

Tools/APS/example_dgemm

Example Intel Application Performance Snapshot: dgemm

- Prepare environment

```
module purge
module add \
    compiler/intel/2022 \
    numlib/mkl/2022
```

- Build dgemm benchmark

```
icc -O2 -qopenmp -xHost -ipo \
    -DUSE_MKL -lmkl_intel_lp64 -lmkl_intel_thread -lmkl_core \
    timing.c stats.c matrix_common.c dgemm.multithread.c -o dgemm
```

- Set up APS environment on cluster HoreKa

```
# Standalone
source /software/all/toolkit/Intel_OneAPI/vtune/latest/apsvars.sh
```

```
# or as part of Intel VTune
module add devel/vtune/2023
```

- Set up APS environment on cluster BwUniCluster 2.0

```
source /software/all/toolkit/Intel_OneAPI/vtune/latest/apsvars.sh
```

- Run benchmark dgemm with APS

```
export OMP_NUM_THREADS=76
export MKL_NUM_THREADS=76
export KMP_AFFINITY="verbose,granularity=core,respect,scatter"
```

```
aps dgemm -n 8000
```

Number of repetitions set to 30. Overwrite with command line option -m.

Matrix size: 8000

Repeat multiply 30 times.

Alpha = 1.000000

```

Beta =      1.000000
Allocating Matrices...
Allocation complete, populating with values...
Performing multiplication...
Calculating matrix check...

```

```

=====
|| E ||_∞:      0.000000E+00
-> Solution check PASSED successfully.
Memory for Matrices: 1464.843750 MB
Multiply time:      6.406013 seconds
FLOPs computed:     3072384000000.000000
Min GFLOP/s:        4469.474308 GF/s
Max GFLOP/s:        4930.761834 GF/s
Average GFLOP/s:    4799.236530 GF/s
Std. dev. GFLOP/s:  657.393548 GF/s
Median GFLOP/s:     4846.325150 GF/s
MAD GFLOP/s:        28.129912 GF/s
=====

```

Intel(R) VTune(TM) Profiler 2023.1.0 collection completed successfully. Use the "aps --

- Generate APS report

```
aps --report <...>/aps_result_20230526
```

Loading 100.00%

| Summary information

```

|-----
Application                  : dgemm
Report creation date         : 2023-05-26 11:30:53
OpenMP threads number per Process: 76
HW Platform                  : Intel(R) Xeon(R) Processor code named Icelake
Frequency                    : 2.39 GHz
Logical core count per node  : 152
Collector type               : Driverless Perf per-process counting
Used statistics              : <...>/aps_result_20230526

```

| Your application might underutilize the available logical CPU cores
| because of insufficient parallel work, blocking on synchronization, or too much I/O.

```

|
Elapsed Time:                6.61 s
SP GFLOPS:                   0.00
DP GFLOPS:                   4603.66
Average CPU Frequency:       2.61 GHz
IPC Rate:                    2.57
Serial Time:                 0.55 s      8.27% of Elapsed Time
OpenMP Imbalance:            0.00 s      0.03% of Elapsed Time

```

Physical Core Utilization:	91.50%
Average Physical Core Utilization:	69.50 out of 76 Physical Cores
Memory Stalls:	17.60% of Pipeline Slots
Cache Stalls:	8.40% of Cycles
DRAM Stalls:	2.90% of Cycles
Average DRAM Bandwidth:	N/A
Data for this metric is not collected since it requires system-wide	
performance monitoring. Make sure the sampling driver is properly installed on	
your system: https://software.intel.com/en-us/vtune-amplifier-help-sep-driver .	
Otherwise, enable a driverless Perf-based sampling collection by setting the	
/proc/sys/kernel/perf_event_paranoid value to 0 or less.	
NUMA:	49.10% of Remote Accesses
A significant amount of DRAM loads was serviced from remote DRAM. Wherever	
possible, consistently use data on the same core, or at least the same	
package, as it was allocated on.	
Vectorization:	100.00%
Instruction Mix:	
SP FLOPs:	0.00% of uOps
DP FLOPs:	100.00% of uOps
Packed:	100.00% from DP FP
128-bit:	0.00%
256-bit:	0.00%
512-bit:	100.00%
Scalar:	0.00% from DP FP
Non-FP:	0.00% of uOps
FP Arith/Mem Rd Instr. Ratio:	3.77
FP Arith/Mem Wr Instr. Ratio:	401.67
Memory Footprint:	
Resident:	1577.00 MB
Virtual:	7031.00 MB

Graphical representation of this data is available in the HTML report: <...>/aps_report