STAR Goals for Magellan
12 Mar 2010

- People
  - Leve Hajdu, BNL, grid developer
  - Jerome Lauret, BNL, STAR Computing Leader
  - Doug Olson, LBNL, STAR-OSG liaison…
  - Others at BNL and LBNL
  - Collaboration with Kate Keahey/NIMBUS at ANL
- Primary data center is bnl.gov
- Major analysis facility is nerc.gov
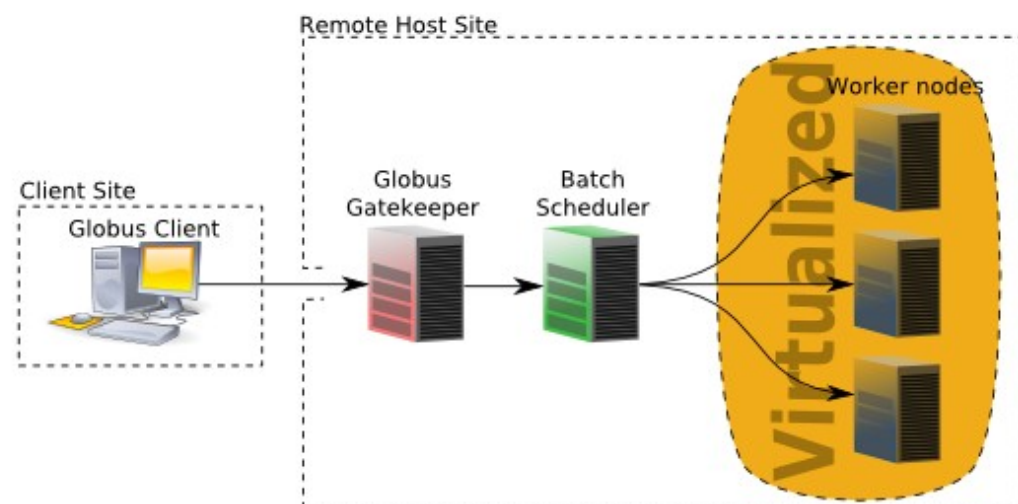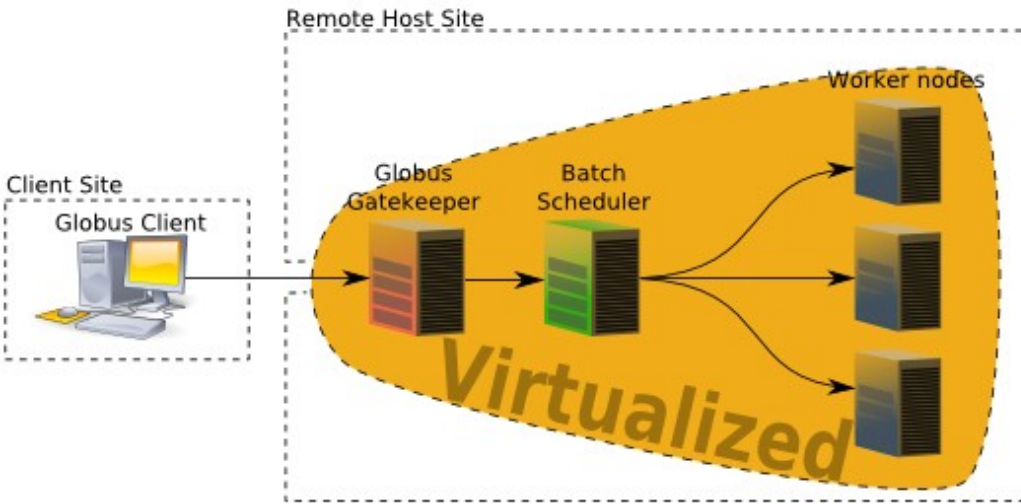- www.star.bnl.gov

