
Driver/Toolkit/Samples
NVIDIA

Access to system

- ssh login.rc.fas.harvard.edu
- Username “cfest###” where ### is your three digit number.
- Password is “P@Ssword!### where ### is your three digit number.
- Enter 6 digit code

Access to system

- `Login with previous instructions`
- `module load hpc/cuda-5.0.35`
- `cp -r ~jbentz/computefest .`
- Use the modules environment system
  - `module load <ModuleName>` to load a module
  - `module list` to show which modules are loaded
- Use batch system. Every job must be submitted to the queue.
  - `sbatch`—submit a job to the queue
  - `squeue`—show all jobs in the queue
  - `mdel`—delete a job from the queue
- https://rc.fas.harvard.edu/odyssey-quickstart-guide/

Software

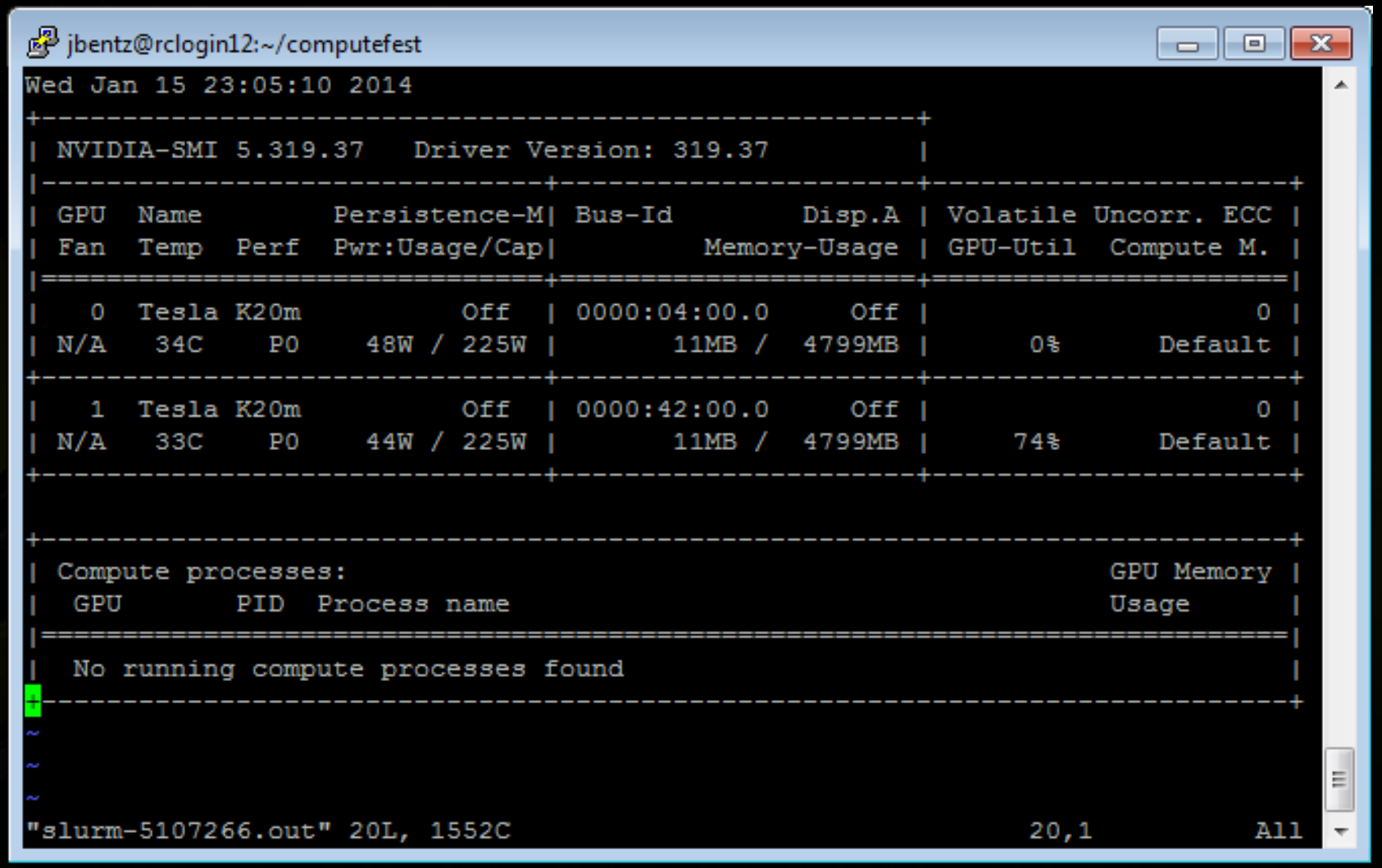
- GPU Driver
- CUDA toolkit
  - Includes all the software necessary for developers to write applications
    - Compiler (nvcc), Libraries, Profiler, Documentation
- SDK
  - Not strictly required but a good idea for ensuring your system is running properly.
  - Many examples with code samples illustrating lots of the important programming constructs and techniques.
- www.nvidia.com/getcuda Above software from NVIDIA is free

