Characterizing Task Usage Shapes in Google's Compute Clusters

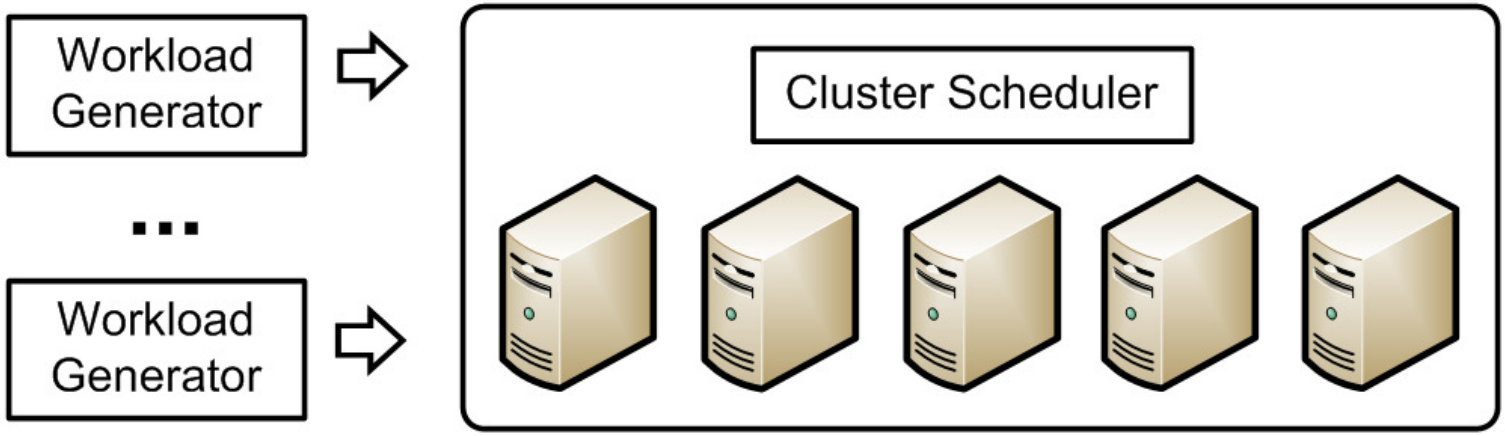
Qi Zhang[1], Joseph L. Hellerstein[2], Raouf Boutaba[1]
[1]University of Waterloo, [2]Google Inc.

Introduction

- Cloud computing is becoming a key component of today's IT infrastructure
- Characteristics of Cloud computing
  - Extremely large scale infrastructure and workloads
  - Diversity in workload composition
    - User facing vs. batch applications (e.g. MapReduce)
  - Different performance objectives
- Workload management becomes a challenging problem in cloud computing environments
  - Need to understand the impact of management activities on workload performance
    - E.g. scheduler change and capacity upgrade

Motivation

- Using *performance benchmarks* to assess the impact of management activities
- Existing approach: use historical traces as performance benchmark
  - Advantage: high accuracy
  - Disadvantage: expensive; only simulates performance in the past



Motivation

- Need to construct workload models
- In this work, we create a workload model of *task usage shapes* that describes task resource consumption at run-time
- The accuracy of our model is evaluated by its ability to reproduce the performance characteristics of real workloads
  - Key performance metrics: Task wait time and Resource utilization
- Our Previous work
  - Workload characterization at a medium-grained level[1]
  - Not clear if the model is sufficient for predicting workload performance

