# Towards Characterizing Cloud Backend Workloads: Insights from Google Compute Clusters

Asit K. Mishra, Chita R. Das, The Pennsylvania State University University Park
Joseph L. Hellerstein, Walfredo Cirne, Google Inc.

## Contribution

--describes an approach to workload classification and its application to the Google Cloud Backend

  --the Google Cloud Backend workload is a collection of tasks, each of which executes on a single machine in a compute cluster

  --employ a medium-grain approach to task classification

--method for workload classification

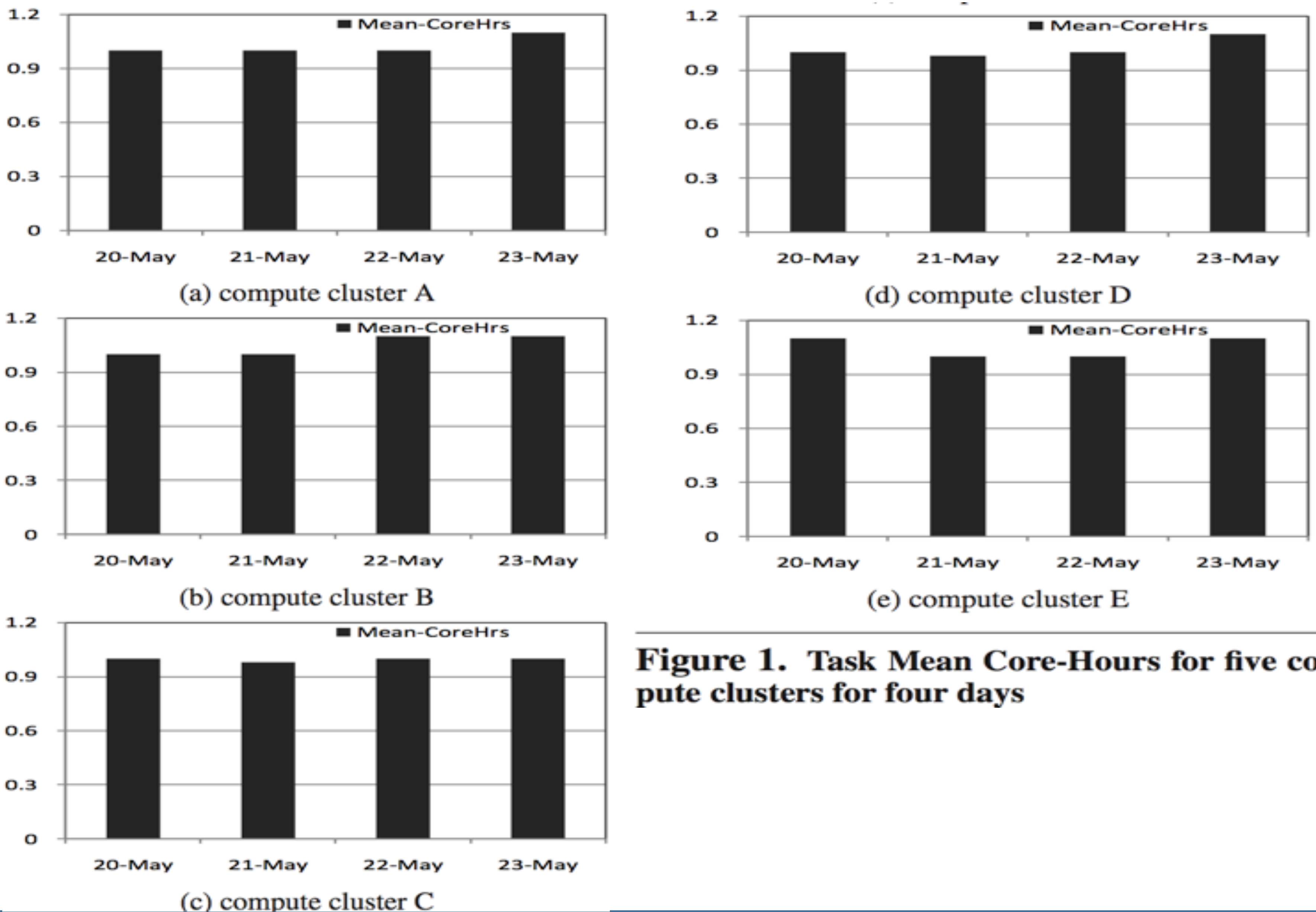
--based on the notion of qualitative coordinates, glean several insights about the Google Cloud Backend

# Coarse-Grain Task Classification

--there is a single workload for all tasks
  --The data consist of records collected from five Google production compute clusters over four days
  --A record reports on a task’s execution over a five minute interval

--compute the mean and standard deviation of resource usage
  --CPU usage is reported as the average number of cores used by the task over the five minute interval, similar to memory

--define a multi-dimensional representation of task resource usage
  --time in seconds
  --CPU usage in cores
  --memory usage in gigabytes
  --use metrics that combine time with a resource dimension: hours, core-hours, GB-hours




