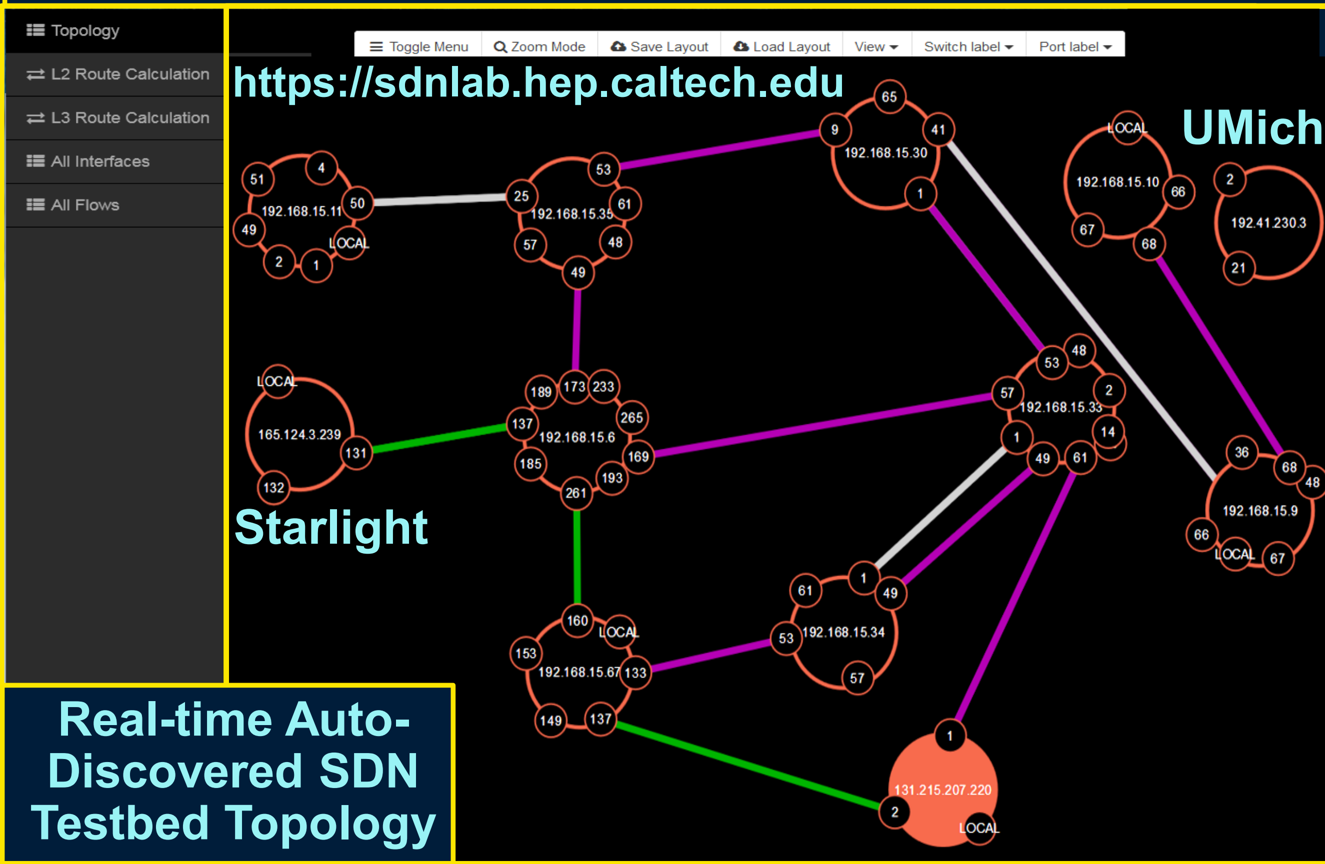
29 100G NICs; Two 4 X 100G and Two 3 X 100G DTNs; 9 32 X100G Switches

Smooth Single Port Flows up to 170G; *120G over the WAN.* With Caltech's FDT TCP Application http://monalisa.caltech.edu/FDT

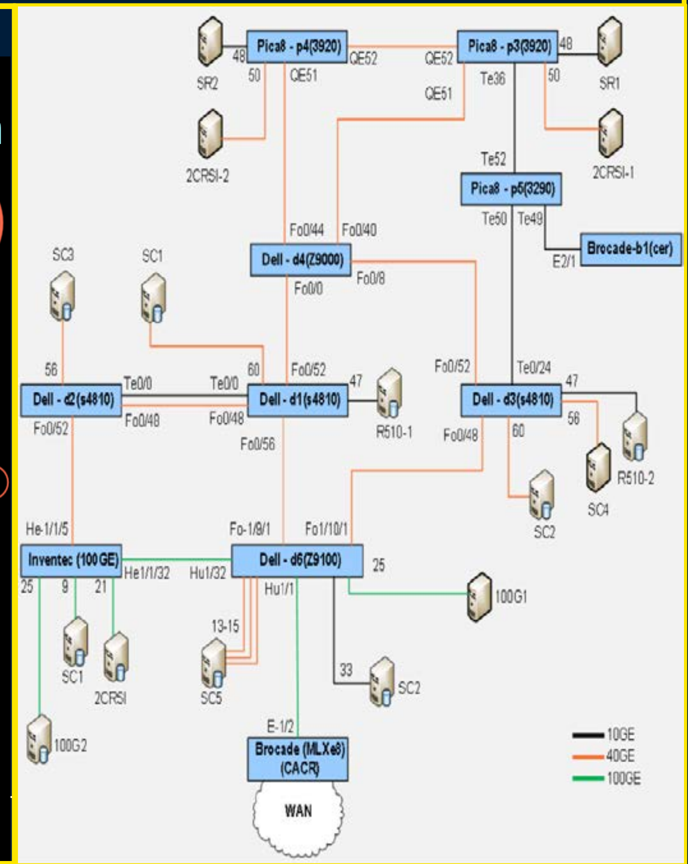
SDN State of the Art Development Testbed

Caltech, Fermilab, StarLight, Michigan; + CERN, Amsterdam, Korea

- ❑ 11 Openflow switches: **Dell, Pica8, Inventec, Brocade**
- ❑ Many 40G, N X 40G, 100G Servers: **Dell, Supermicro, 2CRSI, Echostreams; and 40G and 100G Network Interfaces: Mellanox, QLogic**
- ❑ Caltech Equipment **funded through the** NSF DYNES, ANSE, CHOPIN projects, and vendor donations



Real-time Auto-Discovered SDN Testbed Topology



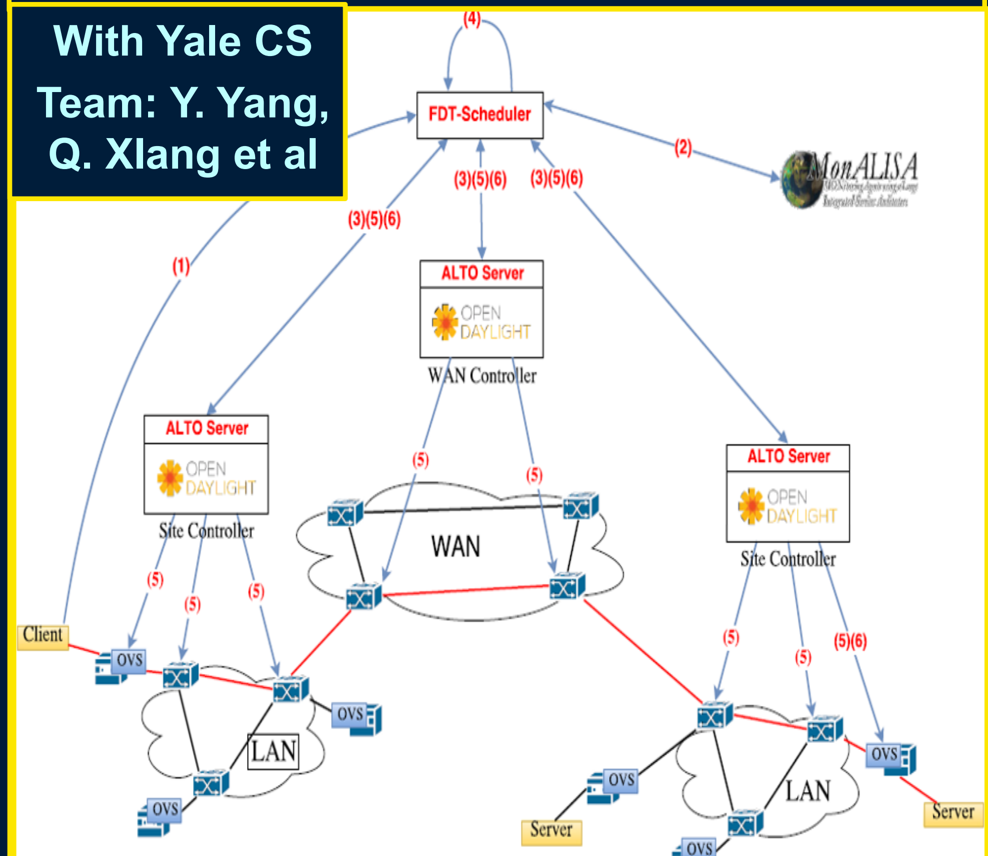
Next Generation "Consistent Operations"

Site-Core Interactions for Efficient, Predictable Workflow

- Key Components: (1) OVS at edges to stably limit flows (2) Application Level Traffic Optimization (ALTO) in Open Daylight for end-to-end optimal path creation, coupled to flow metering and high watermarks set in the network core
- Real-time flow adjustments triggered as above
- Optimization using "Min-Max Fair Resource Allocation" (MFRA) algorithms on prioritized flows
- Flow metering in the network fed back to OVS edge instances; changes applied to ensure smooth progress of flows end-to-end
- High Water Marks to protect the world's R&E networks

Consistent Ops Paradigm applied to file transfers with ALTO, OVS and MonALISA FDT Schedulers

With Yale CS Team: Y. Yang, Q. XIang et al



Demos: Internet2 Global Summit in May; SC16 in November

Bringing Pre-Exascale and Exascale LCFs

Into the Global Dynamic Ecosystem

Exascale Ecosystems
for Next-Generation Data Intensive Sciences

- The opportunity for HEP (CMS example):
  - CPU needs will grow 65 to 200X by HL LHC
  - Dedicated CPU that can be afforded will be an order of magnitude less; even after code improvements on the present trajectory
- DOE ASCR/HEP Exascale Workshop June 2015:
  - Exposed the favorable LCF outlook + issues
- Short term Goal: Making such systems a grid resource for CPU using data resident at Fermilab Tier1 and US Tier2s
- Important Long Term benefits
  - Folding LCFs into a global ecosystem for data intensive sciences
  - Building a modern coding workforce
  - Shaping the future architecture and operational modes of Exascale Computing Facilities

Leadership

3 Pilot Programs with Argonne

1. MIRA as a CMS grid resource
2. Precise NLO generators on Mira with new more efficient methods
3. DTN and process design for 100G+ data transfers

