

Geant4 MT Performance

Soon Yung Jun (Fermilab)

21st Geant4 Collaboration Meeting, Ferrara, Italy

Sept. 12 - 16, 2016

Introduction

- Geant4 multi-threading (Geant4 MT) capabilities
 - Event-level parallelism
 - Available since 10.0
 - Status (see Andrea's talk in Plenary 7)
- Readiness for large-scale computing?
 - Validation (not a scope of this talk)
 - Performance
- Basic performance metrics
 - Event throughput (weak scaling)
 - Memory reduction
- Scopes of this talk
 - MT performance on different hardware platforms
 - Profiling results

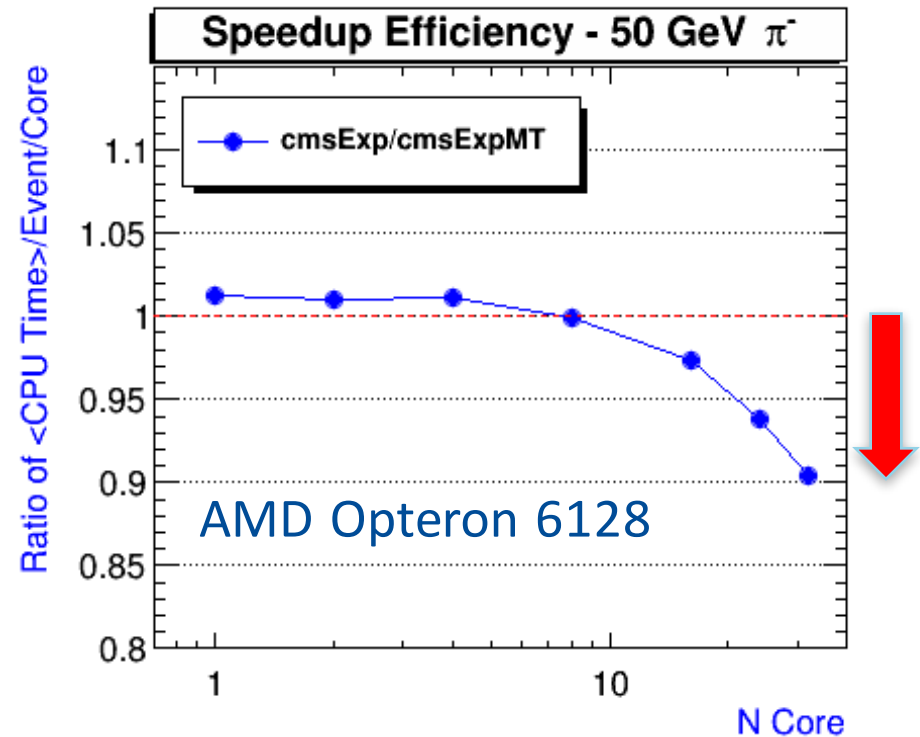
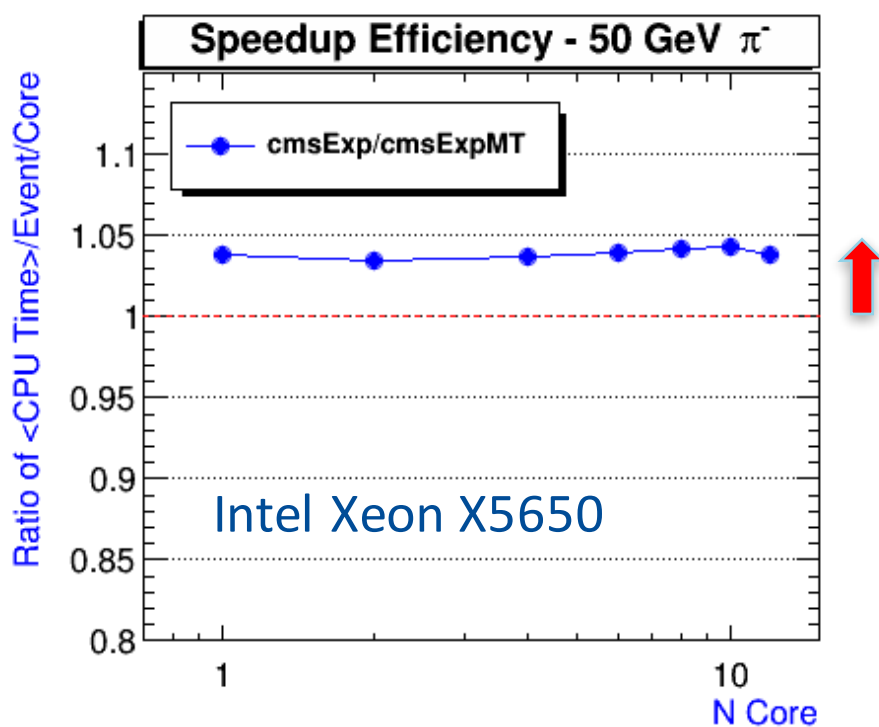
Performance Profiling Experiments

- Application : a standalone CMS detector simulation
 - the CMS geometry (gdml)
 - a volume based magnetic field map excerpted from CMSSW
 - single particle samples (50 GeV pi-,e-) and PYTHIAH \rightarrow ZZ
 - cmsExp (sequential) and cmsExpMT (multi-threading)
- Platform tested for this talk
 - Intel Xeon X5650: dual-socket 6-core (total 12 cores), 12GB
 - AMD Opertron 6128: quad-socket 8-core (total 32 cores), 64GB
 - Intel Xeon Phi 5110P (MIC, Knight's Corner): 60 cores, 8GB
 - Intel Xeon Phi (Knight's Landing), 64 cores, 96GB+16MCDRAM
- Profiling tools
 - OpenSpeedshop (OSS) v2.2
 - Intel VTune Amplifier XE (VTune) 2016

MT Performance on General Purpose CPUs: Intel vs. AMD

- Event throughput = the number of event processed/time

$$\text{Speedup efficiency: } \epsilon(N\text{threads}) = \frac{\text{Throughput}(\text{Sequential})}{\text{Throughput}(N\text{threads})} \times N\text{threads}$$



- What to understand (Geant4 10.2.r06)
 - MT is (sometimes) faster than sequential
 - Degradation as the number of threads increases in AMD

Profiling Comparison: Intel Xeon

- OSS compare: Sequential vs. MT with 1 thread (% of time)
 - Reported time: 1951 (s1) vs. 1878 (t1) seconds for 1028 events of 50 GeV pions (10.2.r06)

```
[openss]: Legend: -c 2 represents db_intel_pi-50_1/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_intel_pi-50_0/cmsExp-pcsamp.openss
-c2,% of -c4,% of Function (defining location)
CPU Time CPU Time
```

6.190079	6.168036	G4PhysicsVector::Value(double, unsigned long&) const (
3.899106	3.807265	cmsExpMagneticField::GetVolumeBaseBfield(double const*
3.309547	3.215778	G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vect
2.335641	2.266614	__ieee754_atan2 (libm-2.12.so)
1.707284	1.512450	G4VDiscreteProcess::PostStepGetPhysicalInteractionLeng
1.568036	1.977405	G4SteppingManager::DefinePhysicalStepLength() (libG4tr
1.484518	1.488923	G4Navigator::ComputeStep(CLHEP::Hep3Vector const&, CLH
1.420893	1.413265	G4VoxelNavigation::ComputeStep(CLHEP::Hep3Vector const
1.404688	1.857006	G4SteppingManager::Stepping() (libG4tracking.so: G4Ste
1.348542	1.255817	G4UniversalFluctuation::SampleFluctuations(G4MaterialC

t1 s1

- A hint of difference in SteppingManager, but not conclusive
- Need to cross-check the number of steps/tracks (by the particle type)

Profiling Comparison: AMD Opertron

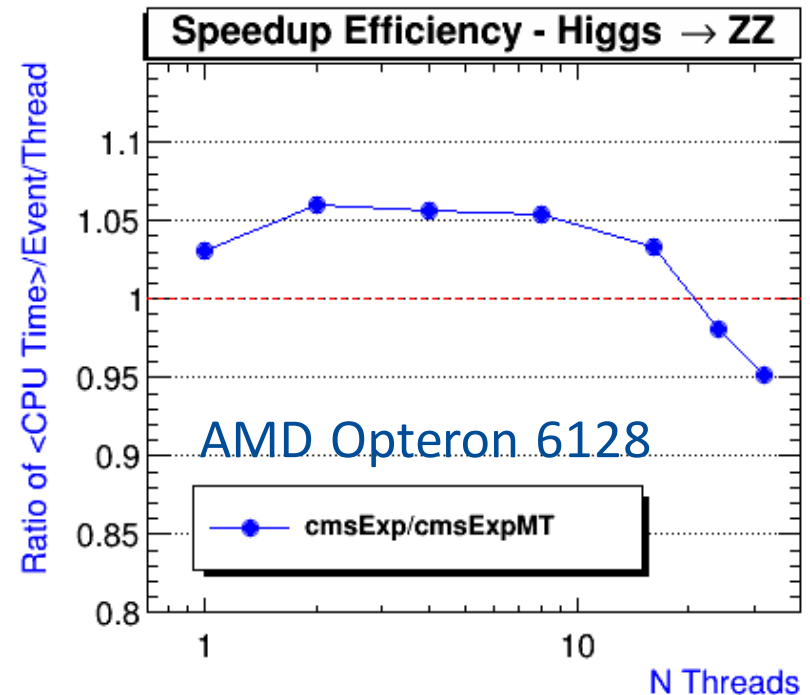
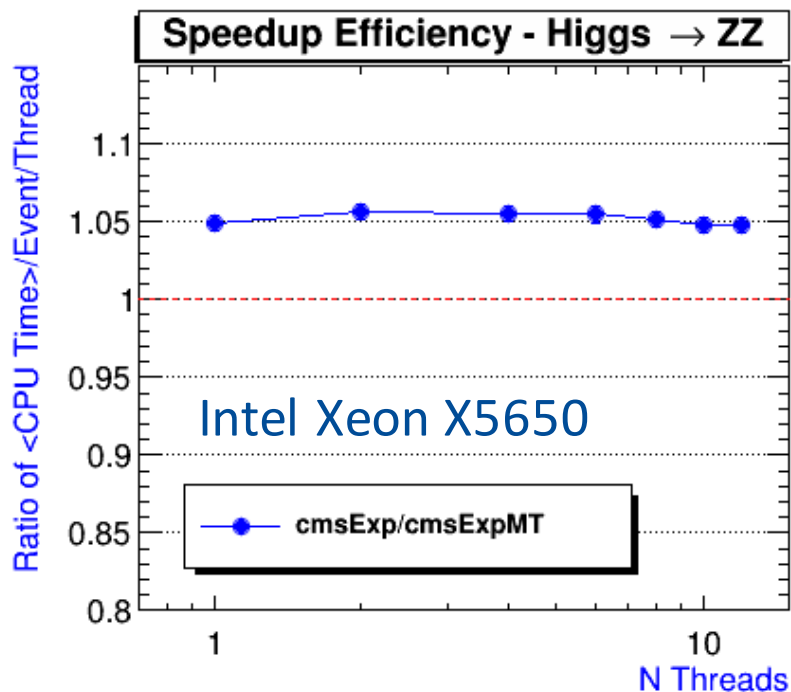
- OSS compare: 32 threads vs. 1 thread (% of time)
 - Experiment with 1028 events of 50 GeV pions (10.2.r06)

```
[openss]: Legend: -c 2 represents db_amd_pi-_50_32/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_amd_pi-_50_1/cmsExpMT-pcsamp.openss
-c 2, % -c 4, % Function (defining location)
CPU Time CPU Time
  t32    t1
5.996050 6.249268 G4PhysicsVector::Value(double, unsigned long&) const
3.835966 3.081700 G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vect
3.777269 4.086459 cmsExpMagneticField::GetVolumeBaseBfield(double const
2.354695 0.528181 G4ParticleChangeForTransport::UpdateStepForAlongStep(
2.338718 2.512992 _ieee754_atan2 (libm-2.12.so)
2.100942 1.179980 G4Navigator::ComputeStep(CLHEP::Hep3Vector const&, CL
1.708470 1.828667 G4SteppingManager::DefinePhysicalStepLength() (libG4t
1.639153 1.369797 G4CrossSectionDataStore::GetCrossSection(G4DynamicPart
1.464572 1.437773 G4VoxelNavigation::ComputeStep(CLHEP::Hep3Vector const
1.423322 1.427549 G4SteppingManager::Stepping() (libG4tracking.so: G4St
```

