

QUIZ
SK5004
JARINGAN DAN PENGOLAHAN DATA PARALEL

Mohammad Rizka Fadhli
20921004

22 December 2021

SOAL 1

Jawablah pertanyaan berikut dengan singkat dan jelas!

Soal a. Jelaskan perbedaan antara program yang dijalankan secara serial dan program yang dijalankan secara *parallel*!

Jawab Program yang dijalankan secara serial berarti program tersebut memproses *task* satu-persatu secara berurut (*sequencial*). Sedangkan program yang dijalankan secara *parallel* berarti program tersebut bisa memproses *multiple tasks* dalam satu waktu secara bersamaan menggunakan *processors* yang berbeda-beda.

Soal b. Jelaskan dengan singkat perbedaan SISD, SIMD, MISD, dan MIMD!

Jawab *Single Instruction Single Data Stream* (SISD) : suatu komputer yang tidak memiliki cara untuk melakukan paralelisasi terhadap instruksi atau data.

Single Instruction Multiple Data Stream (SIMD) : suatu komputer yang mampu memproses banyak aliran data dengan hanya satu instruksi, sehingga operasi yang dilakukan adalah operasi paralel.

Multiple Instruction Single Data Stream (MISD) : suatu komputer yang dapat melakukan banyak instruksi terhadap satu aliran data.

Multiple Instruction Multiple Data stream (MIMD) : suatu komputer yang memiliki beberapa prosesor yang bersifat *autonomous* yang mampu melakukan instruksi yang berbeda pada data yang berbeda.

Soal c. Jelaskan dengan singkat perbedaan mendasar penggunaan OpenMP dan OpenMPI untuk paralelisasi!

Jawab Salah satu perbedaan mendasarnya adalah cara menulis kode programnya. Karena OpenMP adalah cara pemrograman pada sistem *shared memory* sedangkan OpenMPI adalah cara pemrograman pada sistem *distributed memory*.

Soal d. Sebutkan modul selain mpi4py pada *python* yang dapat digunakan untuk paralelisasi!

Jawab Kita bisa menggunakan CUDA dan OpenCL.

Soal e. Selain menggunakan paralelisasi dengan MPI, kecepatan pemrosesan pada *python* dapat dilakukan dengan cara apa saja?

Jawab Kita bisa menggunakan prinsip *multi processing* memanfaatkan operasi *array* alih-alih melakukan *looping*.

Soal f. Jelaskan dengan singkat perbedaan komunikasi *point-to-point* dan kolektif pada MPI!

Jawab *Point-to-point communication* adalah bentuk komunikasi pada *message passing* di mana ada sepasang proses yang saling bekerjasama. Pada MPI, secara sederhana dapat digunakan MPI_Send atau MPI_Recv sebagai definisi proses mengirimkan dan menerima suatu data.

Collective communication adalah bentuk komunikasi antara semua proses dalam sebuah grup proses. Pada MPI, secara sederhana dapat digunakan MPI_Bcast (broadcast) yang melakukan pengiriman data dari sebuah proses ke semua proses lainnya pada grup atau MPI_Reduce untuk melakukan kebalikannya.

SOAL 2

Pertanyaan pada soal ini untuk menguji pemahaman anda tentang *server linux multiprocessor* dan teknis penggunaannya untuk menyelesaikan perhitungan numerik dengan paralelisasi menggunakan MPI.

Soal a Nama host (hostname) server yang digunakan pada kuis ini adalah _____(1), dengan alamat IP _____(2). Sistem operasi yang digunakan pada server ini adalah _____(3) versi _____(4). Server ini memiliki motherboard dengan socket sebanyak _____(5) dengan processor sebanyak _____(6) core, _____(7) thread dengan merk dagang _____(8), kecepatan processor _____(9) GHz dan cache _____(10)(MB). Kapasitas total penyimpanan data pada server ini sebesar _____(11) GB dan RAM total sebesar _____(12)MB. Program MPI yang digunakan pada server ini untuk menghitung perhitungan numerik secara paralel adalah _____(13) dengan modul _____(14) Python versi _____(15)

Jawab

1. ubuntu
2. 69.30.233.90
3. Linux Ubuntu
4. 20.04.3 LTS
5. 40 CPUs
6. 10 per socket
7. 2 per core
8. Intel(R) Xeon(R) CPU E5-2660 v2
9. 2.20 GHz
10. L2 cache: 5MB dan L3 cache: 50MB
11. 32 GB di `dev` tapi sepertinya *per user* mendapatkan *storage* sebesar 6.3 GB (*screenshot* terlampir)
12. 62.9 GB
13. Slurm
14. `mpi4py`
15. 3.8.10

Filesystem	Size	Used	Avail	Use%	Mounted on
udev	32G	0	32G	0%	/dev
tmpfs	6.3G	1.5M	6.3G	1%	/run
/dev/sda3	3.6T	1.4T	2.1T	41%	/
tmpfs	32G	0	32G	0%	/dev/shm
tmpfs	5.0M	0	5.0M	0%	/run/lock
tmpfs	32G	0	32G	0%	/sys/fs/cgroup
/dev/sda2	946M	75M	823M	9%	/boot
tmpfs	6.3G	0	6.3G	0%	/run/user/1011
tmpfs	6.3G	0	6.3G	0%	/run/user/1018
tmpfs	6.3G	0	6.3G	0%	/run/user/1012
tmpfs	6.3G	0	6.3G	0%	/run/user/1015
total	3.8T	1.4T	2.2T	40%	-

Figure 1: Screenshoot df -h --total

Soal b. Secara *default*, *server* ini belum bisa menjalankan program MPI. Apa yang harus dilakukan agar program MPI pada *server* ini dapat dijalankan?

Jawab Kita harus meng-*install* MPI terlebih dahulu di OS Linux pada *server* tersebut. Kemudian meng-*install library* `mpi4py` di Python 3 yang hendak digunakan.

Soal c. Nama program yang digunakan untuk me-*manage resource* dan mengatur antrian *job* perhitungan pada *server* ini adalah _____(1). Buatlah contoh `file.sh` untuk menjalankan perhitungan numerik secara paralel di server ini dengan sistem antrian! Perintah yang digunakan untuk mengeksekusi `file.sh` agar *job* perhitungan masuk ke sistem antrian adalah _____(2)

Jawab

1. Slurm
2. `sbatch file.sh`

Contoh `file.sh`

```
#!/bin/bash
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=720:00:00
python3 program.py
```