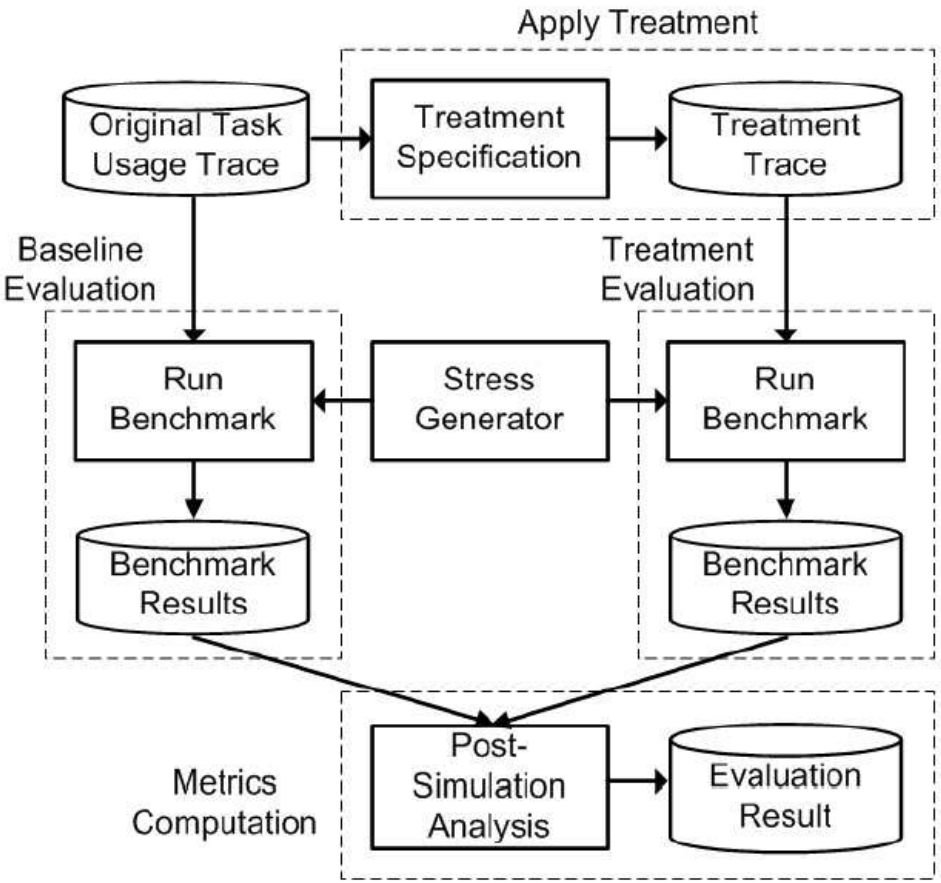
[1]"Towards Characterizing Cloud Backend Workloads: Insights From Google Compute Clusters," A. Misra et al., Sigmetrics Performance Evaluation Review, 2010.

Dataset Description

Compute Cluster	No. of machines	Type 1 (%)	Type 2 (%)	Type 3 (%)	Type 4 (%)
A	10000s	3.12	0.26	3.14	93.47
B	1000s	1.46	0.86	2.52	95.16
C	1000s	4.54	0.34	4.67	90.45
D	1000s	5.86	2.42	31.77	59.95
E	1000s	39.26	1.48	34.27	24.99
F	10s	1.23	0.2	72.93	25.64

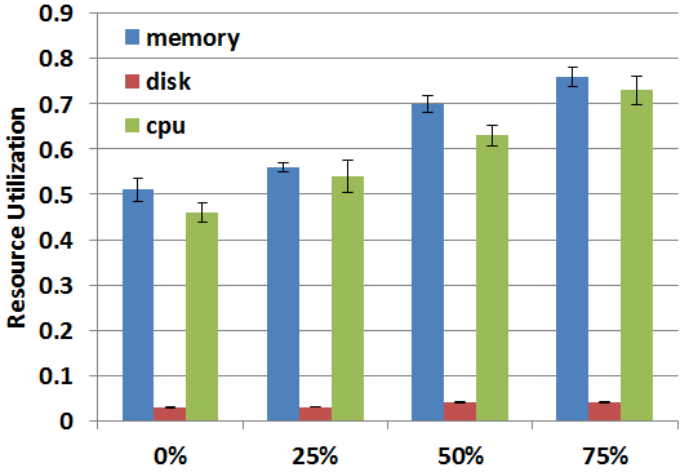
- Workload traces from 6 clusters for 5 days
- 4 types of tasks
  - Type 1: high priority user-facing tasks
  - Type 4: low priority batch tasks
  - Type 2 and 3 stand between Type 1 and Type 4