- Clear signs of difference in G4Navigator and ParticleChangesForTransport::UpdateStepForAlongStep

H → ZZ : Intel vs. AMD

- Speedup efficiency (ϵ) as the number of threads (10.2.r06)
 - The number of events processed = 50 x Nthreads



Profiling Comparison: Intel Xeon

- OSS compare: Intel sequential vs. MT 1 thread (% of time)
 - Experiments with 50 events of $H \rightarrow ZZ$ (10.2.r06)

```
[openss]: Legend: -c 2 represents db_intel_higgs_0/cmsExp-pcsamp.openss
[openss]: Legend: -c 4 represents db_intel_higgs_1/cmsExpMT-pcsamp.openss
-c2,% of -c4,% of
CPU Time CPU Time Function (defining location)

6.315451 6.463411 G4PhysicsVector::Value(double, unsigned long&) const
2.718451 2.977623 G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vect
2.489600 2.606530 cmsExpMagneticField::GetVolumeBaseBfield(double const*
2.327372 2.049718 G4SteppingManager::DefinePhysicalStepLength() (libG4tr
2.047615 1.520584 G4SteppingManager::Stepping() (libG4tracking.so: G4Ste
1.754896 1.365109 G4Transportation::AlongStepGetPhysicalInteractionLengt
1.731395 1.863428 G4UniversalFluctuation::SampleFluctuations(G4MaterialC
1.641349 1.640354 G4Mag_UsualEqRhs::EvaluateRhsGivenB(double const*, dou
1.615017 1.691205 G4PolyconeSide::DistanceAway(CLHEP::Hep3Vector const&,
1.563666 1.831518 G4VDiscreteProcess::PostStepGetPhysicalInteractionLeng
```


Profiling Comparison: AMD Opertron

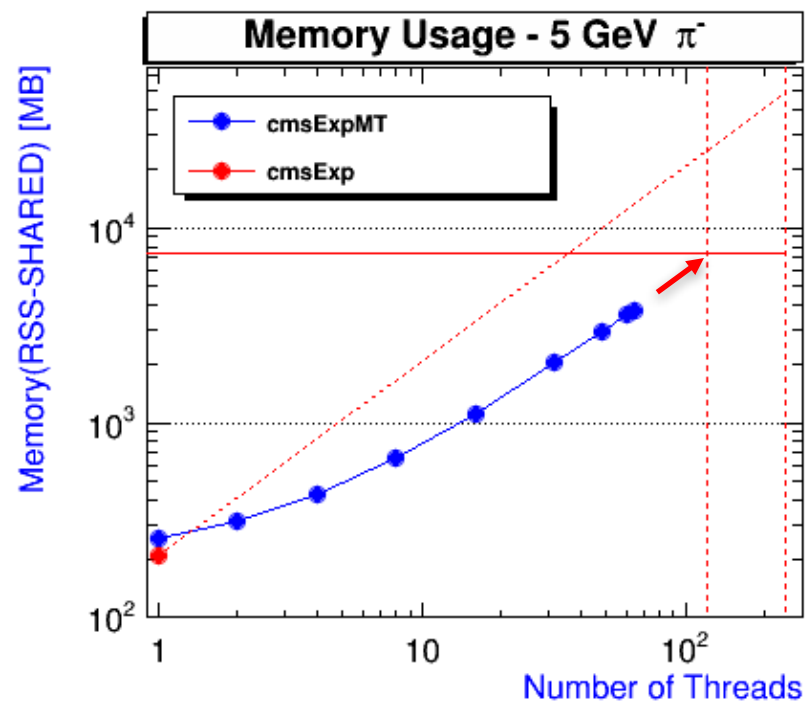
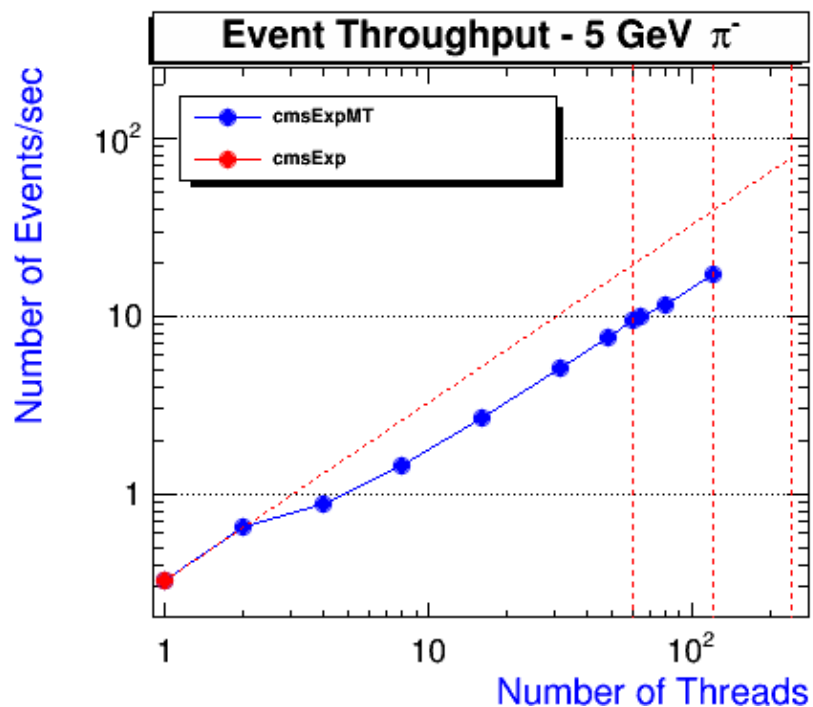
- OSS compare: MT 2 threads vs. MT 32 thread (% of time)
 - Experiments with 50xNthreads events of H->ZZ (10.2.r06)

```
[openss]: Legend: -c 2 represents db_amd_higgs_2/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_amd_higgs_32/cmsExpMT-pcsamp.openss
-c2,% of CPU Time  -c4,% of CPU Time  Function (defining location)
6.487911  6.162949  G4PhysicsVector::Value(double, unsigned long&) const
2.878457  3.341192  G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vec
2.781600  2.523859  cmsExpMagneticField::GetVolumeBaseBfield(double const
1.973615  1.997436  G4SteppingManager::DefinePhysicalStepLength() (libG4t
1.961492  1.804446  G4PolyconeSide::DistanceAway(CLHEP::Hep3Vector const&
.....
1.177427  1.735634  G4Navigator::ComputeStep(CLHEP::Hep3Vector const&, CL
0.872323  1.318057  G4TouchableHistory::GetVolume(int) const (libG4digits
0.749323  1.228093  G4Navigator::LocateGlobalPointWithinVolume(CLHEP::Hep
0.711796  1.712017  G4ParticleChangeForTransport::UpdateStepForAlongStep
0.691357  0.665238  CLHEP::RanecuEngine::flatArray(int, double*) (libG4cl
```

- Again hints of difference: adding counters for the number of steps tracks by the particle type for MT