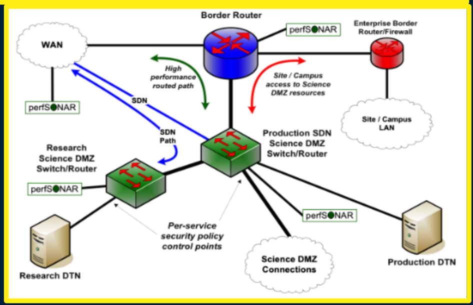
Pilot with Argonne: Operational Architecture for LCFs Work for (LHC and Other) Data Intensive Applications

➡ Developments targeting the CPU Needs at LHC Run3 and HL LHC

- ❒ Developing system architectures in hardware + software that meet the needs
  - ✸ Edge clusters with petabyte caches
    - ✸ Input + output pools: ~10 to 100 Pbytes
  - ✸ A handful of proxies at the edge
    - ✸ To manage and focus security efforts
  - ✸ Extending Science DMZ concepts
    - ✸ Enabling 100G to Tbps SDNs with Edge/WAN Coordination
  - ✸ Identifying + matching HEP units of work to specific sub-facilities adapted to the task
  - ✸ Site-Network End-to-End Orchestration
    - ✸ Efficient, smooth petabyte flows over 100G then 400G (2018) then ~1 Tbps (2021) networks
  - ✸ *Machine Learning to Optimize the Workflow*

Leadership

Next Gen Science DMZ



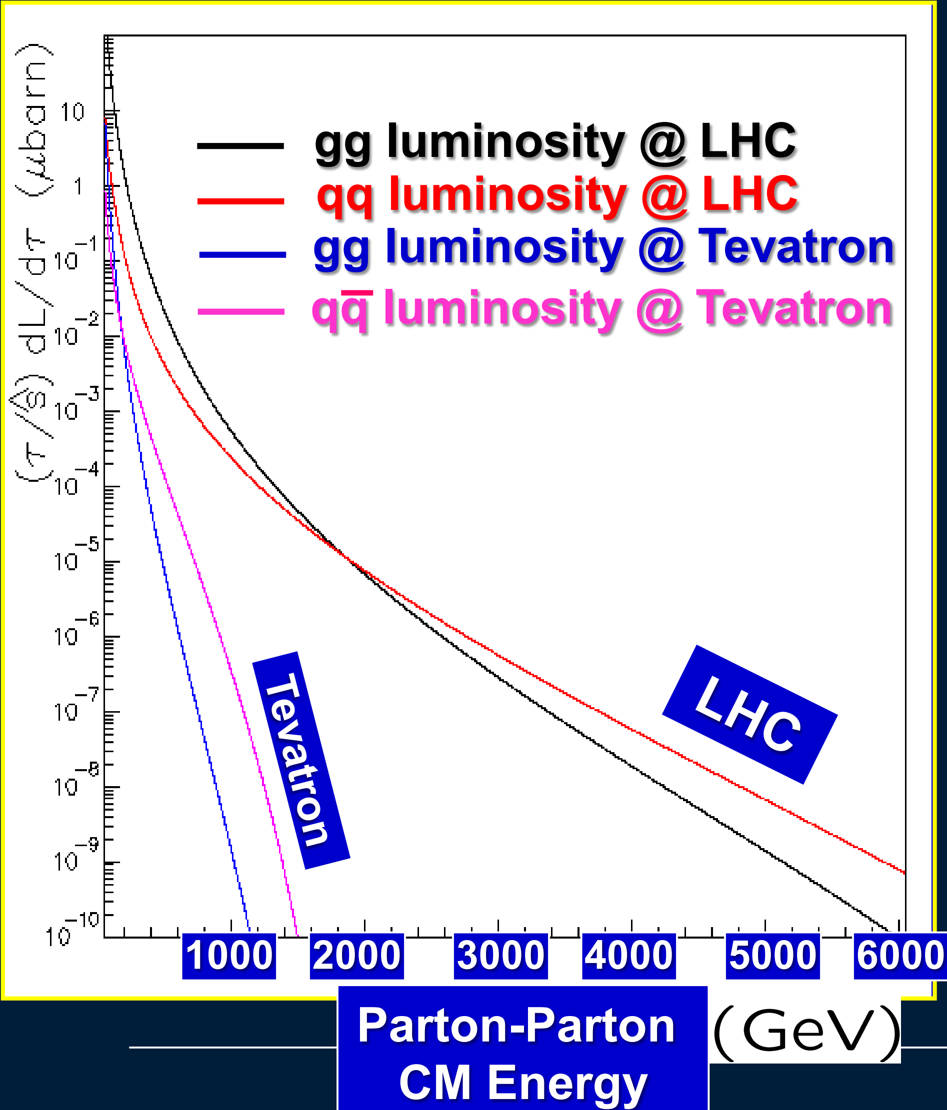
Networks and LCFs for HEP and Exascale Science: Our Journey to Discovery

- Run 1 brought us a centennial discovery: the Higgs Boson
- Run 2 will bring us (at least) greater knowledge, and perhaps greater discoveries: Physics beyond the Standard Model.
- *Advanced networks will continue to be a key* to the discoveries in HEP and other fields of data intensive science and engineering
- Technology evolution *might* fulfill the short term needs
- A new paradigm of global SDN networks should emerge during LHC Run2 (in 2015-18) to address the needs, together with
- *New approaches + a new class of global networked systems to handle Exabyte-scale data, with a focus on ECFs are needed [building on LHCONE, DYNES, ANSE, OliMPS; SDN NGenIA + SENSE]*
- Wide deployment of such systems by ~2023 will be:
  - Essential to meet the challenges at the LHC and HL-LHC
  - A game-changer with the potential to shape both research and daily life: dealing with *truly-Big Data*
- The ongoing Caltech – Fermilab – ESnet partnership, and the comprehensive vision, are the keys to future success

Summary

- *Advanced networks will continue to be a key* to the discoveries in HEP and other data intensive fields of science and engineering
- *Near Term and Decadal Challenges must be addressed: Greater scale, complexity and scope*
- *New approaches + a new class of software driven networked systems to handle globally distributed Exabyte-scale data are being developed*
- Deeply programmable, agile software-defined networks (SDN) are a key ingredient of NGenIA
- Adapting Exascale Computing Facilities to meet the highest priority needs of data intensive science, including high energy physics as a first use case (to be followed by others) will empower the HEP community to make the anticipated next and future rounds of discoveries

The LHC Mission: Opening a Realm of High Energies and a New Era of Discovery



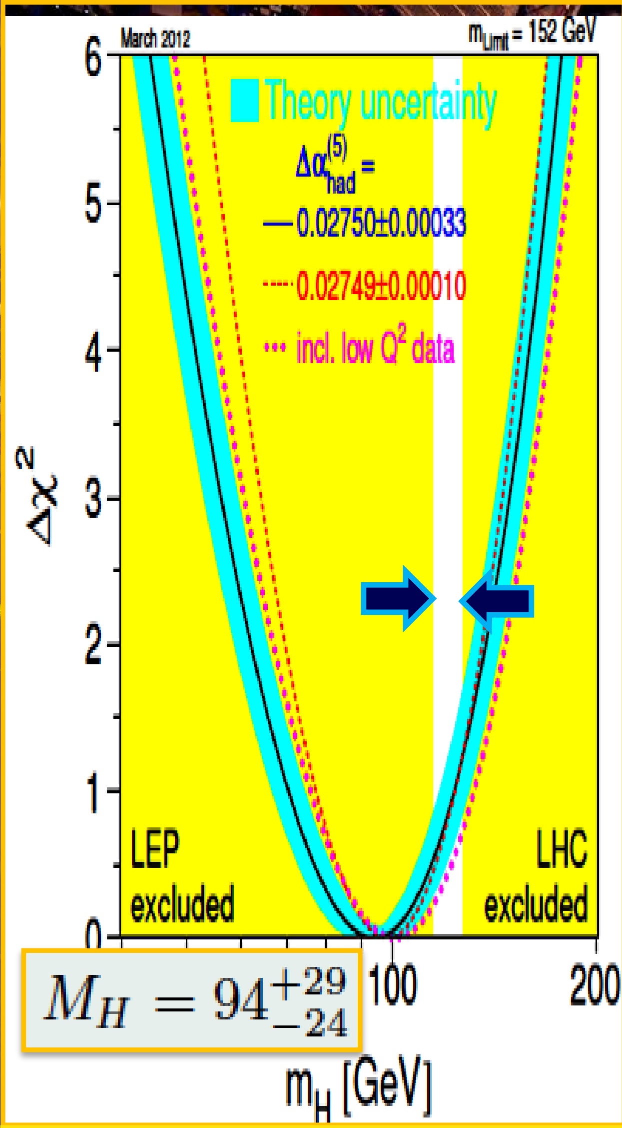
- The LHC is a **Discovery Machine**
- The first accelerator to probe deep into the Multi-TeV scale
- Its mission is ***Beyond the SM***
- There are many reasons to expect new physics

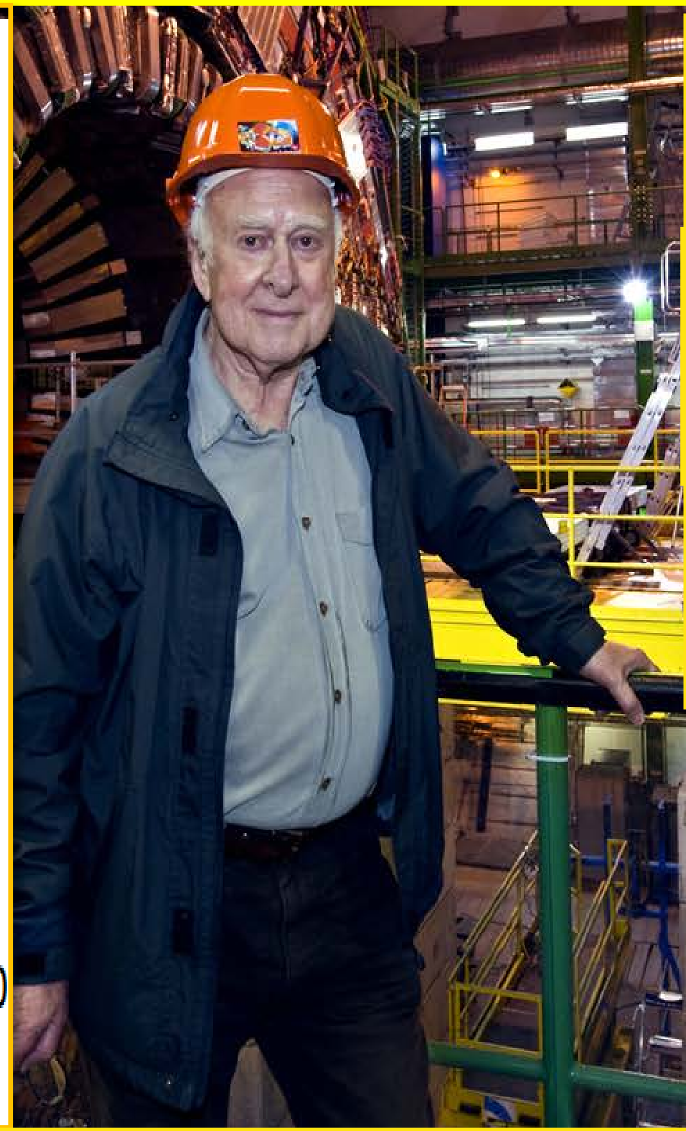
SUSY, Substructures, *Graviton Resonances, Black Holes, Low Mass Strings, … the Unexpected*