(a) compute cluster A

(b) compute cluster B

(c) compute cluster C

(d) compute cluster D

(e) compute cluster E

**Figure 1. Task Mean Core-Hours for five compute clusters for four days**

--there is consistency in mean core-hours from day to day on the same compute cluster

--there is little difference in mean core-hours between compute clusters

--in almost all compute clusters, tasks consume approximately 1 core-hour of CPU

# Coarse-Grain Task Classification

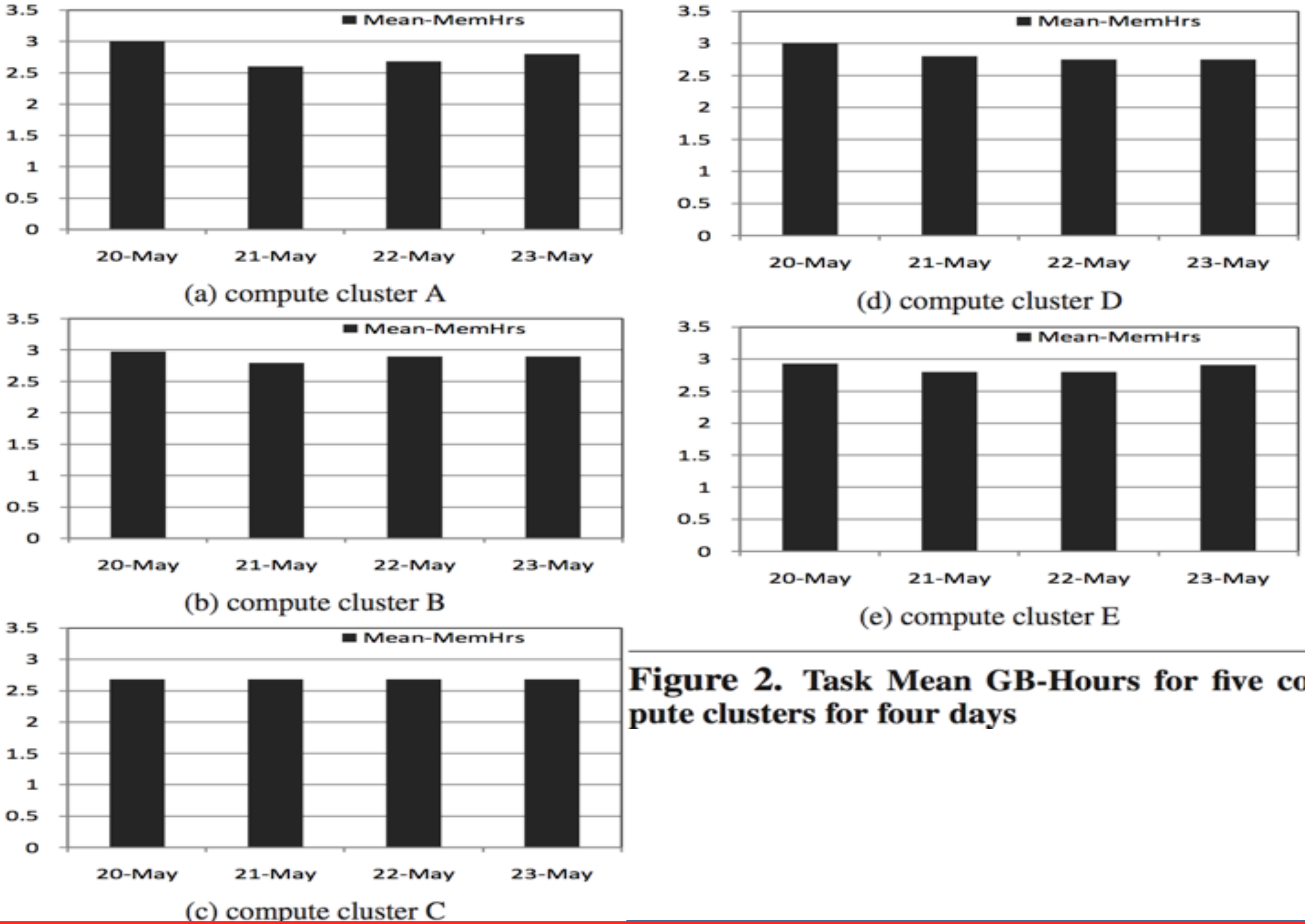



**Figure 2. Task Mean GB-Hours for five compute clusters for four days**

--there is consistency in mean core-hours from day to day on the same compute cluster

--there is little difference in mean core-hours between compute clusters

--in almost all compute clusters, mean GB-hours is about 3.0

--This observation remains unchanged if we use the median value instead of the mean