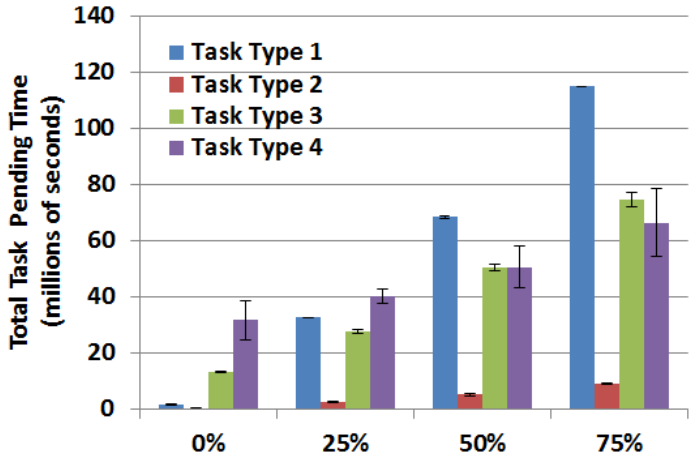
Experiment Methodology



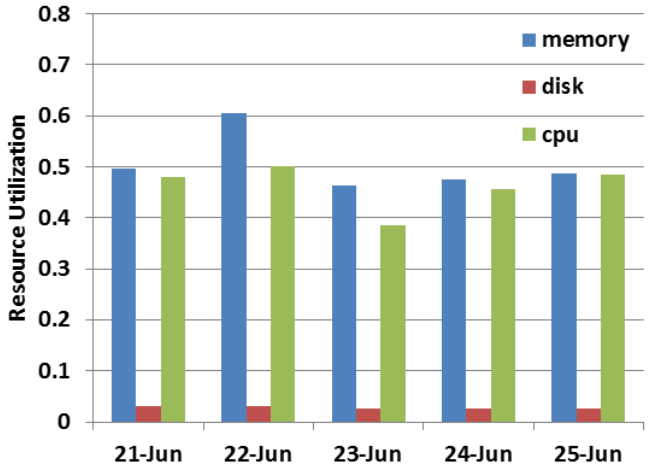
- Performance Metrics:
  - Task wait time
  - Resource utilization
- Stress Generator increases the workload intensity by randomly removing a fraction of the machines