STAR Goals for Magellan

- The primary motivation of STAR is to be able to deploy/instantiate a full system image including STAR-specific software in order to perform reliable data analysis and simulations.
  - This can be either or both of virtual machines or booting on bare metal, with the deciding factor being practical technical issues and performance.
- Next 4 slides illustrate VM/cloud modes that STAR has tried, followed by two slides of specific goals for Magellan.
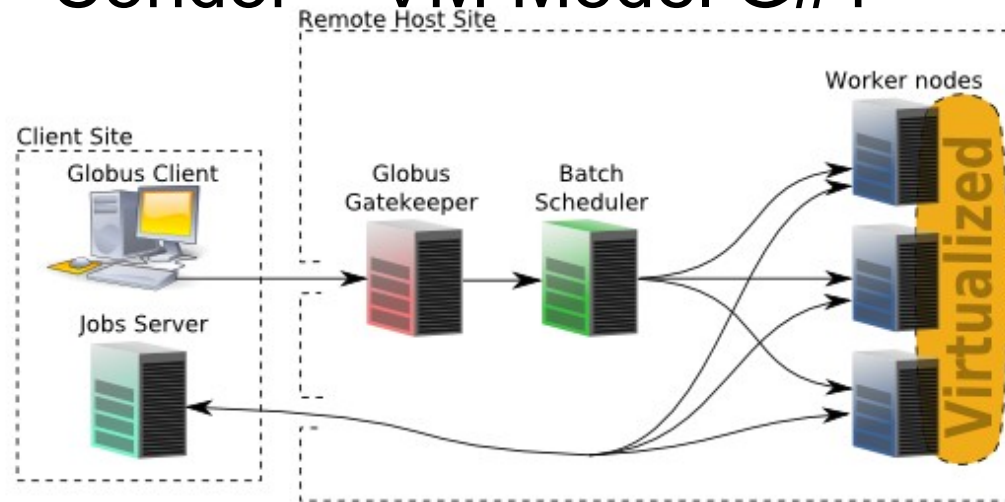
3 Models
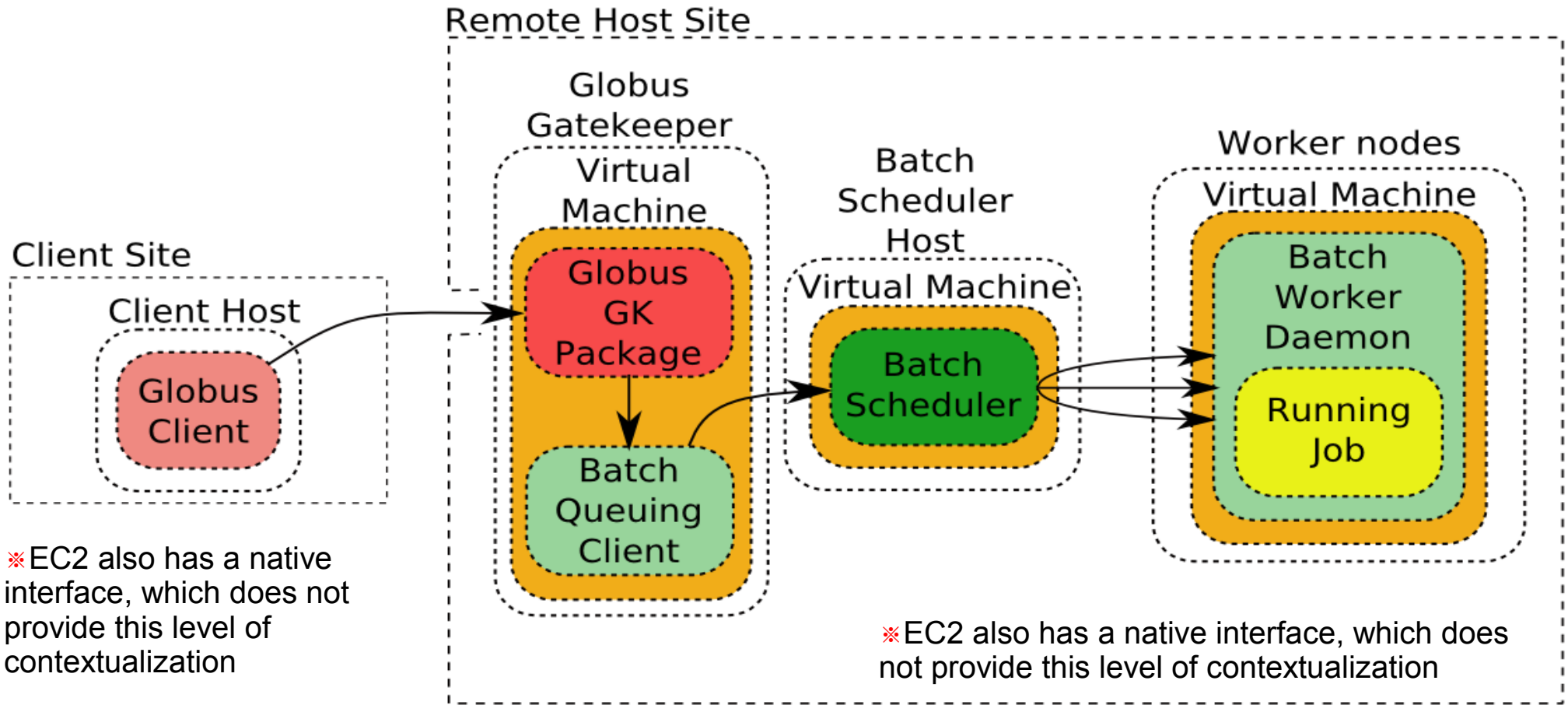
Amazon EC2 with Nimbus Interface   **Clemson Model CI#1**