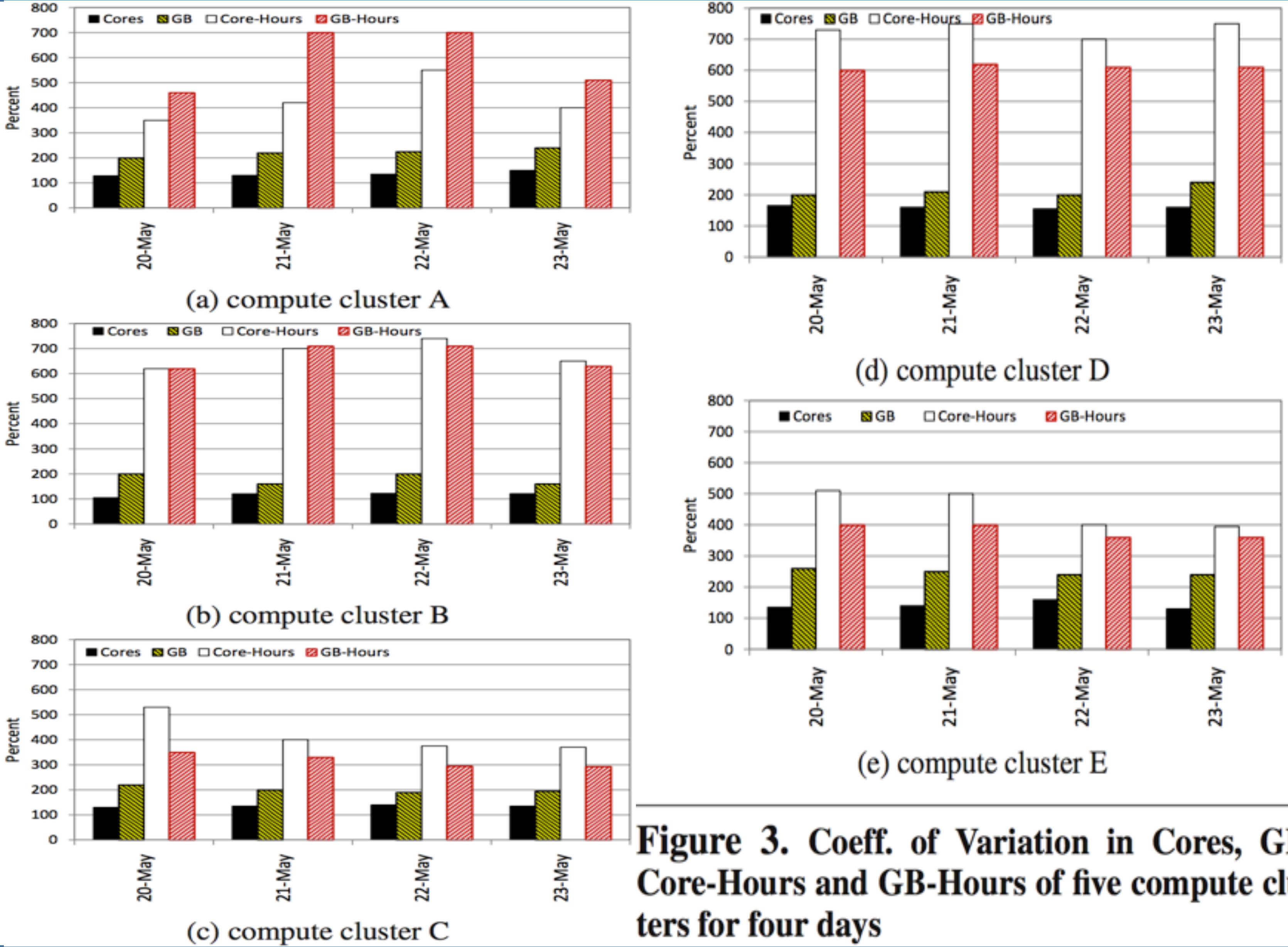



**Figure 3. Coeff. of Variation in Cores, GBs, Core-Hours and GB-Hours of five compute clusters for four days**

--quantify variability in terms of coefficient of variation (CV), the ratio of the standard deviation to the mean (often expressed as a percent)

--guideline is that CV should be much less than 100%

--it results in excessive variability in mean core-hours and GB-hours that makes it even more difficult to draw inferences about resource usage

# Methodology for Constructing Task Classifications

--objective is to construct a small number of task classes such that tasks within each class have similar resource usage
--qualitative coordinates are small, medium, and large
--there are three workload dimensions and each dimension has three qualitative coordinates
   --there are potentially 27 workloads

# Methodology for Constructing Task Classifications

--must address three challenges in this method

- --tasks within a workload should have very similar resource demands as quantified by their within class CV for each workload dimension
- --method must provide a way to compute numeric breakpoints that define the boundaries between qualitative coordinates for each workload dimension
- --minimize the number of workloads

