CMS High Level Trigger Timing

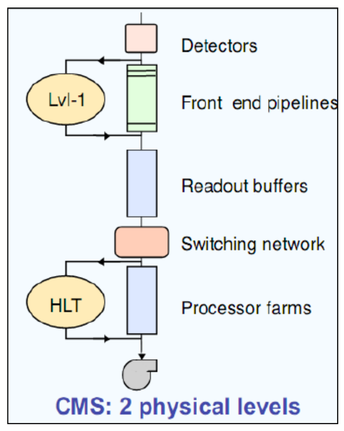
Clint Richardson, Boston University

on behalf of the CMS Collaboration

Introduction

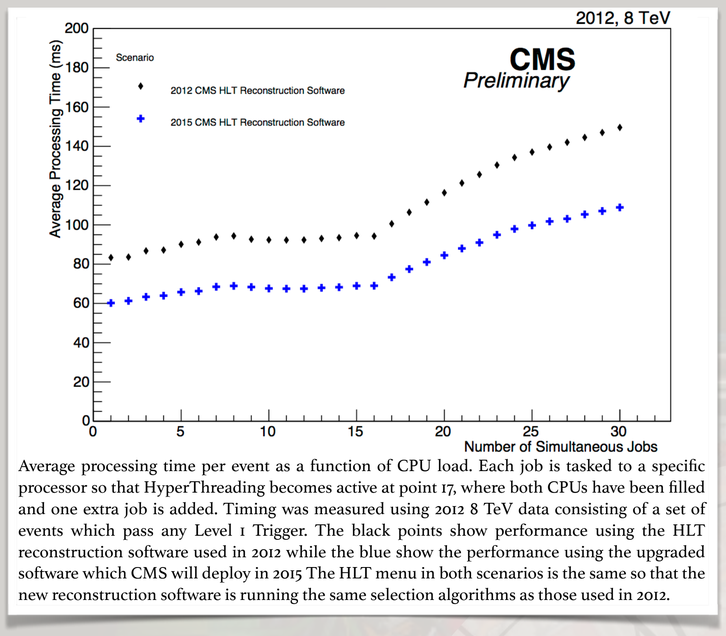
The CMS experiment uses a two-level triggering system in order to select which events are worth saving for offline reconstruction (picture right). The first, Level-1, consists of custom electronics and hardware and reduces the LHC input rate from 40 MHz down to 100 kHz. The second level, the High Level Trigger (HLT) is a collection of software algorithms running on commercial processors (HLT CPU farm) designed to select events in real time and reduces the 100 kHz input rate to roughly 1 kHz for offline storage and reconstruction. The increase in center-of-mass energy and instantaneous luminosity at which the LHC will run in 2015, together with changes to both HLT software and hardware, made deriving an accurate estimate for HLT performance a clear need for CMS during the long shutdown (LS1) of the LHC. In order to obtain an accurate estimate, the software, hardware, and conditions were studied in detail so that the effect of each could be clearly measured. Tests were carried out on multiple CPU generations including the Westmere and Sandy Bridge based machines used during the 2012 run, an Ivy Bridge based machine, and three different Haswell based machines. The full configuration of the HLT CPU farm in 2015 can be seen in the table below.

The effects of Intel's TurboBoost and HyperThreading were taken into account in order to derive an accurate representation of machine performance. In order to isolate these effects from the ones listed above, tests were either carried out across a full spectrum of CPU load, tasking jobs to individual cores and scanning from 1 job to a full machine, or carried out at a previously studied benchmark configuration.



Performance Improvements From New Software

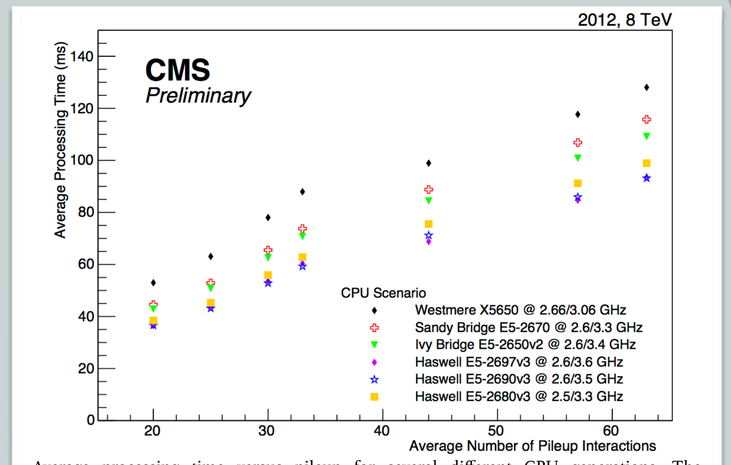
- During LS1 CMS undertook considerable effort to optimize the performance of HLT software in terms of speed of reconstruction.
- The figure to the right shows the performance improvement from the use of 2015 HLT reconstruction software.
- Timing was measured using 2012 8 TeV data with an average of 30 pileup collisions on the Sandy Bridge E5-2670 machines.
- The 2015 reconstruction software gives an across the board improvement of roughly 25%.



