Università
della
Svizzera
italiana

Faculty
of Informatics

Bachelor Thesis

May 13, 2021

Understanding and Comparing Unsuccessful Executions in Large Datacenters

Claudio Maggioni

Abstract

The project aims at comparing two different traces coming from large datacenters, focusing in particular on unsuccessful executions of jobs and tasks submitted by users. The objective of this project is to compare the resource waste caused by unsuccessful executions, their impact on application performance, and their root causes. We will show the strong negative impact on CPU and RAM usage and on task slowdown. We will analyze patterns of unsuccessful jobs and tasks, particularly focusing on their interdependency. Moreover, we will uncover their root causes by inspecting key workload and system attributes such asmachine locality and concurrency level.

Advisor
Prof. Walter Binder
Assistant
Dr. Andrea Rosá

Advisor's approval (Prof. Walter Binder): Date:

Introduction (including Motivation)

State of the Art

- Introduce Ros'a 2015 DSN paper on analysis
- Describe Google Borg clusters
- Describe Traces contents
- Differences between 2011 and 2019 traces

Project requirements and analysis

(describe our objective with this analysis in detail)

Analysis methodology

Technical overview of traces' file format and schema

Overview on challenging aspects of analysis (data size, schema, avaliable computation resources)

Introduction on apache spark

General workflow description of apache spark workflow

The Google 2019 Borg cluster traces analysis were conducted by using Apache Spark and its Python 3 API (pyspark). Spark was used to execute a series of queries to perform various sums and aggregations over the entire dataset provided by Google.

In general, each query follows a general Map-Reduce template, where traces are first read, parsed, filtered by performing selections, projections and computing new derived fields. Then, the trace records are often grouped by one of their fields, clustering related data toghether before a reduce or fold operation is applied to each grouping.

Most input data is in JSONL format and adheres to a schema Google profided in the form of a protobuffer specification[1].

On of the main quirks in the traces is that fields that have a "zero" value (i.e. a value like 0 or the empty string) are often omitted in the JSON object records. When reading the traces in Apache Spark is therefore necessary to check for this possibility and populate those zero fields when omitted.

Most queries use only two or three fields in each trace records, while the original records often are made of a couple of dozen fields. In order to save memory during the query, a projection is often applied to the data by the means of a .map() operation over the entire trace set, performed using Spark's RDD API.

Another operation that is often necessary to perform prior to the Map-Reduce core of each query is a record filtering process, which is often motivated by the presence of incomplete data (i.e. records which contain fields whose values is unknown). This filtering is performed using the .filter() operation of Spark's RDD API.

The core of each query is often a groupBy followed by a map() operation on the aggregated data. The groupby groups the set of all records into several subsets of records each having something in common. Then, each of this small clusters is reduced with a .map() operation to a single record. The motivation behind this computation is often to analyze a time series of several different traces of programs. This is implemented by groupBy()-ing records by program id, and then map()-ing each program trace set by sorting by time the traces and computing the desired property in the form of a record.

Sometimes intermediate results are saved in Spark's parquet format in order to compute and save intermediate results beforehand.

General Query script design

Ad-Hoc presentation of some analysis scripts (w diagrams)

[1] Google 2019 Borg traces Protobuffer specification on Github

Analysis (w observations)

machine_configs

Refer to figure 1.

Observations:

- machine configurations are definitely more varied than the ones in the 2011 traces
- some clusters have more machine variability

machine_time_waste

Refer to figures 2 and 3.

Observations:

- Across all cluster almost 50% of time is spent in "unknown" transitions, i.e. there are some time slices that are related to a state transition that Google says are not "typical" transitions. This is mostly due to the trace log being intermittent when recording all state transitions.
- 80% of the time spent in KILL and LOST is unknown. This is predictable, since both states indicate that the job execution is not stable (in particular LOST is used when the state logging itself is unstable)
- From the absolute graph we see that the time "wasted" on non-finish terminated jobs is very significant
- Execution is the most significant task phase, followed by queuing time and scheduling time ("ready" state)
- In the absolute graph we see that a significant amount of time is spent to re-schedule evicted jobs ("evicted" state)
- Cluster A has unusually high queuing times

task_slowdown

Refer to figure 4

Observations:

- Priority values are different from 0-11 values in the 2011 traces. A conversion table is provided by Google;
- For some priorities (e.g. 101 for cluster D) the relative number of finishing task is very low and the mean slowdown is very high (315). This behaviour differs from the relatively homogeneous values from the 2011 traces.
- Some slowdown values cannot be computed since either some tasks have a 0ns execution time or for some priorities no tasks in the traces terminate successfully. More raw data on those exception is in Jupyter.
- The % of finishing jobs is relatively low comparing with the 2011 traces.

spatial_resource_waste

Refer to figures 5 and 6.

