CMS Computing Plans

Maria Girone

for the Offline and Computing groups

ECFA Preparatory Workshop, September 5th, 2014

Offline and Computing

- Talking concretely about Offline and Computing 10 years in advance is somewhat academic. CMS has been writing a technical proposal during LS1 and what follows is mostly based on the draft document
  - Disruptive technology changes can completely alter the landscape
    - We are counting on it
  - HEP doesn't drive the technology
    - We are a small piece of a much larger market. Industrial areas have surpassed us, and even other sciences are catching up
  - Offline and Computing take an incoming rate of events and complexity has to be handled within a resource envelope
    - Constrained on either end

Resource Growth

- Based on the technology section of the WLCG Computing Model Evolution document we currently we see 25% processing capacity and 20% storage increase per year for the same money
  - This results in a doubling every 3 years for CPU and 4 years for storage, so a factor of roughly 8 and 6 by the timescale of Run4
    - This assumes flat funding, which would be the best scenario we could hope for
      - The plot below has no predictive power, but no exponential growth up to now

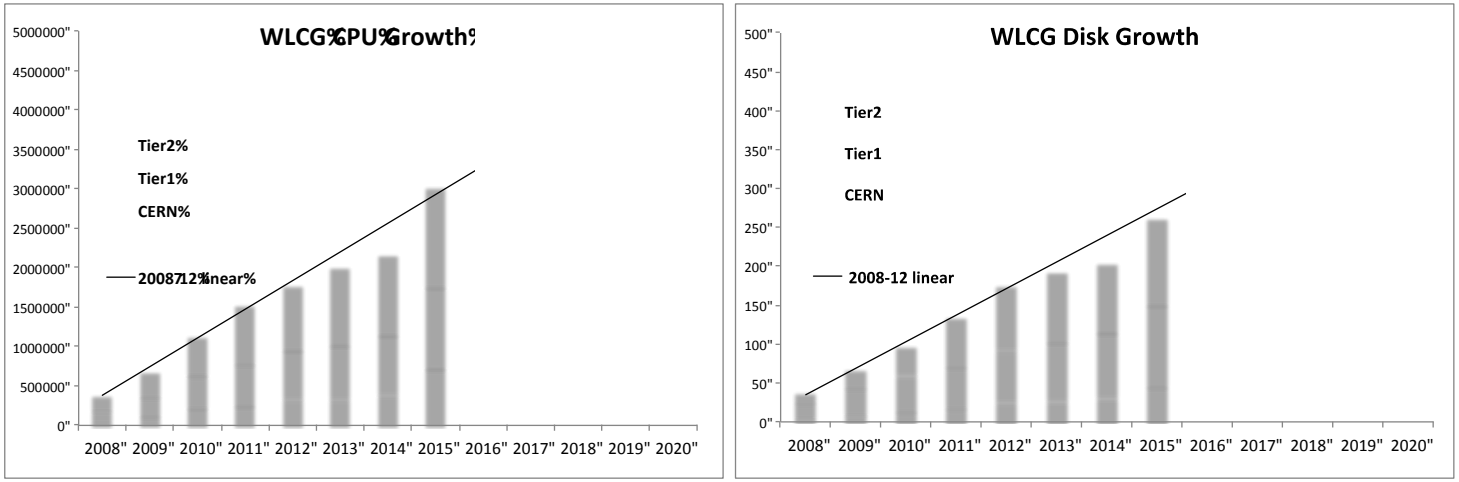


Figure 6.1: Shows the CPU and disk growth through the first 7 years of the program.

Resource Needs

- CMS is planning for 5-7.5kHz of data in Run4. In this scenario CMS would collect 25B-37B raw events per year

  - Estimating from the current software and using the upgrade simulation we see that each of these events is more complicated to reconstruct and larger than the events we will collect in 2015

	Detector	Pile-up (Ave./crossing)	Reconstruction time (Ratio to Run 2)	AOD size (Ratio to Run 2)
Run3	Phase 1	50	4	1.4
	Phase-II	140	20	3.7
Run4	Phase-II	200	45	5.4

Size of the processing problem

- Factoring in the trigger rate and taking a weighed average of the data and simulation tasks we see the computing challenge is 65-200 times worse than Run2

Detector	HLT output rate (kHz)	Total
Phase 1	1	3
Phase-II (140)	5	65
Phase-II (200)	7.5	200

Scale of computing resource needs relative to Run 2 including the increase in projected HLT output rate

- Anticipating a factor of 8 in CPU improvements and a factor of 2 in code improvement, we have a deficit of a factor of 3-15

- Anticipating a factor 6 in storage improvements and having by Phase II events 4-5 times larger, we have still a deficit of 4-5 in storage

