

# CMS: Accounting opportunistic resources in APEL

J. Flix

*w/feedback from: C. Grandi, O. Gutsche, D. Hufnagel,  
A. Pérez-Calero Yzquierdo*

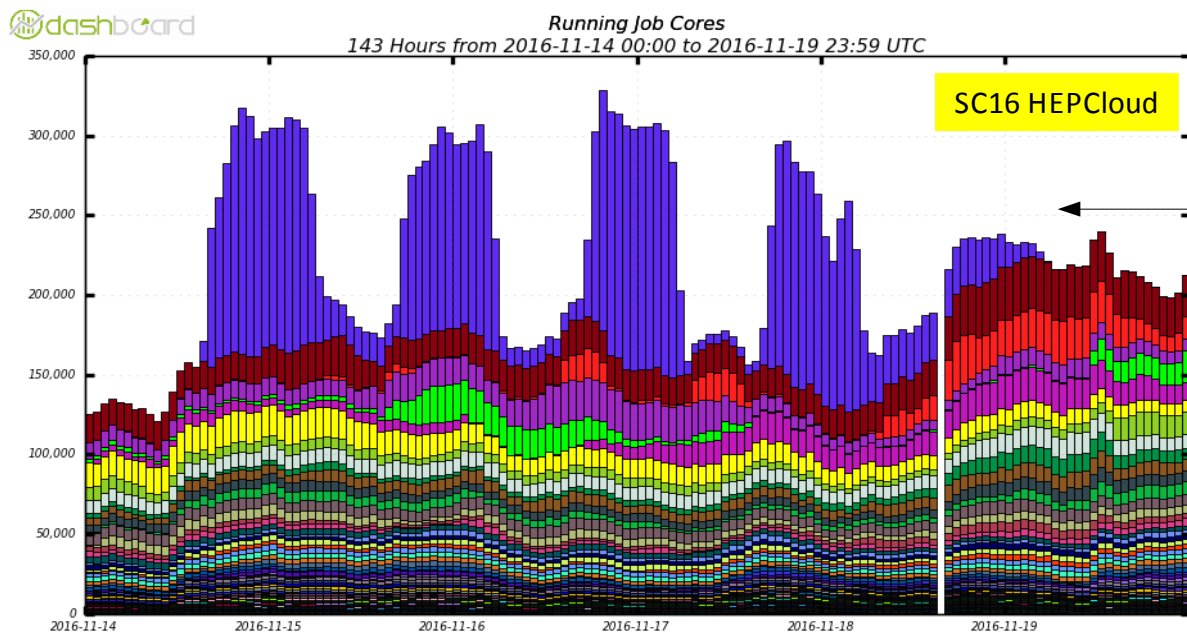
[jflix@pic.es](mailto:jflix@pic.es) / WLCG Accounting TF meeting (9<sup>th</sup> Feb. 2017)

# CMS Opportunistic resource usage

- Exploitation of HPC centres and commercial clouds has been a priority in the CMS Computing Program for at least a couple of years

→ Transparent use of NERSC resources @US (Edison, Cori-1, Cori-2)

→ AWS @US, Google Cloud Platform @US, Aruba @IT, ongoing Microsoft Azure



Using the FNAL HEPCloud facility w/HTCondor to send bursts of CMS simulation jobs to GCP

The bursts were approx. of the same size of the whole CMS Computing at all the Tiers!  
(doubled the capacity of the CMS HTCondor global pool)

- T3\_US\_HEP.Cloud
- T3\_US\_NotreDame
- T2\_US\_Nebraska
- T1\_DE\_KIT
- T2\_UK\_London\_IC
- T2\_FR\_GRIF\_IRFU
- T3\_US\_OSG
- T1\_ES\_PIC
- T2\_BE\_UCL
- T1\_US\_FNAL
- T2\_CH\_CERN
- T2\_US\_Caltech
- T2\_DE\_RWTH
- T2\_UK\_London\_Brunel
- T1\_FR\_CCIN2P3
- T2\_ES\_IFCA
- T2\_RU\_JINR
- T2\_CH\_CSCS
- T0\_CH\_CERN
- T2\_DE\_DESY
- T2\_US\_Purdue
- T2\_IT\_Bari
- T1\_RU\_JINR
- T2\_EE\_Estonia
- T2\_ES\_CIEMAT
- T3\_TW\_NCU
- T2\_UK\_SGrid\_Bristol
- T2\_US\_Wisconsin
- T2\_US\_Florida
- T2\_US\_MIT
- T2\_US\_Vanderbilt
- T2\_IT\_Legnaro
- T2\_UK\_SGrid\_RALPP
- T2\_BR\_SPRACE
- T2\_IT\_Rome
- T2\_RU\_IHEP
- T2\_CH\_CERN\_HLT
- T1\_IT\_CNAF
- T2\_US\_UCSD
- T2\_BE\_IHHE
- T1\_UK\_RAL
- T2\_IT\_Pisa
- T2\_FR\_GRIF\_LLRL
- T3\_UK\_ScotGrid\_GLA
- ... plus 45 more