We do not know what we will find

Nature is More Subtle

State of the Higgs on July 1 2012





LEP Precise Electroweak Data (Indirect)

M_H < 152 GeV (95% CL)

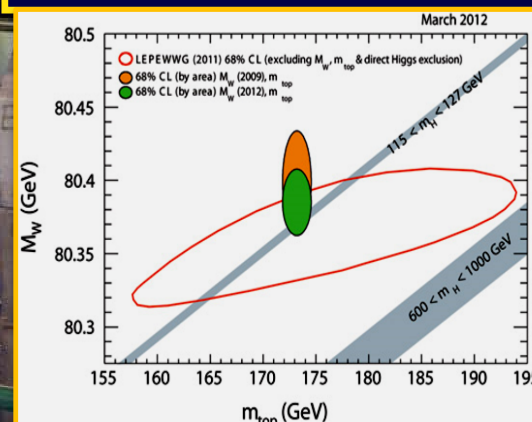
Direct Searches:

LEP: M_H > 114.4 GeV

Fermilab Exclusion 162 - 166 GeV (95%CL)

Direct Searches at CMS (by Dec. 2011)

127 – 600 Excluded



***Closing In:* Only a Narrow 13 GeV Gap Remained**

The Higgs at Last: Signatures

"The delicate, rare fingerprints of the Higgs Boson"

Michael Riordan, Guido Tonelli and Sau Lan Wu
Scientific American **307**, 66 - 73 (2012) Published online: 18 September 2012
doi:10.1038/scientificamerican1012-66

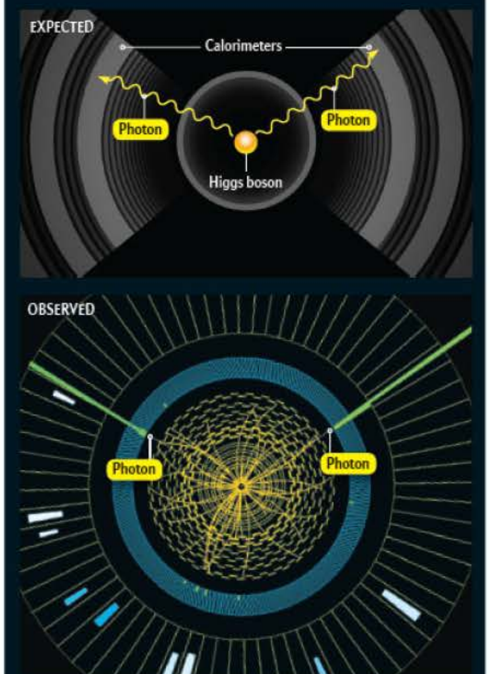
FINDINGS

The Delicate, Rare Fingerprints of the Higgs

The Higgs boson is an extremely unstable particle that quickly decays via a number of different processes, or "modes." Unfortunately, many decay modes are indistinguishable from the thunderous din of ordinary background events that result from 500 million proton-proton collisions every second. The ATLAS and CMS experiments are designed to spot the occasional interesting events that might come from the Higgs decay and throw much of the rest away. The drawings below show four of the most important decay modes that experiments use to search for the Higgs, along with images of actual Higgs-like signals that CMS observed in the 2011 and 2012 runs. (Because the discovery is statistical in nature, no single event can be used as definitive proof.)

Photons

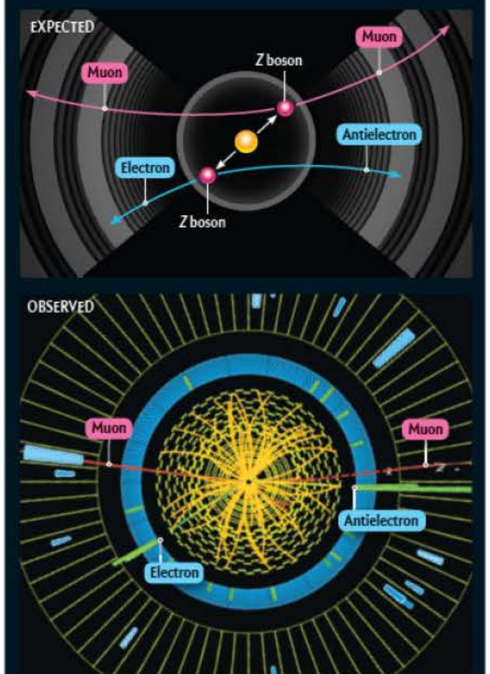
Each detector includes multiple calorimeters, devices for measuring the energy of particles. The innermost calorimeter is particularly alert for photons. These are absorbed in the calorimeter and create tiny electrical signals. If a Higgs decays into two photons, the detector can measure their total energy at extremely high accuracy, which helps to precisely reconstruct the mass of the newly found particle.



H ➔ γγ

Z Bosons

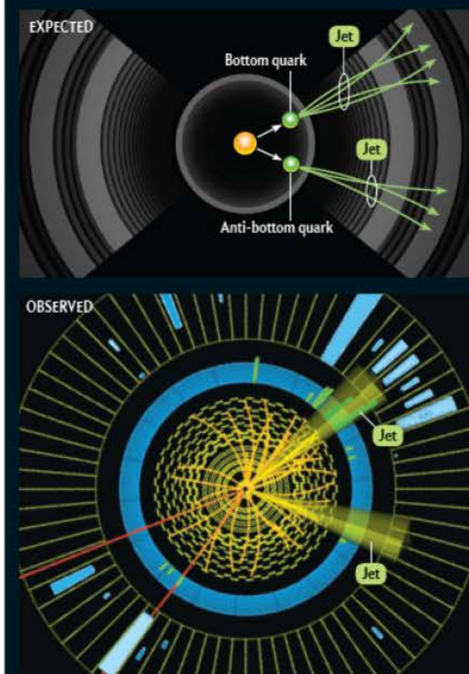
The Higgs may decay into a pair of *Z* bosons, each of which can decay into an electron paired with an oppositely charged antielectron or two muons. An inner tracker and calorimeter measure the electrons, while muons fly out, leaving footprintlike tracks as they go. High magnetic fields bend the path of electrons and muons during their trip, allowing for a high-resolution measurement of their energy and the original Higgs mass.



H ➔ ZZ ➔ e+e- μ+μ−

Bottom Quarks

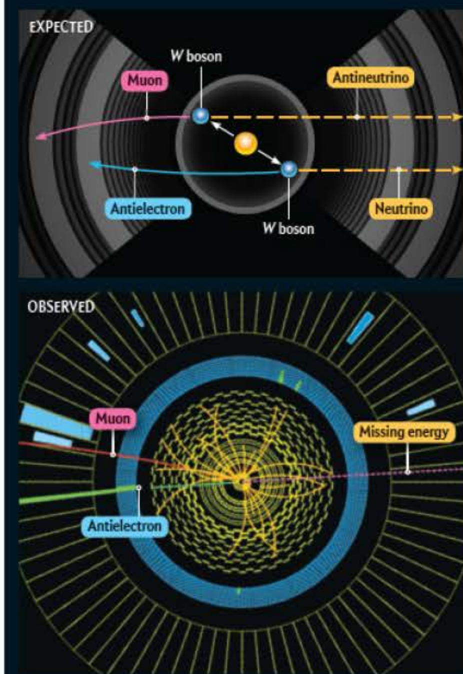
The Higgs can also decay to a bottom quark and its antiparticle, each of which decays into a tight "jet" of secondary particles called hadrons (composite particles made of quarks). These hadrons fly through the detector's inner layers and deposit their energy in the outer calorimeters. Unfortunately, many ordinary collisions also generate jets of hadrons from bottom quarks, which makes it difficult to separate these Higgs events out from the background.



HZ ➔ bb + 2 Leptons

W Bosons

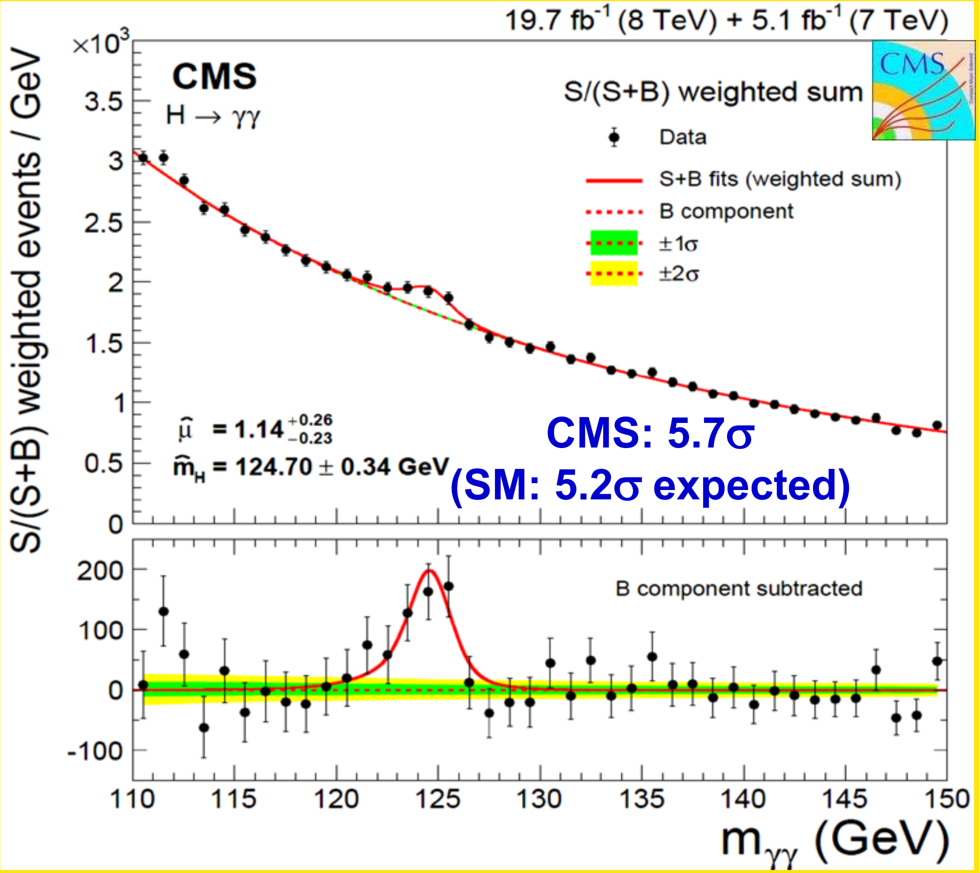
The Higgs can also decay to two *W* bosons, each of which can decay into an electron, antielectron or muon, plus a neutrino or antineutrino. Neutrinos are nearly impossible to detect—they fly out of the detector as if they were never there, taking with them some of the event's energy. Researchers use this missing energy to infer their presence, but the missing energy also prevents them from accurately reconstructing the mass of the original Higgs boson.

