worksheet.md

CSE141L Lab 3 Caching Optimizations

Name: _____ Student ID: _____

Instructions

- Complete this worksheet while reading/working through the lab write up. The worksheet doesn't make sense without the lab.
- The point values are listed for each question. Altering the size of the cells will cost you 1 point. There are 75 points total for the write up portion of the lab.

Cache and dataset characteristics

P1 (4pt) Find out the dimensions (number of data elements) of the following tensors/vectors used in `fc_layer_t::activate` for the cifar100 dataset and fill the following table

Tensor/Vector	Number of Data Elements
`in`	_____
`out`	_____
`weights`	_____
`activation_input`	_____

P2 (4pt) Calculate the size (in Bytes) of the following tensors/vectors used in `fc_layer_t::activate` for the cifar100 dataset and fill the following table

Tensor/Vector	Size in Bytes
`in`	_____
`out`	_____
`weights`	_____
`activation_input`	_____

P3 (4pt) Refer to the lecture slides to find the cache sizes for the skylake processor we are using for this class.

Cache level	Size in Bytes
L1 Cache	_____
L2 Cache	_____
L3 Cache	_____

P4 (4pt) How much of each of these data structures used in `fc_layer_t::activate()` will fit in the L1 and L2 cache?

tensor	% that'll fit in L1	% that'll fit in L2

tensor	% that'll fit in L1	% that'll fit in L2
`in`	_____	_____
`out`	_____	_____
`weights`	_____	_____
`activation_input`	_____	_____

Understanding Tensor_t

Given `tensor_t<double> foo(tdsize(4,3,5,7))` , answer the following (double are 8 bytes):

P1 (1pt) How many elements are there in `foo` ?

P2 (1pt) What's the linear index of element (1,1,1,1)?

P3 (1pt) How far apart are elements that differ by 1 in each dimension?

dim.	distance in bytes	distance in linear index
x		
y		
z		
b		

Tier 1: Reordering and Tiling loops in `fc_layer_t::activate`

P1 (4pt) Change the order of loops from `b i n` to `b n i` in `fc_layer_t::activate` and report the speedup.

Speedup after loop reordering : _____

P2 (4pt) Block the loop `n` in `fc_layer_t::activate` with the tile sizes 1, 2, 4, 8, 16 and fill out the table below.

Dataset	Step size	Blocked implementation time	Speedup vs step size == 1
cifar100	1	_____	_____
cifar100	2	_____	_____
cifar100	4	_____	_____
cifar100	8	_____	_____
cifar100	16	_____	_____

P3 (4pt) In a single line graph, plot the speed up against the different block sizes for blocking the loop `n` in `fc_layer_t::activate` . Block size is the independent vairable.

```
Your graph here
```

P4 (4pt) Consider the blocksize which gave maximum speedup in the previous question P4 and fill out the following table

1. Base implementation time : _____
2. Implementation time of your optimized solution : _____
3. Base implementation L1 misses : _____
4. Your fastest solution L1 misses : _____

Tier 2: Optimizing calc_grads

P1 (4pt) Change the order of loops from `b i n` to `b n i` in the the triply-nested loop in `fc_layer_t::calc_grads` and report the speedup.

Speedup after loop reordering : _____

P2 (4pt) Block loop `n` in the the triply-nested loop in `fc_layer_t::calc_grads` with different step sizes and fill out the following table.

Function	Step size	Base implementation time	Blocked implementation time	Speedup
calc_grads	_____	_____	_____	_____
calc_grads	_____	_____	_____	_____
calc_grads	_____	_____	_____	_____
calc_grads	_____	_____	_____	_____
calc_grads	_____	_____	_____	_____

P3 (4pt) In a single line plot, plot performance vs. block size for blocking the loop `n` in the the triply-nested loop in `fc_layer_t::calc_grads` and return block size that gives maximum speedup. Block size is the independent vairable.

```
Your graph here
```

Best block size : _____

Tier 3: Applying More Optimizations

P1 (5pt) Give a brief description of two additional loops you tried blocking. Report the speedup you achieved for each one.

```
Your answer here
```

P2 (5pt) Give a brief description of an additional optimization you implemented to speedup training.

```
Your answer here
```

Lab Reflection

Follow this link 24 hours before or after the due date to fill out the reflection survey. It is worth 5% of your lab grade.

```
https://forms.gle/VUkpAdC6gUQC94Fd8
```

Addendum

(2pt) For your best version of each function, write out loop

```
nesting order and tile sizes you used.

For example, for
`do_stabilize_tile_y_1()` in `example/stabilize.cpp` you would
write.


        this_frame, pixel_yy[2], offset_x, offset_y, pixel_x, pixel_y
```

The `[2]` means that the `pixel_yy` loop is broken into tiles of size 2.

Note There is not a correct answer, we just want to know what you did.

activate()

calc_grads()

fix_gradients()