Maximum: 328,207, Minimum: 0.00, Average: 220,262, Current: 212,372

<https://cloudplatform.googleblog.com/2016/11/Google-Cloud-HEPCloud-and-probing-the-nature-of-Nature.html>

# CMS Opportunistic resource usage

- Exploiting HPCs/Clouds offers the ability to flexibly absorb peaks in processing requests without acquiring in-house resources
- It is generally acknowledged that the exploitation of opportunistic resources cannot in general be part of the current WLCG pledged resources mechanism
  - But accepted that these resources will be part of the overall picture in the future, although there is no well defined model across LHC experiments yet
  - The current efforts to plug CMS workflows into HPCs and commercial clouds is not negligible for the moment (weeks or months to fully integrate a system)
- We should have mechanisms to account for the work done in these facilities
  - CMS is adopting internal T3\_Country\_Opportunistic

<b>Tier 0</b>	<a href="#">T2_CH_CERN</a>	<a href="#">T2_IT_Bari</a>	<a href="#">T2_UA_KIPT</a>	<a href="#">T3_CH_PSI</a>	<a href="#">T3_KR_KISTI</a>	<a href="#">T3_US_Brown</a>	<a href="#">T3_US_Princeton_ICSE</a>
<a href="#">T0_CH_CERN</a>	<a href="#">T2_CH_CERN_AI</a>	<a href="#">T2_IT_Legnaro</a>	<a href="#">T2_UK_London_Brunel</a>	<a href="#">*T3_CH_Volunteer</a>	<a href="#">T3_KR_KNU</a>	<a href="#">T3_US_Colorado</a>	<a href="#">T3_US_PuertoRico</a>
<b>Tier 1</b>	<a href="#">T2_CH_CERN_HLT</a>	<a href="#">T2_IT_Pisa</a>	<a href="#">T2_UK_London_IC</a>	<a href="#">T3_CN_PKU</a>	<a href="#">T3_KR_UOS</a>	<a href="#">T3_US_Cornell</a>	<a href="#">T3_US_Rice</a>
<a href="#">T1_DE_KIT</a>	<a href="#">T2_CH_CERN_Wigner</a>	<a href="#">T2_IT_Rome</a>	<a href="#">T2_UK_SGrid_Bristol</a>	<a href="#">T3_CO_Uniandes</a>	<a href="#">T3_MX_Cinvestav</a>	<a href="#">T3_US_FIT</a>	<a href="#">T3_US_Rutgers</a>
<a href="#">T1_ES_PIC</a>	<a href="#">T2_CH_CSCS</a>	<a href="#">T2_KR_KNU</a>	<a href="#">T2_UK_SGrid_RALPP</a>	<a href="#">T3_ES_Oviedo</a>	<a href="#">T3_NZ_UOA</a>	<a href="#">T3_US_FIU</a>	<a href="#">T3_US_SDSC</a>
<a href="#">T1_FR_CCIN2P3</a>	<a href="#">T2_CH_CSCS_HPC</a>	<a href="#">T2_MY_SIFIR</a>	<a href="#">T2_US_Caltech</a>	<a href="#">T3_FR_IPNL</a>	<a href="#">T3_RU_FIAN</a>	<a href="#">T3_US_FNALLPC</a>	<a href="#">T3_US_TACC</a>
<a href="#">T1_IT_CNAF</a>	<a href="#">T2_CN_Beijing</a>	<a href="#">T2_MY_UPM_BIRUNI</a>	<a href="#">T2_US_Florida</a>	<a href="#">T3_GR_Demokritos</a>	<a href="#">T3_RU_MEPHI</a>	<a href="#">T3_US_FSU</a>	<a href="#">T3_US_TAMU</a>
<a href="#">T1_RU_JINR</a>	<a href="#">T2_DE_DESY</a>	<a href="#">T2_PL_NCP</a>	<a href="#">T2_US_MIT</a>	<a href="#">T3_GR_IASA</a>	<a href="#">T3_TH_CHULA</a>	<a href="#">*T3_US_HEPCloud</a>	<a href="#">T3_US_TTU</a>
<a href="#">T1_RU_JINR_Disk</a>	<a href="#">T2_DE_RWTH</a>	<a href="#">T2_PL_Swierk</a>	<a href="#">T2_US_Nebraska</a>	<a href="#">T3_HR_IRB</a>	<a href="#">T3_TW_NCU</a>	<a href="#">T3_US_JHU</a>	<a href="#">T3_US_UB</a>
<a href="#">T1_UK_RAL</a>	<a href="#">T2_EE_Estonia</a>	<a href="#">T2_PL_Warsaw</a>	<a href="#">T2_US_Purdue</a>	<a href="#">T3_HU_Debrecen</a>	<a href="#">T3_TW_NTU_HEP</a>	<a href="#">T3_US_Kansas</a>	<a href="#">T3_US_UCD</a>
<a href="#">T1_UK_RAL_Disk</a>	<a href="#">T2_ES_CIEMAT</a>	<a href="#">T2_PT_NCG_Lisbon</a>	<a href="#">T2_US_UCSD</a>	<a href="#">T3_IN_PUHEP</a>	<a href="#">*T3_UK_GridPP_Cloud</a>	<a href="#">T3_US_MIT</a>	<a href="#">T3_US_UCR</a>
