Elastic Cloud Computing in Pharmacometrics: Usage Data and Strategies for Efficient Workflows

William Knebel, Marc R. Gastonguay, Dan Polhamus, Jeffrey T. Hane

Metrum Research Group LLC; Tariffville, CT USA

OBJECTIVES

- Quantify Amazon Elastic Cloud Computing (EC2) usage patterns for a group of M&S scientists
- Estimate Amazon EC2 usage patterns for groups of up to 64 users
- Summarize strategies for effective use of Amazon EC2 resources for pharmacometric analyses

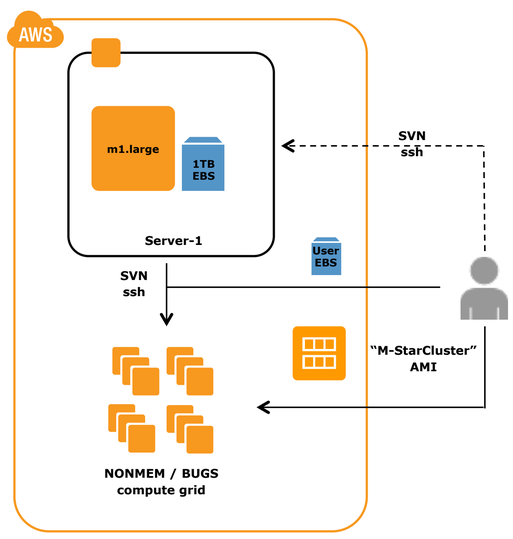
BACKGROUND - Current State of Computation for Modeling and Simulation (M&S)

- The extent and complexity of computation required for model-based drug development applications continues to intensify, while pressure mounts to shorten R&D timelines.
- Software and algorithms continue to develop, but often times, computational requirements are more intensive for advanced methods (e.g. MCMC, SAEM)
- Amazon EC2 offers an almost unlimited amount of computing power. However, optimal usage of EC2 requires a good understanding of the relationship between EC2 availability/resources, modeling strategies, and scientist teams

BACKGROUND - Typical Usage of Pharmacometrics Software with Cloud Computing

- Validated software (in this example: NONMEM®, OpenBUGS, and R) installation, and OS reside on a virtual machine image
- Each scientist in a group is responsible for spinning-up their own cloud-based cluster (Figure 1)
  - Can include multiple cores with same machine image, on as-needed basis
  - Cluster is up only intermittently - e.g. on the day(s) of any analysis work, and then shut down

BACKGROUND - Scalable Pharmacometrics Platform in the Cloud



Production Module

Workflow involves transactions with project source control server (the repository)

User EBS volume workspace

Projects are launched by individuals, from workstations, all over ssh

User–initiated grids using a StarCluster AMI base image

Scalable to >100 instances

Figure 1: *Scalable Pharmacometrics Platform in the Cloud*

METHODS - Cloud Usage

Determination of EC2 usage patterns (Figure 2)

- Usage patterns for 1 scientist and a group of 14 M&S scientists over a six month time period were captured to assess sustained and peak usage.
- Actual data from 14 users were used to simulate, via resampling, the usage pattern for a group of 64 users.

METHODS - Pharmacometric Project

Summary of actual pharmacometric project (Figure 3)

- Modeling of three endpoints requiring numerical integration of differential equations
- Simulation-based evaluation of Phase 2 and 3 trial design options
- Utilized parallel processing at 3 levels: within-run, across trial simulation replicates, across project objectives by 6 individuals on the M&S team

