HW3-EC Report

I didn't test the oneapi program on the cloud platform due to time constraints, but I was able to log in to my remote host and execute the job, see the last part of the screenshot. But all the test data is run local laptop.

The processor of Laptop: intel core i7-1165G7(Products formerly Tiger Lake)

Using the below website, we got the result of the processor used.
https://www.intel.com/content/www/us/en/products/sku/208921/intel-core-i71165g7-processor-12m-cache-up-to-4-70-ghz-with-ipu/specifications.html

System: Ubuntu 22.0.4 LTS
Core: 4, sockets: 1

Mode	Type	Size	Threads	OneAPIThroughput (GFlops/s)	PThreadThroughput (GFlops/s)	HPLThroughput (GFlops/s)
matrix	single	small	4	1.7430	9.09930	
matrix	single	medium	4	1.7625	8.03652	
matrix	single	large	4	1.9238	7.37670	
matrix	double	small	4	1.3034	9.93898	4.3642
matrix	double	medium	4	1.3931	10.26627	19.09
matrix	double	large	4	1.9432	8.73408	77.141
matrix	single	small	N	4.2680	8.41014	
matrix	single	medium	N	7.4198	8.06341	
matrix	single	large	N	3.4950	6.0308	
matrix	double	small	N	3.5612	11.4837	
matrix	double	medium	N	3.3867	10.5635	
matrix	double	large	N	1.6345	8.8881	

For my laptop, N = 8.

Due to assignment's time constraints, I was not able to retest the performance of HPL on my laptop, but from the data available, it can be seen that the number of threads using hardware is better than manually specifying the number of threads. For HPL benchmark, the throughout is better than my implementation of matrix multipilition.

In addition, from the data, we can see that the performance of the simple version of oneapi is worse than our previous pthread version of matrix multiplication, mainly because we are using the simplest loop parallel scheme, if we want to do optimization, we should consider using ND-Range kernel, which can divide the instances into different types of groups and map them precisely to the hardware platform, thus improving performance.

Design

1.Complier
It needs to be executed each time you start the command line(**source /opt/intel/oneapi/setvars.sh)**, and the compiled code of the makefile is modified for the dpc++ compiler.

Makefile M

```
CC=icpx
CFLAGS=-fsycl

build: cpubench

test: cpubench
	bash runbench.sh

cpubench: cpubench.cpp
	$(CC) $(CFLAGS) $< -o cpubench

clean:
	rm -rf cpubench


#  source /opt/intel/oneapi/setvars.sh
# icpx -fsycl ./matrix.cpp -o ./matrix
```

2.Thread

The test file is **. /perftest.sh**. When the thread parameter is -1, the optimal number of threads is started, and the simplest version of the DPC++ parallel_for function is used, where oneAPI gives the optimal thread solution based on the current hardware situation.

When the thread count is 4, we give the parallel_for function the current thread count of 4, so the internal operation uses a block-optimized version of matrix multiplication (consistent with HW3)

```
// depend on hardware
if (isAuto) {
    q.parallel_for(range(size, size), [=](auto index) {
        int i = index[0];
        int j = index[1];

        long sum = 0;
        for (int k = 0; k < size; k++) {
            sum += origin1[i][k] + origin2[k][j];
        }
        origin3[i][j] = sum;
    }).wait();
// depend on config the number of thread
} else {
    int dis = size / num_threads;
    // thread
    q.parallel_for(range<1>(num_threads), [=](id<1> index) {
        int start = index * dis;
        int end = (index + 1) * dis;
        if (end > size) end = size;
        int i, j, k;
        int bsize = 16;
        for (i = start; i < end; i += bsize) {
```

3. Automatical test shell script:

Makefile M $ perftest.sh ×

$ perftest.sh

```
1   # use 4 thread to test oneapi implementation of matrix mu
2   ./cpubench 20 single 1000 4 true
3   ./cpubench 20 single 4000 4 true
4   ./cpubench 20 single 16000 4 true
5   ./cpubench 20 double 1000 4 true
6   ./cpubench 20 double 4000 4 true
7   ./cpubench 20 double 16000 4 true
8   # use bigest hardware thread to test oneapi implementati
9   ./cpubench 20 single 1000 -1 true
10  ./cpubench 20 single 4000 -1 true
11  ./cpubench 20 single 16000 -1 true
12  ./cpubench 20 double 1000 -1 true
13  ./cpubench 20 double 4000 -1 true
14  ./cpubench 20 double 16000 -1 true
```

DEVOPS:

