HETEROGENEITY AND DYNAMICITY OF CLOUDS AT SCALE: GOOGLE TRACE ANALYSIS

Charles Reiss (UC Berkeley), Alexey Tumanov(CMU), Greg Ganger (CMU), Randy H. Katz (UC berkeley), Michael A. Kozuch (Intel Labs)

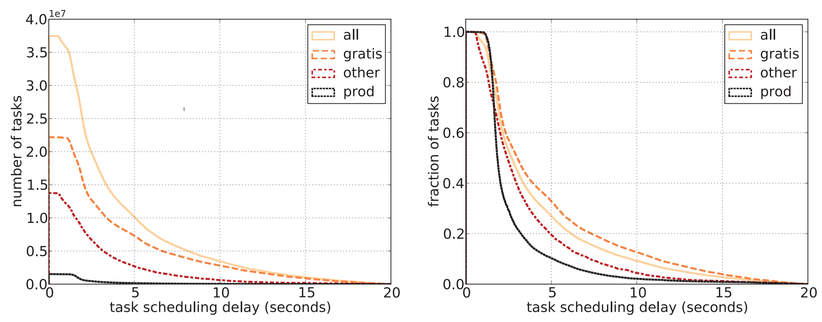
TRACE OVERVIEW

- Massive heterogeneous clouds on the rise
  - Variation in both hardware and software
- Little info available about such clouds
  - Making systems research difficult
- Google has released some trace/usage data
  - Covers one month on one 12500-machine "cell"
  - Includes resource requests, allocations, usage, constraints

Trace Characteristic	Value
time span	29 days
jobs run	650,000
number of active users	925
tasks submitted	25M
scheduler events	143M
compressed size	40 GB

TASK SCHEDULING WITH CONSTRAINTS

- Scheduling latency
  - Distributed over a wide range:
    - Min:0, median:3s, max:20.3days
  - Heavy-tailed distribution
  - Correlated with priority
  - Highest priority tasks are hard to schedule

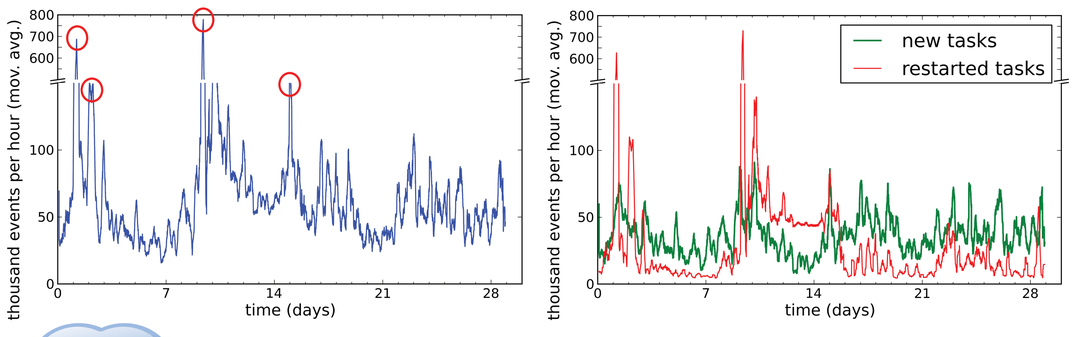


- Constraints are challenging to accommodate
  - Constrained tasks spend 46% more time in the queue
  - Scheduling latency correlates with anti-affinity

priority group	mean scheduling delay (s)
gratis (0 - 1)	242.9
other (2 - 8)	202.3
production (9 - 11)	52.2
9 - 10 only	5.4

DYNAMICITY

- 100K+ scheduler decisions/hour
- Short-duration tasks
- Resubmissions (22M): evictions, fail/restart, kill/restart



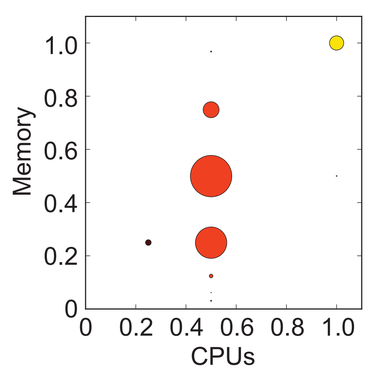
GOOGLE CLUSTER HETEROGENEITY

- Machine configuration variation
  - Cores & memory are independently normalized

