Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

Outlook on LHC Run 2

Oliver Gutsche

CMS workshop at OSG All Hands Meeting 2015

23 March 2015

LHC Run 1 - The Higgs and a lot more

The Nobel Prize in Physics 2013
François Englert, Peter Higgs

The Nobel Prize in Physics 2013

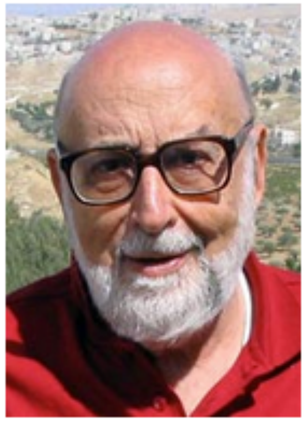


Photo: Pnicolet via Wikimedia Commons

François Englert

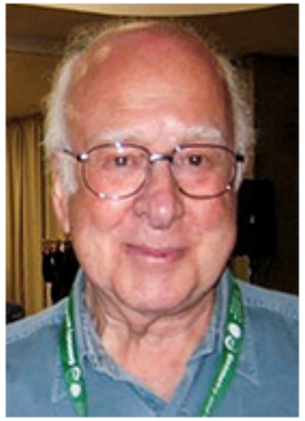
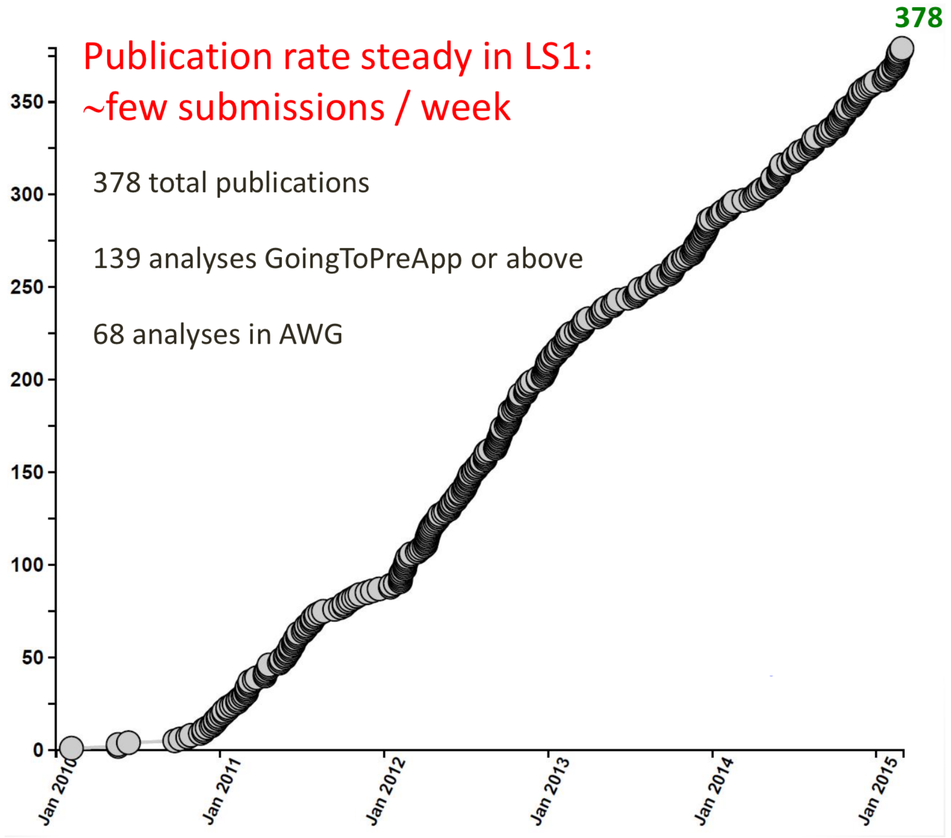


Photo: G-M Greuel via Wikimedia Commons

Peter W. Higgs

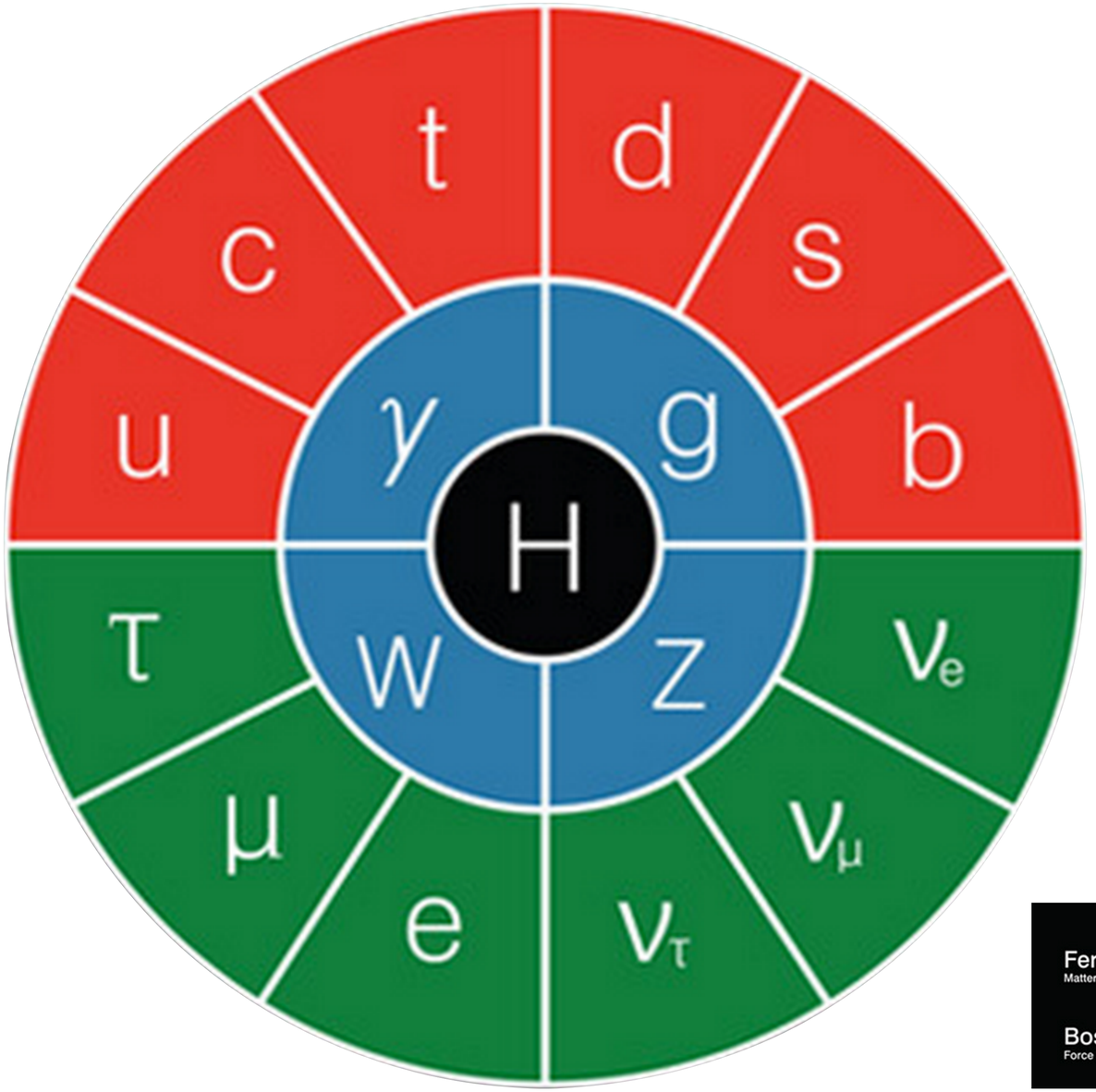
The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs ***"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"***

- CMS and ATLAS discovered the Higgs Boson in 2012
- Over 350 physics results were published on LHC Run 1 data
  - More to come!

