MoSes Projections on a Condor Grid

Sara Triplett, ASA

Jason Vigil

PACIFIC LIFE

Overview

- MoSes modeling environment
- Pacific Life's workload
- History
- Today
- Going forward

MoSes Modeling Environment

- Data-intensive financial projections
  - Up to 300,000 data points
  - 1,000 stochastic scenarios
- Intensive mathematical calculations
- Windows 2003/2008/2008 R2 environment

Pacific Life's Workload

- Quarterly, monthly, and ad-hoc projections
- Attribution analysis
- Sensitivity testing
- "Scramble" requests
- "Whoopsies" projections
- Continuously increasing needs

History – MoSes Master/Worker

- MoSes Master/Worker Process
  1. Set up model on Master machine
  2. Launch desired number of workers
  3. Run projection
- Practical limitations on number of workers
- One Master per projection
- Ties up resources on the Master machine(s)
- Model(s) cannot be used during this time

Today – CycleServer and Condor

- Process
  1. Create MoSes “seed”
  2. Run CycleServer “submit” process
- No need for multiple “Master” machines
- No manual launching of workers
- Models are free for use during these runs
- Trivial to tweak a projection and resubmit
- Trivial to duplicate previous projections

Common Use Cases

- Rerunning problem jobs without starting over
- Burst submissions for strategic timing
- Robust attribution analysis
- Quick sensitivity testing

Our Evolving Needs

- Original grid size – 100 nodes (1Q 2009)
- Today – 1024 nodes (4Q 2011)
- ~50% increase in capacity every 6 months.
- “High Memory” nodes
- Retiring old servers to “last resort” usage
- Amazon Cloud
- Off-site in-house server pool

Looking Ahead

- Expanding to desktops
- Better control of workflow through policy
- More sophisticated DAGs
- Using Condor to distribute other workloads
- Exponential increase in projection needs
- Advances in technology

The Future of Technology

- Virtualize the OS for disaster recovery, rapid provisioning and workload migration
- GP/GPU computing to support nested simulations and significantly reduce host count
- Infrastructure as a Service to achieve necessary capacity and minimize footprint in company-owned datacenters