# Methodology for Constructing Task Classifications

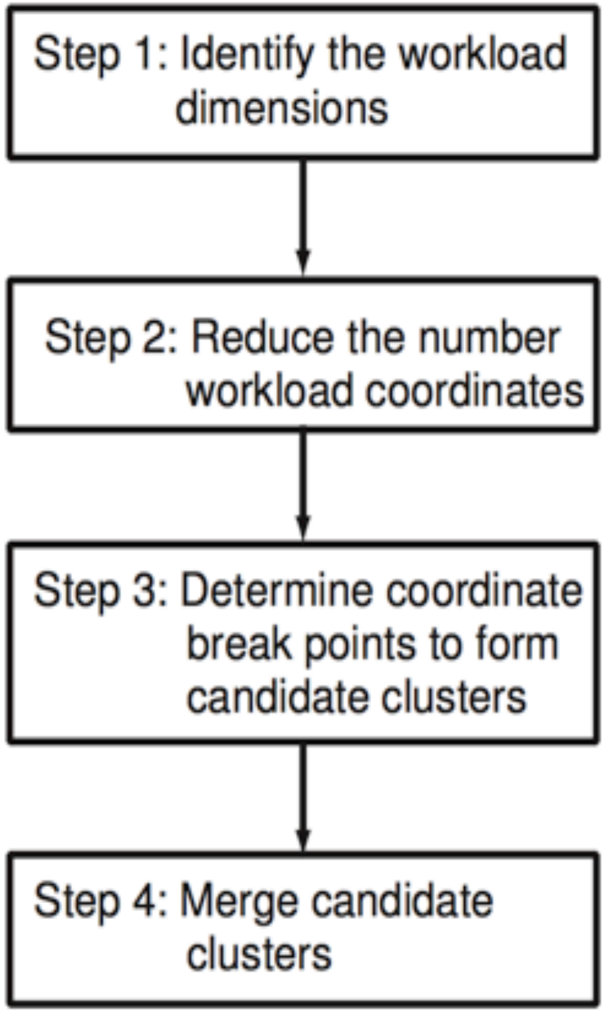



**Figure 4. Methodology for constructing task classifications**

--second step constructs preliminary task classes that have fairly homogeneous resource usage. Using the workload dimensions as a feature vector and applying an off-the-shelf clustering algorithm such as k-means

--determines the break points for the qualitative coordinates of the workload dimensions

--Involves combining "adjacent" preliminary task classes. Adjacency is based on the qualitative coordinates of the class

--the workload smm is adjacent to sms and sml in the third dimension

--Two preliminary classes are merged if the CV of the merged classes does not differ much from the CVs of each of the preliminary classes. Merged classes are de-noted by "*"

# Classification and Resource Characterization for Google Tasks--Task Classification

--With three workload dimensions, this results in 27 preliminary task classes.

--In the analysis of task durations, we observed that that task duration is bimodal: small, large

--only consider the qualitative coordinates small and large for the duration dimension

--reduces the number of preliminary task classes from 27 to 18

--k-means to the re-scaled data to calculate 18 task classes for each compute cluster

| Preliminary Class | Duration(Hours) | CPU (cores) | Memory (GBs) |
|---|---|---|---|
| 1 | Small 0.0833 | Small 0.08 | Small 0.48 |
| 2 | Small 0.0834 | Small 0.19 | Medium 0.67 |
| 3 | Small 0.0956 | Small 0.18 | Large 1.89 |
| 4 | Small 0.0888 | Medium 0.34 | Small 0.38 |
| 5 | Small 0.4466 | Medium 0.47 | Medium 0.81 |
| 6 | Small 0.4166 | Medium 0.38 | Large 1.04 |
| 7 | Small 0.4366 | Large 1.23 | Small 0.44 |
| 8 | Small 0.8655 | Large 0.98 | Medium 0.91 |
| 9 | Small 0.4165 | Large 1.39 | Large 1.54 |
| 10 | Large 18.34 | Small 0.12 | Small 0.48 |
| 11 | Large 19.34 | Small 0.16 | Medium 0.85 |
| 12 | Large 22.23 | Small 0.16 | Large 1.66 |
| 13 | Large 22.83 | Medium 0.38 | Small 0.38 |
| 14 | Large 19.34 | Medium 0.28 | Medium 0.77 |
| 15 | Large 16.89 | Medium 0.41 | Large 1.76 |
| 16 | Large 17.57 | Large 1.89 | Small 0.48 |
| 17 | Large 22.23 | Large 2.34 | Medium 0.97 |
| 18 | Large 20.81 | Large 2.22 | Large 2.09 |

**Table 1. Clustering results with 18 task classes (numbers represent mean value for each dimension)**

| Qualitative Coordinate | Duration (Hours) | CPU (cores) | Memory (GBs) |
|---|---|---|---|
| **Small/Small** | 0 -2 | 0 - 0.2 | 0 - 0.5 |
| **Medium** | | 0.2 - 0.5 | 0.5 - 1 |
| **Large/Large** | 2-24 | 0.5 - 4 | > 1 |

**Table 2. Breakpoints for Small, Medium and Large for duration, cpu and memory**

| Final Class | Duration(Hours) | CPU (cores) | Memory (GBs) |
|---|---|---|---|
| **1: sss** | Small | Small | Small |
| **2: sm*** | Small | Med | **all** |
| **3: slm** | Small | Large | Small+Med |
| **4: sll** | Small | Large | Large |
| **5: lss** | Large | Small | Small |
| **6: lsl** | Large | Small | Large |
| **7: llm** | Large | Med+Large | Small+Med |
| **8: lll** | Large | Med+Large | Large |