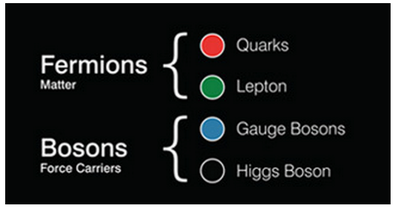


Fermilab

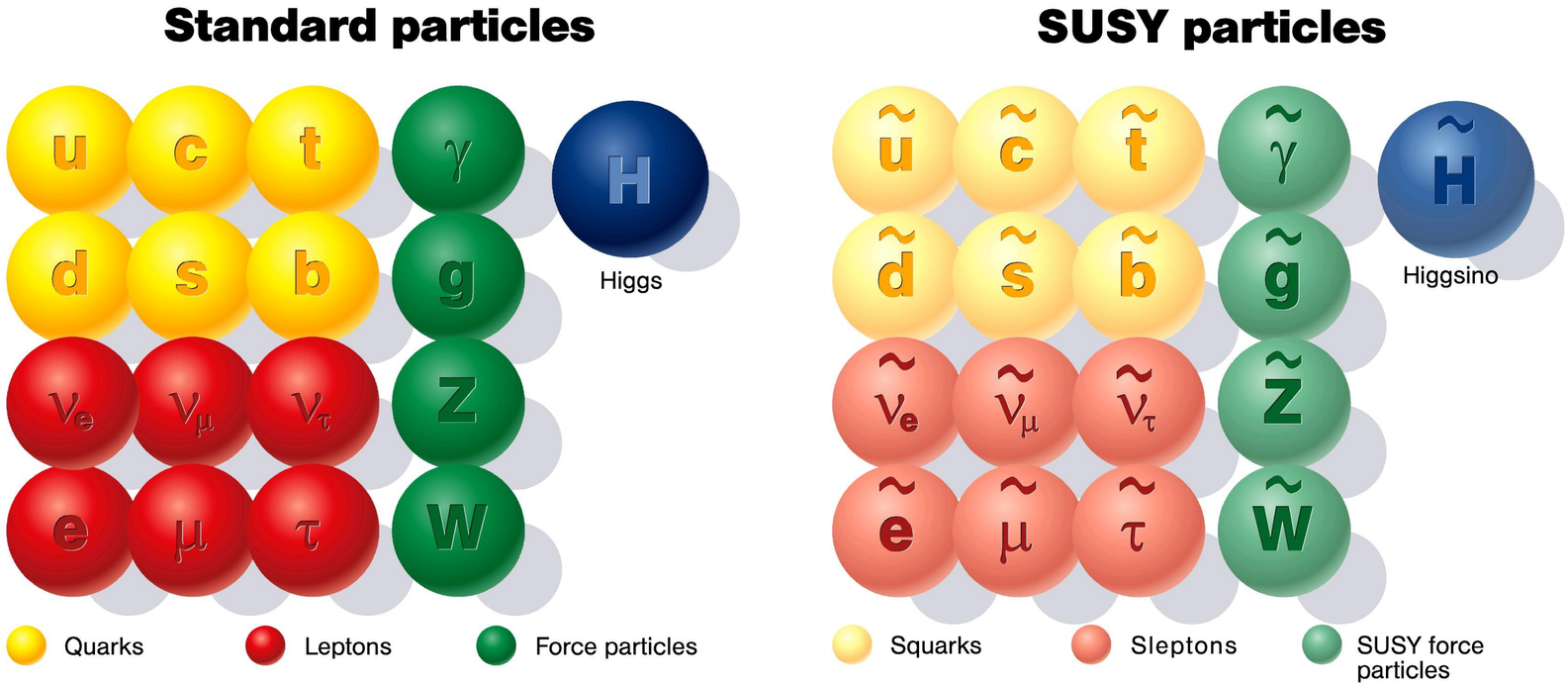
Standard Model of Particle Physics



- What will be “hot” in LHC Run 2
  - **Higgs-Physics**
    - Measuring the properties of the Higgs-Boson
    - Is there only one Higgs-Boson or are there more?