RESULTS

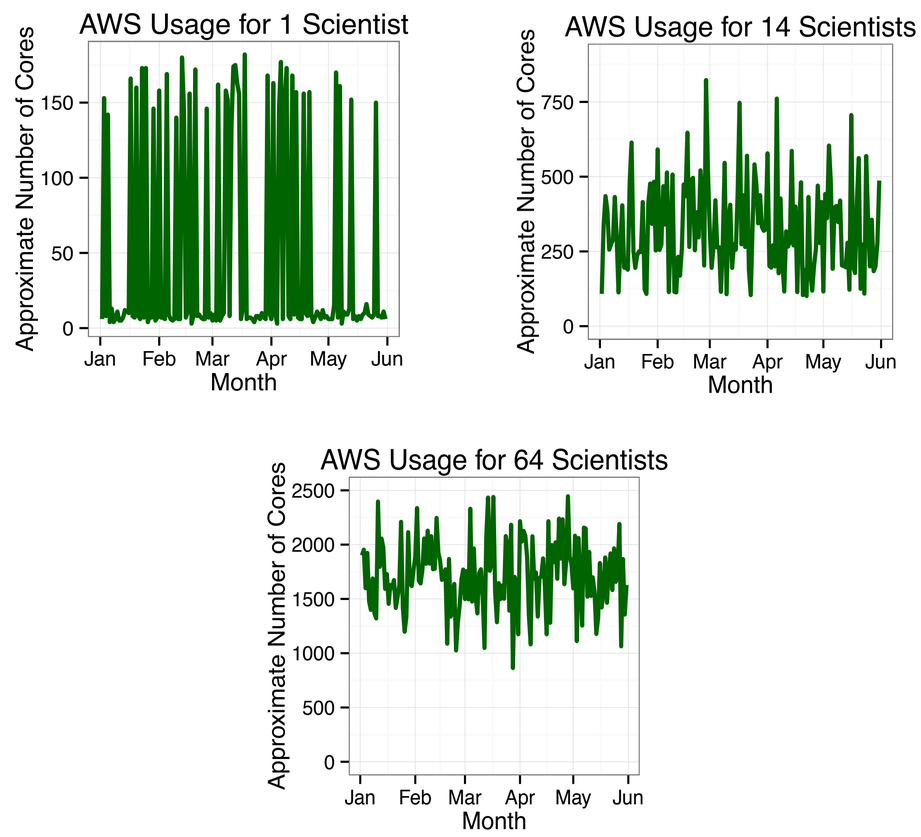


Figure 2: *Usage pattern for 1, 14, and 64 users*

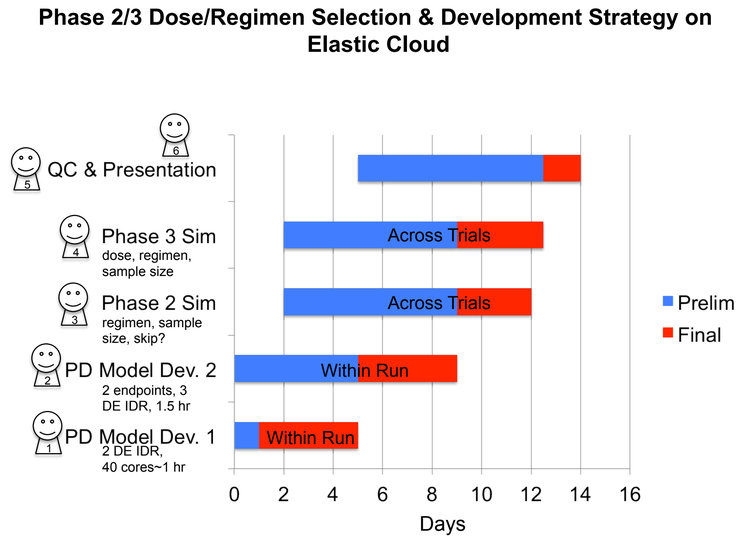


Figure 3: *Resource optimization for pharmacometric project*

BENEFITS AND CHALLENGES - Resource Optimization

Benefits	Challenges
•Timely, efficient completion of project	• Requires good communication and teamwork
• Modeling and simulation not limited by compute resources	• May require revised workflow when moving into cloud
• Allows for parallelization at three levels (run, simulation replicate, team member)	• Upfront planning of tasks and resources (scientists and computing resources)

CONCLUSIONS

- Usage patterns for EC2 and associated software resources, were characterized by peaks and valleys in utilization over time.
- EC2 and software utilization were proportional to the number of users, not maximum number of available cores.
- EC2 virtually eliminated computation time from the critical path for completion of the typical pharmacometrics project.
- Team-based project strategies, with parallel task and computation implementation, maximize the potential utility of EC2 for pharmacometrics workflows.

Abstract#I-55: Presented at Population Analysis Group in Europe Annual Meeting; Alicante, Spain; 11-June 2014