Advanced Computer Architecture

Monsoon 2023, IIIT-H
Suresh Purini

1. On vector instruction sets
   (a) Write a scalar program to compute dot product of two vectors.
   (b) Write a vectorized version of the dot product program.
   (c) Write a program using vector instructions to compute the dot product of elements at even indices ($\Sigma_{i=0}^{n/2} x_{2i} * y_{2i}$).
   (d) Write a program to compute the memory bandwidth of your system.
   (e) Compute peak GFlops/sec for your system.
   (f) For the above problems a), b), c), check if the programs are memory bound or compute bound.
2. on memory hierarchies and performance counters
   (a) Write a matrix-matrix ($N \times N$) multiplication program using the usual 3d-loop and recursive strassen multiplication.
   (b) Using performance monitors compute the number of instructions executed, CPI, L1/L2/L3 cache hit ratios for various matrix dimensions $N = 2^k, 4 \leq k \leq 14$.
   (c) Write a tiled version of matrix multiplication. Using cache grind for 32 KB L1 cache, 1 MB L2 cache, line size64 bytes, find the best tile size which gives best hit ratio for matrix dimension N = 2048. Compare the hit ration using recursive matrix multiplication. Compare the instruction count in both the cases.
3. For the following C-program using cache grind find the optimal cache configuration for a 2-level inclusive fully-associative LRU cache hierarchy (default read and write policies) when input size n is 4000 the following input sizes (Line size: 64 Bytes)

```
for (i = 0; i < n - 1; i++) {
    S1: A[i] = 0
    for (j = 1; j < n - 1; j++)
S2: A[i] += 0.33*(B[i][j] * X[j])
}
```

   Plot a 3D graph with x-axis as L1 cache size, y-axis as L2 cache size and z-axis as miss ratio.