

Cost-Effective HPC: The Community or the Cloud?

A. Carlyle, S. Harrell, P. Smith

Dec 1, 2010

CloudCom 2010

Background

- In July 2010 Amazon announced powerful “Cluster Compute Instances” (CCI) in EC2
- Replace the large, up-front, fixed cost of HPC cluster environments with smaller, variable costs based upon actual usage

Motivation

- Is it more cost-effective for Research University HPC centers to continue to maintain hardware resources...
- Or to begin adopting commercial cloud products like EC2 CCI for core scientific research?

High Performance Computing

- Use of parallel processing for running scientific applications, especially above a TeraFLOP (as of 2010).
- Common users are scientists, engineers, etc. in academic institutions



Purdue's "Community Clusters"

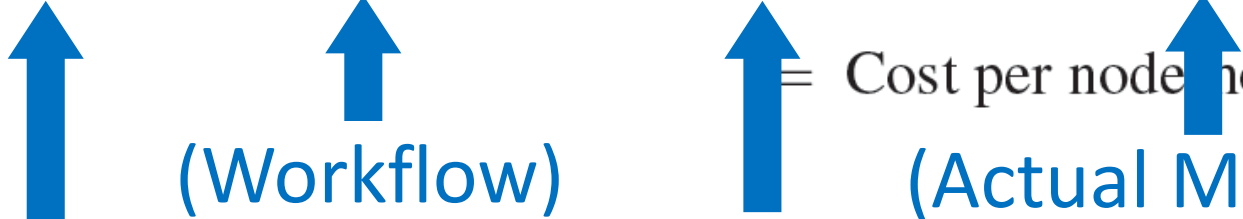
- Commodity Linux Clusters running job scheduling software
- PI pays lump sum for access contract, but costs are subsidized by Purdue
- Unlike commercial cloud fees, actual per node-hour costs to users are a function of their hardware utilization, etc.

Methodology

- Produce “apples-to-apples” cost comparisons from actual usage patterns
 - Factor for user utilization rates
 - Factor for user workflow style
 - Factor for system performance differences
- Apply simple formula across all Purdue HPC user groups

“EC2-Equivalent” Cost

$$\left(\frac{\# \text{ months}}{\# \text{ core-hrs}} \right) \left(\frac{\# \text{ core}}{1 \text{ node}} \right) \left(\frac{1 \text{ hour (Cluster)}}{\text{hour (EC2)}} \right) \left(\frac{\text{Month cost}}{1 \text{ month}} \right)$$



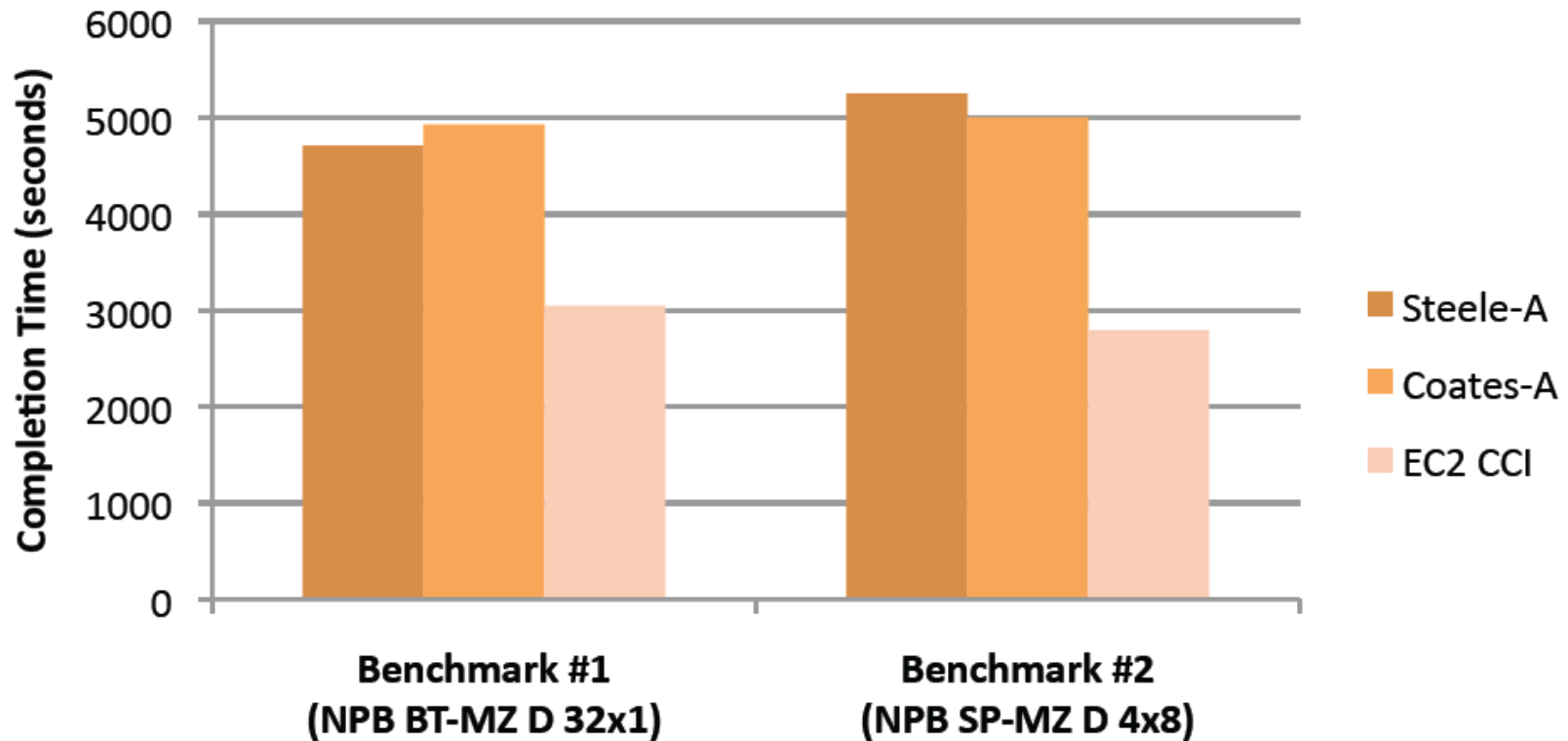
(Utilization) (Workflow) (Performance) = Cost per node hour (Actual Money)

