LA Grid Summit 2011
November 3rd & 4th 2011
Florida Atlantic University

Partnership for International Research and Education
A Global Living Laboratory for Cyberinfrastructure Application Enablement

# A Study of Allocation and Provisioning Policies in Clouds


David Villegas, PhD Student, FIU
**FIU Advisor:** S. Masoud Sadjadi, FIU
**PIRE International Partner Advisors:** Alexandru Iosup, TU Delft; Ignacio M. Llorente, UCM


National Science Foundation

## I. Research Overview and Outcome

### Research Problem

Motivated by Cmeter Cloud benchmarking tool from TU Delft:

- Successor to Grenchmark, a tool to measure performance of different Grids.
- Grenchmark generates synthetic job workloads and sends them to Grid schedulers. Then measures different metrics.
- Cmeter was originally designed to do the same with Infrastructure as a Service Clouds.
- However, IaaS clouds instantiate Virtual Machines, not jobs!
- What is the role of VM instantiation in the job scheduling process?

### Methodology

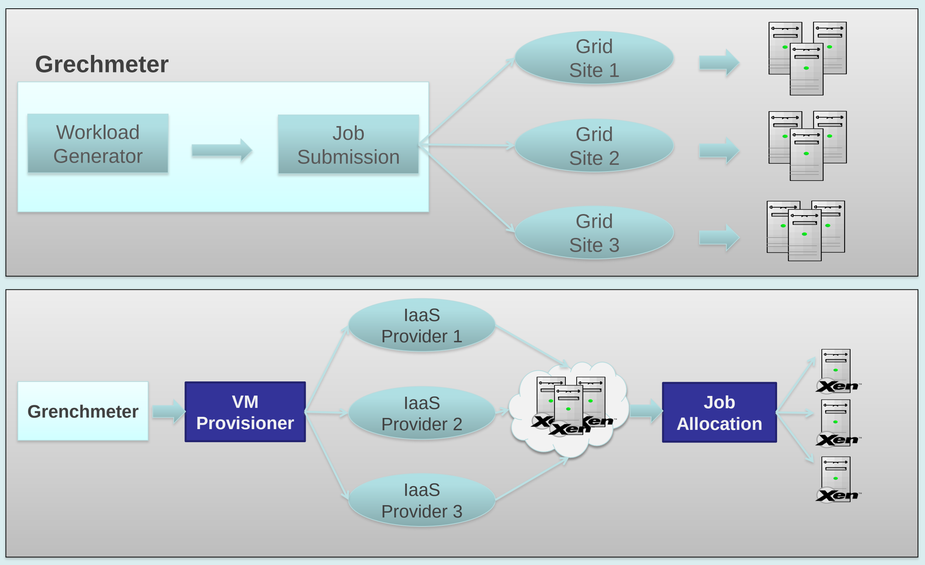
Study the two-phase scheduling model

- Make Cmeter components pluggable
  - Cloud Adaptors (OpenNebula, Eucalyptus, EC2)
  - Provisioning & Allocation policies
- Determine which workloads to test
  - Scientific, Web, Virtual Labs
  - Synthetic vs. Recorded traces
- Choose IaaS Clouds for execution
  - DAS4, FIU's Magellan, Amazon EC2
- Implement different Provisioning & Allocation policies
  - Startup, On-demand, Geometric, FCFS, Multi-queue
- Identify performance metrics for comparison
  - Makespan, Cost, Throughput, Execution ratio

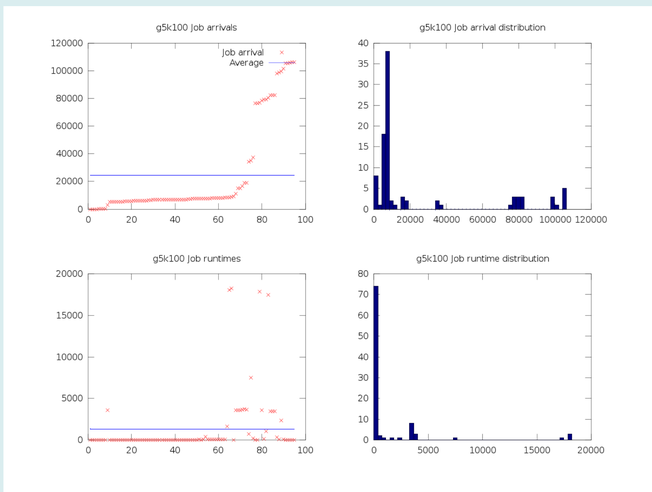
### Research Outcome

- Implemented Skymark Architecture
- Run first experiments with simple policies
- First findings:
  - Impact of VM booting time in job execution ratio (e.g. short jobs)
  - Difficulty to reconcile different goals across policies
    - E.g. Cost Minimization vs. Throughput Maximization
  - Provisioning-Aware Allocation Policies
    - Decouple two phases
    - Establish a communication channel among them

### General Architecture

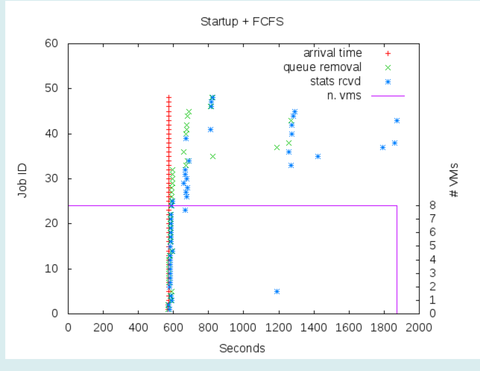



### Initial Results




- Use synthetic and trace-based workloads.
- Traces from Grid5000 archive.
- Slicing, sampling and scaling to achieve a representative (yet manageable) model.
- Distribution of very small jobs and very large jobs.
- Different arrival patterns.

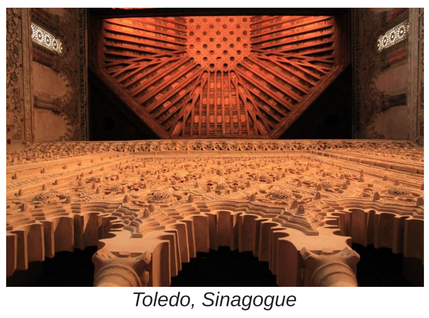
- Run an initial experiment with a subset of g5k workload.
- Implemented Startup Provisioning for 8 VMs and FCFS Allocation.
- Implemented On-Demand + FCFS, run some initial experiments.
- Really bad results: Job Runtime vs. VM booting time.
- Next step: implement multi-queue allocation / provisioning for VM reuse




## II. International Experience


*Toledo, Corpus Christi celebration*


*Universidad Complutense de Madrid*


*Toledo, Sinagogue*


*Universidad Complutense de Madrid*

In this trip, I had the chance to collaborate with scientists from top labs in the fields of cloud computing and scheduling. After returning home, my network had already expanded to new researchers in the OpenNebula team and TU Delft, which have proved to be a great asset in my scientific career.

### Spain

Spain is a very old and diverse country with a rich history. The Complutense University can be dated from the end of the XIII century!
I was lucky to visit Toledo during the celebration of Corpus Christi. The city is full of breathtaking monuments, such as its sinanogue.

### Delft

Delft is located in the south of the Netherlands. It hosts one of the top ranking research institutions in Europe, TU Delft. I loved Delft canals, churches, the city square, and how everybody moves around in bicycles.


*Delft "new" church*


*Delft city square*


*Delft canals*

## III. Acknowledgement


**The material presented in this poster is based upon the work supported by the National Science Foundation under Grant No. OISE-0730065. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.**