Condor – VM Model G#1

Amazon EC2 With Nimbus Interface Model

Remote Host Site

Client Site

Client Host

Globus Client

Globus Gatekeeper
Virtual Machine

Globus GK Package

Batch Queuing Client

Batch Scheduler Host
Virtual Machine

Batch Scheduler

Worker nodes
Virtual Machine

Batch Worker Daemon

Running Job

※EC2 also has a native interface, which does not provide this level of contextualization

※EC2 also has a native interface, which does not provide this level of contextualization

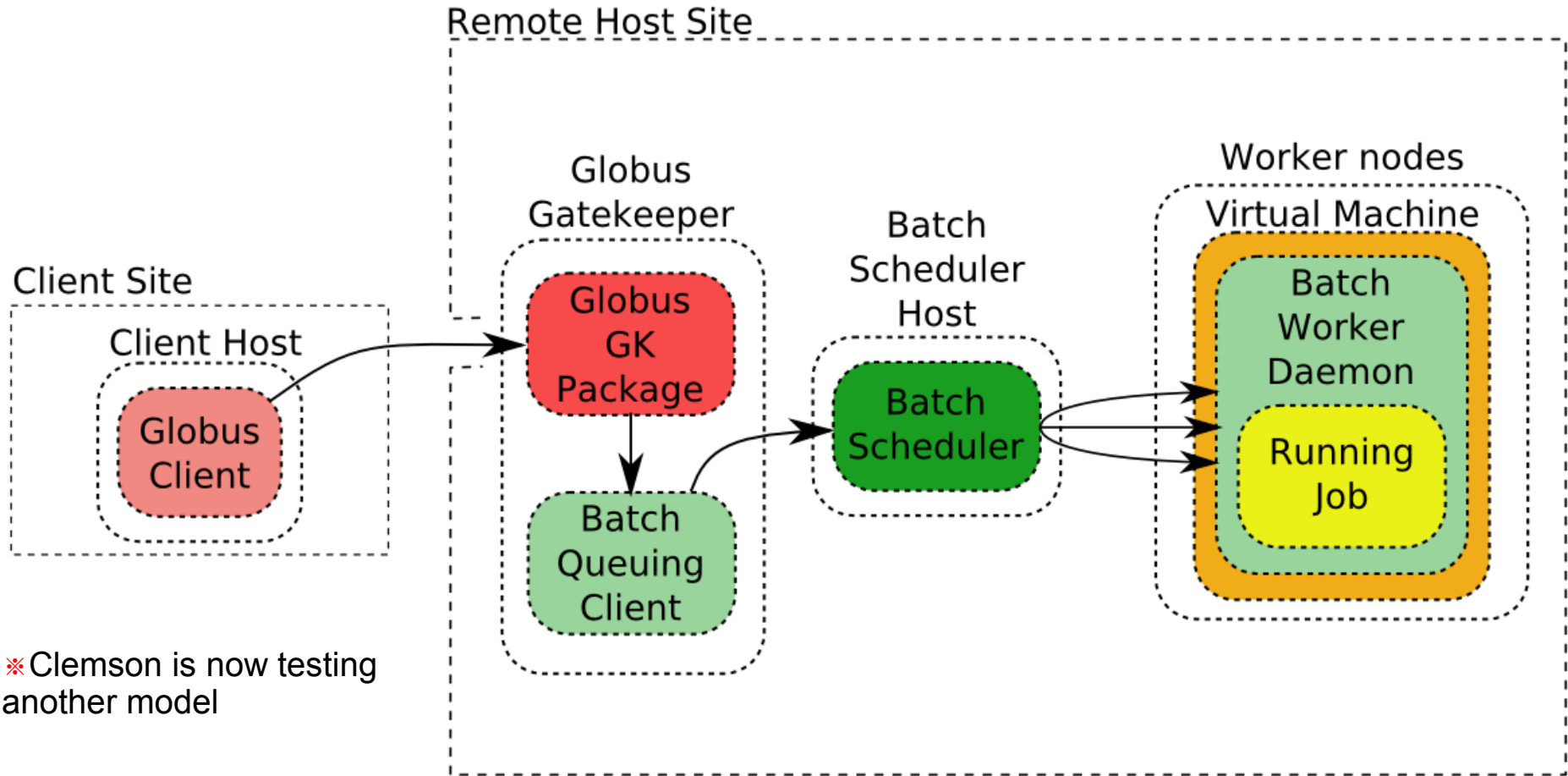| Pro	Con
-Guarantee on the number parallel slots (*not a hard requirement HENP (embarrassingly parallel)*) -Runs one job after the other without needing to boot up a new VM	-Base images need to be provided by host site -Contextualization waste on start-up and shutdown

