CompSci 590.03: Introduction to Parallel Computing

Homework 5 - MPI Programming (11/11/15)

Lecturer: Alvin R. Lebeck *Scribes: Mengke Lian, Kai Fan, Chaoren Liu*

Contribution distribution:

- Mengke Lian: finish all MPI serial, OpenMP, offload versions
- Kai Fan: mainly working on the offload version.
- Chaoren Liu: implement the matmul_offload_mpi.c version.

The dimensionality of square matrix multiplication is set to **8192**.

The baseline result for MPI only program with `-np 4 -ppn 1` is

```
Start copy data to all machines
Finish copy data to all machines, communication takes 0.591476 sec
Start gather data from all machines
Finish gather data from all machines, communication takes 0.070078 sec
Order 8192 multiplication in 131.094273 seconds
```

The result for MPI only program with `-np 128 -ppn 32` is

```
Start copy data to all machines
Finish copy data to all machines, communication takes 7.508244 sec
Start gather data from all machines
Finish gather data from all machines, communication takes 4.786869 sec
Order 8192 multiplication in 10.430095 seconds
```

The result for OpenMp MPI program with `-np 4 -ppn 1` is

```
Start copy data to all machines
Finish copy data to all machines, communication takes 0.549187 sec
Start gather data from all machines
Finish gather data from all machines, communication takes 0.063677 sec
Order 8192 multiplication in 6.989820 seconds
```

The result for offload version OpenMp MPI program (**half-half partition**) with `-np 4 -ppn 1` is

```
Start copy data to all machines
Finish copy data to all machines, communication takes 0.556785 sec
Start gather data from all machines
Finish gather data from all machines, communication takes 0.671248 sec
Order 8192 multiplication in 18.101347 seconds
```

Discussion

For MPI only program, more machines in the setting will cause more communication time, but will potentially reduce the multiplication time. To find the optimal total time, there exists a trade between the communication and program computation.

Table 1: Running Time (sec) on Different Implementation

Implementation	**Copy Time**	**Gather Time**	**Multiplication Time**	**Total**
`MPI(4)`	0.591476	0.070078	131.094273	131.75
`MPI(128)`	7.508244	4.786869	10.430095	22.71
`MPI+OpenMp`	0.549187	0.063677	6.989820	7.60
`Offload+MPI+OpenMp`	0.556785	0.671248	18.101347	19.32

For OpenMp+MPI program, it makes sense for small communication time in the case of only 4 machines, and the thread parallelism within each machine will reduce the multiplication computing time.

For offload version, we first partition the matrix `C` and scatter it into MPI machines, and then partition each part into two parts, one for offload and one for host. We observe that the offload is surprisingly slow, thus for each part of `C`, we can change the portion of the work distributed on offload and host. We found that when we only distribute 1/16 of the work to offload and distribute the rest to host, the multiplication time can reduce to 11 seconds.