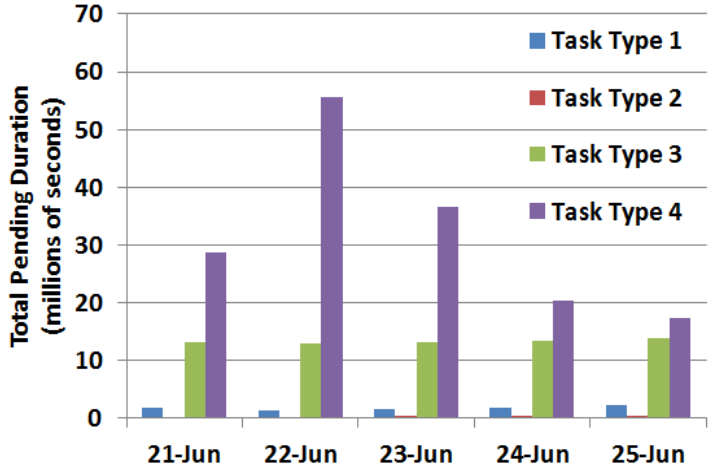
Characteristics of Performance Metrics





Effect of Removing Machines on Performance Metrics for Cluster A

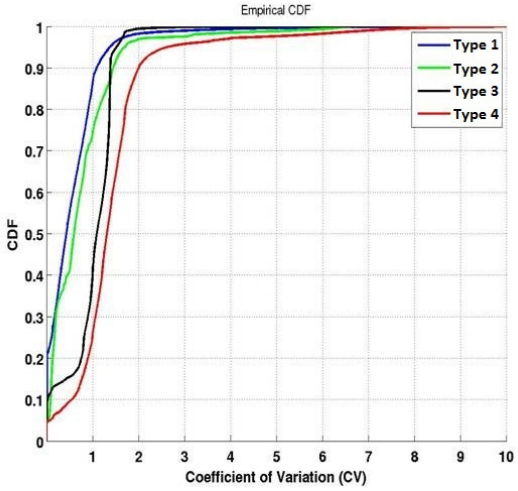




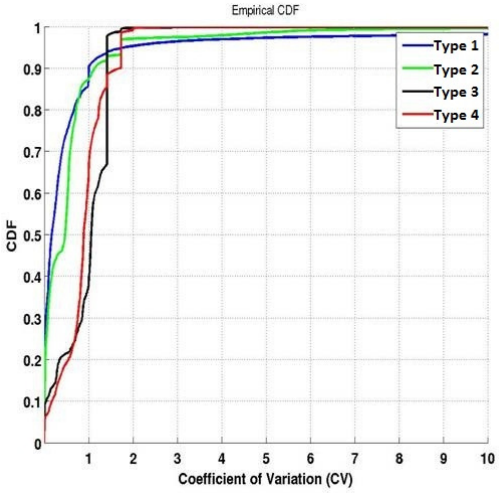
Day-to-day variability of Performance Metrics for Cluster A

Workload Characteristics

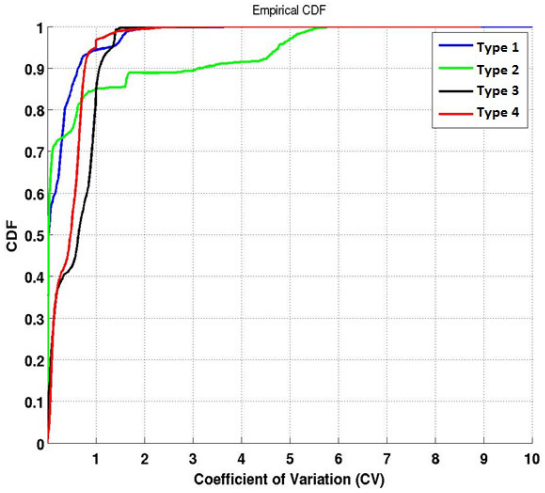
- Most of the tasks have low coefficient of variation
  - CPU has the highest CV, but mean is low
- This suggests that we can simply use the mean usage as a model for capturing workload characteristics
- We call this model the *mean usage model*



