

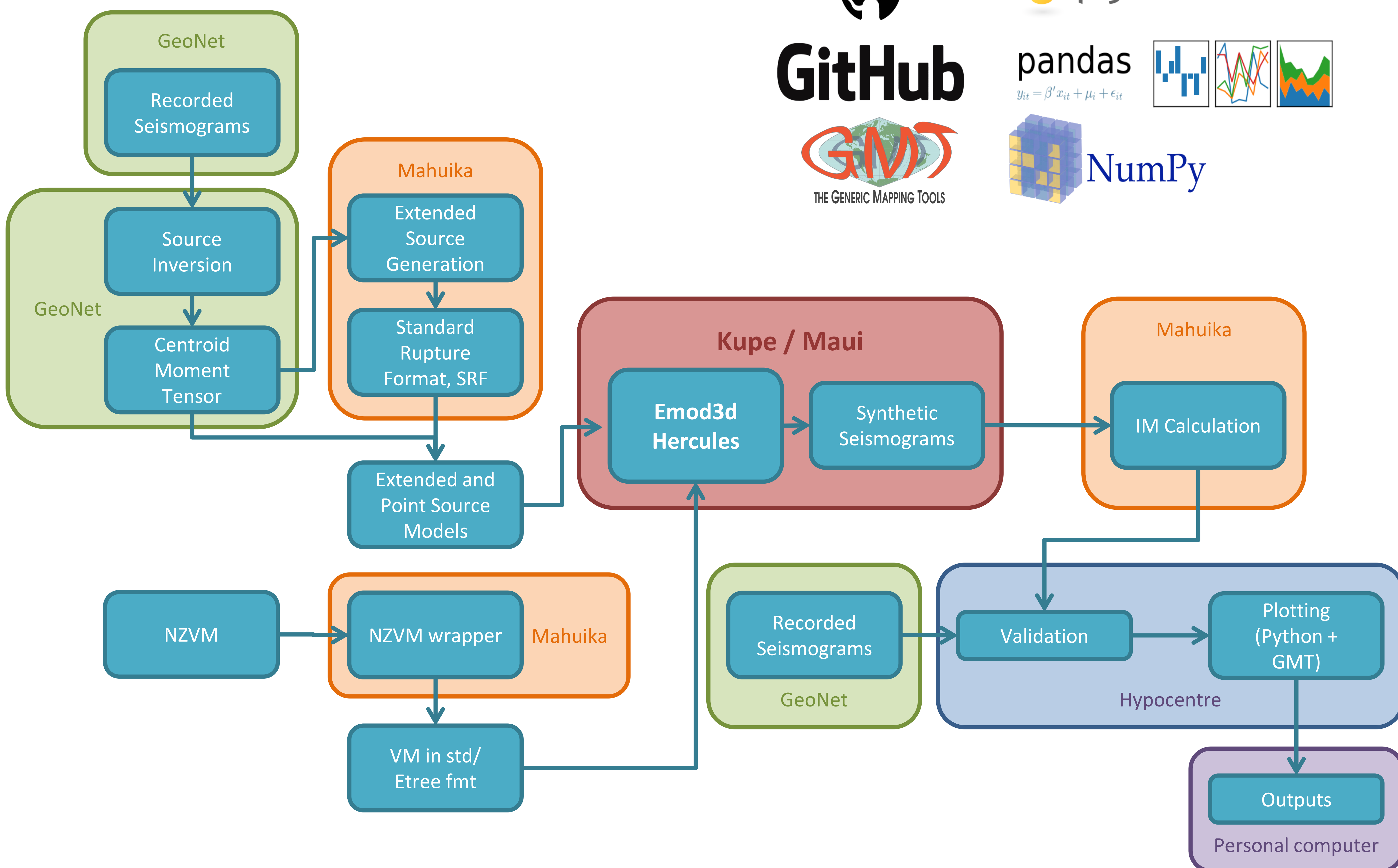
## 1. Background and Objective

Significant portions of QuakeCoRE research require large-scale computationally-intensive numerical ground motion (GM) simulations. The amount of data and complexity of computation make the large-scale simulation practically impossible to run on a researcher's workstation. QuakeCoRE started collaboration with New Zealand eScience Infrastructure (NeSI), the national high performance computing (HPC) provider to gain the necessary computational capacity and execution speed. In the last 12 months transition to the new supercomputing facility has begun. For the results of our workflow, refer to other posters from our research group.

## 3. Software Technologies

We use a multitude of open-source technologies (logos are adjacent). All of our source code for our projects are available on GitHub: <http://github.com/ucgmsim>

## 4. Workflow



## 5. Computational Demands

The synthetic seismograms are calculated on NeSI's largest computing resources: Kupe and Maui. This computation requires much more processing and memory resources than a typical laptop or desktop has available.

- Low Frequency (Emod3d & Hercules) is the main computational workhorse. It takes around 70% of the total compute time for a single simulation.
- High frequency and Broadband calculations (Synthetic Seismograms) take up 25% of the simulation time.
- Source and velocity model generation and Intensity Measure (IM) Calculation require less computation and are run on a smaller compute cluster at NeSI: Mahuika. Both clusters have access to the same storage, which means the large files do not need to be transferred.
- After IM Calculation, the IM files are transferred to UC storage. From there we can do validation analysis, plotting, simulation visualisations and any other downstream analysis.

## 2. Hardware

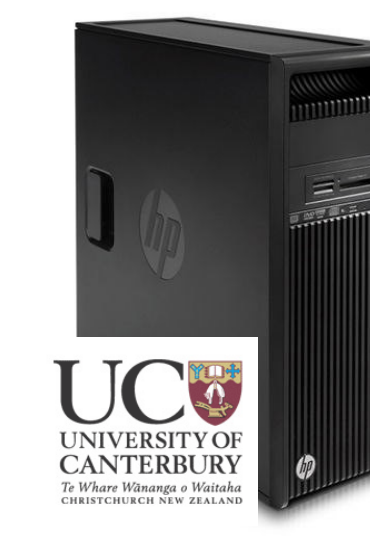
Maui is NeSI's leading capability machine. It is ranked as the world's 321 fastest supercomputer. It has 10,000 times the performance of Hypocentre; our local high performance workstation. Mahuika is an additional capacity resource that is useful for pre/post processing.



Maui / Kupe supercomputer



Mahuika Compute Cluster



Hypocentre - High performance Workstation



## 6. Automation

The above workflow shows how a single simulation is run. The strength of our computational workflow is the ability to be able to run 100s or 1000s simulations as part of a run-group. When running that magnitude of simulations, an automated workflow is necessary to reduce the overhead of managing the computation.

- **Technologies:** When a run group is installed a SQLite database is created which manages the jobs. When there is computational resources available for us to use (queried automatically every 5 minutes), a script (Slurm) is generated that contains the jobs to be run next.
- **Workflow:** The installation script takes a group of runs and associated parameters for source creation. From there, once the run is started, the batches will run. This will start the process up to computing simulated intensity measures.
- **Future Work:** Inevitably when running many differing simulations there will be errors. In our last Cybershake run-group, there were many errors that required manual intervention. As part of the continuing work on this workflow, these errors will either be automatically handled on occurrence or the appropriate checks will be run prior to ensure they do not happen during an automated submission.