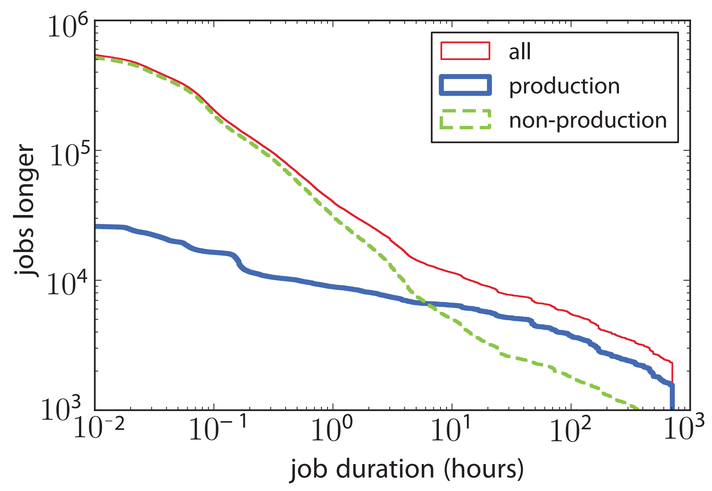
Factor of 4

Count	Platform	CPU	Memory
6732	B	0.50	0.50
3863	B	0.50	0.25
1001	B	0.50	0.75
795	C	1.00	1.00
126	A	0.25	0.25
<100	B and C	(various)	(various)

Three micro-architectures

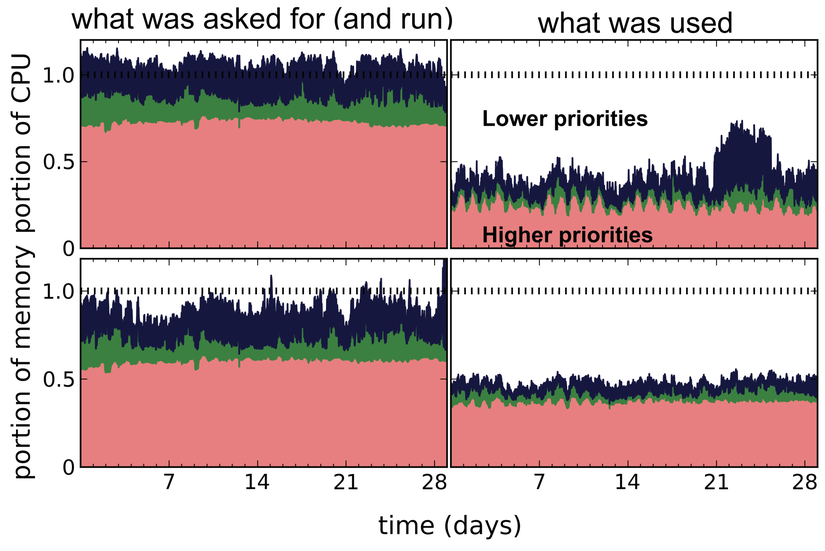


- Resource request variation
  - Large range of task shapes (boulders & sand)
    - 14,000 distinct request shapes (goodbye slots!)
- Job and task duration variation: orders-of-magnitude



UTILIZATION VS. ALLOCATION

- *Average* CPU & memory utilization: 40-60% of allocation
  - Typical non-Google: 7-25%
  - Uses over-commitment & preemption



C.Reiss, A.Tumanov, G.R.Ganger, R.H.Katz, M.A.Kozuch, Heterogeneity and Dynamicity of Clouds at Scale: Google Trace Analysis. In SoCC'12, Oct 2012.

Intel Science & Technology Center for Cloud Computing

Carnegie Mellon University · Georgia Tech · intel · PRINCETON UNIVERSITY · UC Berkeley