Average processing time per event as a function of CPU load. Each job is tasked to a specific processor so that HyperThreading becomes active at point 17, where both CPUs have been filled and one extra job is added. Timing was measured using 2012 8 TeV data consisting of a set of events which pass any Level 1 Trigger. The black points show performance using the HLT reconstruction software used in 2012 while the blue show the performance using the upgraded software which CMS will deploy in 2015 The HLT menu in both scenarios is the same so that the new reconstruction software is running the same selection algorithms as those used in 2012.

HLT Performance Vs. Level of Pileup

- During 2012 the proton-proton collisions at the LHC had an average between 20 and 30 pileup collisions throughout a standard fill.
- In 2015 CMS will see an increase in pileup collisions with an expected average of 40 during main data taking.
- As the figure to the right shows, all CPU generations give qualitatively the same behavior with HLT processing time increasing roughly linearly with the level of pileup.

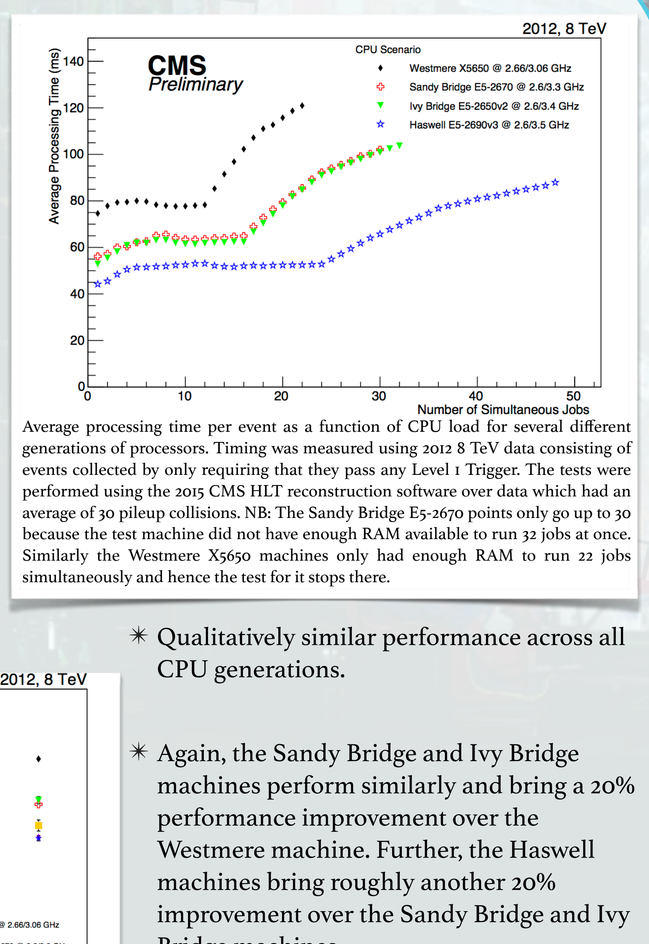


Average processing time versus pileup for several different CPU generations. The performance was measured using 2012 8 TeV data consisting of a set of events passing any Level 1 Trigger. The machines were tested running with one CPU fully loaded without HyperThreading and using the 2015 CMS HLT reconstruction software. The difference in slope between the pileup 20 to 33 points and those between 44 and 63 is due to the fact that the higher pileup runs were taken without out-of-time pileup present.

- The Sandy Bridge and Ivy Bridge processors perform very similarly and bring roughly a 20% performance increase over the Westmere CPU.
- The Haswell based processors deliver another 20% improvement over the Sandy Bridge and Ivy Bridge CPUs.

Performance Improvements From New Hardware

- The HLT CPU farm in 2015 will consist of multiple generations of hardware (see table below for details). In addition we studied the following: the E5-2650v2 which has 8 cores and one NUMA node per CPU, the E5-2690v3 which has 12 cores and one NUMA node per CPU, and the E5-2697v3 which has 14 cores and two NUMA nodes per CPU.
- In order to derive a timing budget for 2015 the performance of all generations needed to be studied in detail.
- When possible, a full scan across all generations was done (right) while when we had limited access to test machines an abbreviated sampling of CPU occupancy was performed (below).



Average processing time per event as a function of CPU load for several different generations of processors. Timing was measured using 2012 8 TeV data consisting of events collected by only requiring that they pass any Level 1 Trigger. The tests were performed using the 2015 CMS HLT reconstruction software over data which had an average of 30 pileup collisions. NB: The Sandy Bridge E5-2670 points only go up to 30 because the test machine did not have enough RAM available to run 32 jobs at once. Similarly the Westmere X5650 machines only had enough RAM to run 22 jobs simultaneously and hence the test for it stops there.