Examine GPU h/w and driver

- `nvidia-smi`
  - -h for help
  - -q for long query of all GPUs
  - PCIe Bus ID
  - Driver Version
  - ECC state
  - Power State/Fans/Temps/Clockspeed

- `sbatch computefest/runit.nvidia-smi`
  - `Open the resulting *out file`

nvidia-smi


jbentz@rclogin12:~/computefest
Wed Jan 15 23:05:10 2014
+--+
| NVIDIA-SMI 5.319.37   Driver Version: 319.37 |
|-------------------------------+----------------------+----------------------+
| GPU Name   Persistence-M| Bus-Id   Disp.A | Volatile Uncorr. ECC |
| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |
|===============================+======================+======================|
| 0 Tesla K20m   Off | 0000:04:00.0   Off | 0 |
| N/A   34C   P0   48W / 225W | 11MB / 4799MB | 0%   Default |
+-------------------------------+----------------------+----------------------+
| 1 Tesla K20m   Off | 0000:42:00.0   Off | 0 |
| N/A   33C   P0   44W / 225W | 11MB / 4799MB | 74%   Default |
+-------------------------------+----------------------+----------------------+
+---+
| Compute processes: GPU Memory |
| GPU   PID Process name   Usage |
|===|
| No running compute processes found |
+---+
~
~
~
"slurm-5107266.out" 20L, 1552C   20,1   All

CUDA toolkit

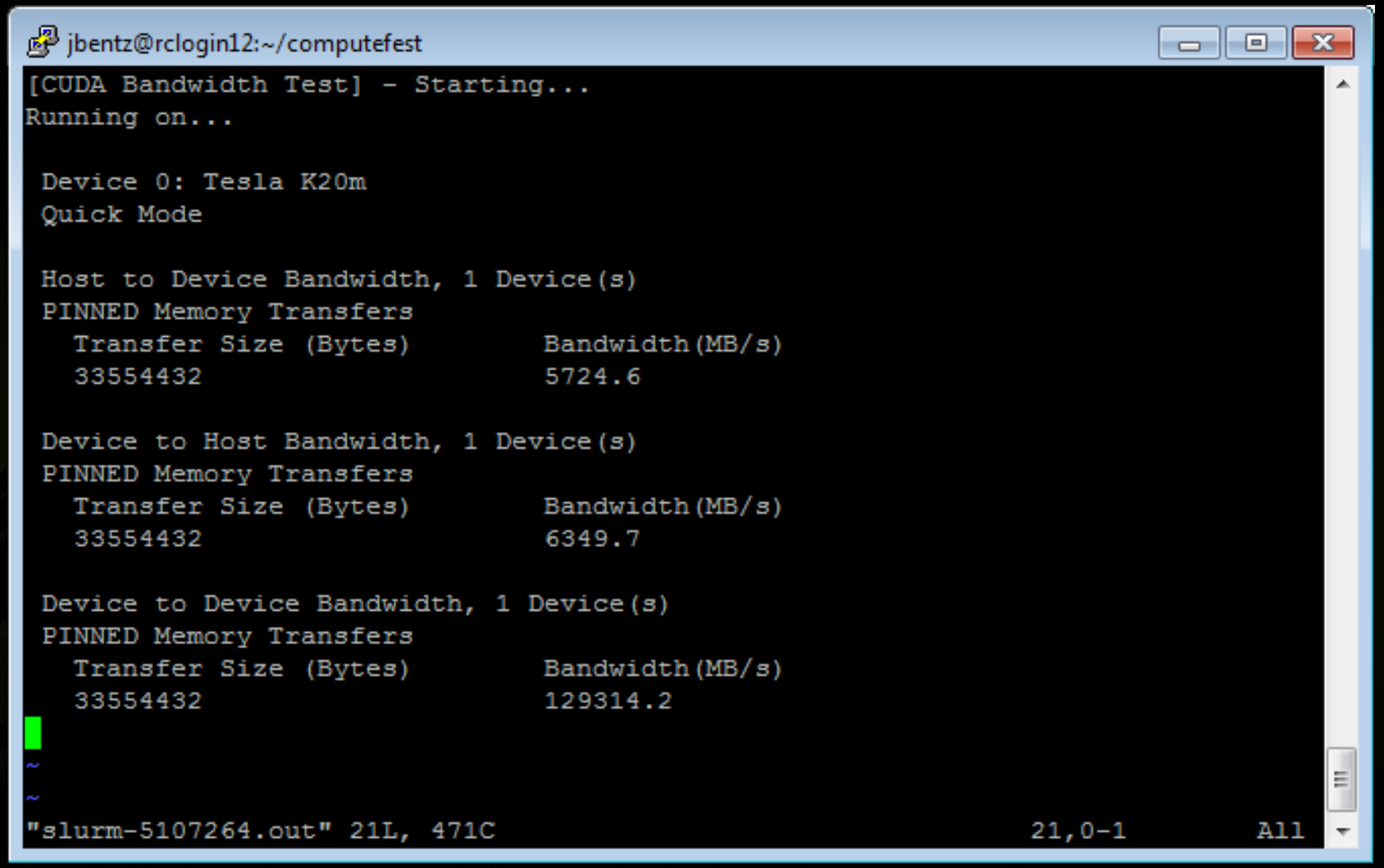
- Compiler (nvcc)
- Libraries
  - BLAS, FFT, sparse, RNG, NPP, OpenCL
- Profiler
  - Visual or command-line profiling available.