**Table 3. Final task classes (workloads)**

# Classification and Resource Characterization for Google Tasks--Assessments

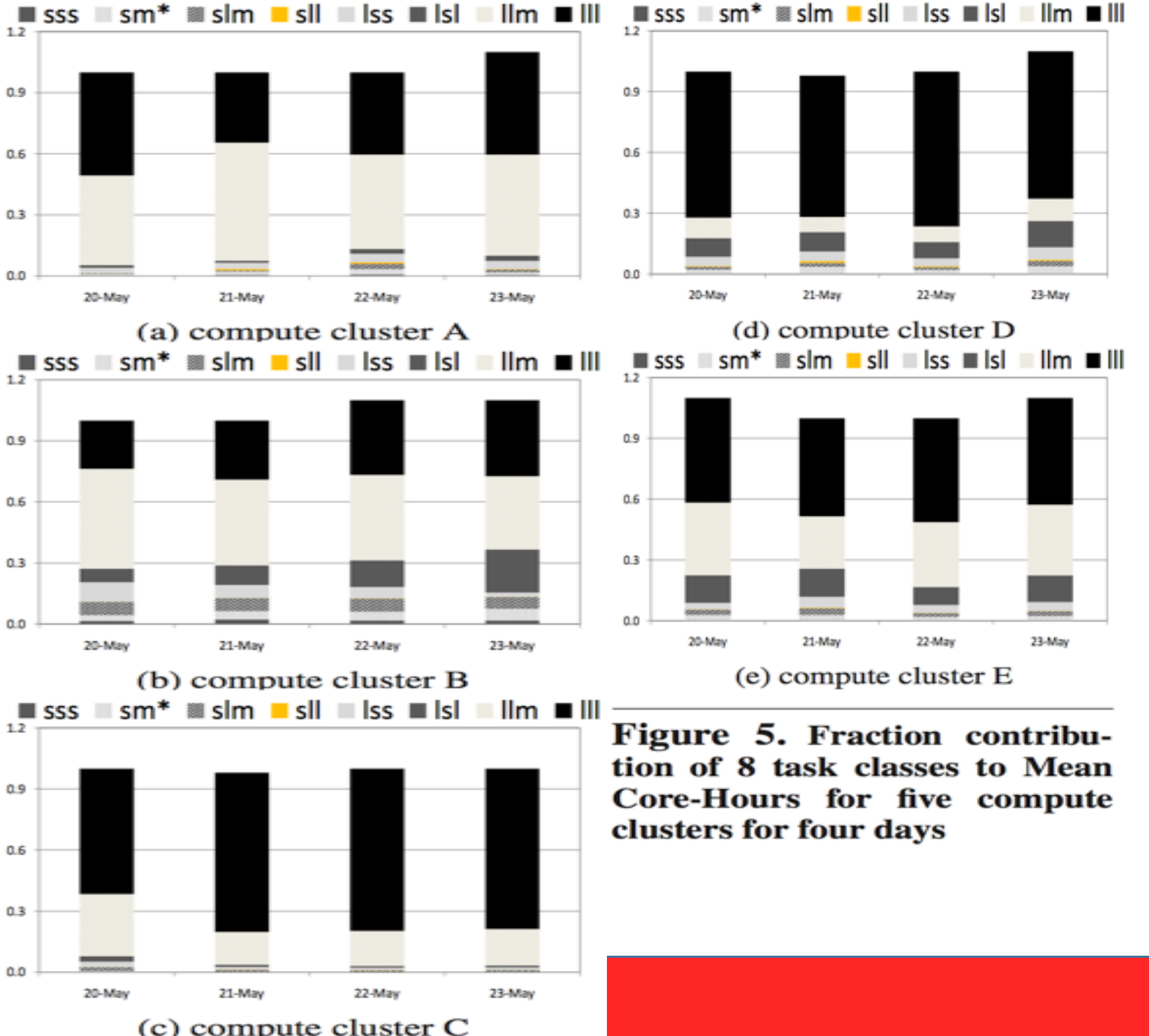
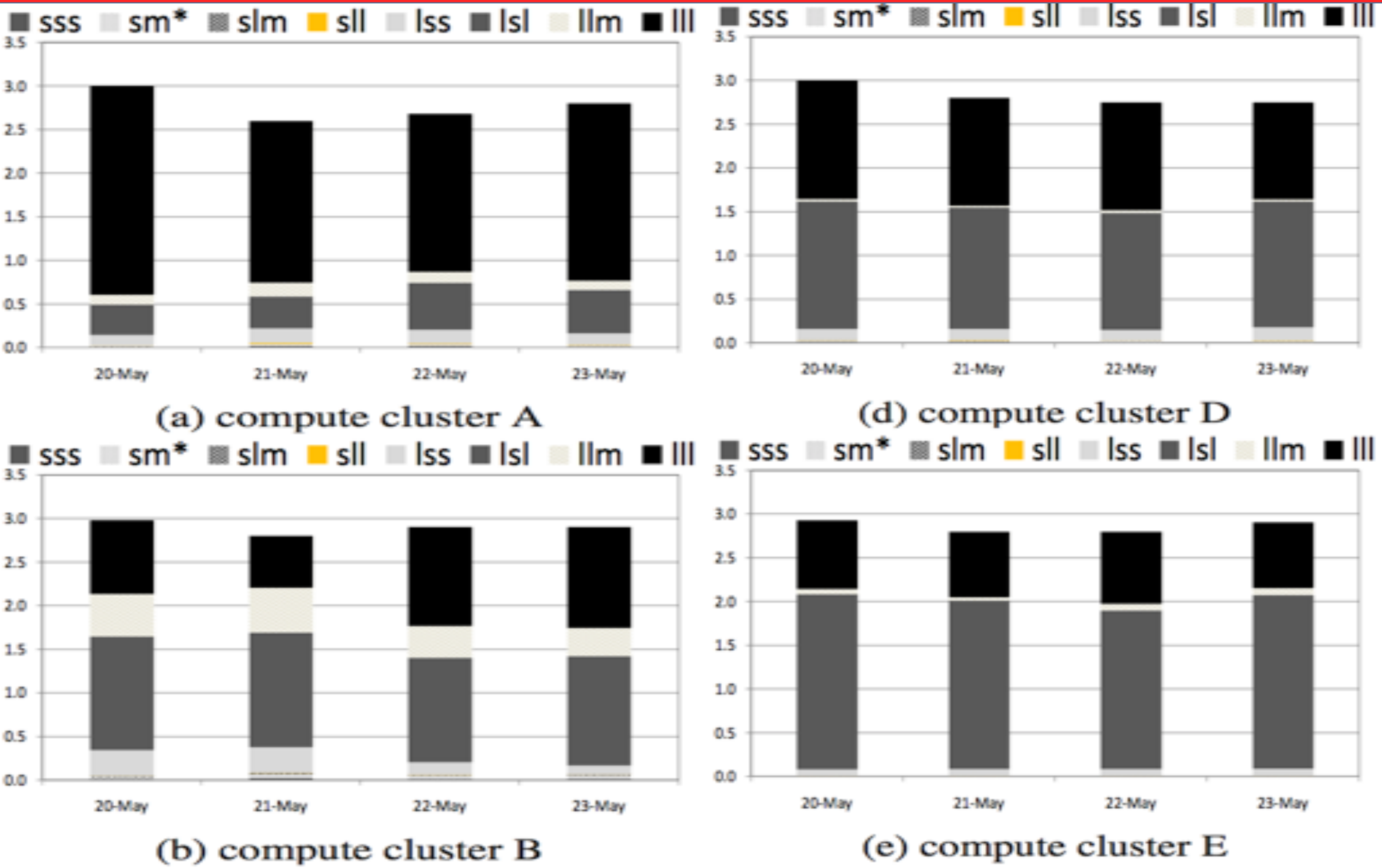