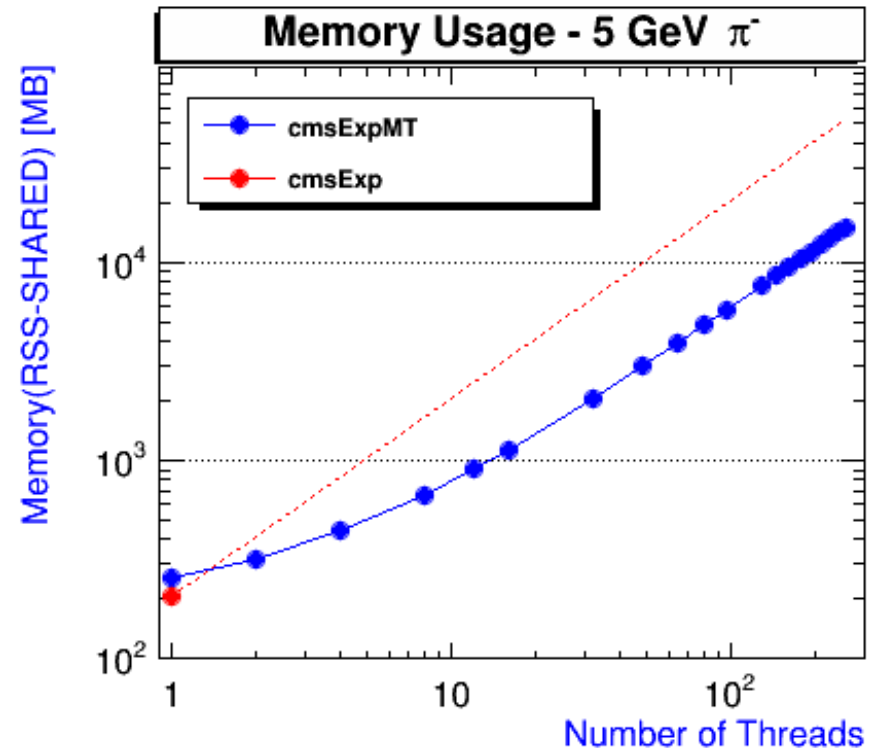
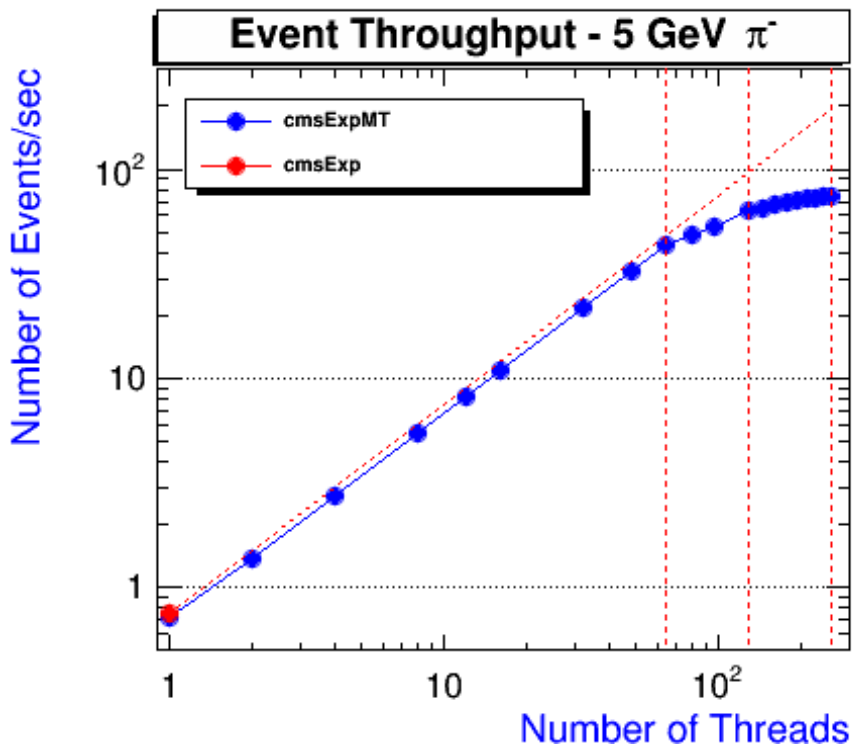
MT performance: Xeon Phi 5110 (MIC, Knight's Corner)

- cmsExp on MIC: 5 GeV π^- (Events = 1028 x N-threads)
 - 60 cores (4 way hyper-threading), 1.03 GHz, 7.8 GB memory
 - Significant scalability loss from N threads = 2 to N threads = 4
 - Hit memory limit (~ 7.3 GB available) @ 120 threads
 - Need to re-measure throughput with physics samples (threshold for the memory limit and the maximal cores to utilize)



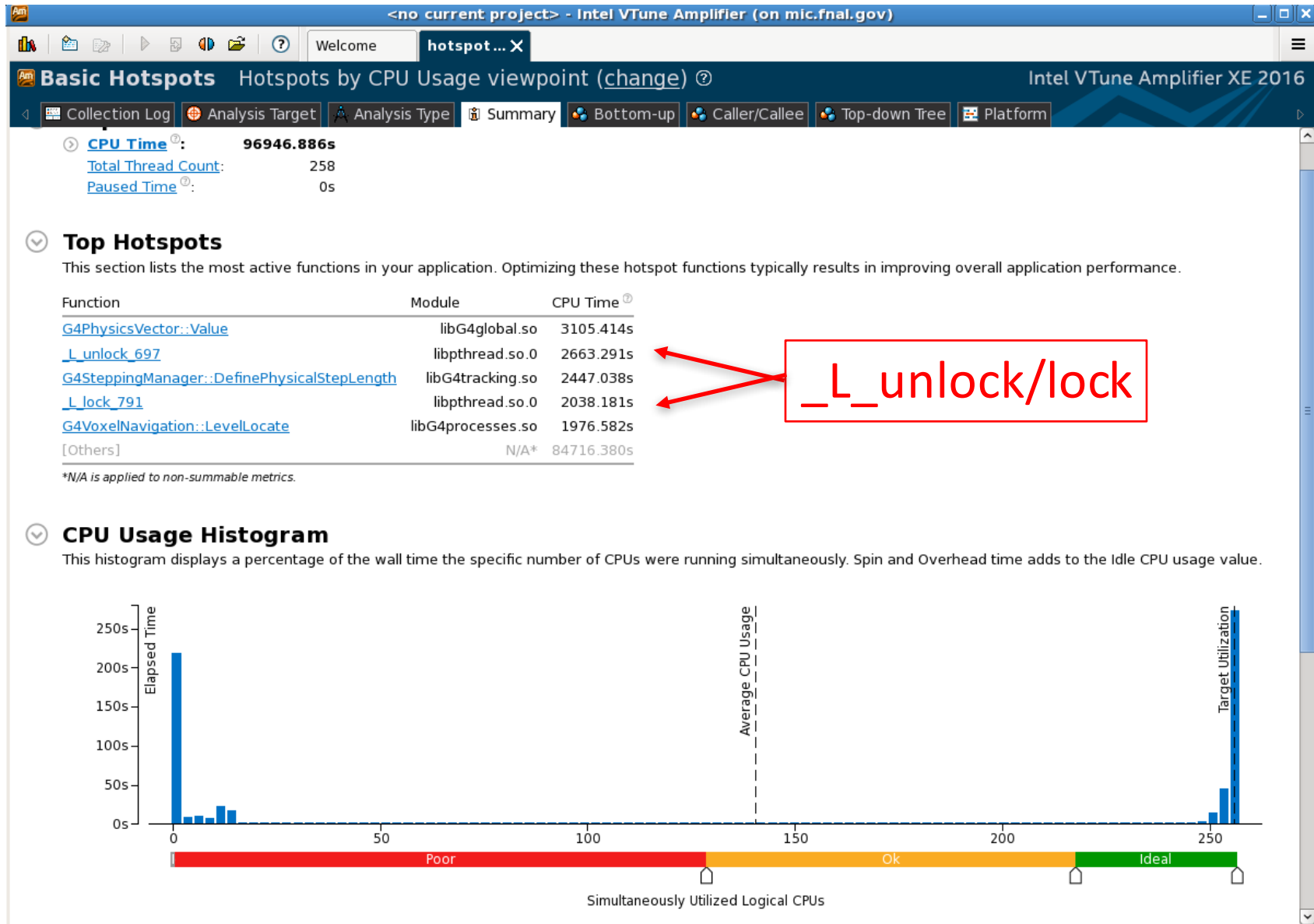
MT Performance on KNL

- Performance on Intel Xeon Phi Processor (Knight's Landing)
 - Developer Edition: Single Socket 1.30 GHz, 64 core
 - MEMORY: 96GB, 2133MHz DDR4, 16GB MCDRAM memory
 - Geant4 10.2.p02 with -xMIC-AVX512
 - Experiment with N-threads x1028 Events of 5 GeV π^- (10.2.r06)