Mean CPU usage=0.239 Cores



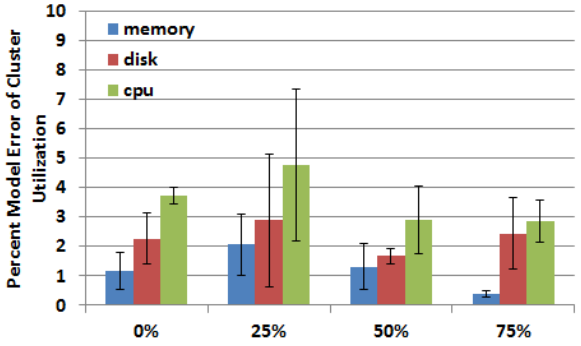
Mean Mem usage=0.219 GB



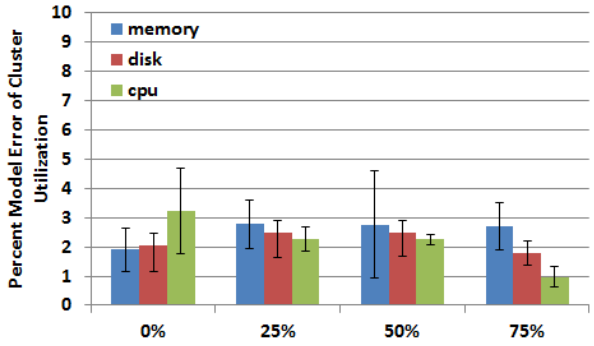
Mean Disk usage=0.33 GB

Distribution of CV for CPU, Memory and Disk for Tasks in Compute Cluster D

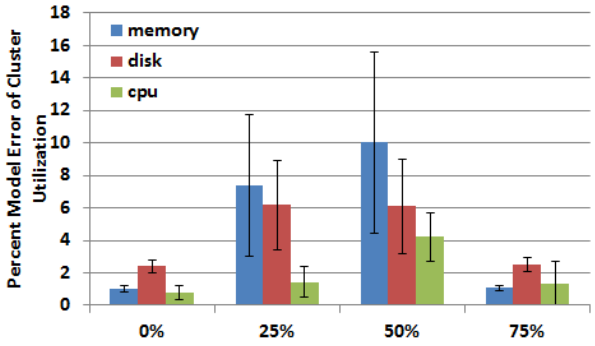
Evaluating the Mean Usage Model: Resource Utilization



