Running FreeSurfer in the cloud

Paul Wighton
FreeSurfer Course
2017/04/03
Disclaimer

- I’m paid a salary by CorticoMetrics LLC

- There’s a million different ways to do this. This is (hopefully) a relatively simple approach
Is the cloud right for you?

• Pros
  – Unlimited resources available
  – No downtime (or hardly ever)
  – No hardware issues (maintenance/warranties)
  – Economies of scale

• Cons
  – IT/Security learning curve
  – Cost (when compared to subsidized HPC)
AWS batch

Product Details

With AWS Batch, you simply package the code for your batch jobs, specify their dependencies, and submit your batch job using the AWS Management Console, CLIs, or SDKs. AWS Batch allows you to specify execution parameters and job dependencies, and facilitates integration with a broad range of popular batch computing workflow engines and languages (e.g., Pegasus WMS, Luigi, and AWS Step Functions). AWS Batch efficiently and dynamically provisions and scales Amazon EC2 and Spot Instances based on the requirements of your jobs. AWS Batch provides default job queues and compute environment definitions that enable you to get started quickly.

https://aws.amazon.com/batch/details/
Terminology

- **AWS S3 Bucket**
  - A place to store files
  - We will keep subject data here

- **AWS EC2 instance**
  - Where the computation will occur
  - AWS batch will manage this for us

- **Docker container**
  - A way to define a runtime environment
    - (think: virtual machine)
  - We will keep/run FreeSurfer binaries and here
  - https://hub.docker.com/r/corticometrics/fs6-aws/
Step 1:
User creates and uploads subject data to be processed to AWS S3 bucket
Step 2: User submits job to AWS batch

recon-all -s bert -all -parallel
Step 3: AWS batch provisions some hardware to run the job and grabs the appropriate docker container.
Step 4:
Docker container copies subject data from bucket
Step 5: recon-all is run inside the docker container
Step 5: output is copied back to bucket

S3://my-fs-subjects/

Flow

AWS batch
Flow

Step 6: Provisioned hardware is released

S3://my-fs-subjects/
Step 7: User downloads processed data from bucket

Flow

S3://my-fs-subjects/

AWS batch
Demo

- github.com/corticometrics/fs-aws