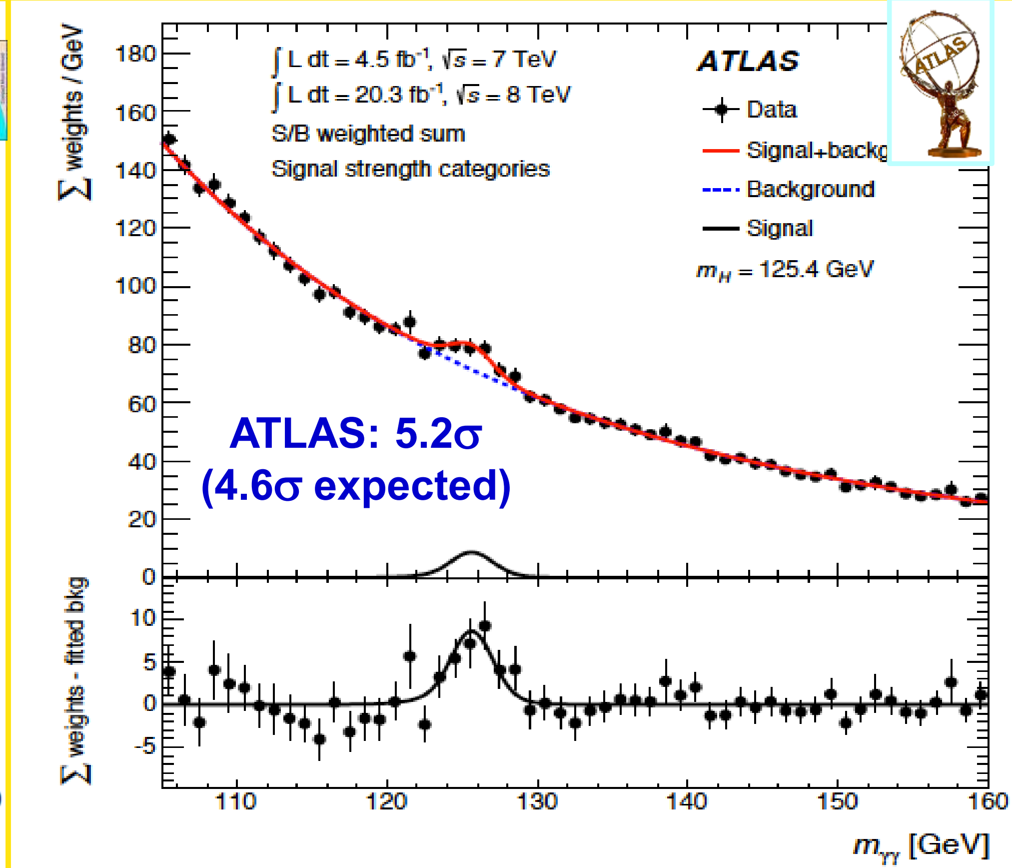


H ➔WW➔Leptons+MET

H ➡ $\gamma\gamma$ at LHC Run 1 (2015)

Enough for Discovery in this channel alone





Arxiv 1407.0558v2 EPJ C74 (2014) 3076

Phys. Rev. D90 (2014) 112015

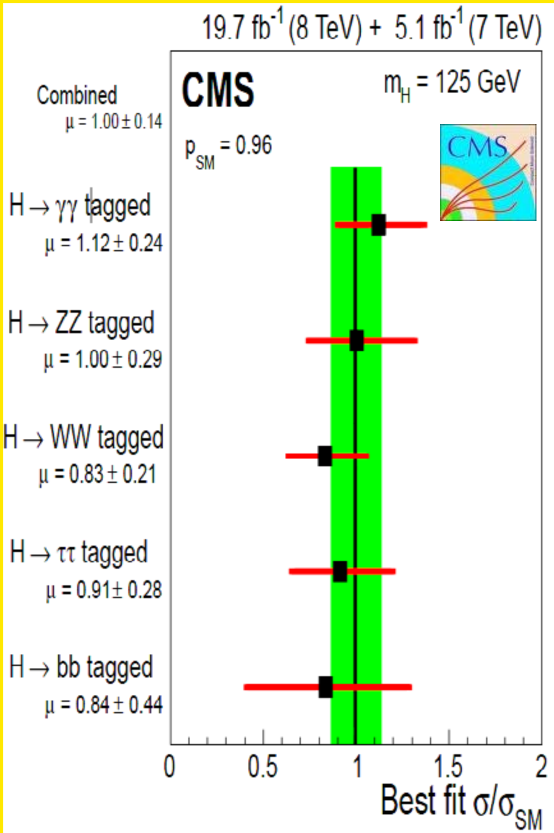
ATLAS and CMS Each Observe a Signal with Local Significance > 5σ

CMS	$\mu(m_H=124.7\text{ GeV}) = 1.14 \pm 0.21\text{ (stat) }^{+0.09}_{-0.05}\text{ (syst)}^{+0.13}_{-0.09}\text{ (theo)}$
ATLAS	$\mu(m_H=125.6\text{ GeV}) = 1.17 \pm 0.23\text{ (stat) }^{+0.10}_{-0.08}\text{ (syst)}^{+0.12}_{-0.08}\text{ (theo)}$

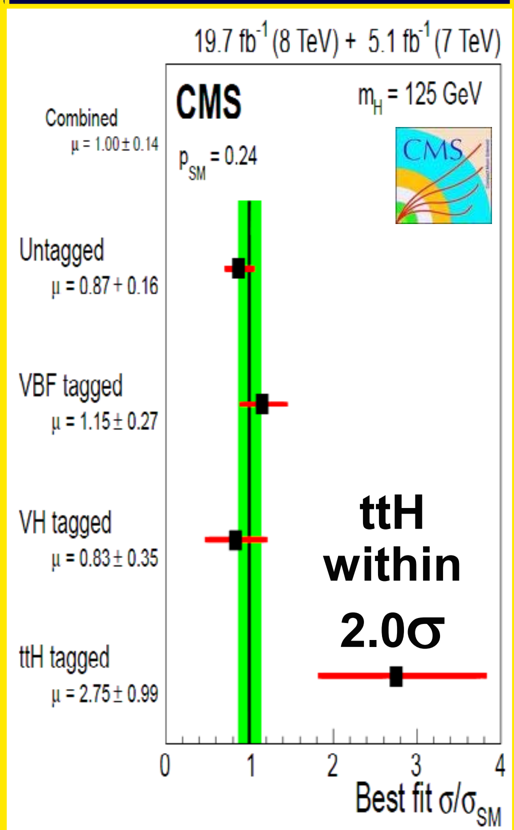
Higgs Signal Strengths $\mu = \sigma/\sigma_{SM}$

Very SM-Like

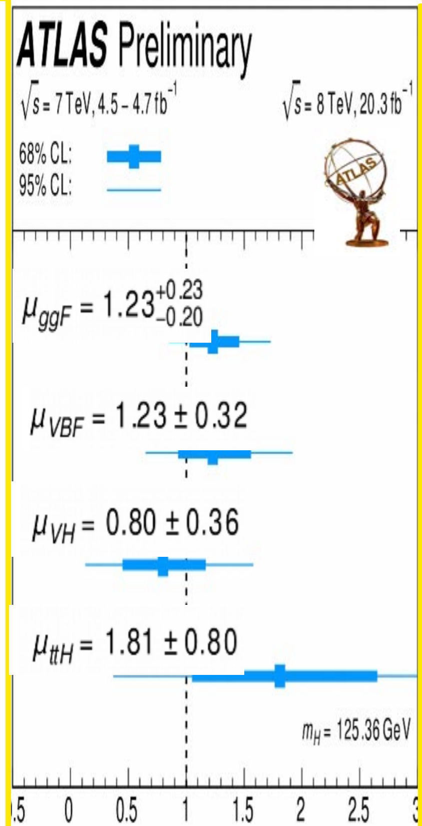
Best Fit σ/σ_{SM} by Decay Mode χ^2/NDF = 1.0/5



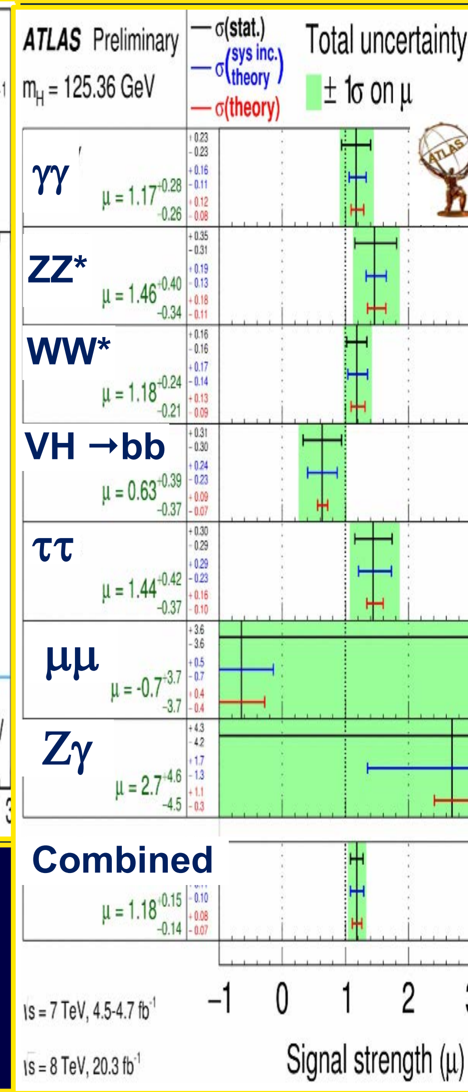
σ/σ_{SM} by Production mode χ^2/NDF = 5.4/4



σ/σ_{SM} by Production Mode



σ/σ_{SM} by Decay Mode



Best-fit overall signal strengths:

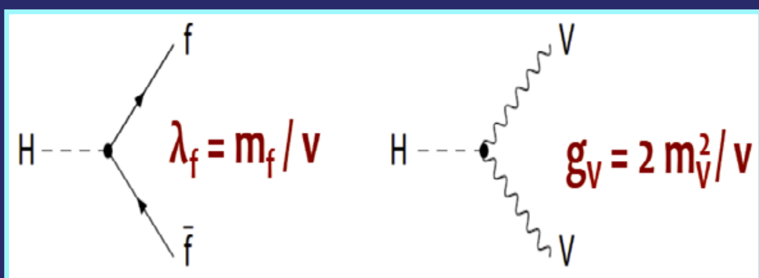
ATLAS $\sigma/\sigma_{SM} = 1.18^{+0.15}_{-0.14}$