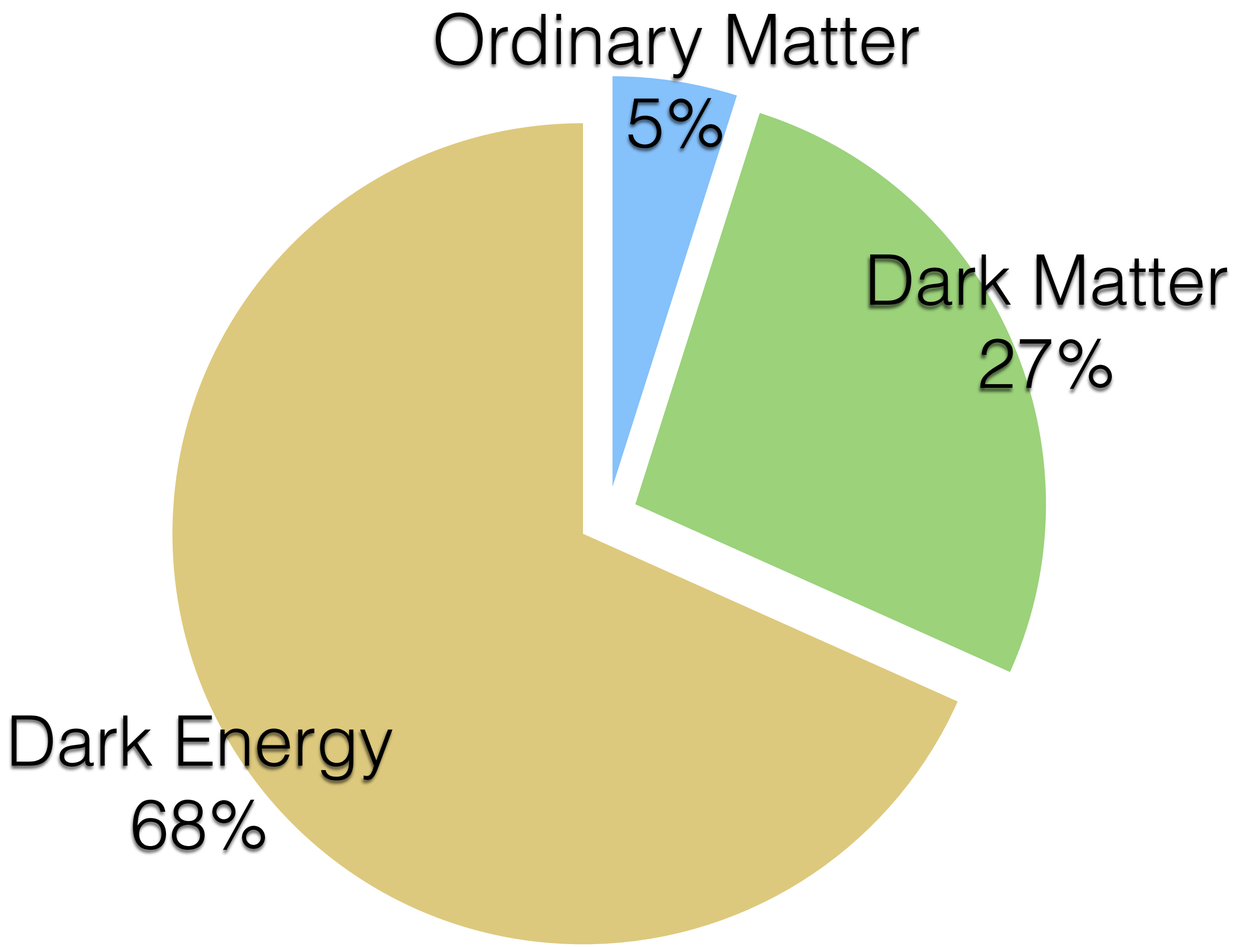


New Physics - Beyond the Standard Model (BSM)



- What will be “hot” in LHC Run 2
  - **Supersymmetry (SUSY)**
    - Double the particle “zoo”
    - Is Supersymmetry real?

New Physics - Beyond the Standard Model (BSM)



- What will be “hot” in LHC Run 2?

  - Dark Matter

    - What is Dark Matter?
    - Can it be explained by a particle?
    - Can we find this particle?

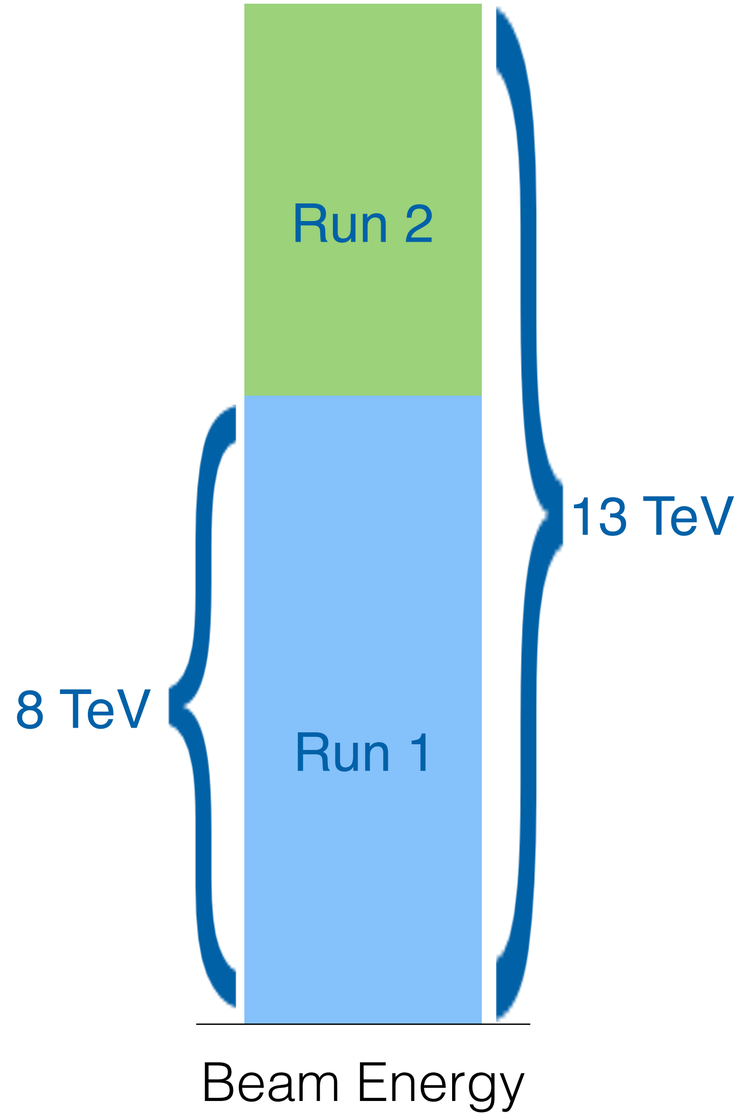
Fermilab

Why are we so excited about LHC Run 2

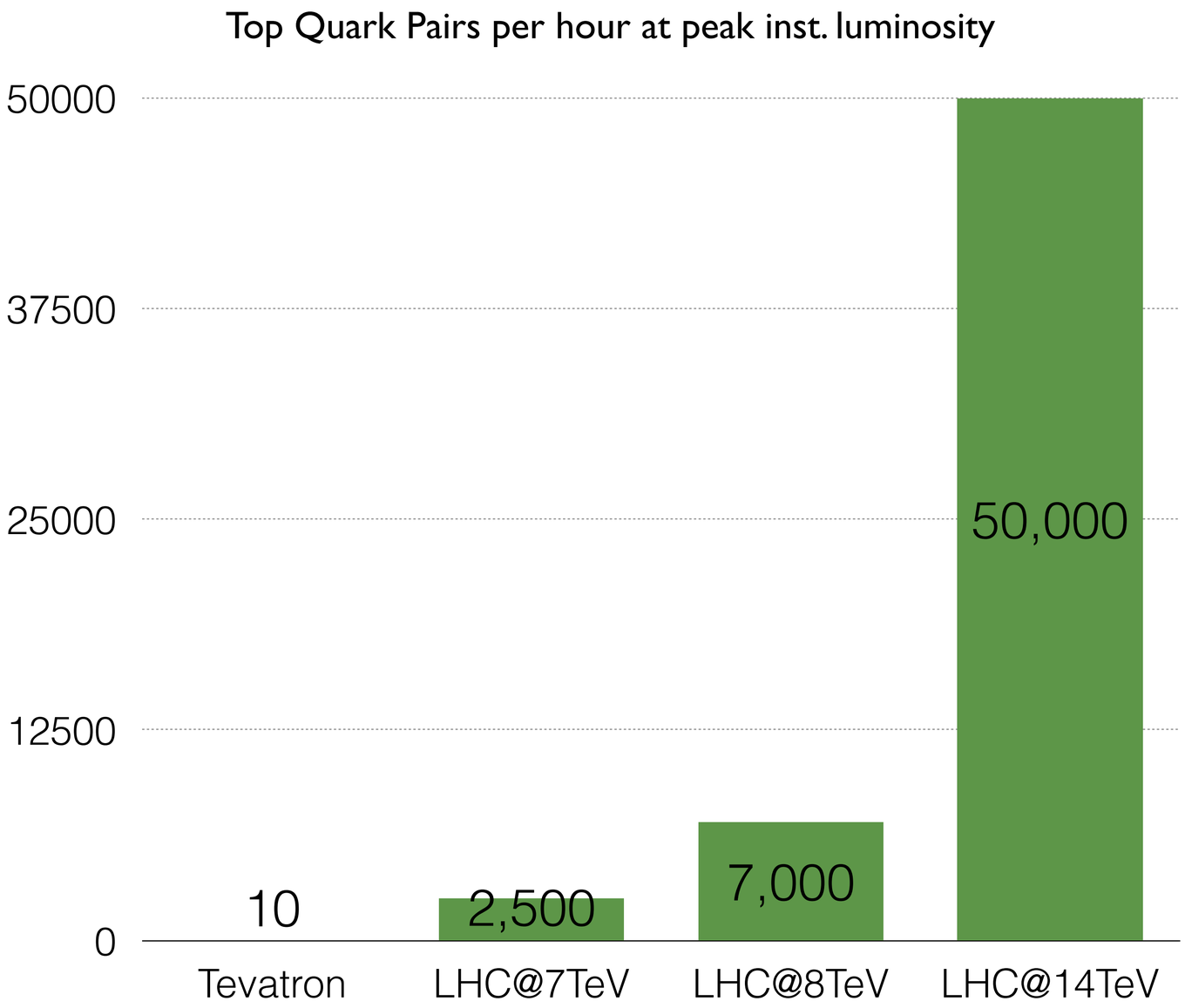
- Energy is increased to 13 TeV!
  - Largest energy increase for many years to come.
  - Reach for new physics increases significantly.