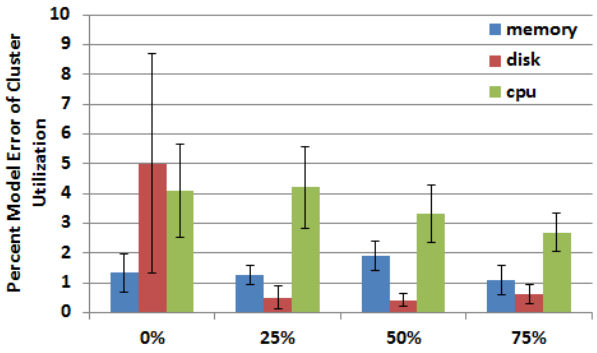
Compute Cluster A



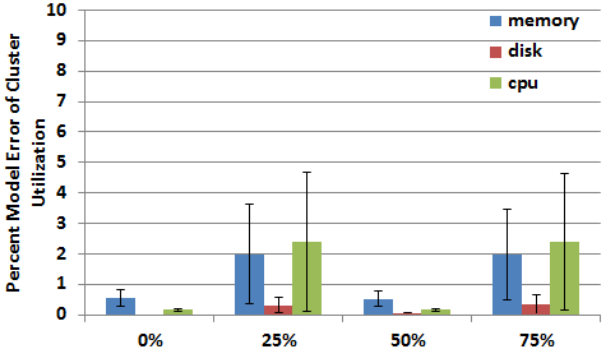
Compute Cluster B



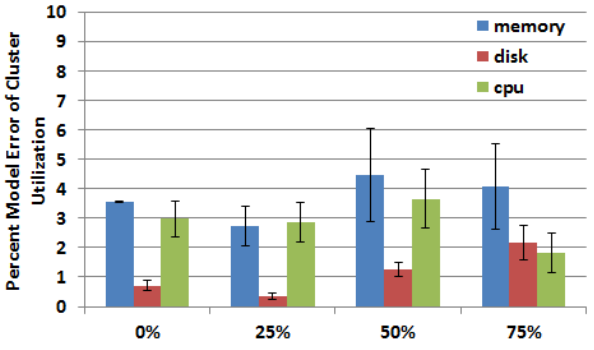
Compute Cluster C



Compute Cluster D



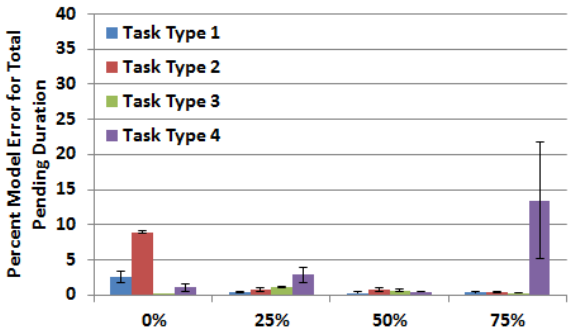
Compute Cluster E



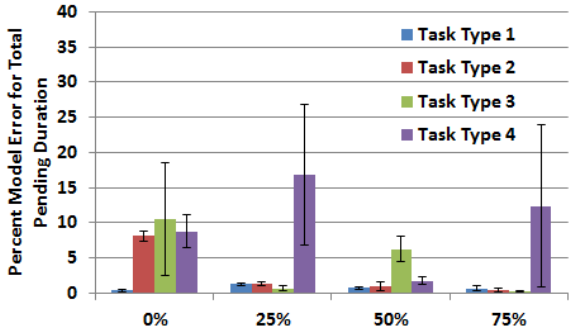
Compute Cluster F

The mean model accurately reproduces resource utilization in each compute cluster.

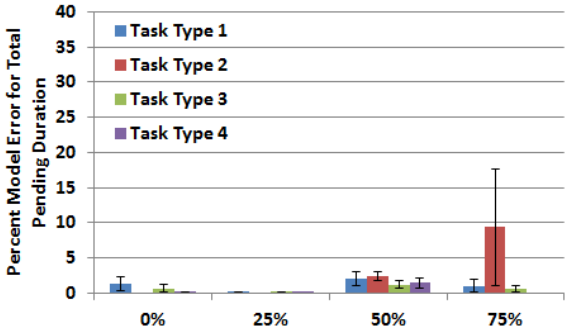
Evaluating the Mean Usage Model: Task Wait Time