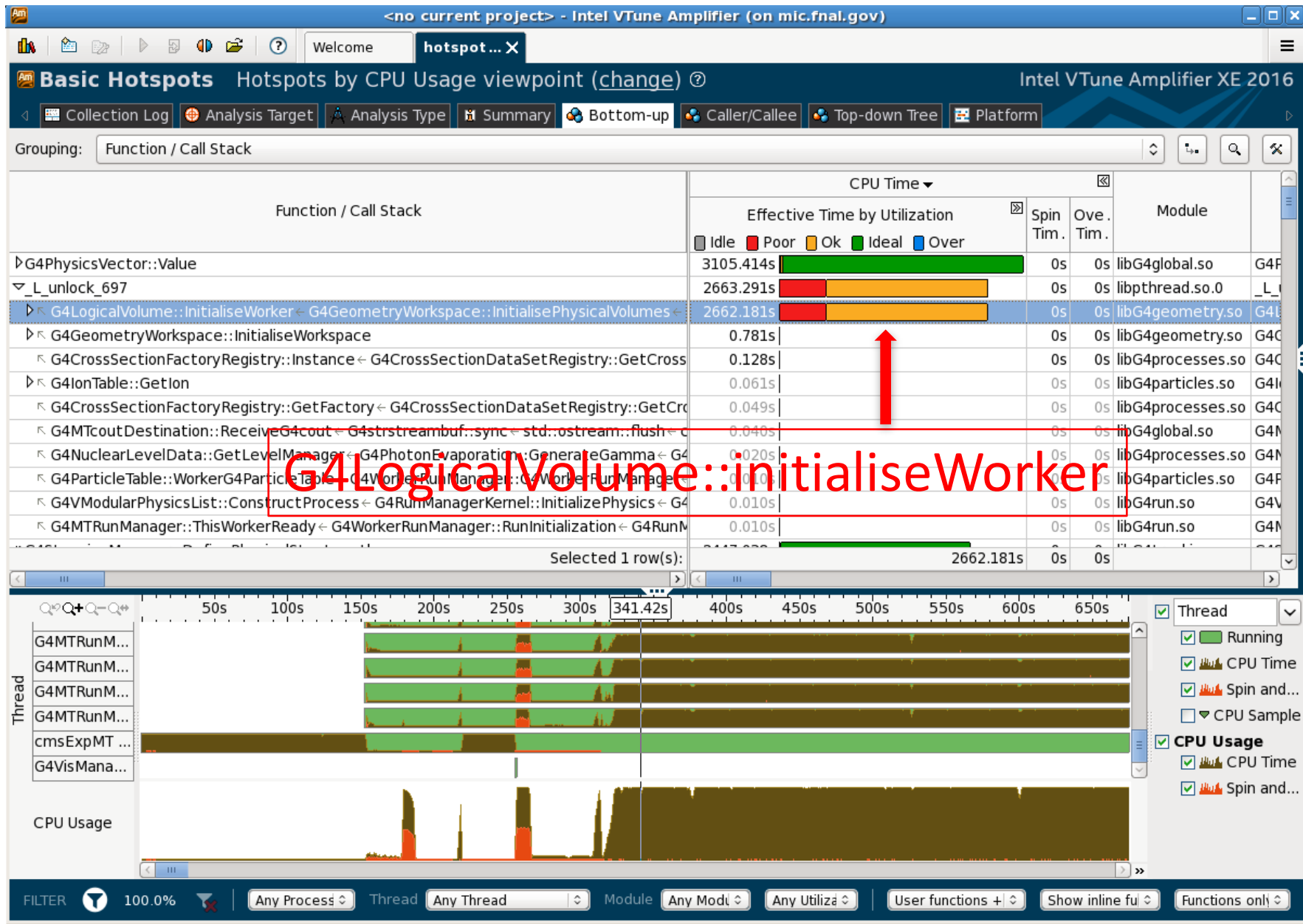


Profiling Results: KNL

- Hotspots with N-threads = 256

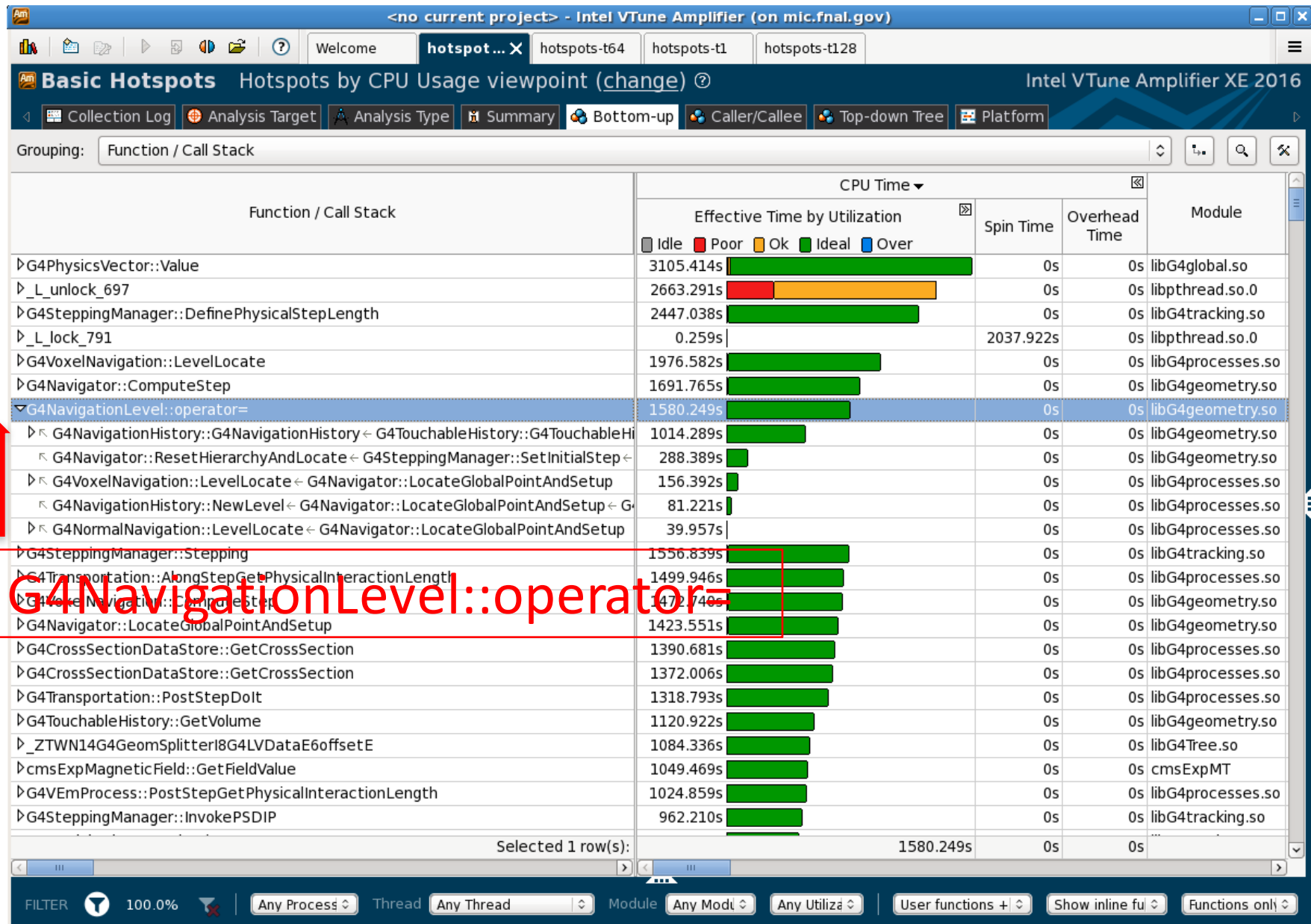


Profiling Results: KNL (N threads = 256)



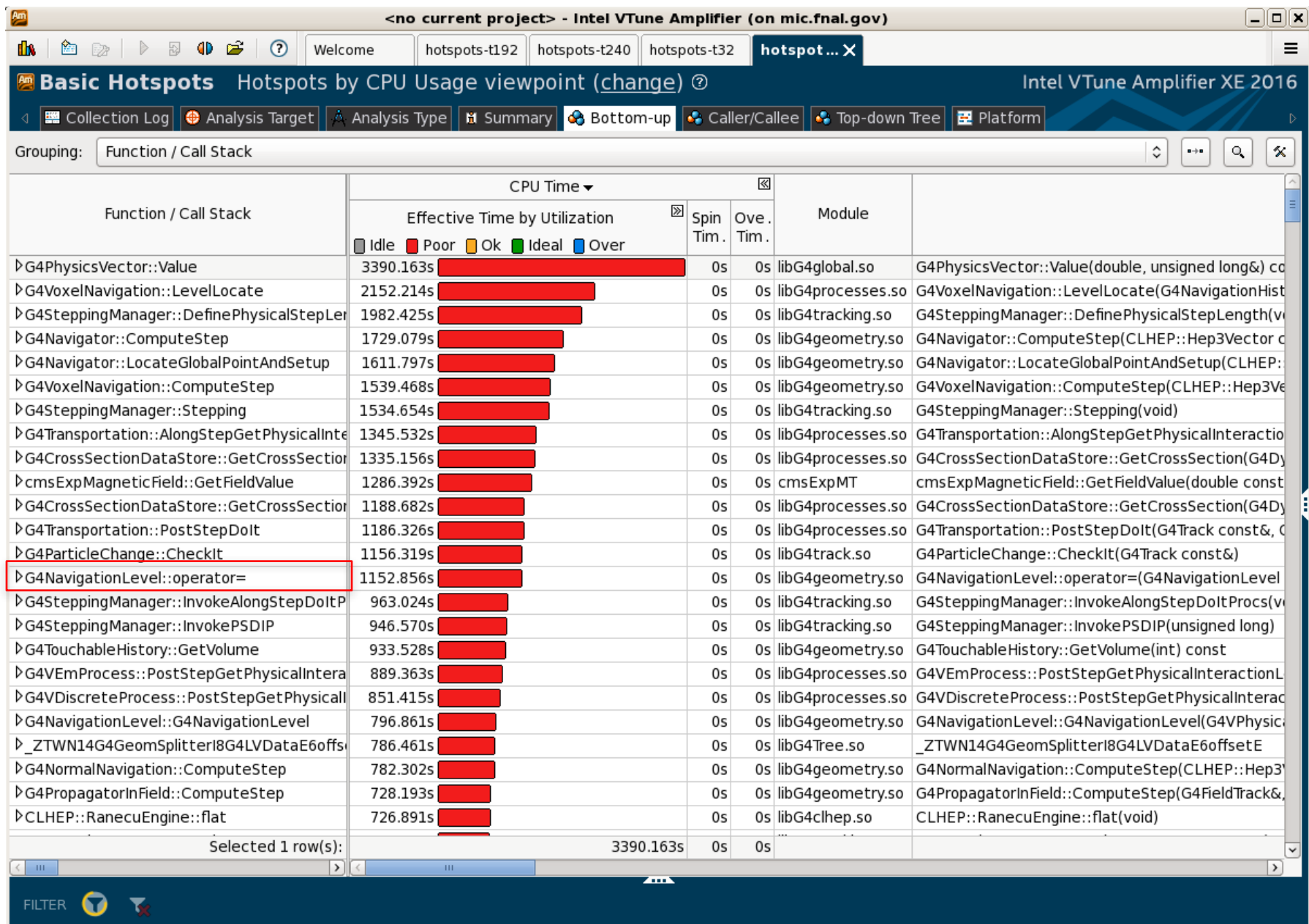
- `_L_lock`: also called by `G4LogicalVolume::initialiseWorker`

Profiling Results: KNL (N threads = 256)



- Also seen with N threads = 198 and 128

Geant4MT: Profiling Result (N threads =128)



Summary

- Reviewed Geant4 MT performance
 - standalone CMS detector simulation (single particle, $H \rightarrow ZZ$)
- Performance on different systems and profiling results
 - No major issues on Intel Xeon
 - Degradation seen on AMD as the number of threads is partially understood
 - Xeon Phi (KNC) shows problems at N-threads > 120
 - Xeon Phi (KNL) shows stable performance
- More tests to understand results of AMD/KNL profiling data
 - Examine stepping information on AMD and KNL (sequential vs. 1-threads and 1-thread vs. N-threads)
 - Test $H \rightarrow ZZ$ on KNL (scalability and memory)