CMS: $\sigma/\sigma_{SM} = 1.00 \pm 0.14$

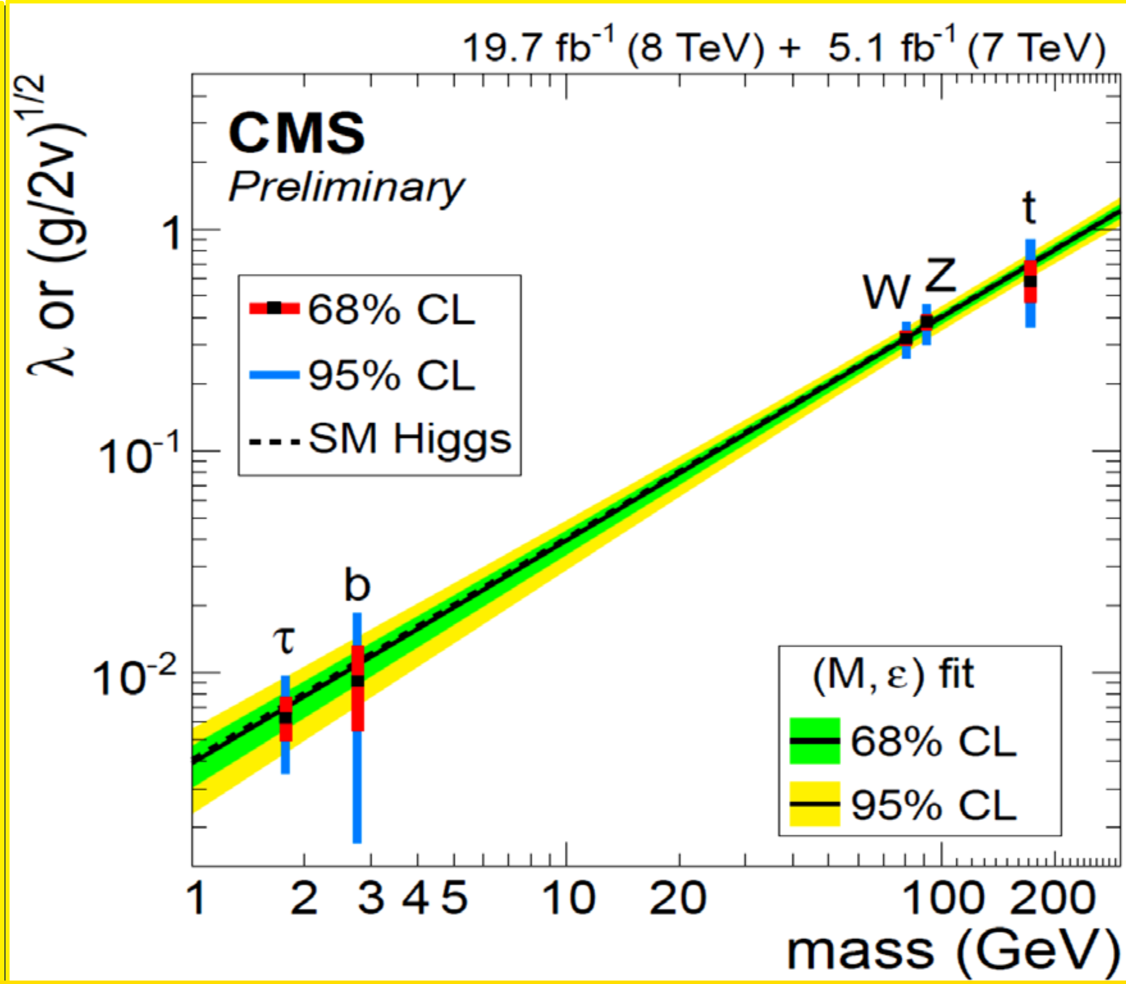
+ H → SUSY H, exotics, Portal DM,

The Couplings vs Mass

- We usually say "the Higgs boson couplings are proportional to the mass of the particle"
- More precisely, the Feynman rules are:



- Plot the couplings vs mass using λ_f and $\mathrm{sqrt}(g_v/2v)$



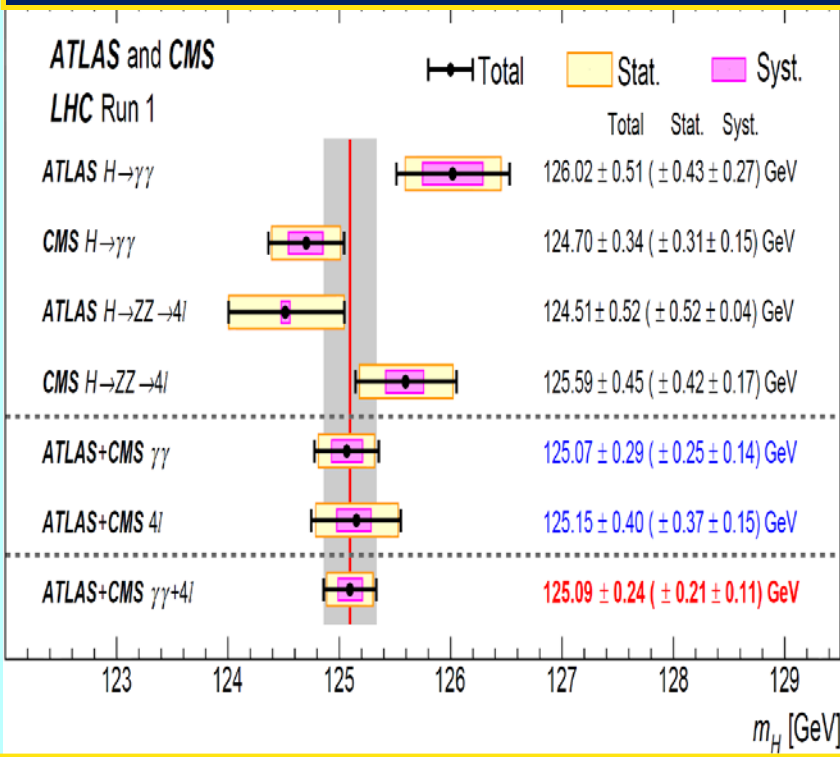
More Data is needed to make precise determinations
Especially for the Fermions: *b, t, τ*

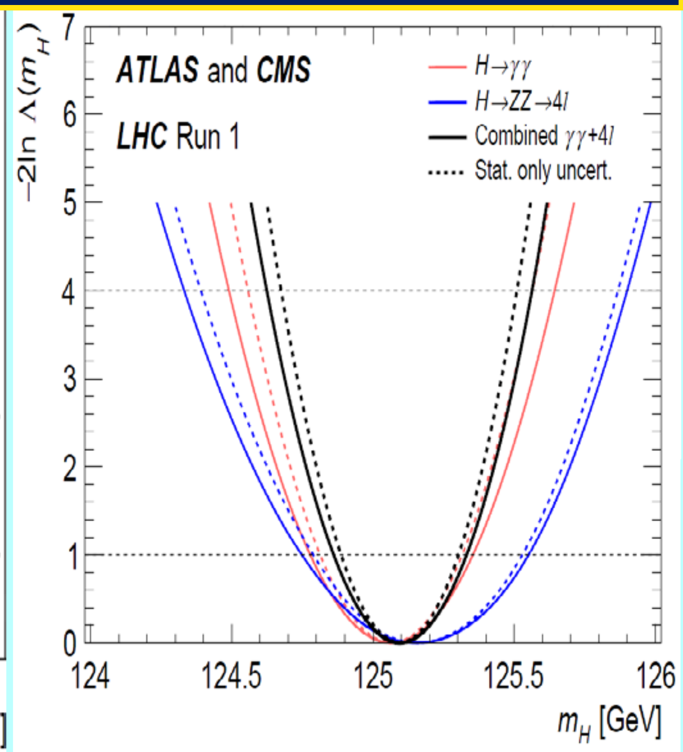
Combined Mass Measurement
from $H \to ZZ \to 4\ell$, $H \to \gamma\gamma$

arXiv:1503.7589

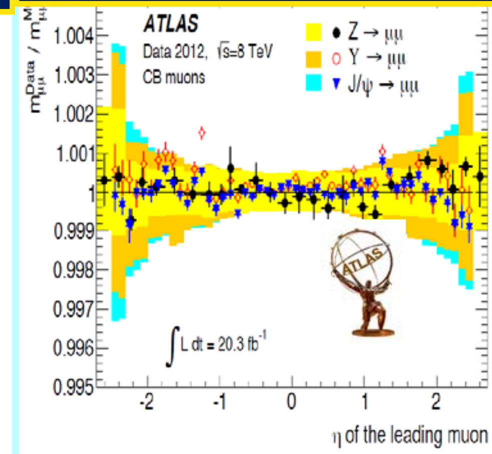
Profiling M_H; μ(ggH,ttH) and μ(VBF, VH) for $\gamma\gamma$; $\mu(4\ell)$ for ZZ

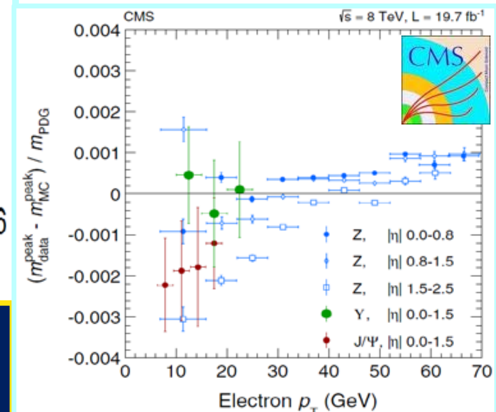
Mass Determined to 0.2% Accuracy !