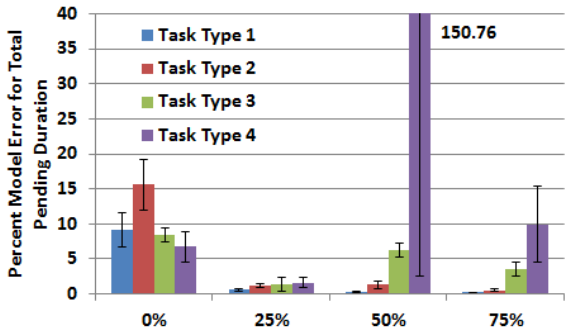
Compute Cluster A



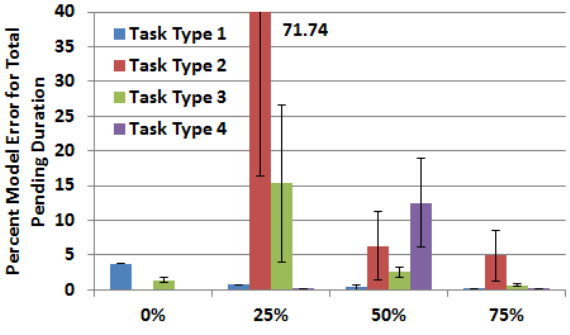
Compute Cluster B



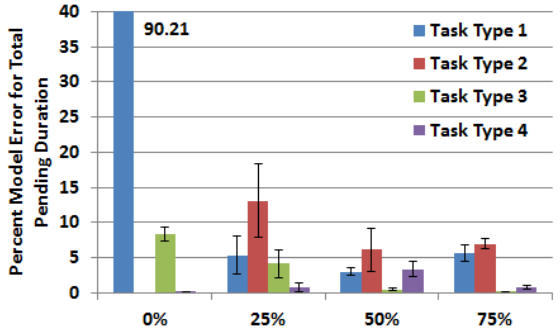
Compute Cluster C



Compute Cluster D



Compute Cluster E



Compute Cluster F

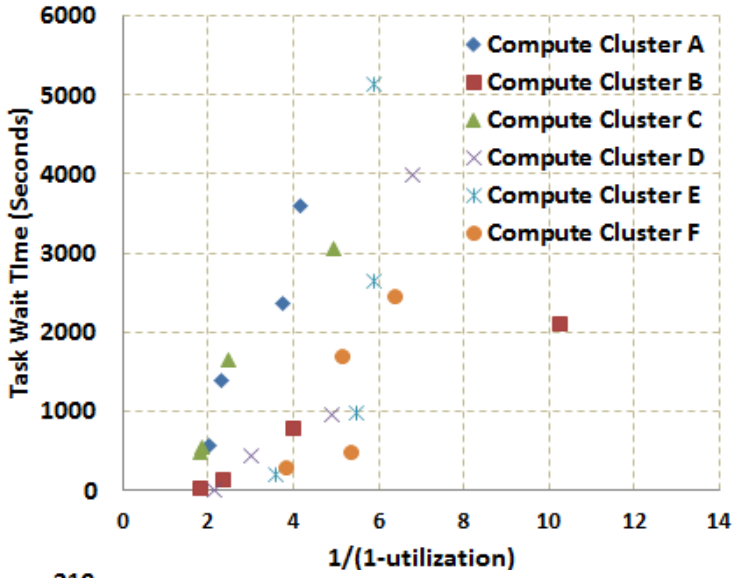
The mean model is accurate for reproducing **average** task wait time

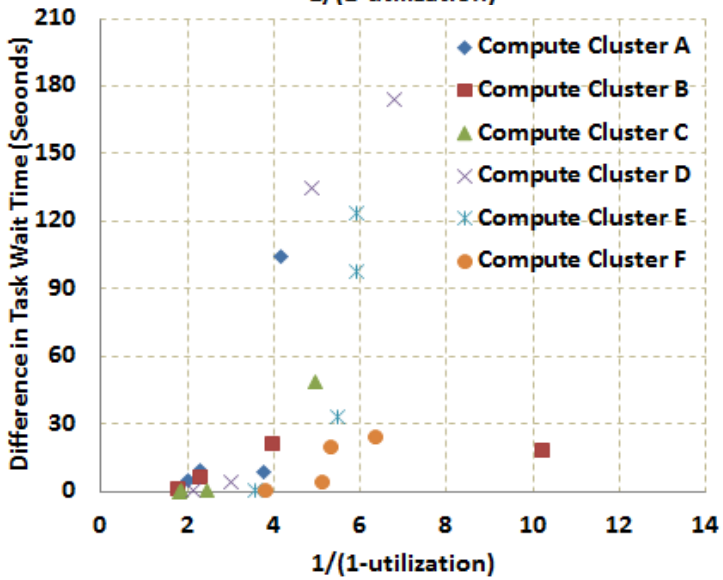
Analysis of Simulation Results

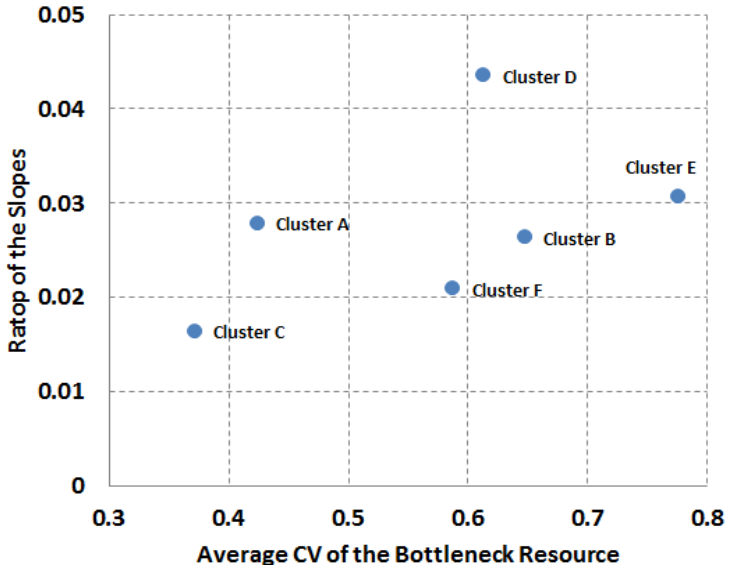
- The mean usage model
  - Performs well for predicting resource utilization for all resource types (5% error)
  - Performs moderately well for predicting task wait time (10 - 20% error on average)