◄-Submitting site is managing everything►

STAR

U.S. DEPARTMENT OF ENERGY

ISGC

Office of Science
U.S. DEPARTMENT OF ENERGY

Open Science Grid

The Clemson Model CI#1

Remote Host Site

Client Site

Client Host

Globus Client

Globus Gatekeeper

Globus GK Package

Batch Queuing Client

Batch Scheduler Host

Batch Scheduler

Worker nodes

Virtual Machine

Batch Worker Daemon

Running Job

※Clemson is now testing another model

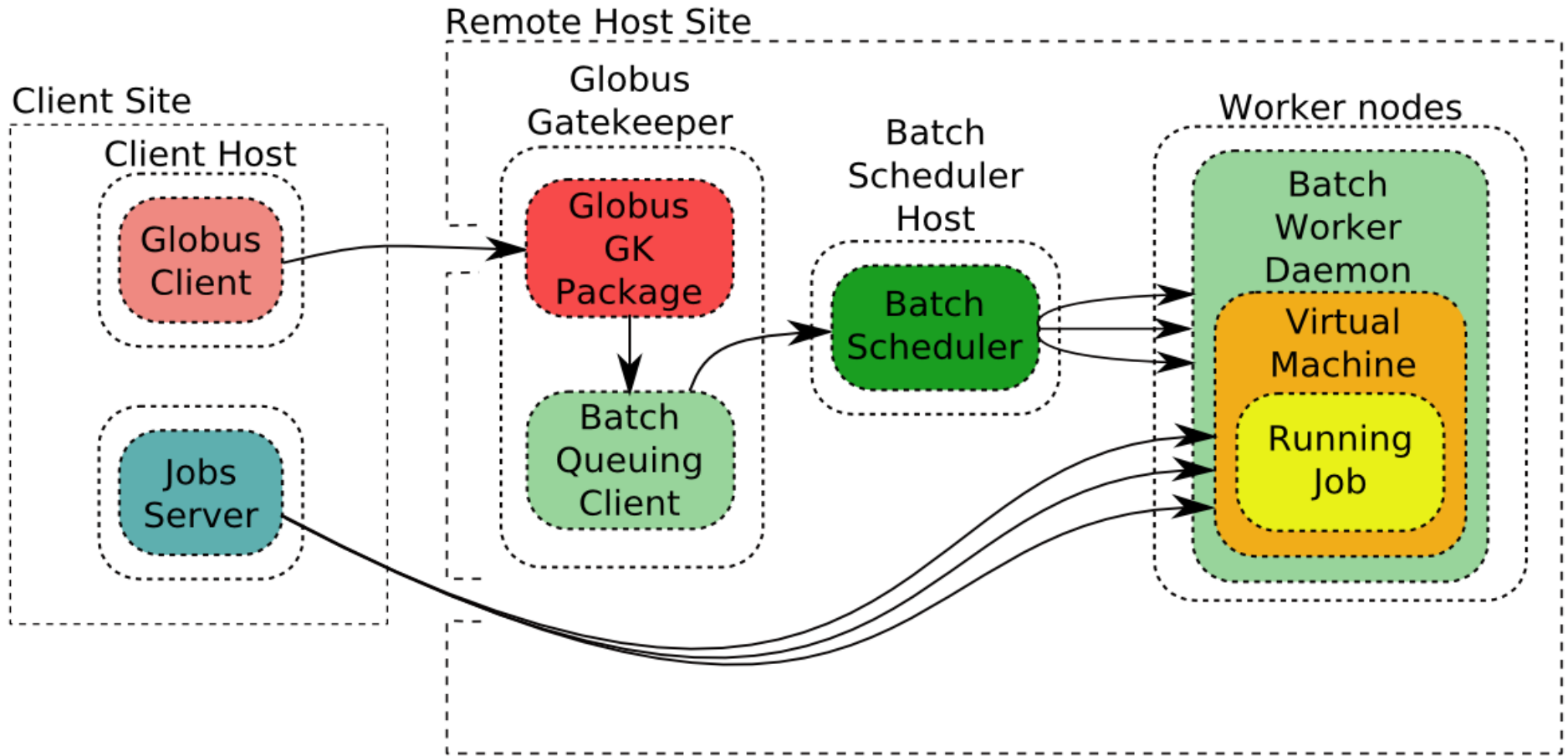| Pro	Con
-Most transparent to the user	-Batch worker MUST be supported by VM OS -Batch worker installed by host site into image (this is a lot of work for the host site)

Condor – VM Model G#1

Pro	Con
-Can run a large variety of images *(No site specific base image needed, no contextualization)*	-User must be trusted to shutdown the VM -User must figure out how to pull job in -Booting for each job is inefficient (multi-job submission framework must be supplied by user)*

STAR Workflows for Magellan

- STAR Workflows: all single processor / job
  - Production: CPU-intensive, GB/job I/O, hours/job, GB/job memory, managed data (simple mapping between data and job)
    - Simulation: KB/job input, GB/job output
    - Reconstruction: GB/job input, 10s MB/job output
    - Embedding: GB/job input, GB/job output
  - User Analysis: mix of I/O intensive, CPU-intensive, large memory
    - Typically GBs/job input, MBs/job output
    - Job length & memory vary per analysis type
    - Data challenge: input data size (# files, GBs) per job job depends on analysis & compute resources
- Platforms
  - Magellan Batch
  - Magellan Eucalyptus
  - Future: Maybe Amazon EC2
  - *Possible application for surge computing*

COLLABORATION: Jeff Porter, Doug Olson, Iwona Sakrejda et al.

STAR Experiments

- Initial work: Use tested distributed model
  - OSG / Globus interface to dynamic cluster of known characteristics: batch system, job slots, resources per job
  - Worker node lifetime spans many jobs
  - Worker node images built with STAR software + worker nodes client tools (as exists on PDSF)
  - Successful production with ~100 node cluster with NIMBUS+EC2 to meet short term needs
- Alternative model for testing: Simple simulation job in a VM
  - Job is packaged within VM to start on image boot
  - Job requests parameters from outside source, ships results when done
  - Job completion triggers image shutdown
  - Initial tests on EC2, larger testing under development at Wisconsin