Samples (free download from nvidia.com)

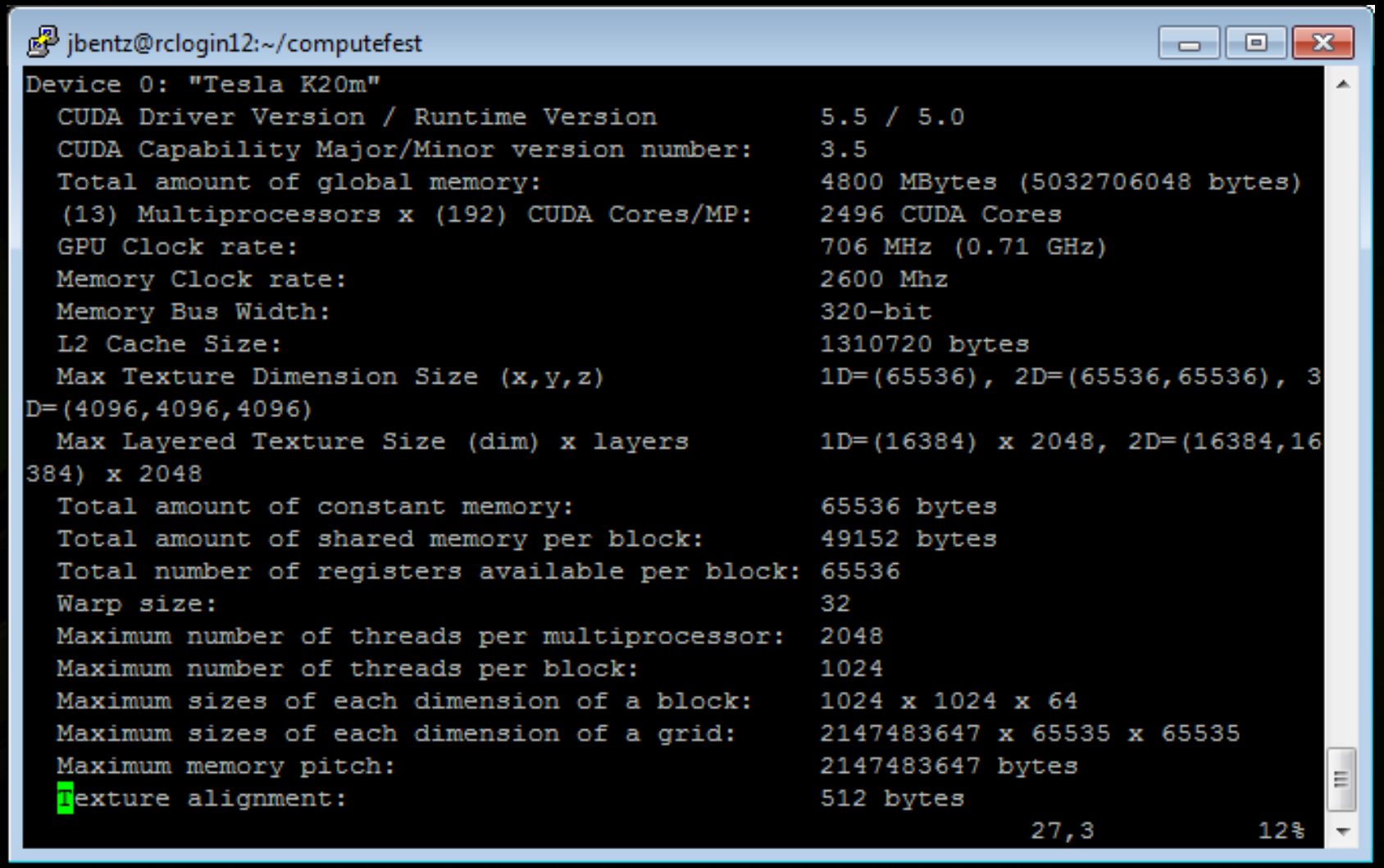
- `~jbentz/sdk/5.0/bin/linux/release`

- **Sample programs to illustrate CUDA and OpenGL programming constructs and algorithms.**
- **Useful diagnostic tests to query the GPU and its performance**