**Figure 5. Fraction contribution of 8 task classes to Mean Core-Hours for five compute clusters for four days**

--the contribution of the task classes to mean core-hours is consistent from day to day for the same compute cluster

--there are significant differences between compute clusters as to the contribution of task classes



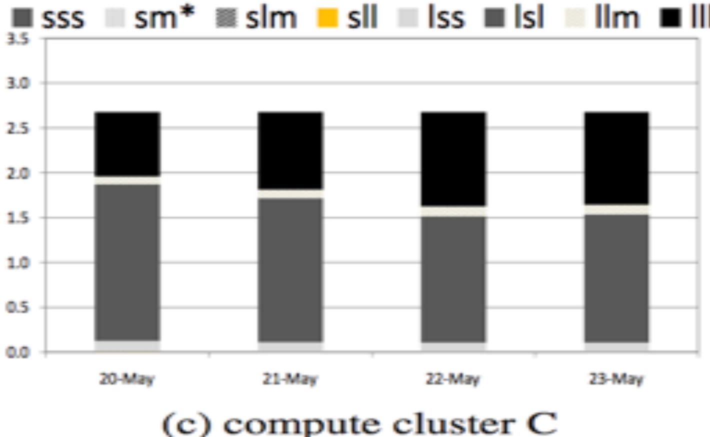




**Figure 6. Fraction contribution of 8 clusters to Mean GB-Hours for five compute clusters for four days**

--contributions by task class to mean GB-hours are quite similar from day to day within the same compute cluster

--there are dramatic differences between compute clusters in terms of the contribution to mean GB-hours by task class