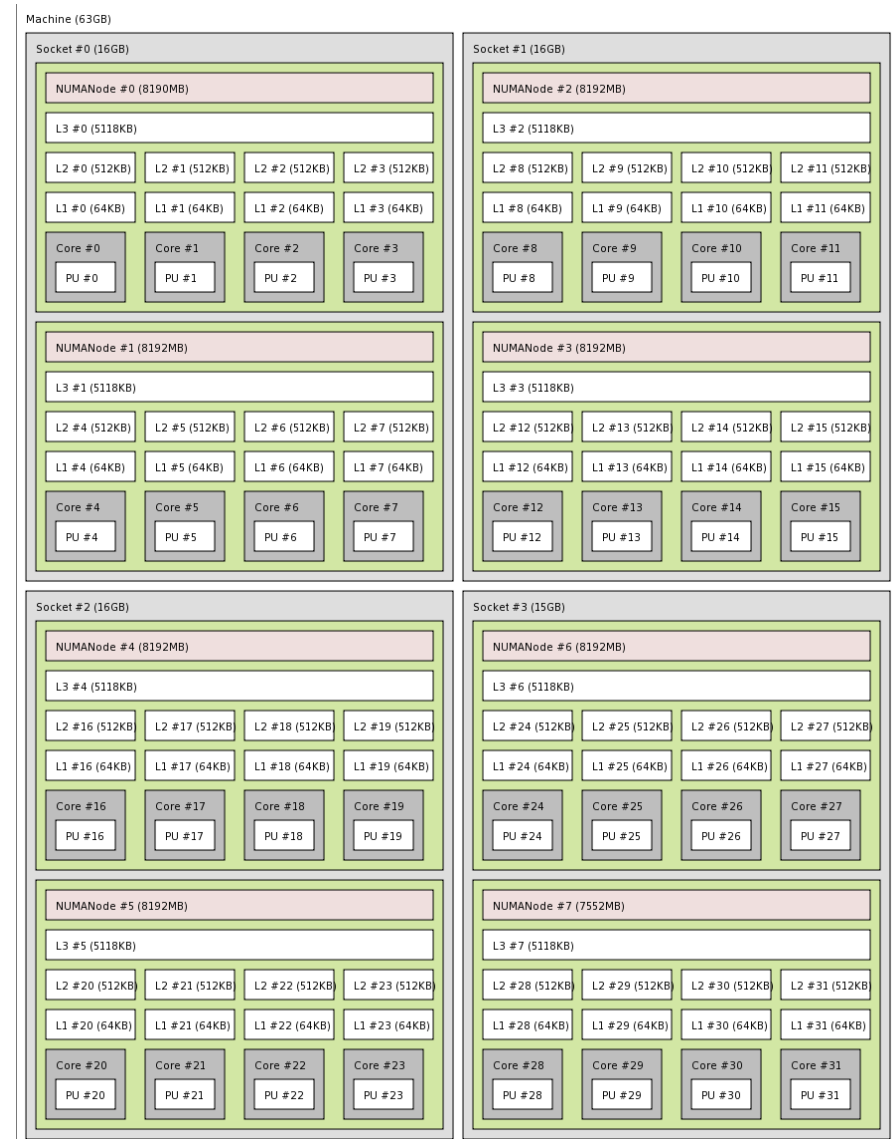
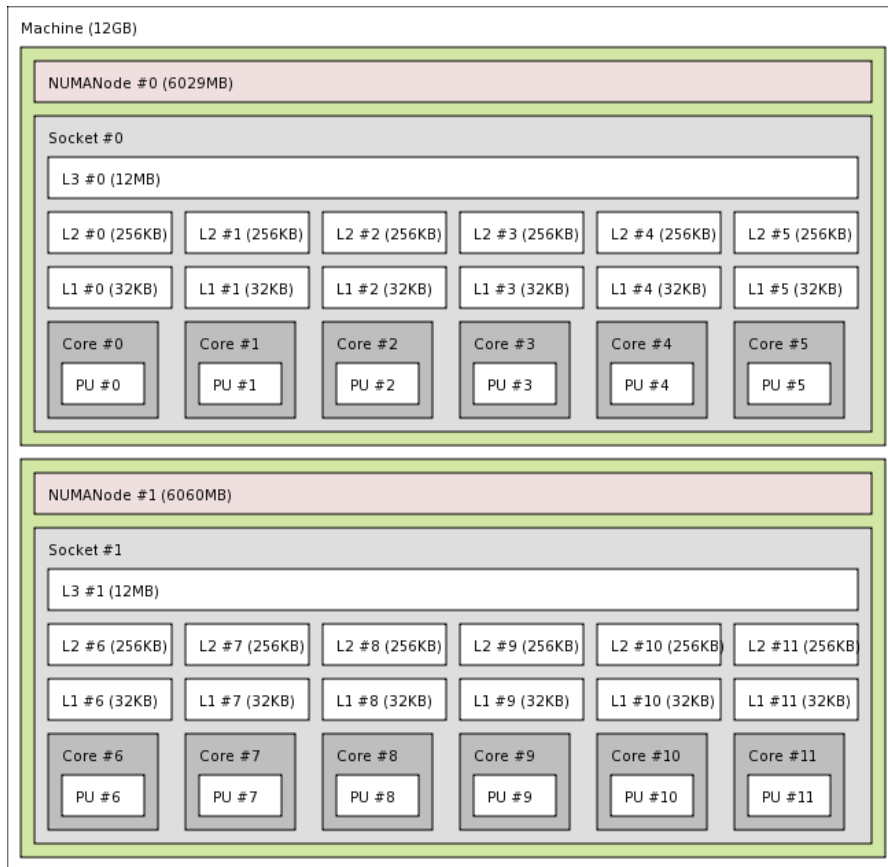
Backup Slides

Intel Xeon vs. AMD Opteron

- NUMA memory nodes, sockets, shared caches cores

Xeon X5650

Opteron 6128HE



Exclusive time: Intel Xeon

- OSS compare: 1 threads vs. 12 thread (% of time)
 - Experiment with 1028 events of 50 GeV pions (10.2.r06)

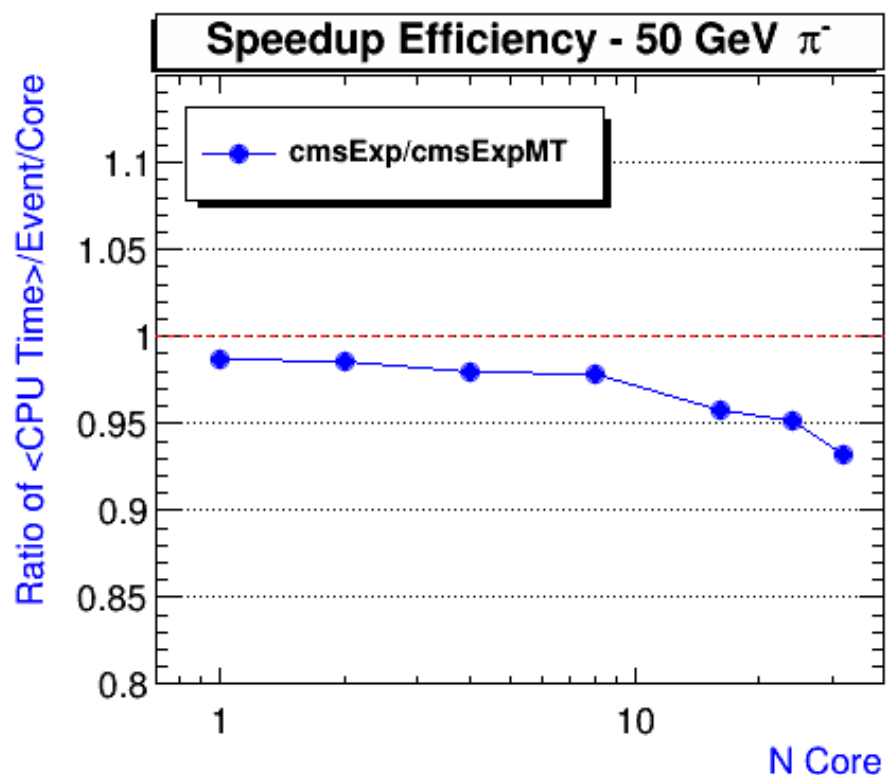
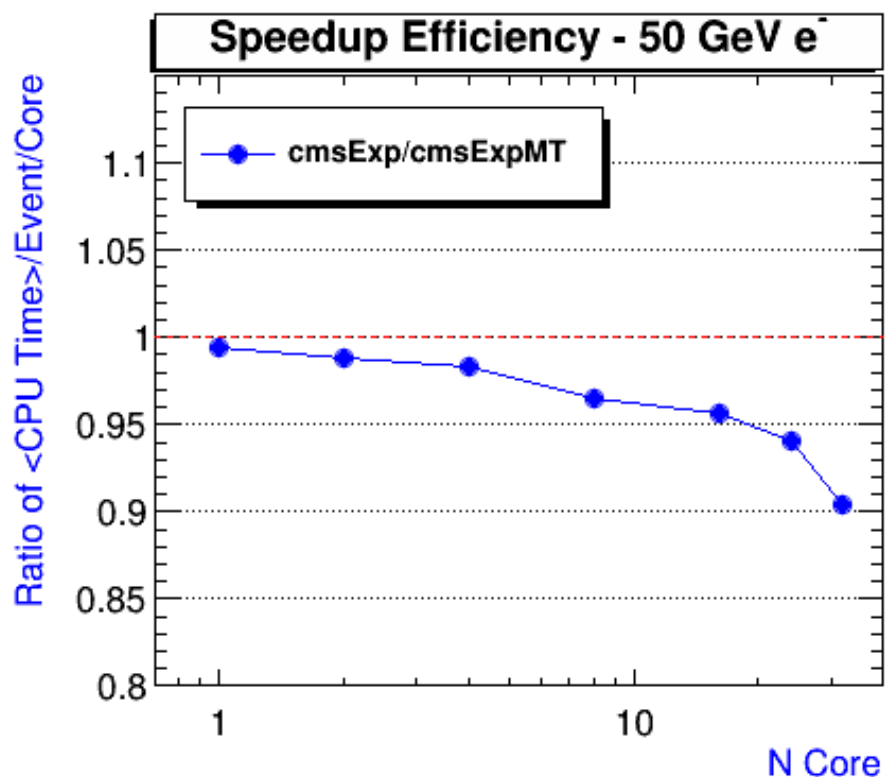
```
[openss]: Legend: -c 2 represents db_intel_pi-_50_12/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_intel_pi-_50_1/cmsExpMT-pcsamp.openss
-c 2, %   -c 4, %   Function (defining location)
CPU Time  CPU Time
6.238938  6.190079  G4PhysicsVector::Value(double, unsigned long&) const (l:
4.018038  3.899106  cmsExpMagneticField::GetVolumeBaseBfield(double const*,
3.537249  3.309547  G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vector
2.378695  2.335641  __ieee754_atan2 (libm-2.12.so)
1.735548  1.707284  G4VDiscreteProcess::PostStepGetPhysicalInteractionLength
1.607452  1.568036  G4SteppingManager::DefinePhysicalStepLength() (libG4trac
1.531425  1.484518  G4Navigator::ComputeStep(CLHEP::Hep3Vector const&, CLHEP
1.474424  1.404688  G4SteppingManager::Stepping() (libG4tracking.so: G4Step
1.456177  1.420893  G4VoxelNavigation::ComputeStep(CLHEP::Hep3Vector const&
1.410107  1.336284  G4VEmProcess::PostStepGetPhysicalInteractionLength(G4Tra
```

t12 t1

- No changes in call paths
- No significant timing perturbation (a good sanity check!)