- Unchartered territory
  - In LHC Run 1, we were looking primarily for the Higgs-Boson
  - Previous experiments narrowed down where to look and what to look for
  - Higgs-Discovery was the culmination of a 50-year long search
    - Started with the theoretical prediction of the mechanism by Higgs, Englert et. al.

- LHC Run 2
  - Several candidates of new physics (Supersymmetry, Dark Matter) but no clear indication what to look for and where
  - Excitement comes from searching for the unknown



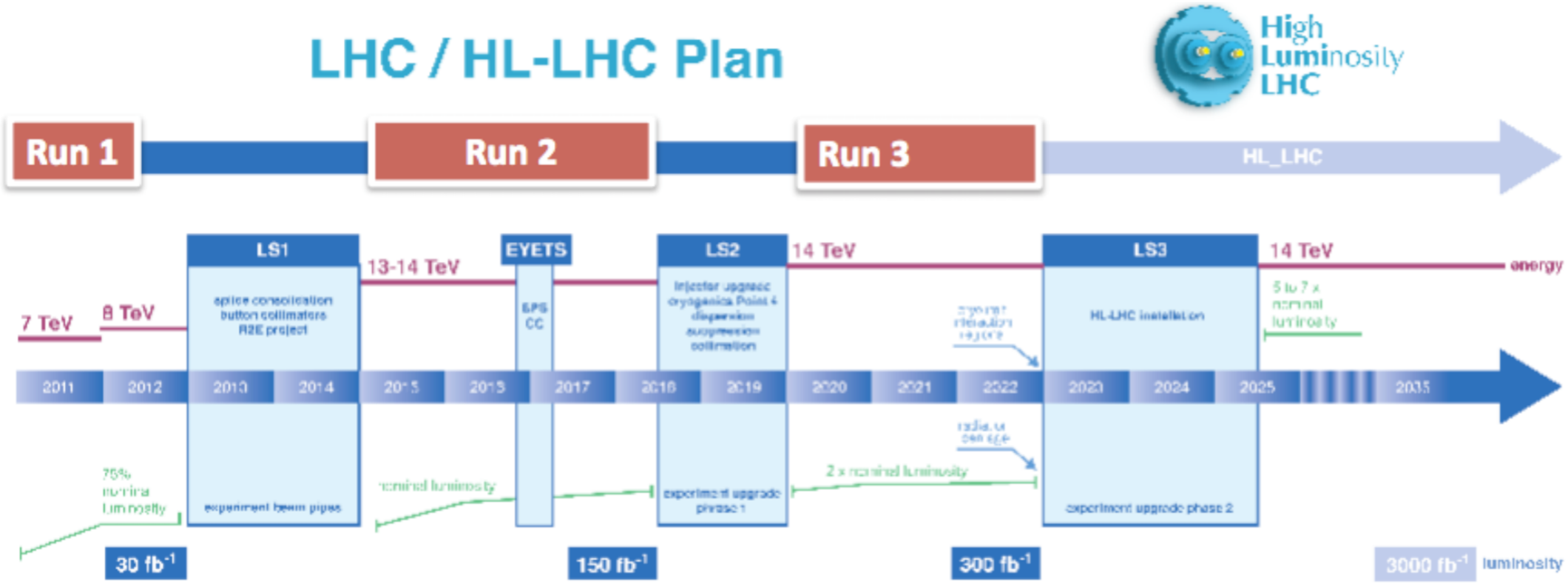
Top quark pair production



- Top quark is the heaviest quark we know
  - Heavier than the Higgs-Boson
- Production rate increase shows the significance of LHC energy boost

cross sections from arXiv:1303.6254: Tevatron ~7pb, LHC@7TeV ~172pb, LHC@8TeV ~246pb, LHC@14TeV ~954pb
peak inst. luminosity: Tevatron: ~$4 \times 10^{32} cm^{-2} s^{-1}$, LHC@7TeV: ~$4 \times 10^{33} cm^{-2} s^{-1}$, LHC@8TeV: ~$8 \times 10^{33} cm^{-2} s^{-1}$, LHC@14TeV: ~$1.5 \times 10^{34} cm^{-2} s^{-1}$ (estimate for 2015)