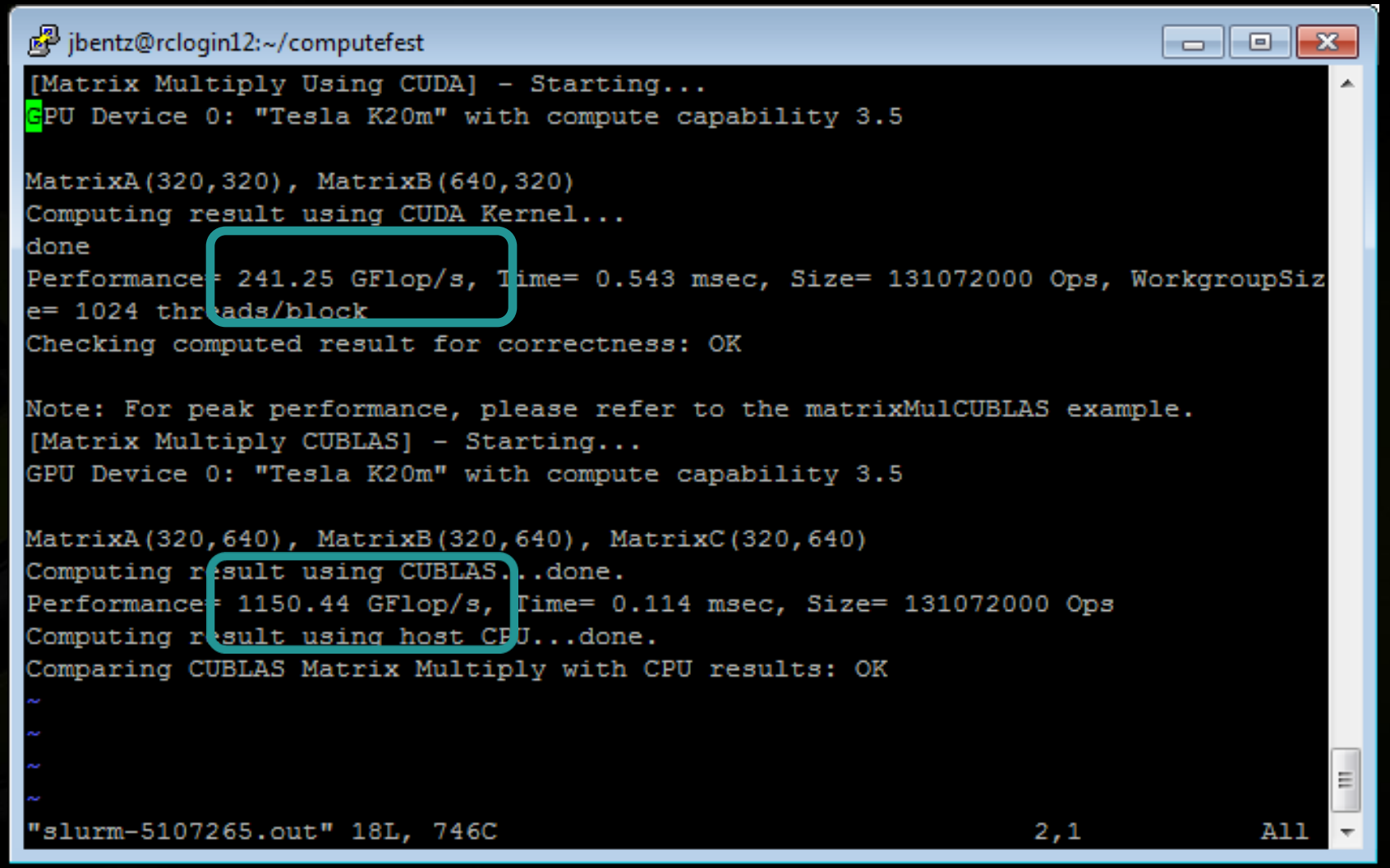
sbatch computefest/runit.bandwidth



sbatch computefest/runit.query



sbatch computefest/runit.matmul



Recap

- Driver
  - `nvidia-smi` to query the GPU hardware and state
- CUDA Toolkit
  - Development tools for GPU programming
- SDK/Samples
  - Sample code as well as diagnostic tests