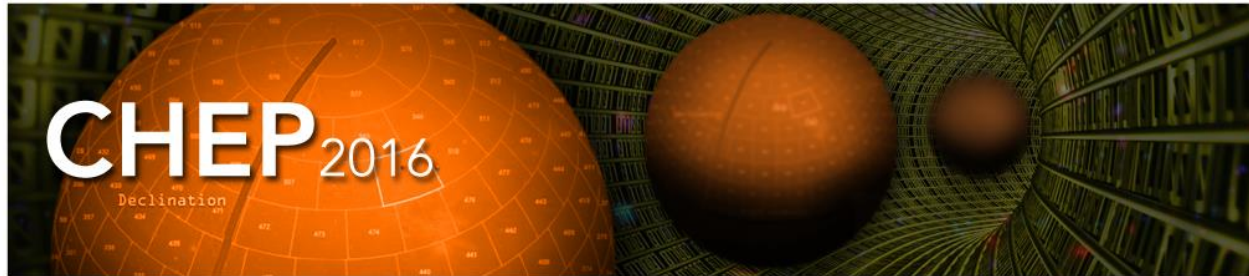


Experience Using Commercial Clouds in CMS

Maria Girone
October 13th, 2016



22nd International Conference on Computing in High Energy and Nuclear Physics, Hosted by SLAC and LBNL, Fall 2016

on behalf of the CMS collaboration

This work has been documented in the note FERMILAB-PUB-16-170-CD

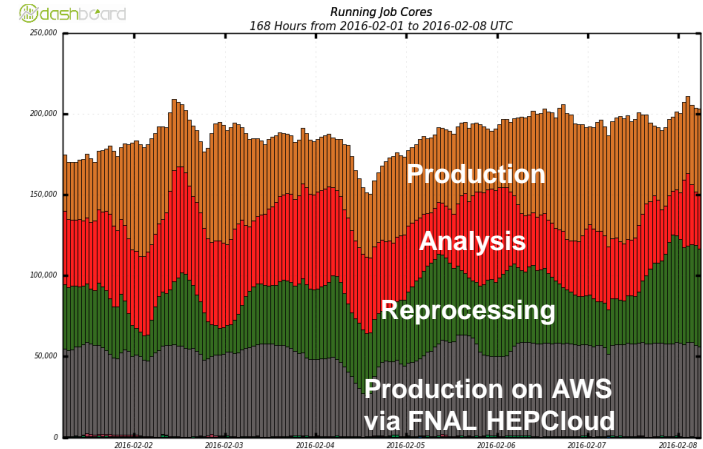
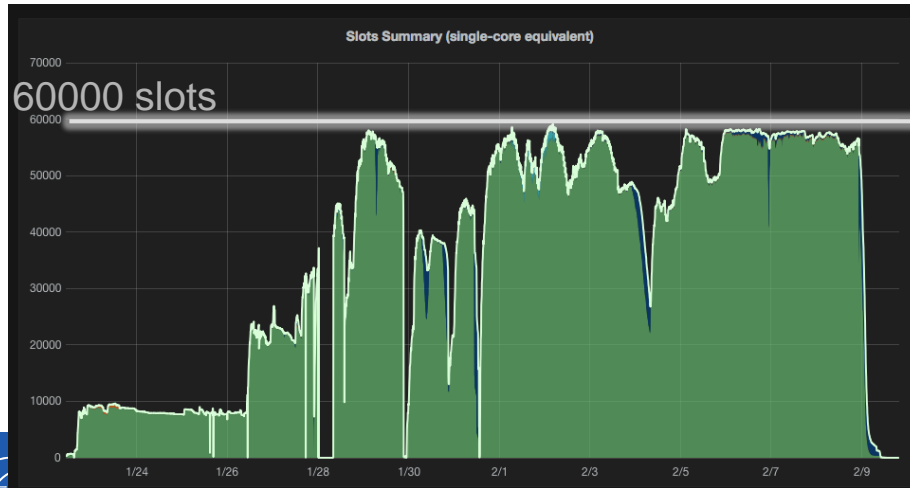
Goals of the CMS AWS Project

- Demonstrate the ability to significantly augment the peak capacity of the CMS production system using commercial computing resources
 - Investigate scheduling for peak
 - Investigate the cost, reliability, and efficiency of large-scale commercial computing resources
- Demonstrate the capability of executing in production any of the CMS centrally organized workflows on AWS
 - Demonstrate scalability up to additional 56000 compute cores, steady state
 - Contribute significantly to a official CMS data production for the 2016 winter conferences
- One of the use cases of the Fermilab HEPCloud effort
 - IaaS through Cloud computing for elasticity, without over-provisioning local resources
- AWS academic 9 to 1 matching grant awarded to CMS from June 2015 to March 2016
 - Size of award based on what it would cost to do one month of large-scale processing
 - Conditional waiver for exporting data if export costs <15% of total monthly bill
 - Use of research networks like ESnet
- These kind of allocations are an interesting resource to expand quickly to pursue an opportunity or fix a problem
 - Insurance if we needed to reprocess or make a new simulation campaign
 - It's not possible to fold contingency into the dedicated resources, but these sort of allocations maybe powerful sources of problem recovery



Scale achieved

- All CMS simulation requests fulfilled for Moriond 2016
 - 2.9 million jobs, 15.1 million wall hours
 - 9.5% badput – includes preemption from spot pricing
 - 87% CPU efficiency
 - 518 million events generated



Outlook

- The tests performed by Fermilab and CMS on AWS have demonstrated that is possible to utilize dynamically provisioned Cloud resources
 - Proven capability to execute efficiently **both** data intensive and CPU intensive workflows
 - HEPCloud facility was able to add by 33% total CMS resources
- Average cost per core-hour
 - On-premises resource: **.9 cents** per core-hour
 - Includes power, cooling, staff
 - Off-premises at AWS: **1.4 cents** per core-hour
 - Ranged up to 3 cents per core-hour at smaller scale
- AWS required the workflows to be carefully chosen and tuned and the workflow system to increase in the dynamic scale
- We could not identify any workflow that could not be executed
- AWS was much more competitive in terms of cost and will be an interesting resource to supplement our dedicated resources