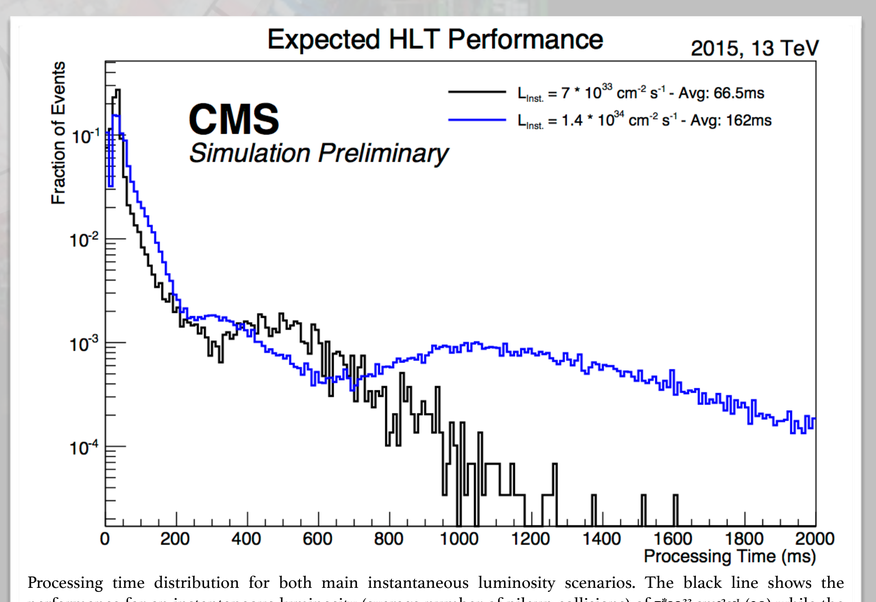
Average processing time for various CPU configuration scenarios. Timing was measured using 2012 8 TeV data consisting of events collected by only requiring a Level 1 Trigger Accept. The meaning of bin labels is as follow: 1job - 1 job on the cpu; 1job NUMA - one job running on each NUMA node; NUMA - running a single cpu with one of its NUMA nodes filled; CPU - running the machine with one of its CPUs fully loaded; Full - running both cpus on the machine fully loaded; 2jobs HT two jobs running on the same core using HyperThreading; NUMA HT - the same as NUMA but doubling the jobs and using HyperThreading; CPU HT - the same as CPU but doubling the number of jobs and using HyperThreading; Full HT - the same as Full but doubling the number of jobs and using HyperThreading.

- Qualitatively similar performance across all CPU generations.
- Again, the Sandy Bridge and Ivy Bridge machines perform similarly and bring a 20% performance improvement over the Westmere machine. Further, the Haswell machines bring roughly another 20% improvement over the Sandy Bridge and Ivy Bridge machines.
- Taking into account relative CPU performance, the configuration of the HLT CPU farm below, and a 100 kHz Level-1 input rate, the farm would be saturated by a process which, when run on its own with the rest of the CPU free, had an average processing time of ~160 ms on the Ivy Bridge CPU (used to measure expected HLT performance in 2015, shown right).

HLT CPU Farm Hardware Configuration in 2015

Year Installed	2011	2012	2015
CPU (Architecture)	X5650 (Westmere)	E5-2670 (Sandy Bridge)	E5-2680v3 (Haswell)
CPUs per Motherboard	2	2	2
Physical Cores per CPU	6	8	12
RAM	24 GB	32 GB	64 GB
Clock Speed (w/ TurboBoost)	2.66 (3.06) GHz	2.6 (3.3) GHz	2.5 (3.3) GHz
Total # of cores in farm	3456	4096	8640

Expected HLT Performance in 2015



Processing time distribution for both main instantaneous luminosity scenarios. The black line shows the performance for an instantaneous luminosity (average number of pileup collisions) of $7*10^{33}$ cm^{-2}s^{-1} (20) while the blue shows performance for $1.4*10^{34}$ cm^{-2} s^{-1} (40). The timing was measured using 13 TeV Monte Carlo simulating proton-proton collisions which were required to pass any Level 1 Trigger and represents expected HLT performance in 2015. The CPU used for measuring this performance was the Ivy Bridge based E5-2650v2 which was configured to run only a single job for each test.

HLT performance expected in 2015 has been estimated using simulation of 13 TeV proton-proton collisions. Key features include:

- Performance shown is with the configuration of one job tasked to one physical core and the rest of the CPU free.
- Cluster of events around 1100 (500) ms for the 40 (20) pileup scenario due to iterative tracking and particle-flow identification algorithms. Here again one sees the effect of pileup on HLT processing time.
- With a timing budget of ~160ms in 2015, the average processing time is well below the timing budget for the pileup 20 scenario.
- The pileup 40 scenario will use the full resources of the HLT CPU farm.