Scale of solutions

- It is unlikely we will get a factor of 5 more money, nor will the experiment be willing to take a factor of 5 less data
  - Big improvements are needed

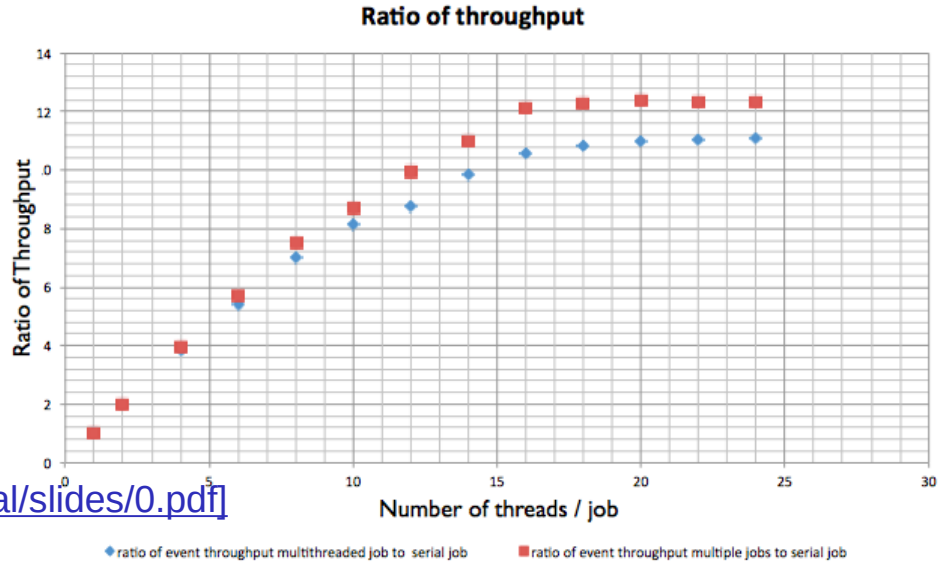
- CMS is investigating many areas. We have begun in some areas and are getting organized in others. We look forward to communicating between groups on promising areas
  - Changes in architectures and technology (see P. Elmer's talk)
  - Code improvements
  - Data reduction and selection techniques
  - Specialized Computer Centers
  - …

Targets

- Roughly 40% of the CMS processing capacity is devoted to task identified as reconstruction
  - Prompt reconstruction, re-reconstruction, data and simulation reco
    - Improving the number of events that can be reconstructed per computing unit per Swiss Franc is the single biggest savings
- ~20% of the offline computing capacity is in areas identified as selection and reduction
  - Analysis selection, skimming, production of reduced user formats
- The remaining 40% is a mix
  - Lot of different activities with no single area to concentrate optimization effort
    - Simulation already has a strong ongoing optimization effort
    - User analysis activities developed by many people
    - Smaller scale calibration and monitoring activities

Improvements in Processing

- In software CMS has been focusing on new architectures for the future
  - Low cost and low power ARM processors, and high performance GPU systems
  - Offline is devoting substantial development effort to improving the ability to run code across multiple cores
    - Very hard to get high efficiency as the number of cores used increases
- CMS has already achieved >99% parallel safe code and has excellent efficiency up to 8 cores



https://indico.cern.ch/event/258092/session/7/contribution/93/material/slides/0.pdf]

Data Reduction and Selection Techniques

- CMS maintains as open as possible triggers and datasets are reduced to optimized selections for particular activities
  - Nearly all the selections are serial passes through the data by users and groups
    - Relies on multiple distributed copies and many reads of each event
- CMS is investigating ways to reduce the amount of computing spent on data reduction
  - Event tags and catalogs can improve the selection speed and efficiency
  - Big Data tools like Map Reduce can make scalable IO and reuse the selection criteria

Specialized Centers

- CMS would like to investigate the scale of improvement in the cost per capacity of using specialized centers for dedicated workflows like reconstruction and event selection
  - If this is the most efficient way of working, it could be a significant change in how we support and provision computing services
    - Not all services and capabilities will be at all sites
  - It would introduce a more heterogeneous and complex system
    - From an operations perspective and from a support and funding perspective

Outlook

- We are facing a large deficit for Run4 due to increased trigger rates and event complexity. Technology evolution alone will not close the gap
  - We either need a huge injection of money or innovative improvements
- Specialized hardware and massively parallel low cost and low power systems have the potential of significantly reducing the cost per processing
  - Important implications for application algorithm developers
- Specialized data reduction centers can reduce the total computing needed for the bulk of analysis by reusing the calculation across users and group
  - Specialized centers with direct I/O reduce the numbers of replicas needed
- CMS is investigating a number of R&D areas, and is seeking for collaborating between experiments, sites, and groups