LHC schedule

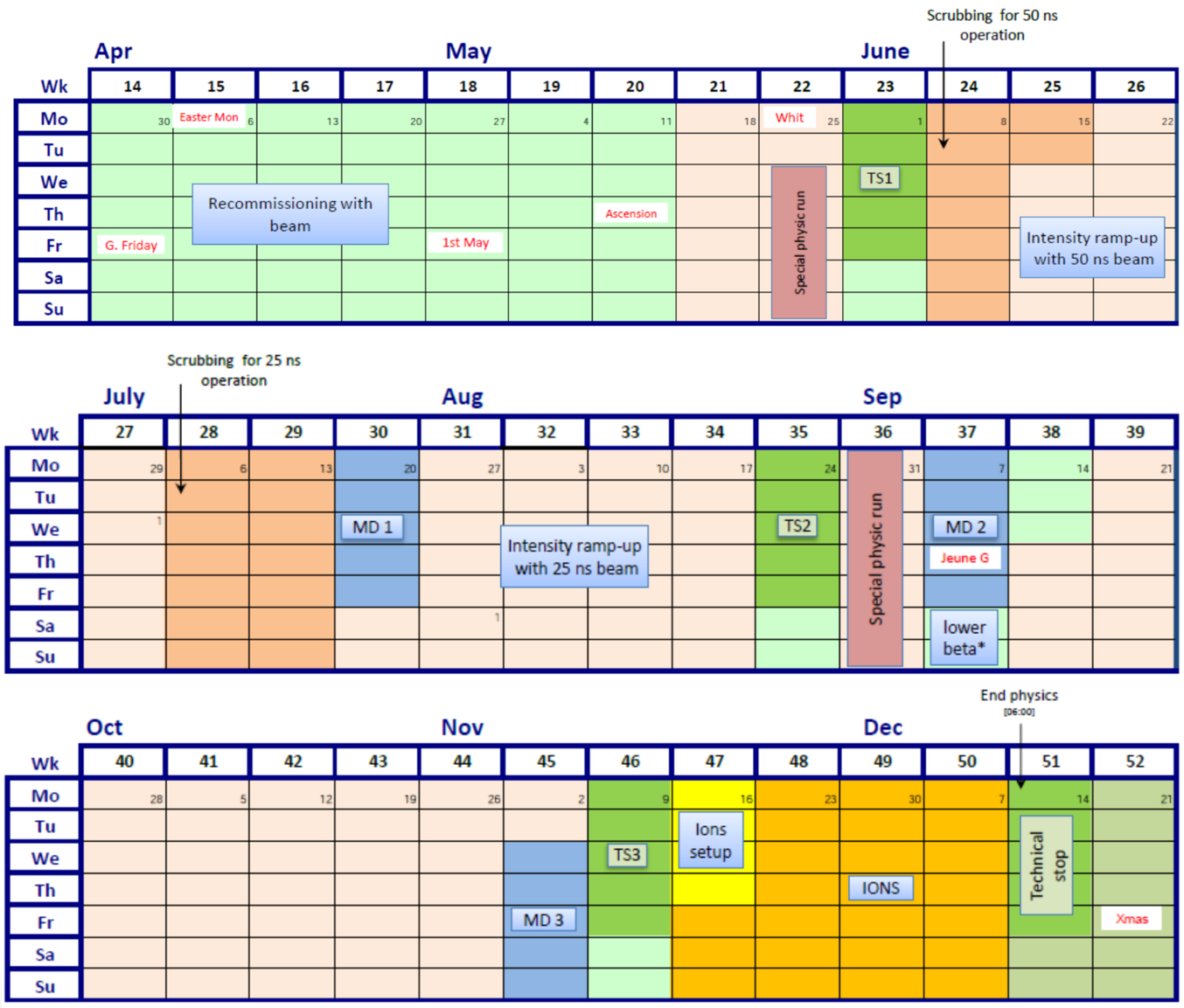


I.Alonso, ECFA 2014

- We are starting LHC Run 2 in 2015!

Fermilab

LHC Run 2 schedule



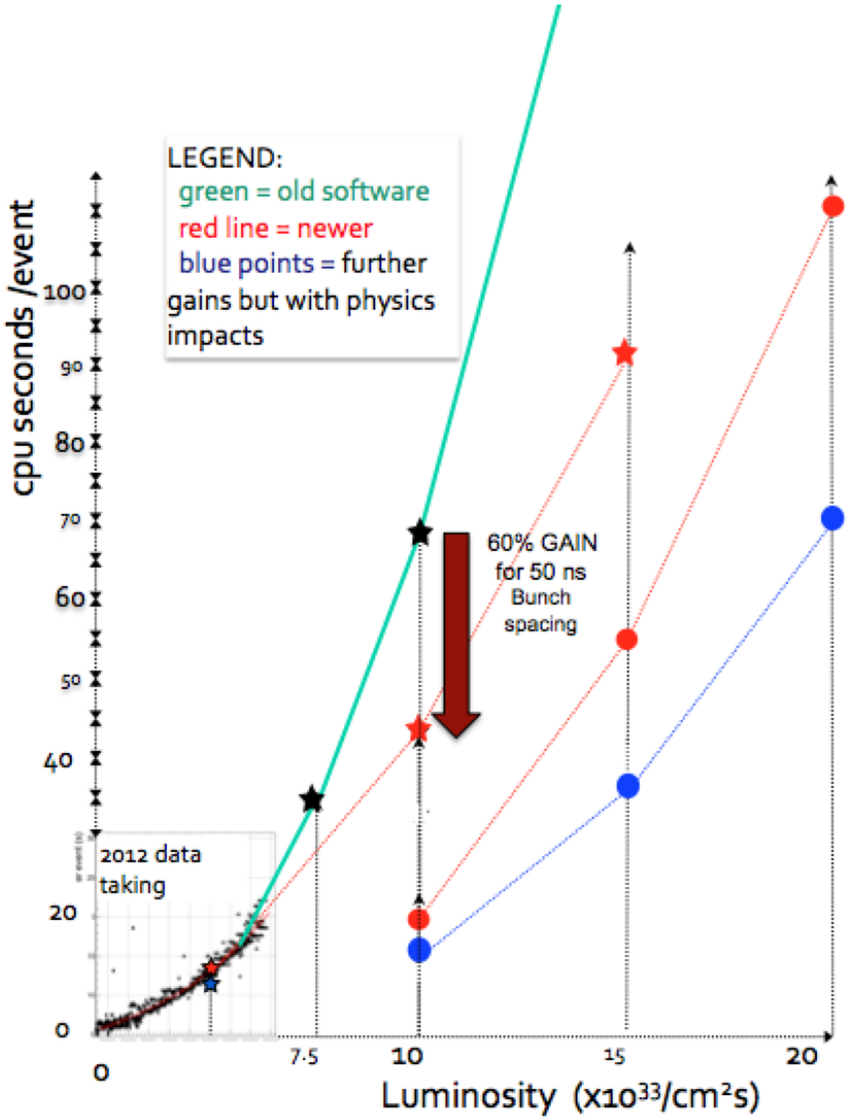
- First beam splashes in CMS expected **March 25th**
- 2015 running periods
  - 50ns running June 17 – July 5 (~1 fb^{-1})
  - 25ns running (“phase 1”) July 26 – August 23 (~3 fb^{-1})
  - 25ns running (“phase 2”) Sept 17 – Nov 3 (~7 fb-1)
- Most of the 2015 data will be recorded late in the year

Numbers

- Reconstruction time increases by factor ~2.5
  - Mainly due to the increase in number of PileUp events, average ~35 PU interactions per collision
- Trigger rate increases by factor ~2.5 to 1 kHz
  - Required to continue the study of the properties of the light Higgs boson discovered in LHC run 1, and the searches for new physics at the new energy frontier
- Take performance of 2012 ➜ CMS would need factor 6 in CPU resources
  - Only getting ~factor 2 ➜ We need to be more efficient and more clever!

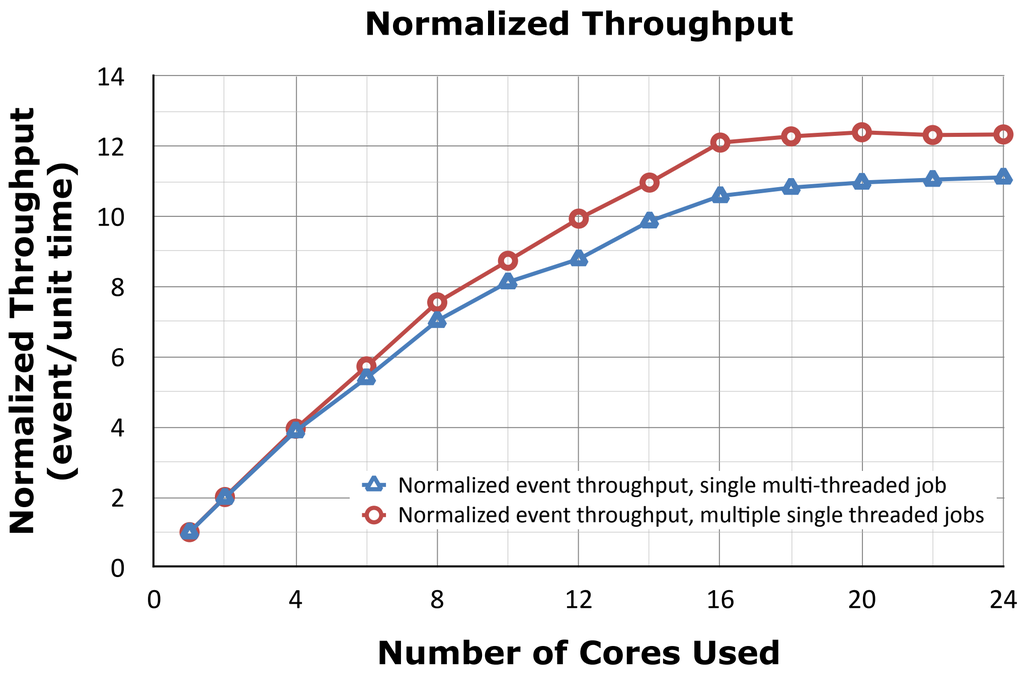
CMSSW

- During LS1, we improved the software significantly
  - re-gained 60% of the factor ~2.5 in processing time increase
- Remaining increase in reconstruction time and increase in trigger rate
  - Will need to run part of PromptReco at Tier-1 sites