Calibration with Z, Υ, J/ψ → ee, μμ





Improvement on syst. uncertainties

- final e, γ, μ calibrations
- final detector simulation

Impressive ±0.2% accuracy:

Statistical uncertainty dominates

M_H Values (ATLAS+CMS)

$H \to \gamma\gamma$: $125.07 \pm 0.25 \pm 0.14$

$H \to 4l$: $125.15 \pm 0.37 \pm 0.15$

Combined Channels:

$M_H = 125.09 \pm 0.21 \pm 0.11$

Impressive

0.1 – 0.3% Mass Scale Accuracy

Anomalous CP Couplings of a Spin 1 or 2 Higgs Using H → V(*)V(*) (V= Z, W, γ) Decays

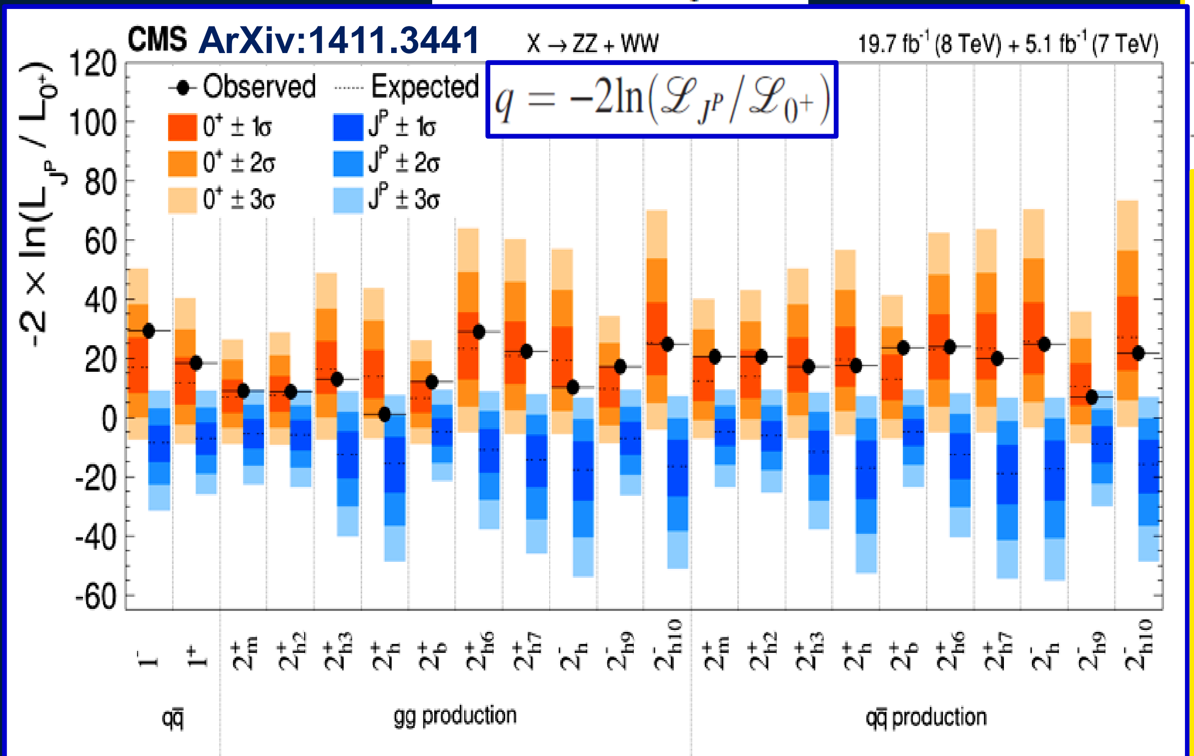
Effective Amplitude Parametrization

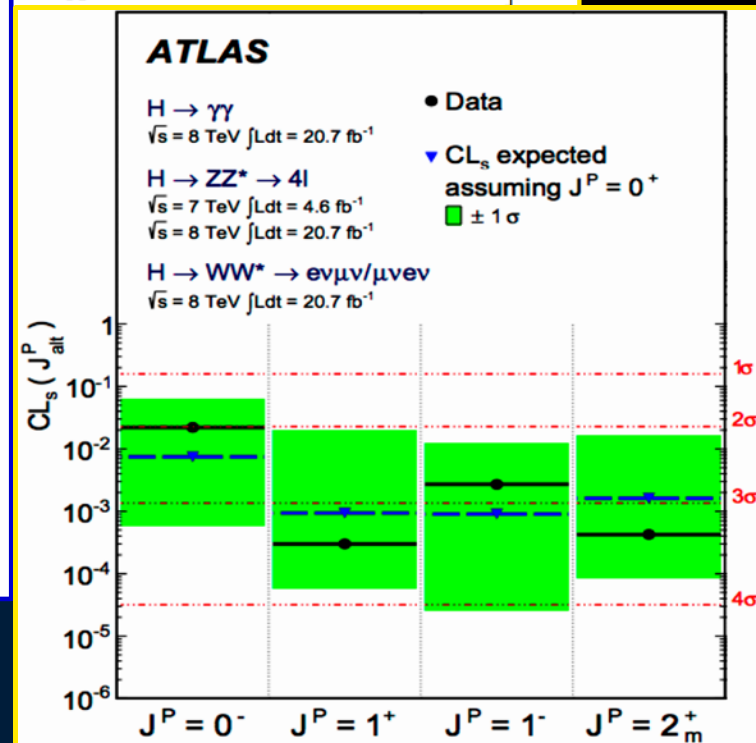
SPIN 1

$$A(X_{J=1} \to V_1 V_2) \sim b_1 \left[(\epsilon^*_{V_1} q)(\epsilon^*_{V_2} \epsilon_X) + (\epsilon^*_{V_2} q)(\epsilon^*_{V_1} \epsilon_X) \right] + b_2 \epsilon_{\alpha\mu\nu\beta} \epsilon^{\alpha}_{X} \epsilon^{*\mu}_{V_1} \epsilon^{*\nu}_{V_2} \tilde{q}^{\beta}$$

SPIN 2

$$\begin{aligned} A(X_{J=2} \to V_1 V_2) \sim \Lambda^{-1} \Big[& 2c_1 t_{\mu\nu} f^{*1,\mu\alpha} f^{*2,\nu\alpha} + 2c_2 t_{\mu\nu} \frac{q_\alpha q_\beta}{\Lambda^2} f^{*1,\mu\alpha} f^{*2,\nu\beta} \\ & + c_3 \frac{\tilde{q}^\beta \tilde{q}^\alpha}{\Lambda^2} t_{\beta\nu} (f^{*1,\mu\nu} f^{*2}_{\mu\alpha} + f^{*2,\mu\nu} f^{*1}_{\mu\alpha}) + c_4 \frac{\tilde{q}^\nu \tilde{q}^\mu}{\Lambda^2} t_{\mu\nu} f^{*1,\alpha\beta} f^{*(2)}_{\alpha\beta} \\ & + m_V^2 \left(2c_5 t_{\mu\nu} \epsilon^{*\mu}_{V_1} \epsilon^{*\nu}_{V_2} + 2c_6 \frac{\tilde{q}^\mu q_\alpha}{\Lambda^2} t_{\mu\nu} (\epsilon^{*\nu}_{V_1} \epsilon^{*\alpha}_{V_2} - \epsilon^{*\alpha}_{V_1} \epsilon^{*\nu}_{V_2}) + c_7 \frac{\tilde{q}^\mu \tilde{q}^\nu}{\Lambda^2} t_{\mu\nu} \epsilon^*_{V_1} \epsilon^*_{V_2} \right) \\ & + c_8 \frac{\tilde{q}^\mu \tilde{q}^\nu}{\Lambda^2} t_{\mu\nu} f^{*1,\alpha\beta} \tilde{f}^{*(2)}_{\alpha\beta} + c_9 t^{\mu\alpha} \tilde{q}_\alpha \epsilon_{\mu\nu\rho\sigma} \epsilon^{*\nu}_{V_1} \epsilon^{*\rho}_{V_2} q^\sigma \\ & + \frac{c_{10} t^{\mu\alpha} \tilde{q}_\alpha}{\Lambda^2} \epsilon_{\mu\nu\rho\sigma} q^\rho \tilde{q}^\sigma \left(\epsilon^{*\nu}_{V_1} (q \epsilon^*_{V_2}) + \epsilon^{*\nu}_{V_2} (q \epsilon^*_{V_1}) \right) \Big], \end{aligned}$$





CMS: J^P Values: many models Other than 0+ Ruled out with ≥ 4σ significance

ATLAS : results on spin/parity (using *H→γγ, ZZ and WW*) Also favor 0+

Where We Go from Here: Precision Measurements of Higgs Spin & CP Properties

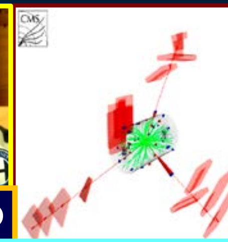
http://dx.doi.org/10.1007/JHEP01(2013)182, CALT-68-2894

Use 4 lepton production & decay configurations to Probe the tensor structure (J, CP) of the couplings of the new particle to elwk gauge bosons ZZ, Zγ, γγ

Using the full Lagrangian

For $H \rightarrow 4l$ and $q\bar{q} \rightarrow 4l$

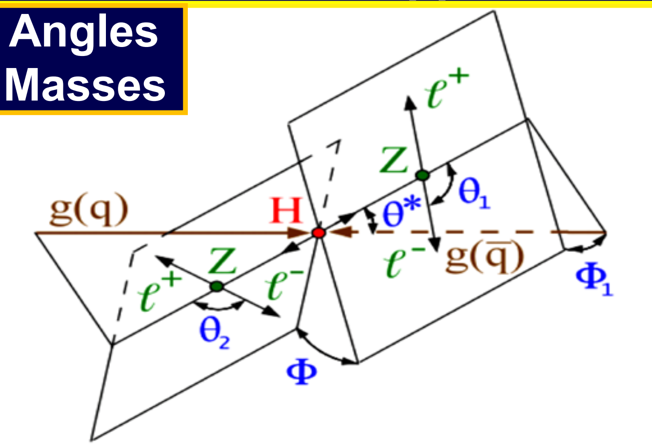
Chen Xie DiMarco

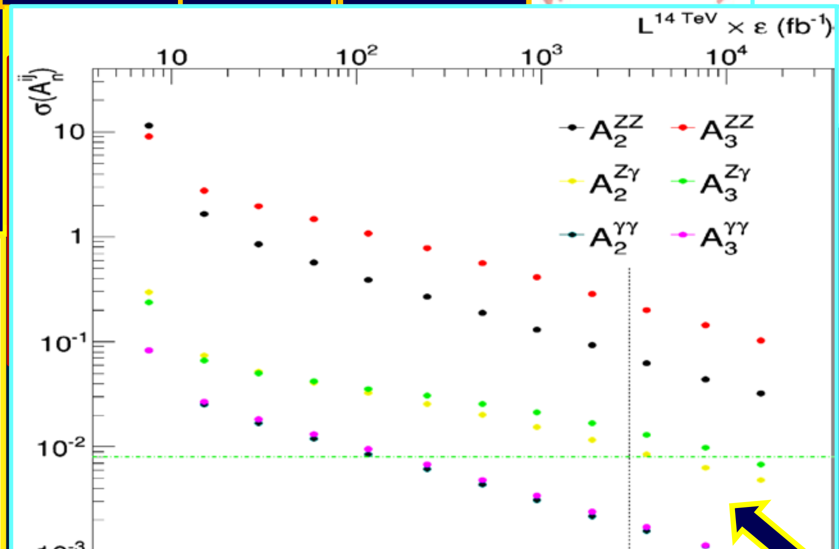


$$\mathcal{L} \sim \frac{1}{v}\varphi\Big(g_h m_Z^2 Z^\mu Z_\mu + g_Z Z^{\mu\nu} Z_{\mu\nu} + \tilde{g}_Z Z^{\mu\nu} \tilde{Z}_{\mu\nu} + g_{Z\gamma} F^{\mu\nu} Z_{\mu\nu} + \tilde{g}_{Z\gamma} F^{\mu\nu} \tilde{Z}_{\mu\nu} + g_\gamma F^{\mu\nu} F_{\mu\nu} + \tilde{g}_\gamma F^{\mu\nu} \tilde{F}_{\mu\nu} + \ldots\Big)$$

$$i\Gamma_{ij}^{\mu\nu} = v^{-1}\Big(A_{1ij} m_Z^2 g^{\mu\nu} + A_{2ij}(k_1 \cdot k_2 g^{\mu\nu} - k_1^\nu k_2^\mu) + A_{3ij}\epsilon_{\mu\nu\alpha\beta} k_1^\alpha k_2^\beta\Big)$$