AMD

- OSS compare: AMD 32 threads vs. 1 thread (% of time)
 - Persistency in difference (by version, by different samples)?
 - Experiments with 1028 events of 50 GeV e^- and π^- (10.2.r07)



Degradation on AMD: Persistency

- OSS compare: AMD 32 threads vs. 1 thread (% of time)
 - Experiment with 1028 events of 50 GeV pi- (10.2.r07)

```
[openss]: Legend: -c 2 represents db_amd_pi-_50_32/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_amd_pi-_50_1/cmsExpMT-pcsamp.openss
-c 2, % -c 4, % Function (defining location)
CPU Time CPU Time
6.003521 5.902264 G4PhysicsVector::Value(double, unsigned long&) const
3.952595 3.383990 G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vec
3.684048 3.856901 cmsExpMagneticField::GetVolumeBaseBfield(double const
2.421680 2.385093 __ieee754_atan2 (libm-2.12.so)
2.024112 1.030340 G4TouchableHistory::GetVolume(int) const (libG4digits
1.877350 1.926192 G4SteppingManager::DefinePhysicalStepLength() (libG4t
1.762748 1.557561 G4VoxelNavigation::ComputeStep(CLHEP::Hep3Vector cons
1.583220 1.659335 G4VEmProcess::PostStepGetPhysicalInteractionLength(G4
1.533902 1.411287 G4SteppingManager::Stepping() (libG4tracking.so: G4St
1.375139 1.416721 G4SteppingManager::InvokePSDIP(unsigned long) (libG4t
```

- G4Navigator::LocateGlobalPointAndSetup is perturbative (consistent with 10.2.r06)

Degradation on AMD: Persistency

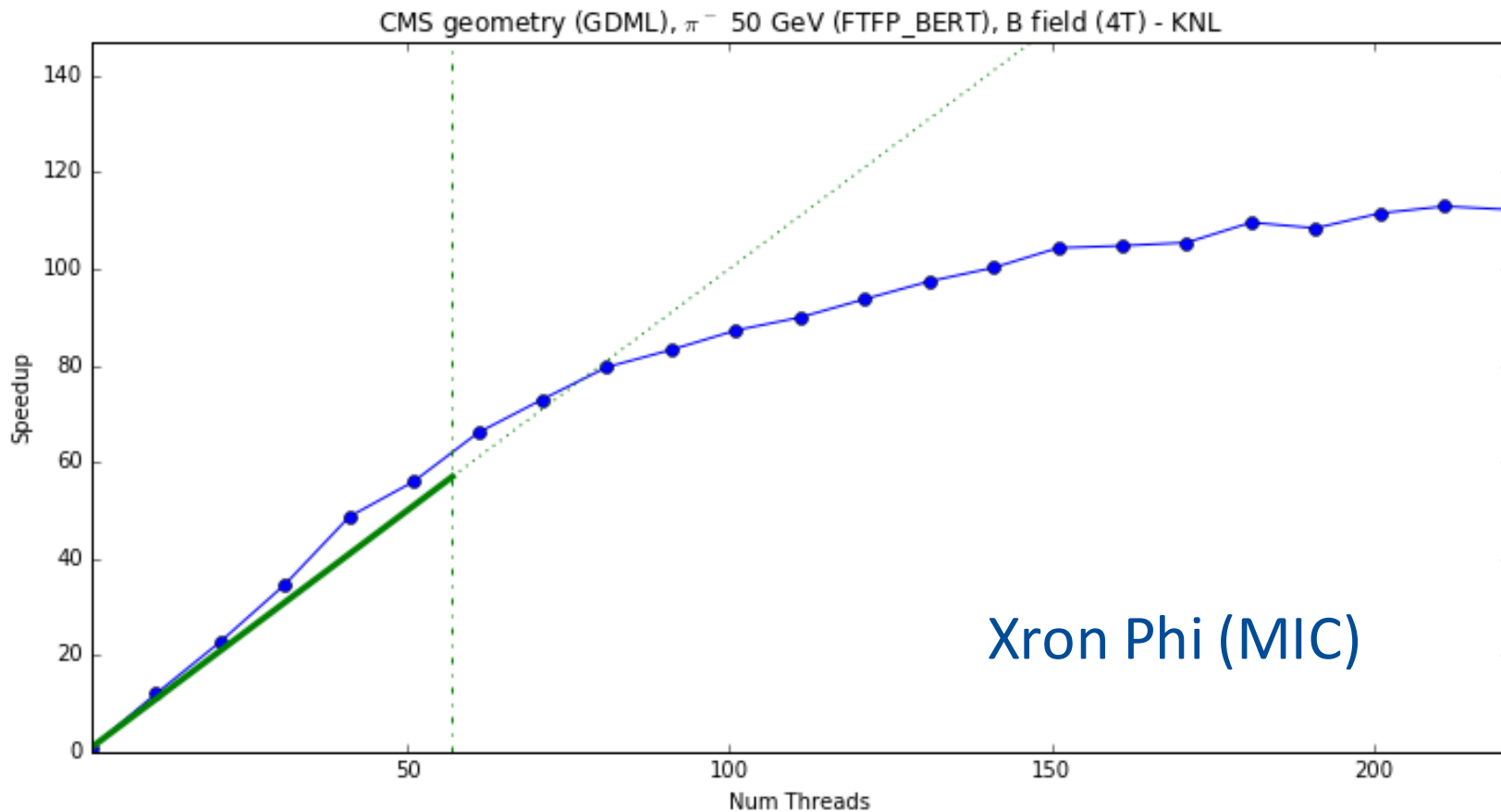
- OSS compare: AMD 32 threads vs. 1 thread (% of time)
 - Experiment with 1028 events of 50 GeV e- (10.2.r07)

```
[openss]: Legend: -c 2 represents db_amd_e-_50_32/cmsExpMT-pcsamp.openss
[openss]: Legend: -c 4 represents db_amd_e-_50_1/cmsExpMT-pcsamp.openss
-c 2, % -c 4, % Function (defining location)
CPU Time CPU Time
6.164769 6.643990 cmsExpMagneticField::GetVolumeBaseBfield(double const
5.847023 6.061244 G4PhysicsVector::Value(double, unsigned long&) const
3.995152 3.008473 G4Navigator::LocateGlobalPointAndSetup(CLHEP::Hep3Vec
2.851305 2.983996 G4UniversalFluctuation::SampleFluctuations(G4Material
2.551915 1.009077 G4TouchableHistory::GetVolume(int) const (libG4digits
2.451316 2.591272 G4VEmProcess::PostStepGetPhysicalInteractionLength(G4
1.982036 2.251656 G4SteppingManager::DefinePhysicalStepLength() (libG4t
1.922636 1.875131 G4UrbanMscModel::SampleCosineTheta(double, double) (l
1.785764 1.304073 CLHEP::RanecuEngine::flat() (libG4clhep.so: RanecuEng
1.664950 1.822590 G4PropagatorInField::ComputeStep(G4FieldTrack&, doubl
```

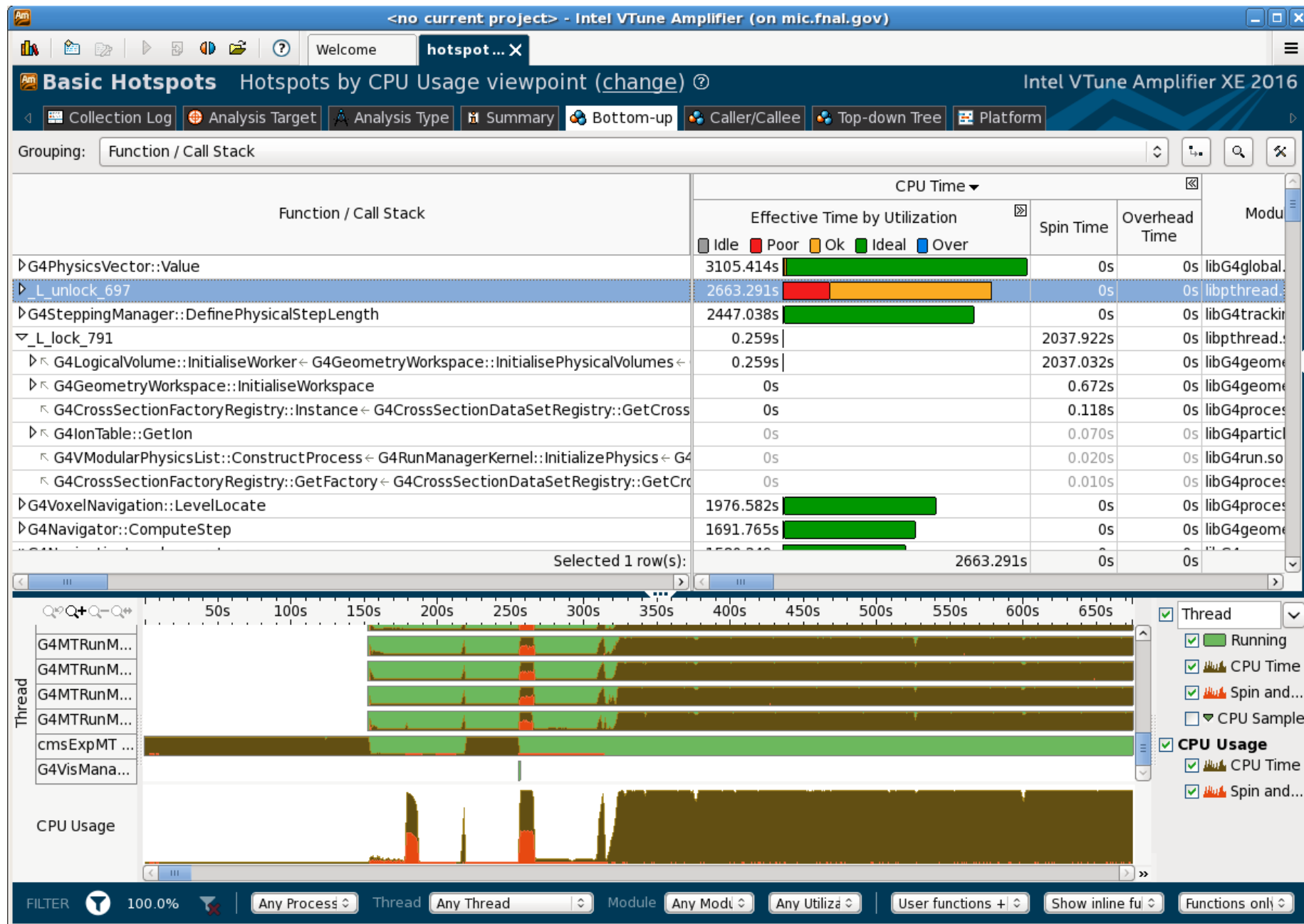
- G4TouchableHistory::GetVolume

Geant4 MT Performance: Xeon Phi (Andrea Dotti)