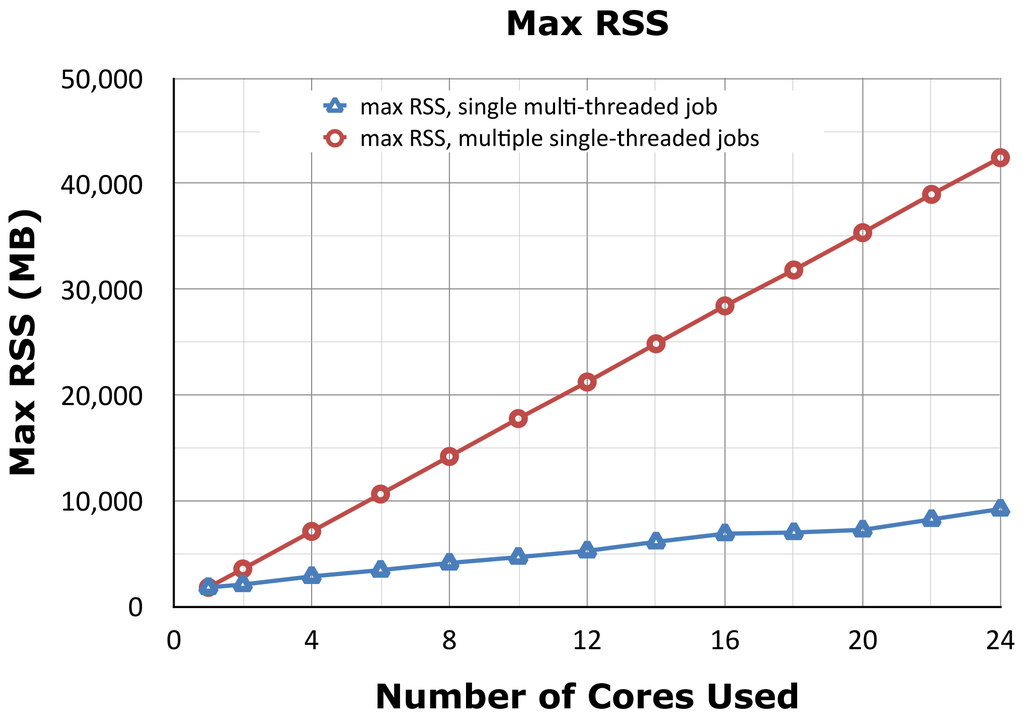


Multi-threading

- Increased trigger rate and increased reco time ➜ lumi sections with many events (high trigger rates) cannot be processed within 48 hours
  - Lumi section (23 sec of data taking, for MC this can be set) still accounting unit for processing and analyses ➜ cannot be split over separate applications
  - Multi-threading allows to process one lumi section on several cores!



95% efficiency compared to single-threaded jobs at **significant memory savings**



Fermilab

PileUp

- Data taking conditions have to be simulated
  - Digitization step adds PileUp by randomly adding pre-simulated MinBias events to the generated physics event
    - PileUp dataset
- Two PileUp mixing modes
  - **Traditional**: pick selected number of MinBias events from MinBias dataset and read from digitization process
    - Can be order of ~100 of events read ➜ high I/O, not immediately suited for Tier-2 sites
  - **Pre-mixing**: pre-mix MinBias events ➜ only one event has to be read per generated physics event
    - Significantly reduced I/O ➜ can be run at Tier-2 sites
- Plan for 2015 is to finalize pre-mixing and use it on Tier-2 level to compensate for Tier-1 level running PromptReco
  - physics validation is fixing the last %-level difference to traditional mixing
  - US-Tier-2 sites prime candidates because you are our most performant and reliable sites
- NB: Fast Simulation now will use similar PileUp mixing strategies as full simulation
  - Fast simulation was 1-2 orders of magnitude faster than full simulation, currently effort ongoing to bring it back to same level of performance
  - This would mean even with pre-mixing significant I/O load ➜ will have to test where to run these workflows

MINIAOD

- Compact high-level data tier (30-50 kB/event) designed to cover the mainstream analyses
  - We expect that this will be the main data tier accessed on Tier-2 sites for analysis

- General overview of size per event for 35 PU
  - RECO: 2.474 MB
    - 6 month window will be kept on Tier-1 disk and not written to disk
    - Mainly used for commissioning and alignment/calibration
  - AOD: 430 kB
  - MiniAOD: 47 kB

Processing Plan

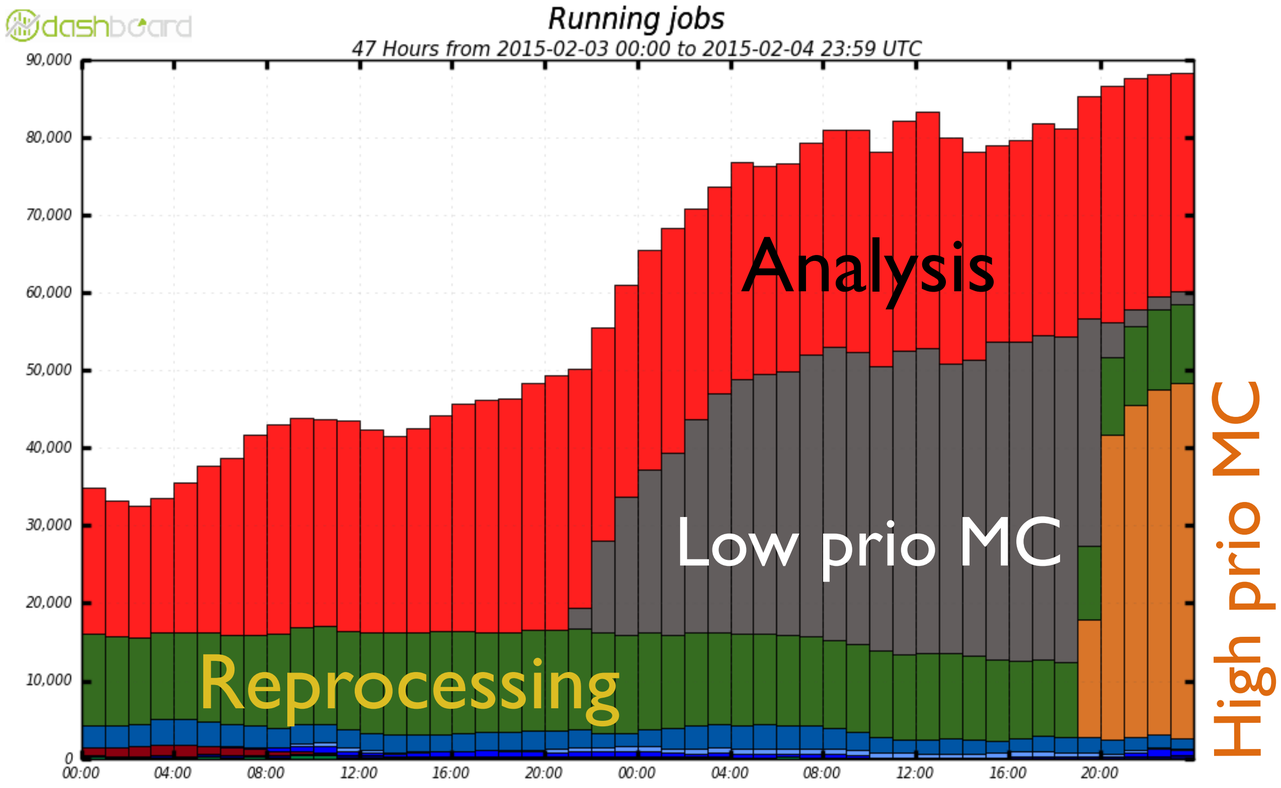
- Expect frequent data re-reconstruction passes throughout the year
  - While we are understanding the data taking conditions and the detector (similar to 2010)
  - At least re-reconstruction passes after the 3 running periods in 2015