<a href="#">T1_US_FNAL</a>	<a href="#">T2_ES_IFCA</a>	<a href="#">T2_RU_IHEP</a>	<a href="#">T2_US_Vanderbilt</a>	<a href="#">*T3_IN_TIFRCloud</a>	<a href="#">T3_UK_London_QMUL</a>	<a href="#">T3_US_Minnesota</a>	<a href="#">T3_US_UCSB</a>
<a href="#">T1_US_FNAL_Disk</a>	<a href="#">T2_FI_HIP</a>	<a href="#">T2_RU_INR</a>	<a href="#">T2_US_Wisconsin</a>	<a href="#">T3_IN_VBU</a>	<a href="#">T3_UK_London_RHUL</a>	<a href="#">*T3_US_NERSC</a>	<a href="#">T3_US_UIowa</a>
<b>Tier 2</b>	<a href="#">T2_FR_CCIN2P3</a>	<a href="#">T2_RU_ITEP</a>	<b>Tier 3</b>	<a href="#">T3_IR_IPM</a>	<a href="#">T3_UK_London_UCL</a>	<a href="#">T3_US_NEU</a>	<a href="#">T3_US_UMD</a>
<a href="#">T2_AT_Vienna</a>	<a href="#">T2_FR_GRIF_IRFU</a>	<a href="#">T2_RU_JINR</a>	<a href="#">T3_BG_UNI_SOFIA</a>	<a href="#">T3_IT_Bologna</a>	<a href="#">T3_UK_SGrid_Oxford</a>	<a href="#">T3_US_NU</a>	<a href="#">T3_US_UMiss</a>
<a href="#">T2_BE_IHHE</a>	<a href="#">T2_FR_GRIF_LL2</a>	<a href="#">T2_RU_PNPI</a>	<a href="#">T3_BY_NCPHEP</a>	<a href="#">T3_IT_Firenze</a>	<a href="#">T3_UK_ScotGrid_ECDF</a>	<a href="#">T3_US_NotreDame</a>	<a href="#">T3_US_UTENN</a>
<a href="#">T2_BE_UCL</a>	<a href="#">T2_FR_IPHC</a>	<a href="#">T2_RU_SINP</a>	<a href="#">T3_CH_CERN_CAF</a>	<a href="#">T3_IT_MIB</a>	<a href="#">T3_UK_ScotGrid_GLA</a>	<a href="#">*T3_US_OSG</a>	<a href="#">T3_US_UVA</a>
<a href="#">T2_BR_SPRACE</a>	<a href="#">T2_GR_Ioannina</a>	<a href="#">T2_TH_CUNSTDA</a>	<a href="#">*T3_CH_CERN_HelixNebula</a>	<a href="#">*T3_IT_Opportunistic</a>	<a href="#">T3_US_ANL</a>	<a href="#">T3_US_OSU</a>	<a href="#">*T3_US_Vanderbilt_EC2</a>
<a href="#">T2_BR_UERJ</a>	<a href="#">T2_HU_Budapest</a>	<a href="#">T2_TR_METU</a>	<a href="#">*T3_CH_CERN_OpenData</a>	<a href="#">T3_IT_Perugia</a>	<a href="#">T3_US_BU</a>	<a href="#">T3_US_Omaha</a>	<a href="#">T3_US_Wisconsin</a>
	<a href="#">T2_IN_TIFR</a>	<a href="#">T2_TW_NCHC</a>		<a href="#">T3_IT_Trieste</a>	<a href="#">T3_US_Baylor</a>	<a href="#">T3_US_Princeton_ARM</a>	

# Opp. resources be accounted in APEL?

*Whether your experiment would be interested that opportunistic resources are accounted in APEL?*

→ Yes

*If yes, what are possible scenarios?*

→ The easiest, accounted as if from an existing CMS site (e.g. site extension) or to a "name" known to CMS (e.g. T3\_Country\_Oppportunistic)

*Whether these opportunistic resources are already accounted in the experiment-specific systems?*

→ through the dashboard and the HTCondor global pool, atm ["run/wall time", "running job cores", "# events"...]

*How/whether benchmarking of such resources performed?*

→ Not implemented/available yet

→ benchmark \* time, storing the CPUtime spent and the benchmark separately → this would allow for adopting another benchmark if it becomes available, hopefully a better one

→ integrating a (fast) benchmark with the jobs, at which point we probably have to give up precision

→ We should have a way to store these values in the experiment-specific system ("dashboard")

*How these resources are described regarding topology?*

→ See answer to 2<sup>nd</sup> question

*Would it be possible to retrieve accounting data for the opportunistic resources from the experiment-specific systems via APIs?*

→ We assume that the information available in the dashboard can be retrieved via its own APIs and sent to APEL. To be verified if what is available is enough for APEL