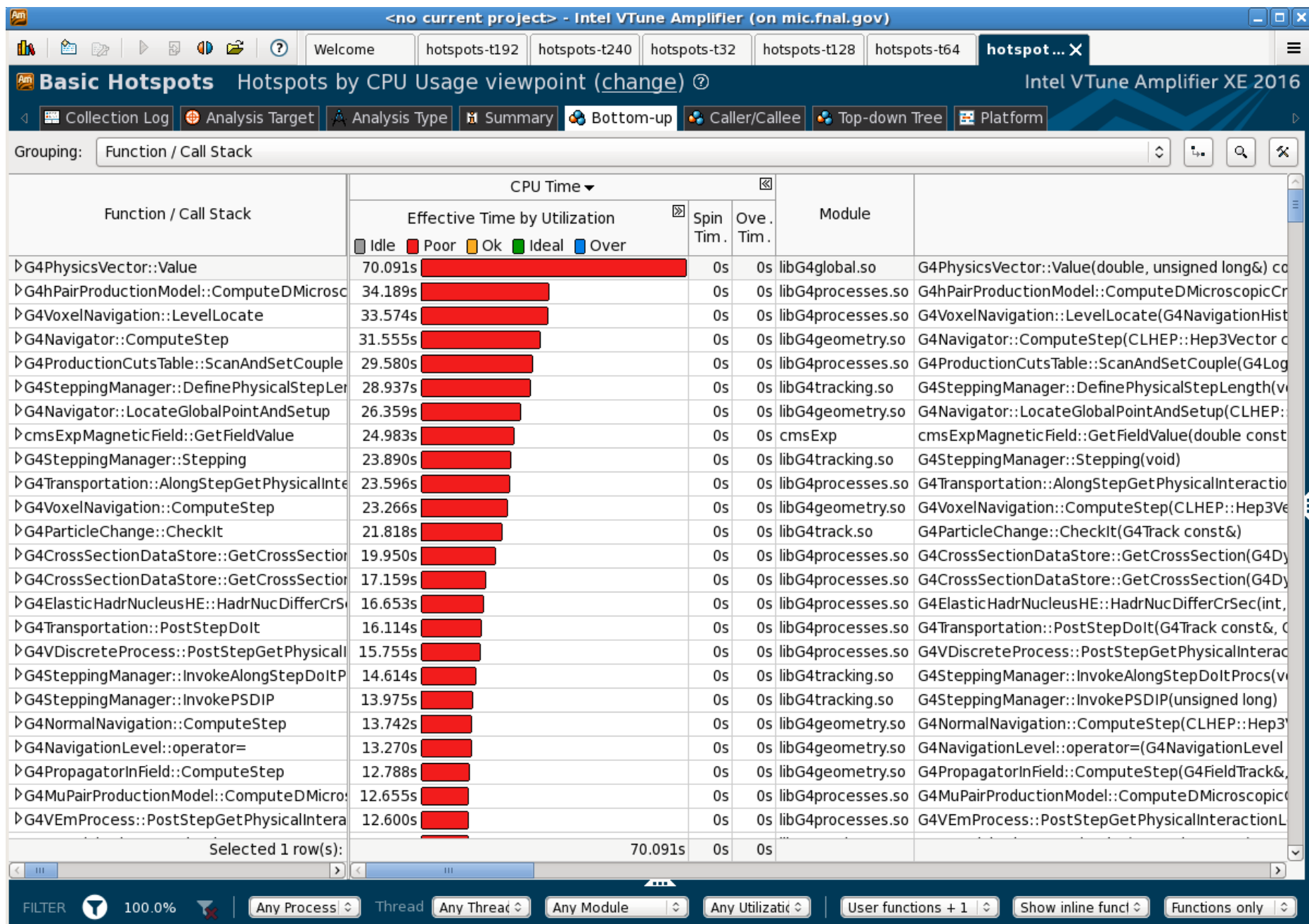
- CMS geometry, **uniform (4T) B-field** (10.2.r06)
- Total number of events processed = **10***(number of threads)
- Intel Xeon Phi 3210A (57 cores), 6GB



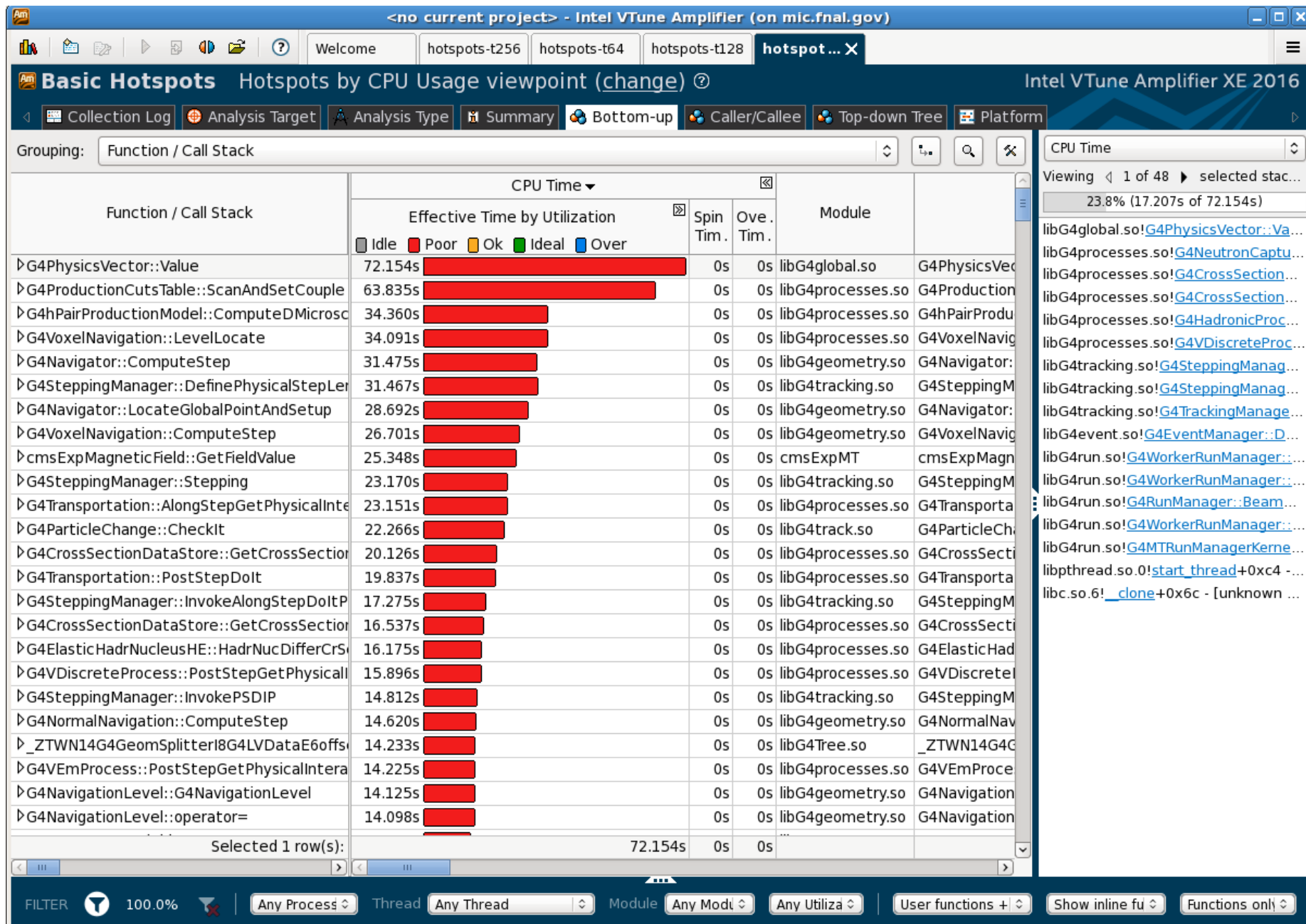
Geant4MT: Profiling Result (N threads = 256)



