Memory and time

Memory needed per iteration per source

In one FWI iteration, for each source we have to,

1. compute the wavefield $u(x,z,t)$,
2. compute the adjoint wavefield $a(x,z,t)$, (a needs u to be computed beforehand),
3. compute $\dot{u}(x,z,t)$,
4. compute the crosscorrelation of u (or $\dot{u}$) and a.

Let $|u|_{Gb}$ be the size in gigabytes of the wavefield $u(x,z,t)$,

$$|u|_{Gb} = 8 \cdot \frac{n_x \cdot n_z \cdot n_t}{10^9},$$

where the factor of 8 is for entries of type double.

Ram storage

In one FWI iteration, for each source,

1. compute the wavefield u and store in ram,
2. compute the adjoint wavefield $a(t_o)$ and cross-correlate with $u(t_o)$ for a fixed time on the fly,
3. compute $\dot{u}$ by taking $u(x_{slice})$ slices and replacing variable u (stored in ram) with entries of $\dot{u}$,
4. compute the adjoint wavefield $a(t_o)$ and cross-correlate with $\dot{u}(t_o)$ for a fixed time on the fly.

In any given time, at most there are

$$|u|_{Gb} + |u_{x_{slice}}|_{Gb} \approx |u|_{Gb}$$

used in ram.

Let n_w be the number of workers available and m_{ram} be the total ram available. In order to not get kicked out of the cluster we must have,

$$|u|_{Gb} \cdot n_w < m_{ram}.$$

This prohibits the number of workers, and thus the number of sources to be computed in parallel.

Disk storage

1. Compute the wavefield $u(t_{slice})$ and store in disk. Repeat until all t has been stored.
2. Compute the adjoint wavefield $a(t_{slice})$. Repeat until all t has been stored.
3. Compute $\dot{u}(t_{slice})$ by taking $u(t_{slice})$ slices and storing in disk. Repeat until all t has been stored.
4. Compute crosscorrelation $\sum_t u(-t_{end\,slice}) \odot a(t_{slice})$ for all t. Time slices are stored in ram one by one.
5. Compute crosscorrelation $\sum_t \dot{u}(-t_{end\,slice}) \odot a(t_{slice})$ for all t. Time slices are stored in ram one by one.

In any given time, at most there are $2 \cdot |u_{t_{slice}}|_{Gb}$ being used in ram: $u_{t_{slice}}$ and $\dot{u}_{t_{slice}}$ for step 3, or $u_{t_{slice}}$ (or $\dot{u}_{t_{slice}}$) and $a_{t_{slice}}$ for step 4 (or 5).

In order to not get kicked out of the cluster we must have,

$$2 \cdot |u_{t_{slice}}|_{Gb} \cdot n_w < m_{ram}$$

used in ram. The time sliced wavefields can be tailored to be small, so more workers can participate. The memory trade off is the amount of disk space available,

$$3 \cdot |u|_{Gb} \cdot n_s < m_{disk}$$

where the factor of three is for $u, \dot{u}$ and a, and n_s is the number of sources.

Benchmark - aka pitch for buying an ssd

If the ssd is more than 2.5 times faster in writing-reading than the one we have now it would make sense to buy it (50min/20min = 2.5, see below).

For one iteration, three source locations and three workers:

- The “ram code” takes 20min and at most 15Gb per worker. This method limits the number of workers.

- The “disk code” takes 50min and at most 6Gb per worker (this amount of memory can be reduced by user). This method does not limit the number of workers, but limits the number of sources via the space on disk.

The wavefield size for this problem is 7.5Gb ($n_x \approx 1100, n_z \approx 200, n_t \approx 4500$), so the “disk method” to make sense for larger problems would need ≈200Gb. The time varies on wether Gabe is using sonic or not.