5 Angles
3 Masses





Sample the full space including detector effects

Quantitative In-Depth Study

Very Computation - ally Intensive 8D Calculations

Made Tractable for the First Time by Yi Chen

Remarkable γγ and Zγ sensitivity

Key Challenge to Tackle with Machine Learning

- ***Reduce computing time***
- ***Release simplifying assumptions***
- ***Bring out more subtle effects***

Yi Chen Thesis: **CP Odd & Even γγ Couplings can be probed to 1% Level Within the First 300-400/fb at the LHC !**

Anomalous CP Couplings of a Spin 0 Higgs Using H → $V^{(*)}V^{(*)}$ (V= Z, W, γ) Decays

Effective Amplitude Parametrization

- a_1: SM CP-even coupling
- Λ_1: BSM Scale (GeV)
- a_2 (a_3): CP even (odd) anomalous Couplings
- Results in cross section fractions f, phases ϕ

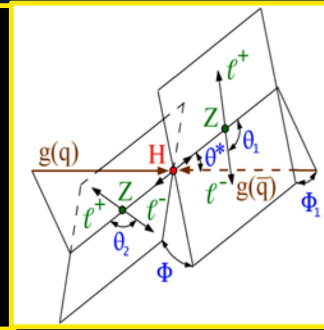
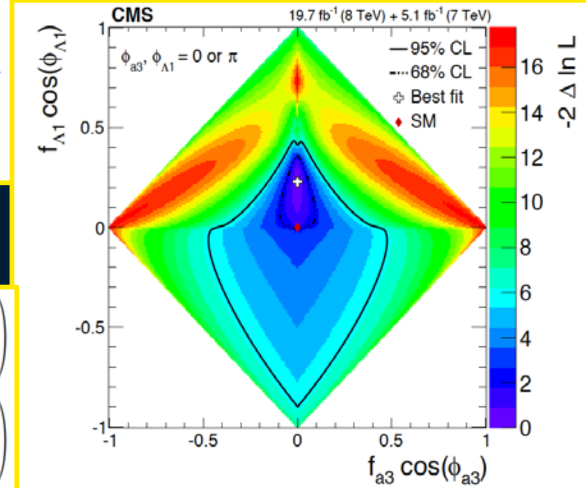


Illustration: $f_{\Lambda 1}$ vs f_{a3} constraints with ϕs 0, π



Starting with Spin-0 :

SM HZZ decay (a1)

Leading momentum dependent correction

$$A(X_{J=0} \to V_1 V_2) = v^{-1}\left(\left[a_1 - e^{i\phi_{\Lambda_1}}\frac{q^2_{Z_1}+q^2_{Z_2}}{(\Lambda_1)^2}\right] m_Z^2 \epsilon^*_{Z_1}\epsilon^*_{Z_2}\right.$$
$$+ a_2 f^{*(Z)}_{\mu\nu} f^{*(Z),\mu\nu} + a_3 f^{*(Z)}_{\mu\nu}\tilde{f}^{*(Z),\mu\nu}$$
$$+ a_2^{Z\gamma} f^{*(Z)}_{\mu\nu} f^{*(\gamma),\mu\nu} + a_3^{Z\gamma} f^{*(Z)}_{\mu\nu}\tilde{f}^{*(\gamma),\mu\nu}$$
$$\left. + a_2^{\gamma\gamma} f^{*(\gamma)}_{\mu\nu} f^{*(\gamma),\mu\nu} + a_3^{\gamma\gamma} f^{*(\gamma)}_{\mu\nu}\tilde{f}^{*(\gamma),\mu\nu}\right),$$

non-SM scalar (a2)

Pseudo-scalars (a3)

SM Zγ

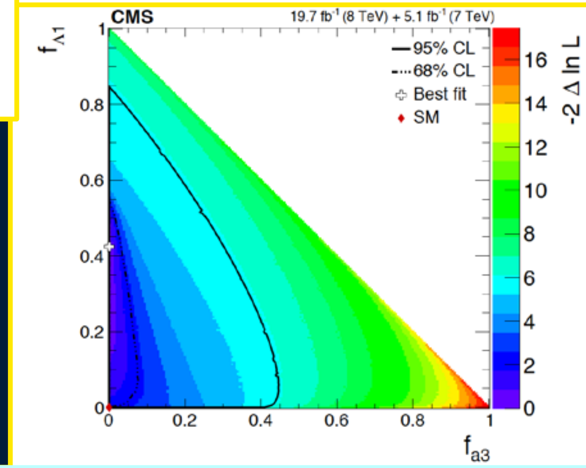
SM $\gamma\gamma$

ArXiv:1411.3441

$$f_{a3} = \frac{|a_3|^2\sigma_3}{|a_1|^2\sigma_1+|a_2|^2\sigma_2+|a_3|^2\sigma_3+\tilde{\sigma}_{\Lambda_1}/(\Lambda_1)^4} \qquad \phi_{a3} = \arg\left(\frac{a_3}{a_1}\right)$$

$$f_{a2} = \frac{|a_2|^2\sigma_2}{|a_1|^2\sigma_1+|a_2|^2\sigma_2+|a_3|^2\sigma_3+\tilde{\sigma}_{\Lambda_1}/(\Lambda_1)^4} \qquad \phi_{a2} = \arg\left(\frac{a_2}{a_1}\right)$$

$$f_{\Lambda 1} = \frac{\tilde{\sigma}_{\Lambda_1}/(\Lambda_1)^4}{|a_1|^2\sigma_1+|a_2|^2\sigma_2+|a_3|^2\sigma_3+\tilde{\sigma}_{\Lambda_1}/(\Lambda_1)^4} \qquad \phi_{\Lambda 1},$$



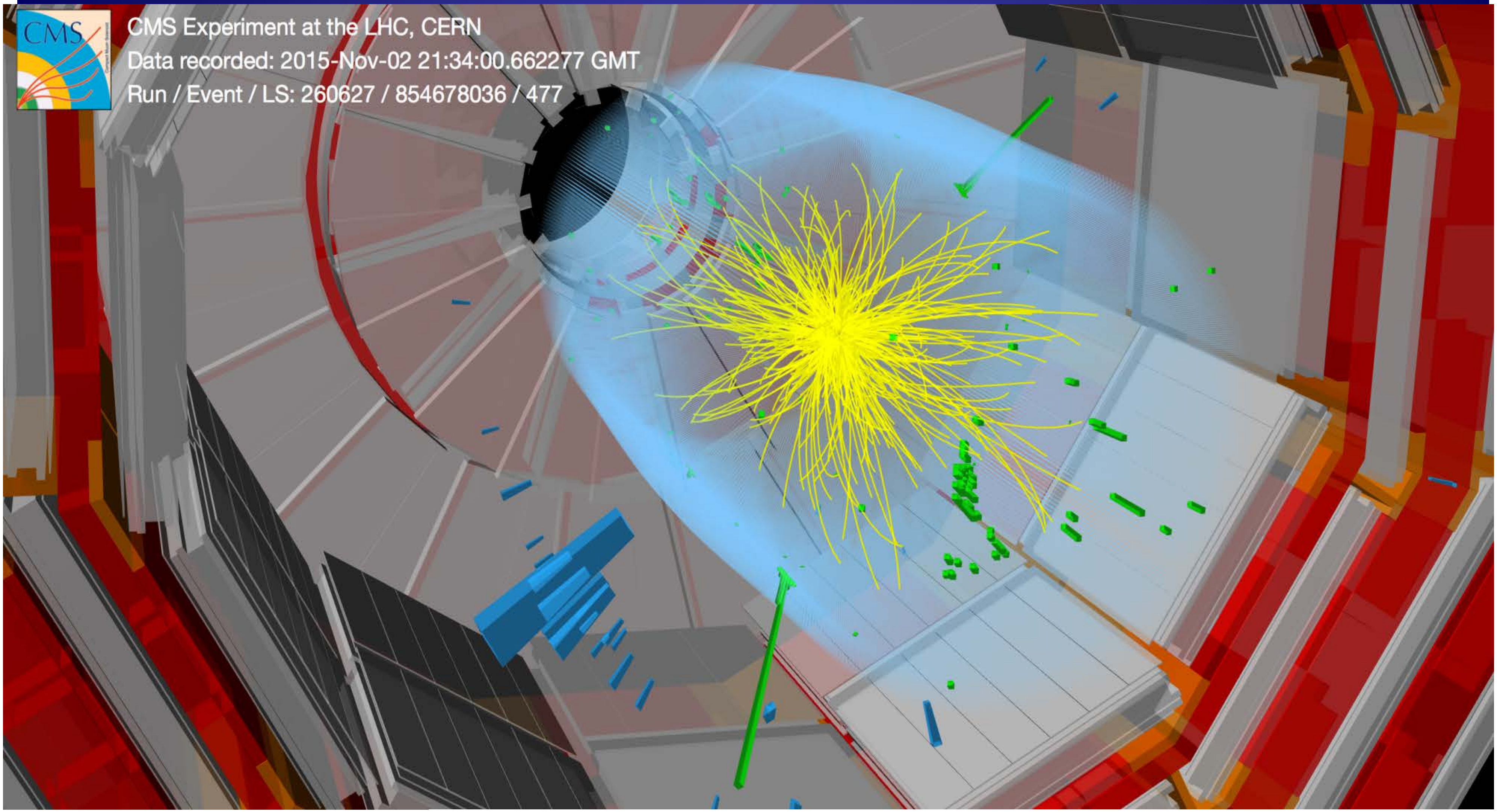
H → Z(γ^*)Z(γ^*) → 4ℓ : Full 8D phase space: (5 angles, M_{Z1}, M_{Z2}, M_{4l})

H → WW → $\ell \nu \ell \nu$: dilepton mass and transverse masses constrain the fractions