- Current plan foresees 4 Billion MC events plus significant amount of FastSim samples
  - Possibly with several re-digitization/re-reconstruction passes following data re-reconstruction passes

- Data PromptReco and re-reconstruction will run Multi-threaded!
  - Digitization not yet multi-threaded at sufficient efficiency, MC digitization/ reconstruction will stay single-threaded for 2015

Global Pool

- Significant simplification of glideinWMS setup for LHC Run 2
  - One single global pool
  - One kind of pilots sent to Tier-2 sites
    - No need for prioritization on CE level, all is done by glideinWMS
- Results in significantly increased flexibility
- Tier-1 sites need to run Multicore pilots
  - Using dynamic partitioning to schedule single-threaded applications
- Tier-2 sites are not required to run Multicore pilots
  - But they can if they want to ➜ example: Purdue



Fermilab

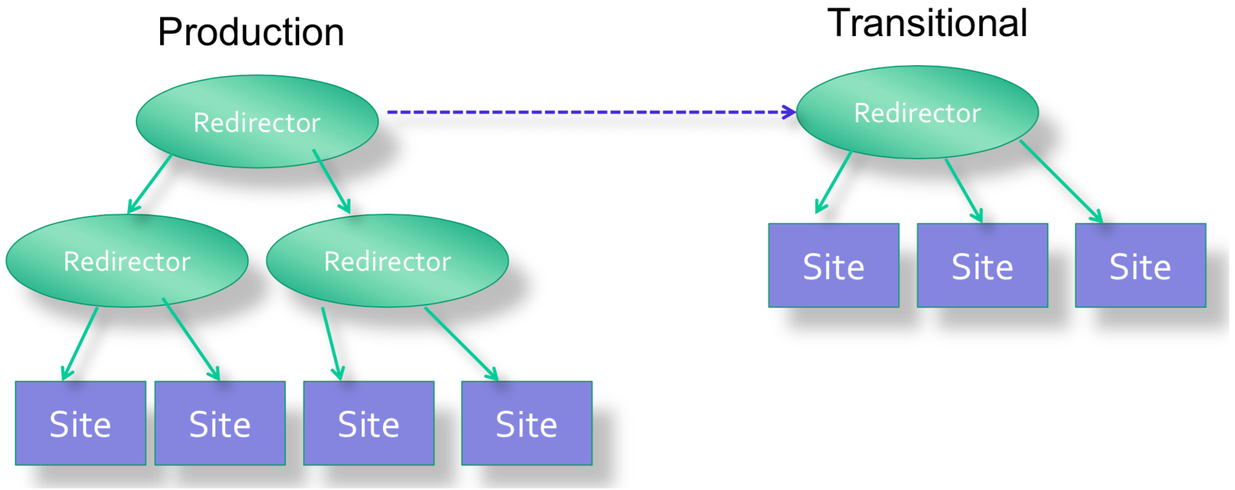
CPU Prioritization

- Prioritization on pledges resources is defined by global CMS
  - This includes opportunistic resources if not specified otherwise
- National VOMS groups can be used to prioritized national resources
  - Pilots with the national VOMS group proxy would be sent and CEs would prioritize appropriately
- New proposal, discussed within the context of enabling scheduling on non-WLCG resources (like a Tier-3):
  - Define list of “local users” per site
    - The list of CMS user names in a flat-formatted list by placing an appropriate file in SITECONF. These users are added to group $SITENAME.
    - The list of groups considered “local” to the site (also placed in SITECONF).
      - Allows T2_US_Nebraska to state “all T1_US_FNAL users are considered local”.
  - At schedd, we would map job->group; pilot would just maintain list of allowed groups.
  - The “local user” pilot is the same DN at all sites. Uses command-line arguments to determine local site.
- In the US, we currently don’t have special prioritization for US users (e.g. national voms group), should we introduce this (again)?
- Please have a look at Brian’s talk at the March 2015 Offline & Computing week:
  - https://indico.cern.ch/event/368377/session/3/contribution/16/material/slides/0.pdf

Data management improvements

- Dynamic Data Placement and Automatic Cache Release is in operation
  - Handles samples in specified quotas at the Tier-2 sites, nominally 60% of pledged disk resources
  - New samples are placed at random and reliable Tier-2 sites by production teams
  - Further replication is handled by DDM

- AAA - Transitional Federation
  - Stabilize xrootd federation and guarantee high level of file availability
  - Joining the production federation:
    - Passes the SAM access test (almost all T1's and T2's do already)
    - Can sustain a 10 Hz file-open rate, 600 simultaneous jobs at 0.25 MB/s (= 1.2 Gbit/s)
  - Move to the transitional federation:
    - AAA-related ticket open for longer than two weeks
    - SAM access test < 50% for two weeks
    - HammerCloud test success rate < 80% for two weeks
      - Test at a rate of ~50 jobs, 200 GB/day
  - Conservative "two months" to complete establishing the transitional federation