- Interpreting model errors
  1. Understand the impact of utilization on performance metrics
  2. Correlate estimation error with theoretical model error (i.e., CV of task usage shapes)

Analysis Result for Task Wait Time

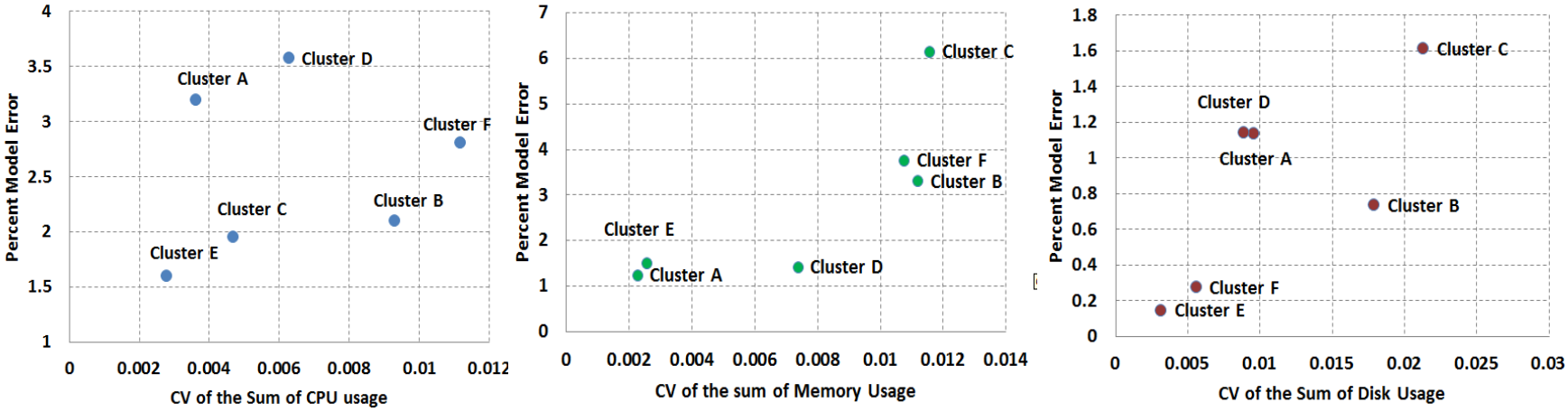






- Both task wait time and difference in task wait time grow exponentially with utilization
- The ratio of growth rate positively correlate with average CV

Analysis Result for Resource Utilization



- Utilization has low impact on model error utilization
- Correlating model error with CV of sum weighted by time
  - Using the fact that variance of the sum is the sum of variance
  - Short task has less impact on model error than long tasks

Conclusion and Future Work

- We studied the problem of deriving characterization models for task usage shapes in Google's compute cloud.
  - For performance forecasting and analysis in hypothetic scenarios
- We show that simply capturing the mean usage of each task (i.e., the mean usage model) is sufficient for capturing workload performance in terms of resource utilization and task wait time
- Future (on-going) work
  - Capture more fine-grained workload characteristics
  - Using clustering algorithms to find more accurate clusters.

Thank you!