1. ssh to connect the oneapi cloud server.

```
Last login: Thu Nov 10 01:04:37 on ttys000
➜  ~ bash ~/Downloads/setup-devcloud-access-171569.txt
The /Users/ziyun/.ssh directory for SSH client configuration already exists.
Appending SSH connection configuration to /Users/ziyun/.ssh/config
Creating the private SSH key /Users/ziyun/.ssh/devcloud-access-key-171569.txt
Important: this file is your access key, keep it safe like you would a password.
Done! Now you can access Intel DevCloud by running ssh devcloud or, if you are behind a proxy, ssh devcloud.proxy
➜  ~ ssh devcloud
The authenticity of host 'ssh.devcloud.intel.com (12.229.61.118)' can't be established.
ED25519 key fingerprint is SHA256:/Dlip01tdMyRmhMDc870Z4Uk7AancwwoTnbb0EZajK0.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ssh.devcloud.intel.com' (ED25519) to the list of known hosts.
The authenticity of host 'devcloud (<no hostip for proxy command>)' can't be established.
ED25519 key fingerprint is SHA256:lS/2XMIIjUjLdbBOVwX+d0HoUE7o/Vp+YH+vIyxO2XQ.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'devcloud' (ED25519) to the list of known hosts.
################################################################################
#
# Welcome to the Intel DevCloud for oneAPI Projects!
#
# 1) See https://devcloud.intel.com/oneapi/ for instructions and rules for
# the OneAPI Instance.
#
# 2) See https://github.com/intel/FPGA-Devcloud for instructions and rules for
# the FPGA Instance.
#
# Note: Your invitation email sent to you contains the authentication URL.
#
# If you have any questions regarding the cloud usage, post them at
# https://software.intel.com/en-us/forums/intel-devcloud
#
# Intel DevCloud Team
#
```

2. Config the remote container ssh-key to my github, and git clone the hw3-ec code. And execute setvars.sh to pre-config current running environment.

```
u171569@login-2:~$ cd ~/.ssh/
u171569@login-2:~/.ssh$ ls
authorized_keys  id_rsa  id_rsa.pub
u171569@login-2:~/.ssh$ cat id_rsa.pub
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABgQC4iygqpvH1e9pw+FoZ/JK24hR6bSSvyK10t8j8EAnZks0jZkZzE+QR431VyMr
8II7TJwZbUuff2mZO4Ypg8+h8eFOHZJNHyVhve+kxBeostYuQpx3tl2leNYQ66yO0n9Byh/1IYtyVQnw/81GIzHJMHoSTCzZj/Y
A1u4fc8MhNfDquIJLfNCfXVxFNv++NGP9SwuhpDjxH0Ug4UbwXoRKrvBjMQKqibgbwoJCIhKzO6YZd68qlS3LhlWMvSdwER5RQe
d4REcpbIDW79n+h1k=
u171569@login-2:~/.ssh$ cd ../
u171569@login-2:~$ git clone git@github.com:datasys-classrooms/cs553-fall2022-hw3-ec-yz-msq-bt.git
Cloning into 'cs553-fall2022-hw3-ec-yz-msq-bt'...
remote: Enumerating objects: 30, done.
remote: Counting objects: 100% (30/30), done.
remote: Compressing objects: 100% (25/25), done.
remote: Total 30 (delta 7), reused 27 (delta 5), pack-reused 0
Receiving objects: 100% (30/30), 14.37 KiB | 1.60 MiB/s, done.
Resolving deltas: 100% (7/7), done.
u171569@login-2:~$ cd cs553-fall2022-hw3-ec-yz-msq-bt/
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ ./cpubench 20 single 1000 -1 true
-bash: ./cpubench: No such file or directory
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ make
icpx -fsycl  cpubench.cpp -o cpubench
make: icpx: Command not found
Makefile:10: recipe for target 'cpubench' failed
make: *** [cpubench] Error 127
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ source /opt/intel/oneapi/setvars.sh

:: initializing oneAPI environment ...
   -bash: BASH_VERSION = 4.4.20(1)-release
   args: Using "$@" for setvars.sh arguments:
:: advisor -- latest
:: ccl -- latest
:: clck -- latest
:: compiler -- latest
:: dal -- latest
:: debugger -- latest
:: dev-utilities -- latest
:: dnnl -- latest
:: dpcpp-ct -- latest
:: dpl -- latest
:: embree -- latest
:: inspector -- latest
:: intelpython -- latest
:: ipp -- latest
:: ippcp -- latest
:: ipp -- latest
:: ispc -- latest
:: itac -- latest
```

3. How to run job on the cloud platform

```
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ git checkout -- .
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ git pull
Updating 507acd3..8f78051
error: The following untracked working tree files would be overwritten by merge:
        job.sh
Please move or remove them before you merge.
Aborting
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ rm -rf job.sh
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ git pull
Updating 507acd3..8f78051
Fast-forward
 job.sh | 15 +++++++++++++++
 1 file changed, 15 insertions(+)
 create mode 100644 job.sh
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ qsub -l nodes=1:gpu:ppn=2 -d . job.sh
2035764.v-qsvr-1.aidevcloud
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ watch -n 1 qstat -n -1
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ cat job
job-result.md    job.sh           job.sh.e2035755  job.sh.e2035764  job.sh.o2035755  job.sh.o2035764
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ cat job-result.md
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ cat job-result.md
u171569@login-2:~/cs553-fall2022-hw3-ec-yz-msq-bt$ cat job.sh.e20357
```