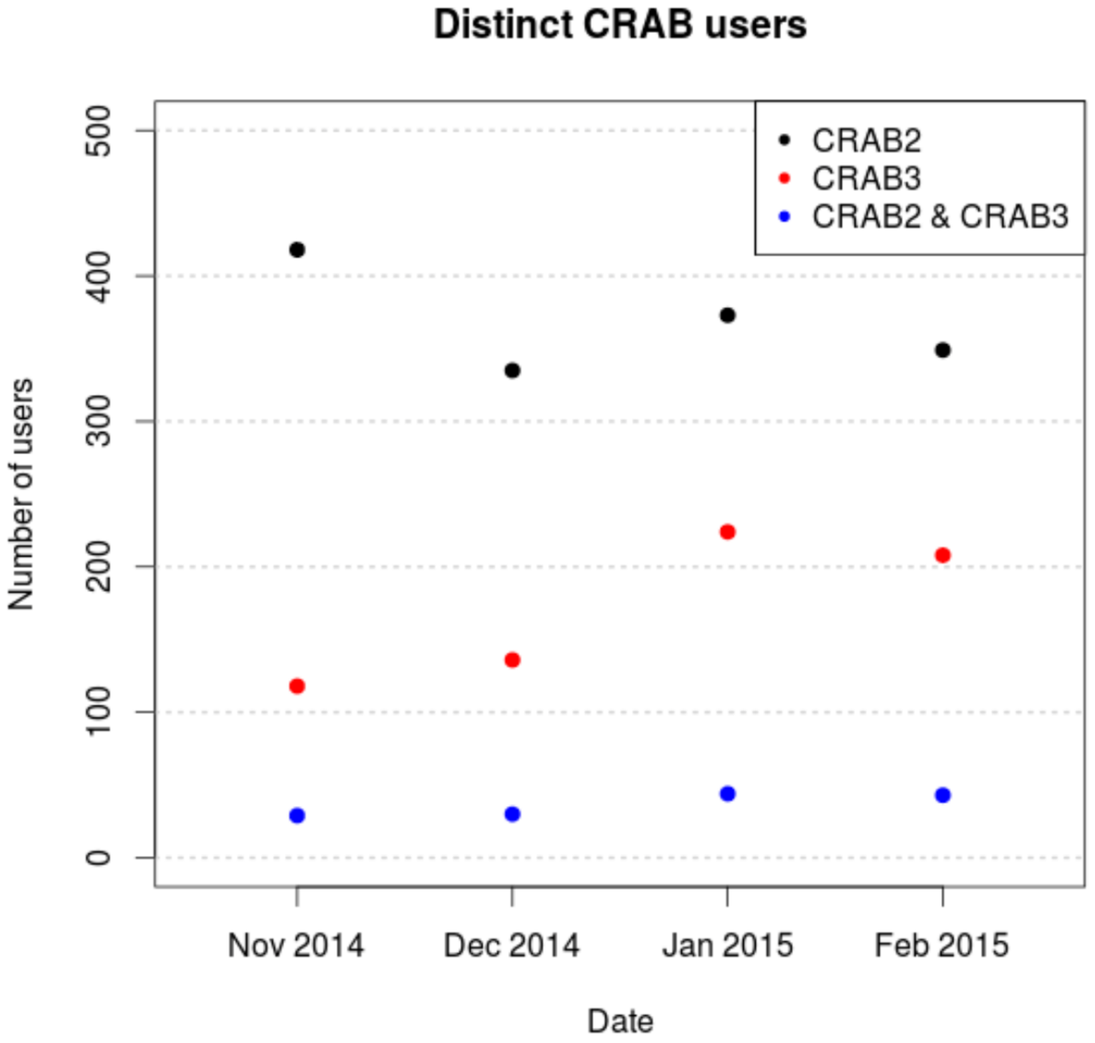
Fermilab

Review of User Data Management in Run 2

- Keep /store/user/<username> at user's home T2 under responsibility of their national CMS organizations
- Introduce a global quota per group for group-produced data. Managed centrally in CMS (possibly through PhEDEx)
- Users clearly expressed the desire to track also non-EDM data in DBS. Propose to:
  - extend CRAB3 to handle this
  - develop a tool that allows users to register an arbitrary directory of files into DBS and PhEDEx
- Some low-cost improvements are proposed for few CRAB stageout cases
  - Give a few files back as quick as possible
  - Allow to retrieve small outputs to submission UI w/o bothering with /store/user @ T2 (a' la Crab2)
  - Help users to merge small outputs into larger files
- Help users to move around data in /store/user
- Make it possible (and easy) to move data from /store/user to /store/result
- Policy for users or groups to ask for a data set created in /store/result to become part of DDM managed data
- Currently working on implementation plan and time schedule

CRAB 3

- CRAB3 is the default system since beginning 2015
- Zero effort for CRAB2: May 1st
- We see no need to coercively move remaining CRAB2 users to CRAB3
- CRAB3 vs CRAB2 users at the level of 40%-60%
  - Still a significant fraction of the users need to migrate



R&D

- CMS is organizing R&D topics lead by Pete Elmer and Daniele Bonacorsi
  - Weekly meeting where new ideas are discussed and status updates are given

- Example, Pisa presented their multi-VO Docker-based setup during March 2015 Offline & Computing Week
  - https://indico.cern.ch/event/368377/session/14/contribution/61/material/slides/1.pdf

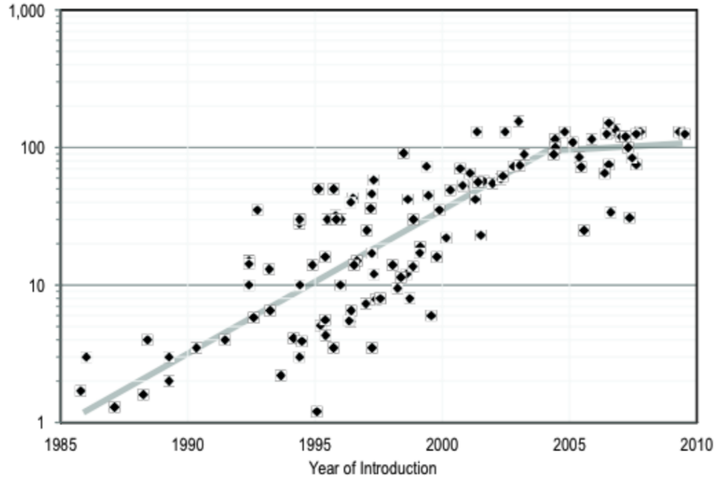
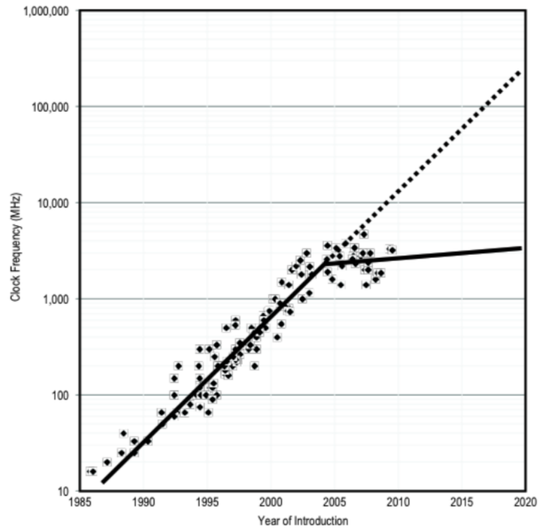


FIGURE A.4 Microprocessor power dissipation (watts) over time (1985-2010).



From: "The Future of Computing Performance: Game Over or Next Level?"

Fermilab

Summary

- LHC Run 2 and especially 2015 will be very exciting for all of us in CMS
  - Essentially, we are taking data with a new machine and have to understand our detector anew
  - Discovery potential is enormous and phase space is large

- During LS1, we all worked hard on tackling the problems because of higher PileUp, increased beam energy and trigger rate
  - A lot of new systems and procedures will be needed to make LHC Run 2 a success