Scenarios: Real phases; floating phases

Best Fit Results very close to SM expectations

Search for diphoton resonances



Diphoton event with m(γγ) = 745 GeV

So Far: No (Clear) Signs of New Massive Particles

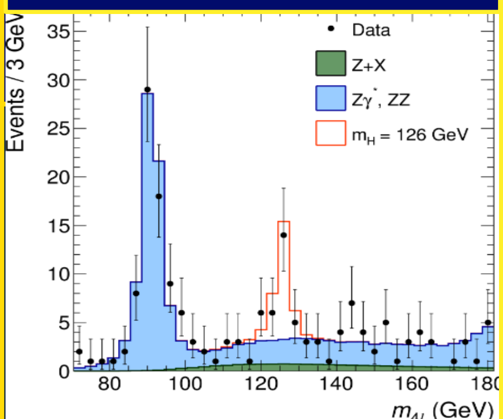
- ❒ 500 Publications using data from 7- and 8- TeV pp and heavy ion collisions
- ❒ Discovery of the Higgs boson, a new baryon, $\Xi^{0*}{}_b$.
  Scores of other results
- ❒ Extension of lower limits on the mass of new particles
- ❒ New Limits on the rate of rare new phenomena

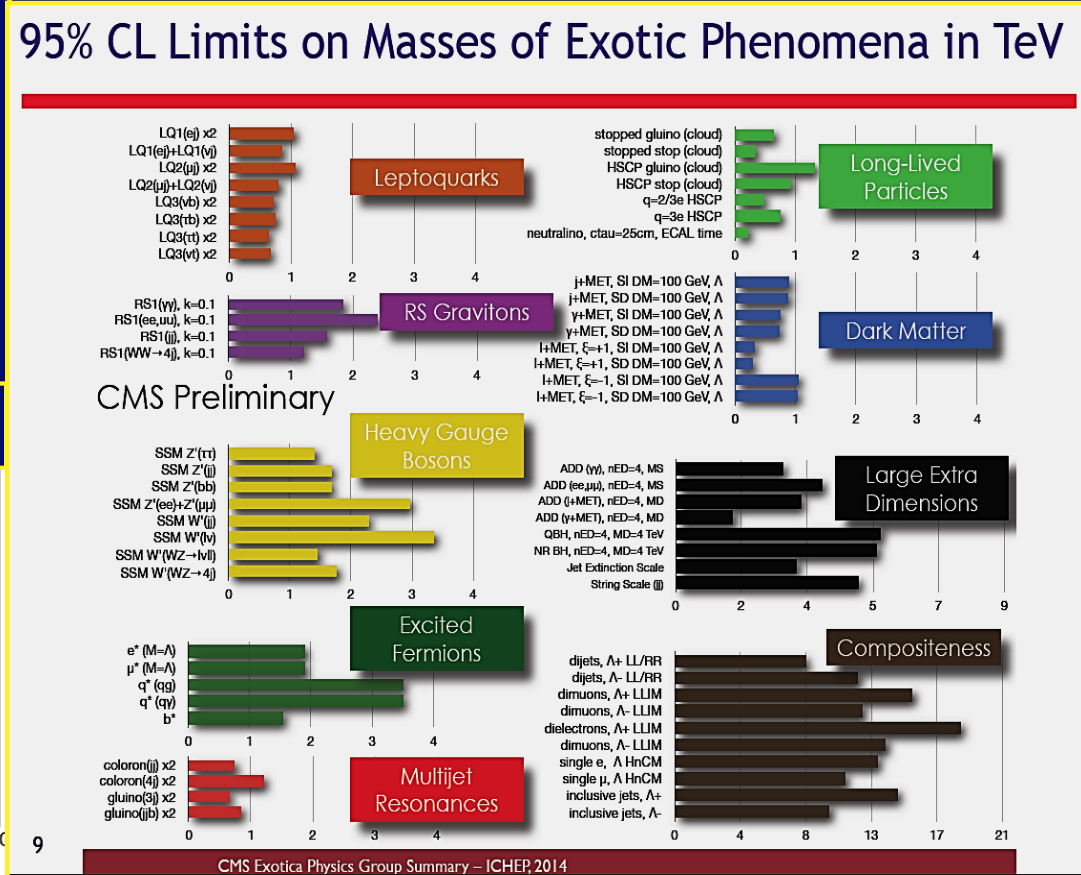
But for New Particles and Phenomena:

- ❒ *We Do Not Know* what we will find
- ❒ A Vast target for Unsupervised ML-Driven, model-independent searches

- ❒ Higgs Channels:
- ❒ We *Knew* what these would look like

$pp \to H \to ZZ^*$





All CMS pubs: http://cms-results.web.cern.ch/cms-results/public-results/publications/

The Outlook

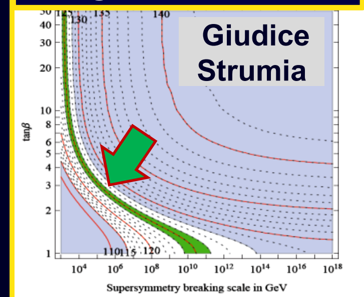
- **SM or not:** the 125 GeV Higgs boson has taken us to the threshold of an era of new physics, with a host of questions
- Natural, Split or High Scale SUSY ?:
  - A nearby 3rd generation at <~1 TeV ?
  - Another nearby scale at ~5-50 TeV ?
- *OR: new singlets, doublets, triplets; new scalars, vectors, composites, extra dim. ?*
- Vacuum (meta)stability ➡ Another new scale at $\sim 10^{10\text{-}12}$ GeV ?
- Neutrino masses (via seesaws or RH ν): A "similar" intermediate scale ?
- *The Discovery has Expanded our Vision*

➡ Run2 : a new horizon to explore and test our ideas: on EWSB and beyond

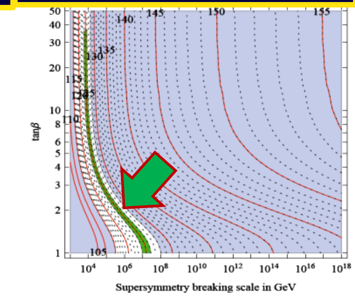
Apologies for all I could not cover

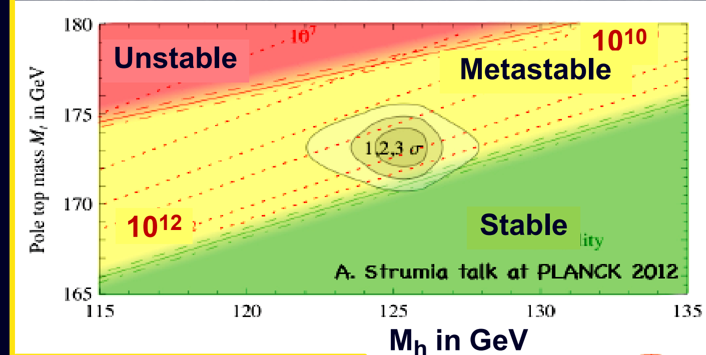
$$M_h^2 \overset{M_A \gg M_Z}{\approx} M_Z^2 \cos^2 2\beta + \frac{3m_t^4}{2\pi^2 v^2}\left[\log\frac{M_S^2}{m_t^2} + \frac{X_t^2}{M_S^2}\left(1 - \frac{X_t^2}{12 M_S^2}\right)\right]$$

High Scale SUSY

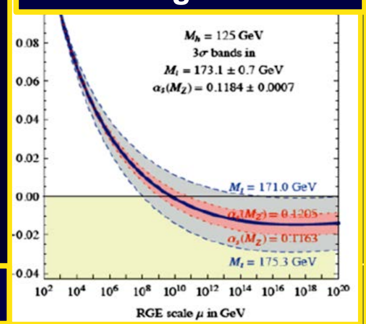


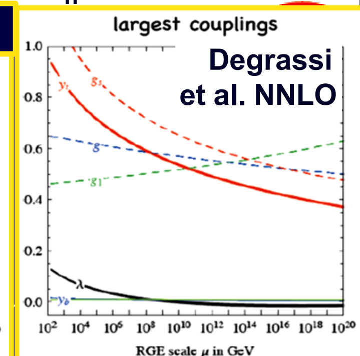
Split SUSY





λ at High Scales





Statistics: Computing Limits
for the Higgs Search

CMS uses the CL_s method to set limits on $\mu = \sigma/\sigma_{SM}$

- Frequentist approach including systematic error evaluation

Likelihood function:

$$\mathcal{L}(data\,|\,\mu,\theta) = \text{Poisson}\,\big(\overbrace{data}^{\text{Observed}}\,\big|\,\underbrace{\mu\cdot s(\theta)+b(\theta)}_{\text{Expected S+B}}\big)\cdot\overbrace{p(\tilde{\theta}|\theta)}^{\text{Systematics}}$$

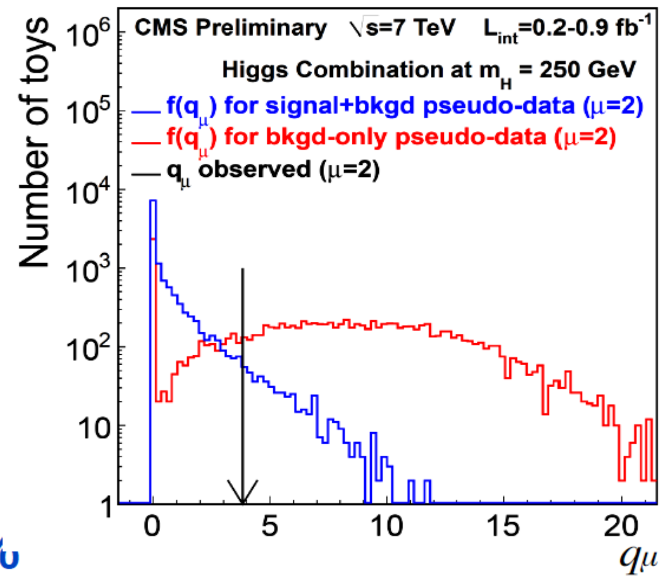
Test statistics:

$$q_\mu = -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}\begin{array}{l}\leftarrow \text{fix } \mu\text{, vary } \hat{\theta}_\mu \\ \leftarrow \text{vary } \hat{\mu} \text{ and } \hat{\hat{\theta}} \\ \quad 0 \le \hat{\mu} \le \mu\end{array}$$

Finally, calculate CL_s (toy MC):

$$CL_s = \frac{P\left(q_\mu \ge q_\mu^{obs}\,|\,\mu s(\hat{\theta}_\mu^{obs}) + b(\hat{\theta}_\mu^{obs})\right)}{P\left(q_\mu \ge q_\mu^{obs}\,|\,b(\hat{\theta}_0^{obs})\right)}$$

95% C.L. is on μ value giving CL_s = 1 - 95%



CERN-CMS Note-2011-005: Procedure for the LHC Higgs Boson Search Combination in Summer 2011

Statistics: Computing Significance for the Higgs Search

To quantify observed excess (above background only hypothesis)

- Same machinery as on previous slide but to test probability of the null hypothesis

Approximate p-value (probability of the null hypothesis):

$$\tilde{p} = \frac{1}{2}\left[1 - \mathrm{erf}\left(\sqrt{q_0^{\mathrm{obs}}/2}\right)\right]$$

where q_0^{obs} is the observed q_μ value for the null hypothesis ($\mu = 0$)

Significance (Z) corresponding to p-value

$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\, dx$$

Probability expressed in σ's of one-sided normal distribution.