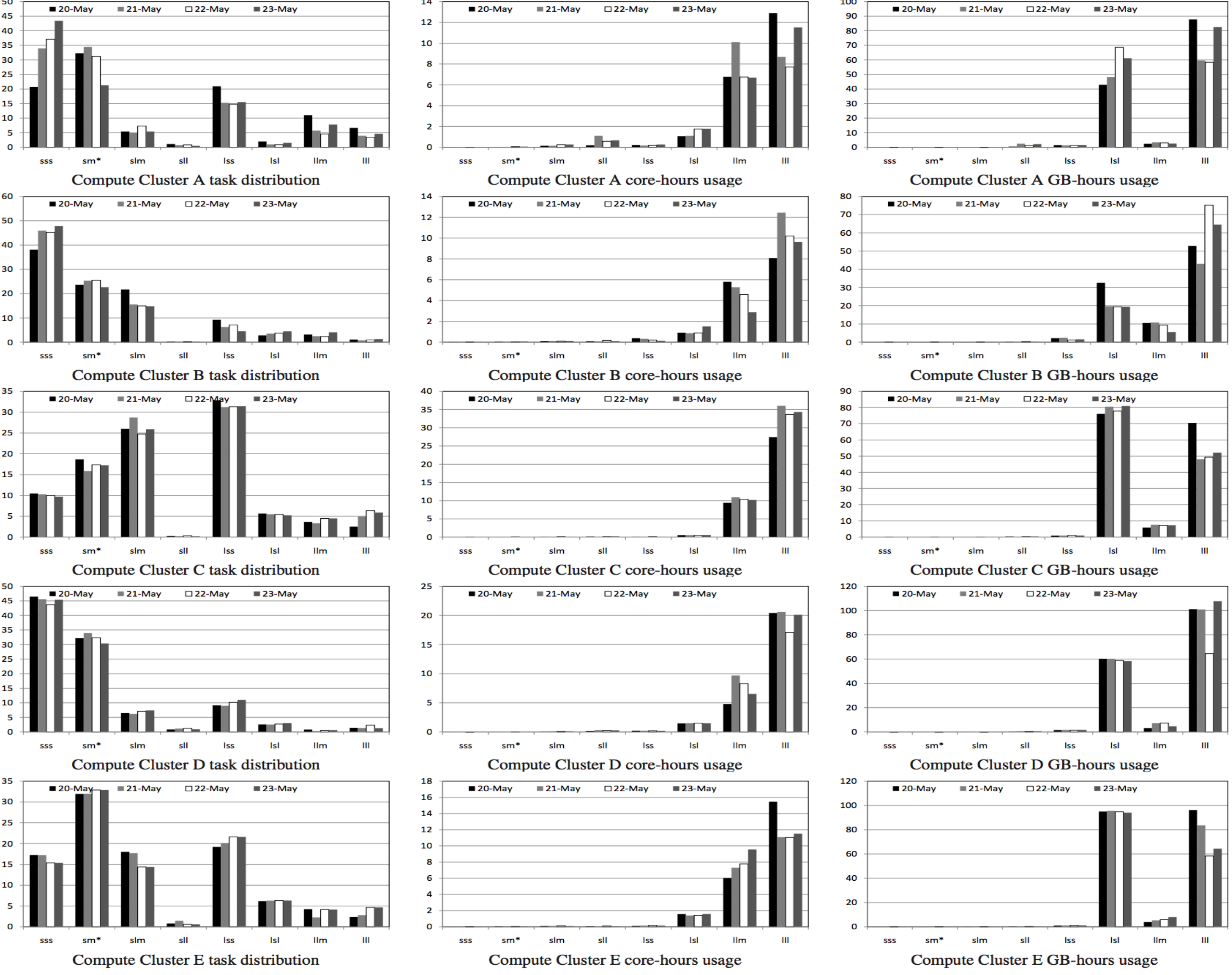



**Figure 7. Percentage distribution of each class and their corresponding resource utilization for five compute clusters from 20 May - 23 May**

--its horizontal axis the 8 final task classes, with bars grouped by day

--Although there is an occasional exception to day to day consistency within the same compute cluster, in general, the bar groups are very similar

--along any column with the same categorical coordinate, we see considerable difference.

--tasks with short duration dominate the task population

--a small number of long running tasks consume most of the CPU and memory


(a) Compute Cluster a, May 20 (b) Compute Cluster a, May 21 (c) Compute Cluster a, May 22 (d) Compute Cluster a, May 23

(a) Compute Cluster b, May 20 (b) Compute Cluster b, May 21 (c) Compute Cluster b, May 22 (d) Compute Cluster b, May 23

(a) Compute Cluster c, May 20 (b) Compute Cluster c, May 21 (c) Compute Cluster c, May 22 (d) Compute Cluster c, May 23

