Initial HPC Migration to Kupe

This is a preliminary document to outline the changes that we have made to the GM simulation work-flow in order to run it on Kupe (HPC 3).

Rewriting the scripts to Slurm

We have taken the base scripts used on Fitzroy and converted them to Slurm.

So far the GM simulation work-flow is completed for the manual interactive submission of a single simulation. We will work to extend this to automated simulations and Cybershake.

There is a very simple installing script for the work-flow that should be improved in the future.

Testing on Pan

We have run initial tests using the new Slurm-based work-flow on Pan.

Some interesting findings:

- For a small computational domain, Pan is much faster than FitzRoy for the LF part. Let's remember that Pan does not have a particularly fast interconnect or file system.
- The HF and BB parts run noticeably faster on the x86 system, removing the need to run them in parallel.

Initial runs on Kupe

Installation

The following list comes from what Jonney needed to do to get the work-flow running on Kupe

- Clone the following projects to the target location where everything will be installed: qcore (https://github.com/ucgmsim/qcore), EMOD3D (https://github.com/ucgmsim/EMOD3D) and slurm_gm_workflow (https://github.com/ucgmsim/slurm_gm_workflow)
- Compile EMOD3D (Using the intel toolchain, even though Kupe seems to be in a weird stage).
- **Fix all the hardcoded paths to /projects/nesi00213** (several still exist on qcore and EMOD3D, at least one on the slurm_gm_workflow).
- Install the workflow using the simple install utility
- Copy Velocity Models, Ruptures and StationInfo from other machine

Initial test run

Once the above has been completed, we are able to run a sample simulation succesfully exactly as a reasearcher would do on FitzRoy (with the limitations noted above).

The first example was the Kelly fault from Cybershake 17p8. We have so far compared the LF parts and they are in perfect agreement as expected. Once the HF and BB parts are done, we will also compare them as needed.

The Kelly fault does not have any execution time information, so Jonney is re-running a simulation that has execution time from Cybershake 17p9.

In term of Core Hours, we have obtain this matrix by running a simulation on Brothers.

Job	Kupe	Fitz	speed-up
Emod3d	14	34	2.43
Post-emod3d	0.016666667	0.4	24
HF	0.5	4.2	8.4
BB	0.1	41	410
	14.6	79.6	5.45

Further testing

To test the scalability, we performed a run on a larger model (February 2011 by Hoby). This model has nx=1400,ny=1200,nz=460.

The following results come from the LF calculation using EMOD3D for this model. The matrix below shows the relations of Cores and mean time for 100 iterations.

Requested cores	Physical cores	Nodes	Mean time for 100 time steps
80	80	2	90.3
128	80	2	129.6
160	80	2	97.47

160	160	4	46.2
256	160	4	65.5
320	160	4	48.6

We will note that Kupe is using hyper-threading by default. If we request N nodes, the best performance will be given by 40*N cores, anything above it seems to penalize the execution time.

Based on the matrix above, we did and estimation of the full run time on Cant 2011 earthquake(with sim_duration=100.0), obtaining

	CPU	time	sec	core hours for the LF part of the simulation
Kupe	160	02:30:00	9200	408.9
Fitz	512	01:50:00	6600	938.7
Speedup				2.3