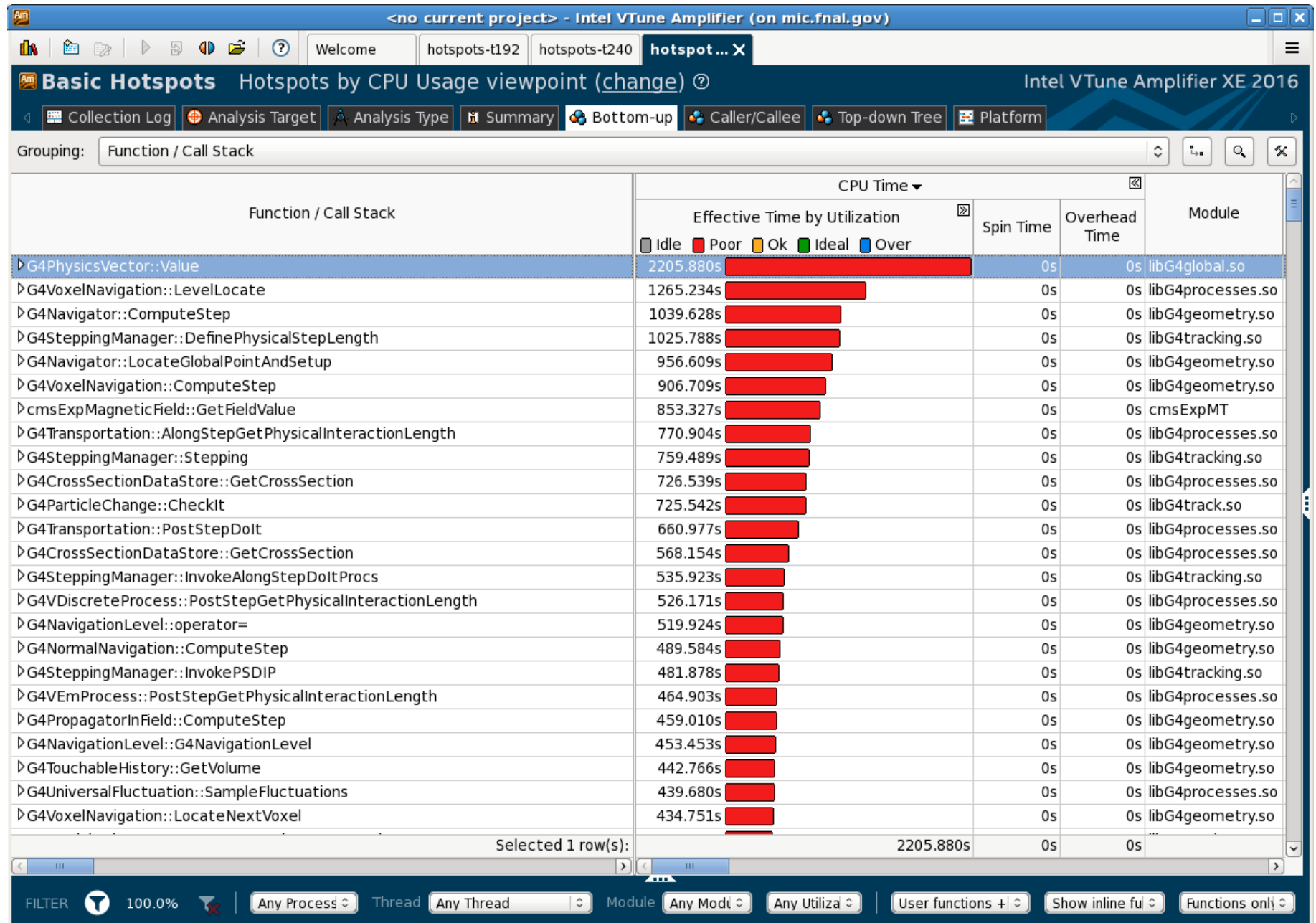
Geant4MT: Profiling Result (Sequential)



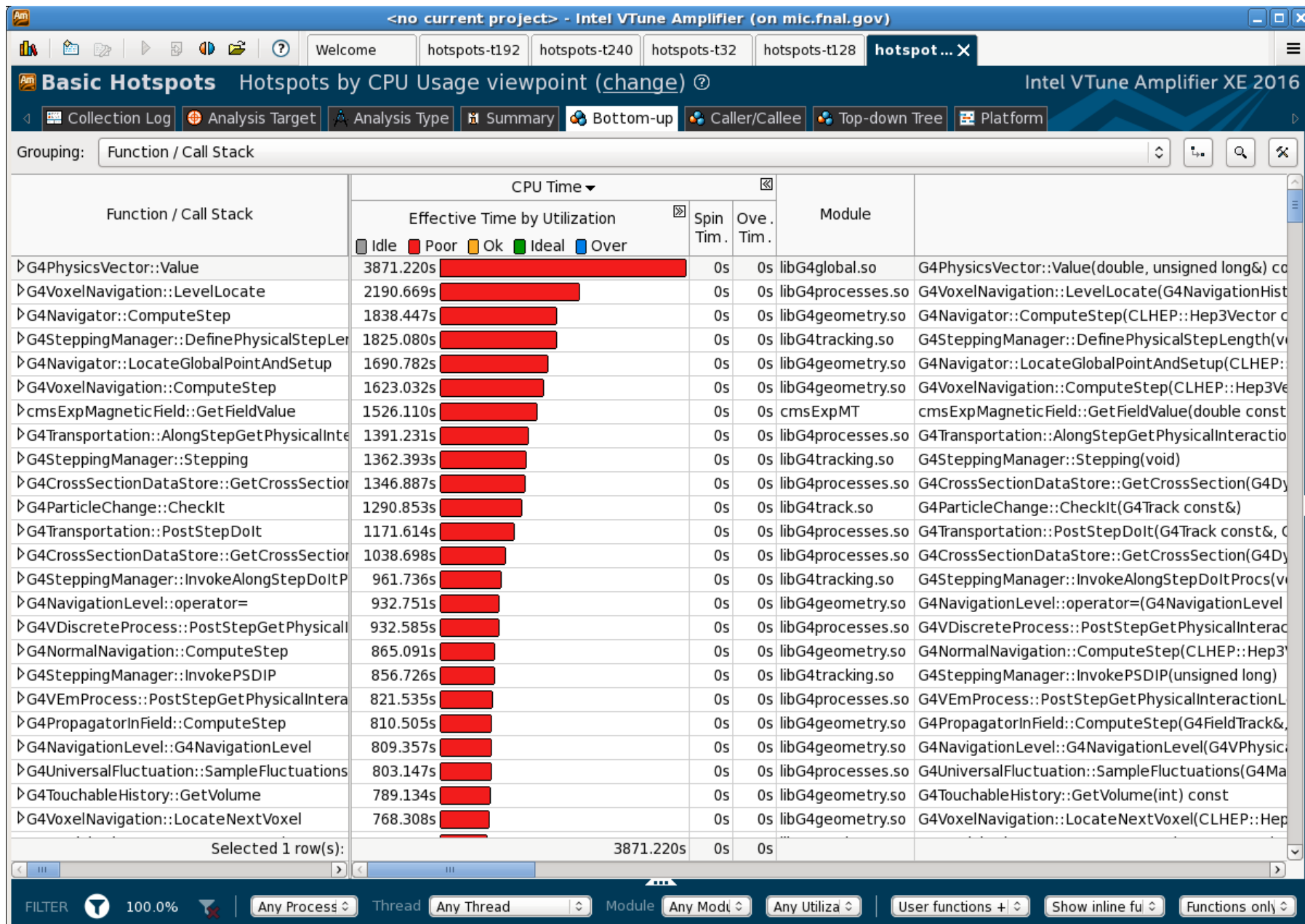
Geant4MT: Profiling Result (N thread = 1)



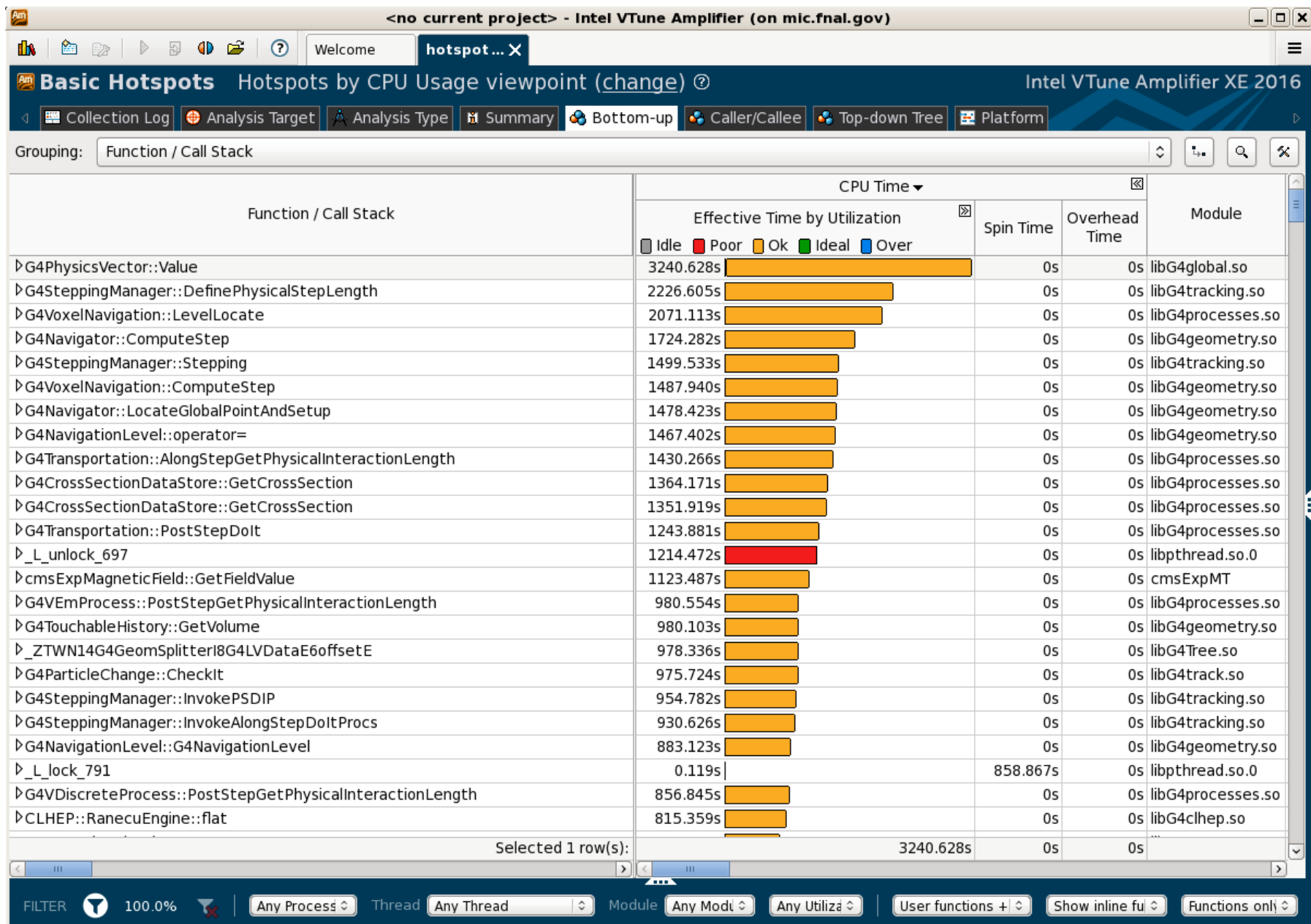
Geant4MT: Profiling Result (N threads = 32)



Geant4MT: Profiling Result (N threads = 64)



Geant4MT: Profiling Result (N threads = 192)



Geant4MT Performance on KNL: icc (16.0.3) vs. gcc (4.9.1)

- Performance on Intel Xeon Phi Processor (Knight's Landing)
 - KNL triples both scalar and vector performance compared with KNC and offers, up to 3.0 TFlop/sec (double) per processor.