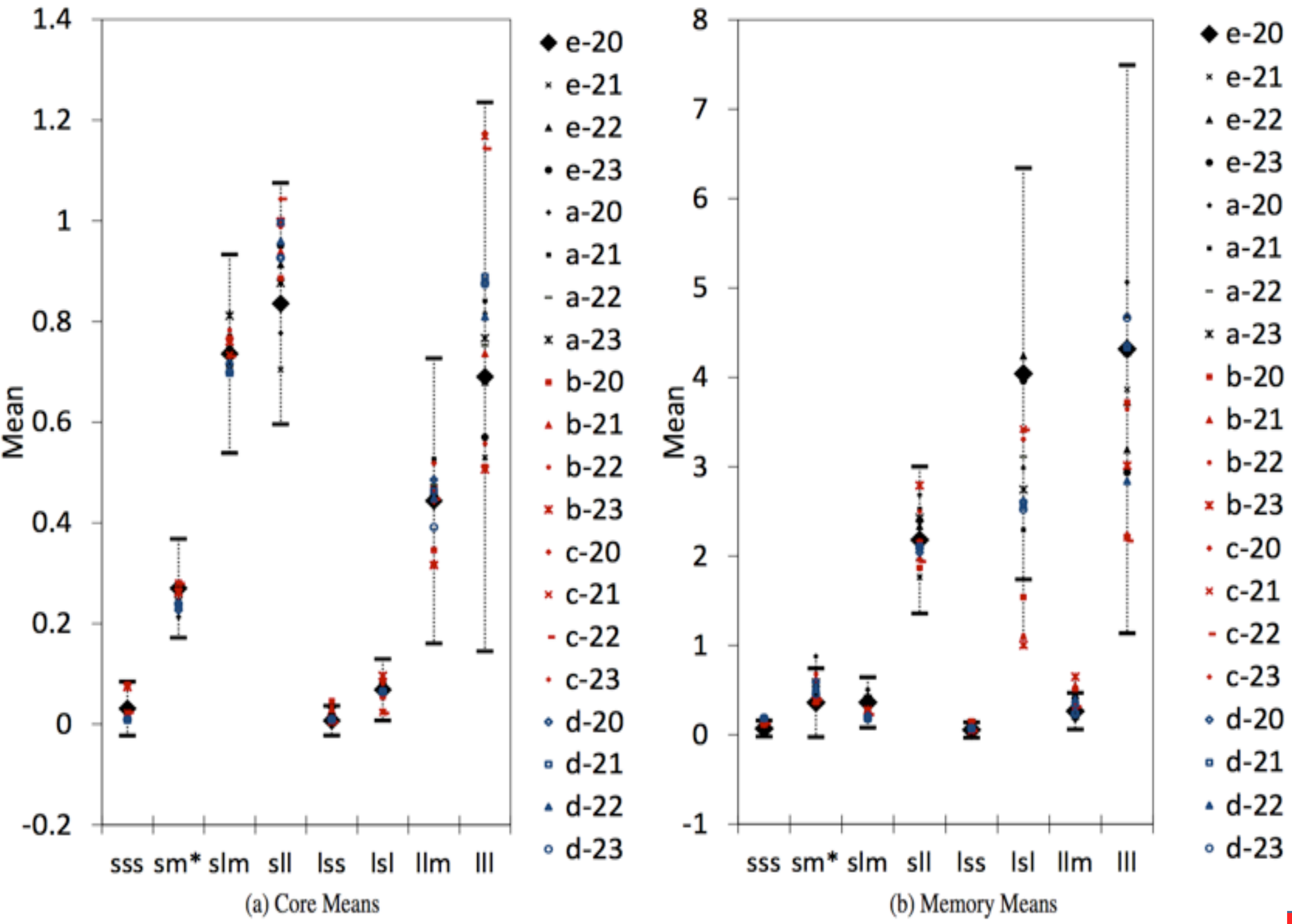
(a) Compute Cluster d, May 20 (b) Compute Cluster d, May 21 (c) Compute Cluster d, May 22 (d) Compute Cluster d, May 23

(a) Compute Cluster e, May 20 (b) Compute Cluster e, May 21 (c) Compute Cluster e, May 22 (d) Compute Cluster e, May 23

**Figure 8. Coefficient of Variation by Task Class**

--CV is always less than 100%, and is consistently less than 50%

--This is a significant reduction from the large CVs in the coarse-grain task classification




--plots mean core-hours by task class for each task-cluster-day studied with error bars for one standard deviation.

--Although task class lll has substantial variability, the mean values of the other task classes are grouped close to the center of the error bars

--a similar behavior for GB-hours

**Figure 9.** Consistency of class means with respect to cpu and memory across compute-cluster-days

# Classification and Resource Characterization for Google Tasks—Insight from task classification

--the duration of task executions is bimodal in that tasks either have a short duration or a long duration

--most tasks have short durations

--most resources are consumed by a few tasks with long duration that have large demands for CPU and memory.

# Application

--capacity planning
  --determines which machine resources must grow by how much to meet future application demands
  --The task classifications developed in this paper provide a way for Google to forecast application growth by tracking changes in task resource consumption by task class

--task scheduling
  --placing tasks on machines to maximize machine utilizations and to meet service level objectives
  --use runtime statistics of tasks on machines to determine which task class they belong to
  --estimate the membership of a newly arrived task based on the (prior) distribution of task cluster memberships