$$\left(\frac{3 \text{ months}}{40,534 \text{ core-hrs}} \right) \left(\frac{8 \text{ core}}{1 \text{ node}} \right) \left(\frac{1 \text{ hour (Coates)}}{0.6 \text{ hour (EC2)}} \right) \left(\frac{\$262}{1 \text{ month}} \right)$$

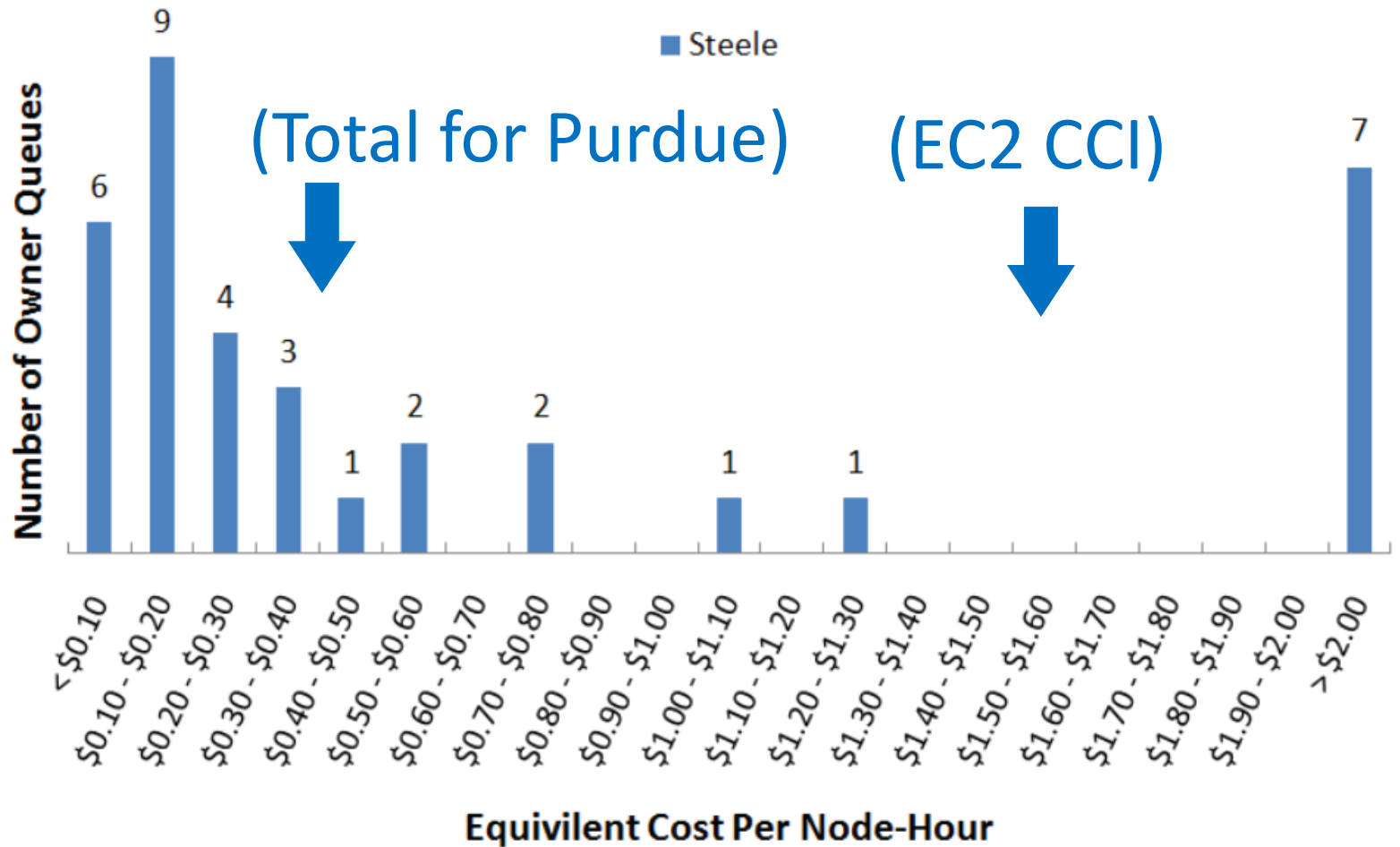
= \$0.26 per node-hour

Performance Results

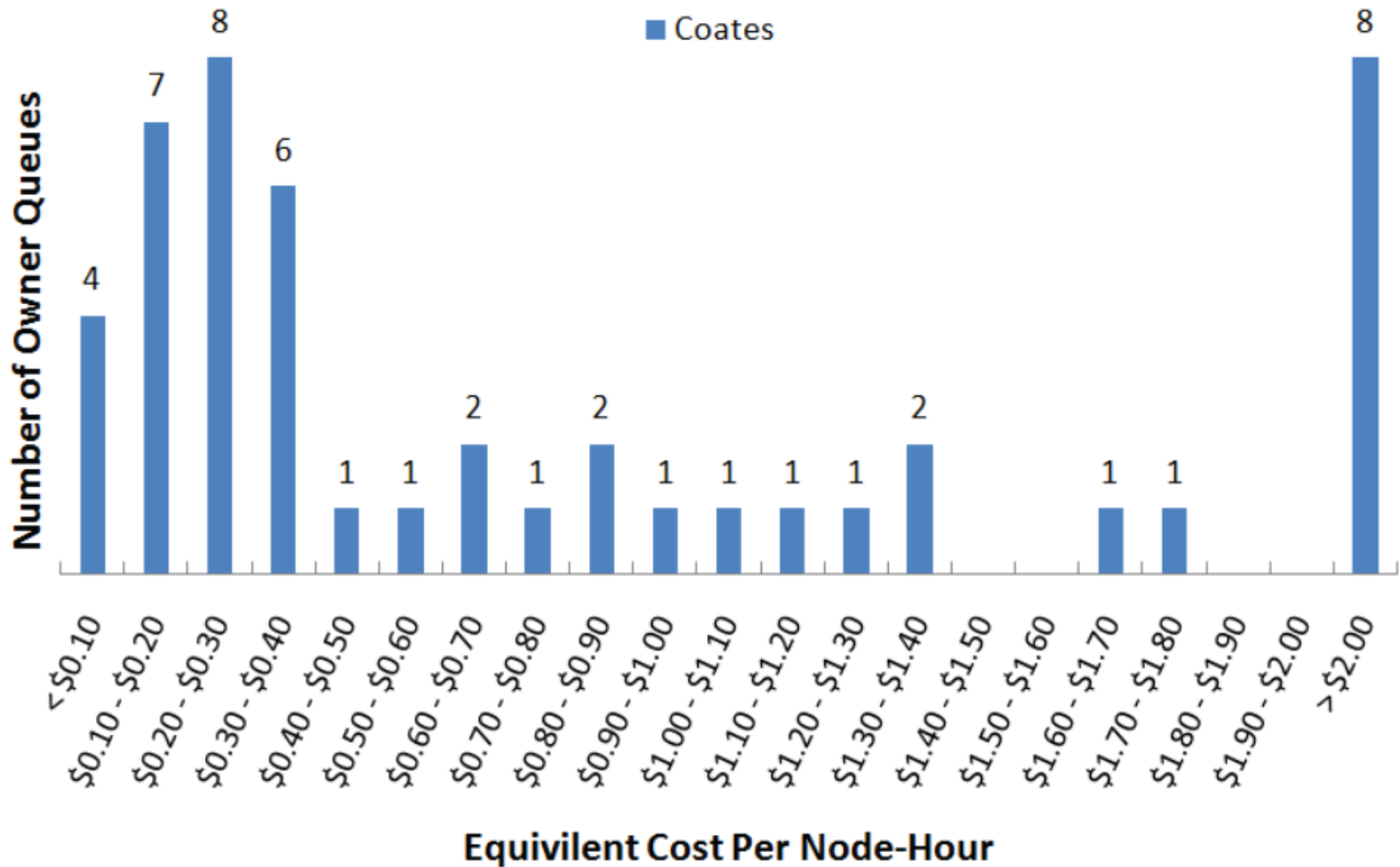
RCAC vs. EC2 CCI Benchmark Timings
(less is better)



Equivalent Cost Results



Equivalent Cost Results



Conclusions

- Currently more cost effective for Purdue to maintain own HPC clusters for most researchers
- 15% of researchers could reduce costs by using EC2
- Useful if EC2 Cluster Compute Instances were made “spot instances” so price could fluctuate

Questions?