Observations:

- Most (mesasured and requested) resources are used by killed job, even more than in the 2011 traces.
- Behaviour is rather homogeneous across datacenters, with the exception of cluster G where a lot of LOST-terminated tasks acquired 70% of both CPU and RAM

figure_7

Refer to figures 7, 8, and 9.

Observations:

- No smooth curves in this figure either, unlike 2011 traces
- The behaviour of curves for 7a (priority) is almost the opposite of 2011, i.e. in-between priorities have higher kill rates while priorities at the extremum have lower kill rates. This could also be due bt the inherent distribution of job terminations;
- Event execution time curves are quite different than 2011, here it seems there is a good correlation between short task execution times and finish event rates, instead of the U shape curve in 2015 DSN

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	8729	1.639218%
1.000000	0.500000	124234	23.329891%
0.591797	0.333496	103013	19.344801%
0.259277	0.166748	78078	14.662260%
0.708984	0.333496	55801	10.478864%
0.386719	0.333496	36237	6.804943%
0.958984	0.500000	31151	5.849843%
0.708984	0.666992	29594	5.557454%
0.386719	0.166748	27011	5.072393%
1.000000	1.000000	12286	2.307187%
0.591797	0.166748	9902	1.859496%
1.000000	0.250000	7550	1.417814%
0.958984	1.000000	3552	0.667030%
0.259277	0.333496	3024	0.567877%
0.591797	0.666992	1000	0.187790%
0.259277	0.083374	634	0.119059%
0.958984	0.250000	600	0.112674%
0.500000	0.062500	54	0.010141%
0.500000	0.250000	34	0.006385%
0.479492	0.250000	12	0.002253%
0.708984	0.250000	6	0.001127%
0.591797	0.250000	4	0.000751%
0.708984	0.500000	2	0.000376%
0.479492	0.500000	2	0.000376%

(a) All clusters

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	1377	1.623170%
0.591797	0.333496	29487	34.758469%
1.000000	0.500000	13440	15.842705%
0.708984	0.333496	12495	14.728764%
0.386719	0.333496	9057	10.676144%
0.386719	0.166748	5265	6.206238%
0.708984	0.666992	4608	5.431784%
1.000000	1.000000	4446	5.240823%
0.591797	0.166748	2484	2.928071%
0.958984	0.500000	1143	1.347337%
0.958984	1.000000	654	0.770917%
1.000000	0.250000	366	0.431431%
0.479492	0.250000	6	0.007073%
0.708984	0.250000	6	0.007073%

(b) A cluster

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	134	0.264812%
0.591797	0.333496	16184	31.982926%
1.000000	0.500000	9790	19.347061%
0.708984	0.333496	8448	16.694992%
0.958984	0.500000	5502	10.873088%
0.708984	0.666992	3832	7.572823%
1.000000	1.000000	2214	4.375321%
0.591797	0.166748	2152	4.252796%
0.386719	0.333496	816	1.612584%
0.958984	1.000000	618	1.221296%
0.591797	0.666992	500	0.988103%
0.386719	0.166748	412	0.814197%

(c) Cluster B

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	1466	2.274208%
0.259277	0.166748	15754	24.439204%
0.386719	0.333496	11104	17.225652%
0.591797	0.333496	10404	16.139741%
0.958984	0.500000	6634	10.291334%
1.000000	0.500000	5654	8.771059%
0.386719	0.166748	3580	5.553660%
0.708984	0.666992	2900	4.498774%
1.000000	1.000000	2736	4.244361%
1.000000	0.250000	2132	3.307375%
0.958984	1.000000	766	1.188297%
0.708984	0.333496	620	0.961807%
0.958984	0.250000	600	0.930781%
0.591797	0.166748	112	0.173746%

(d) Cluster C

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	498	0.794309%
0.591797	0.333496	28394	45.288376%
0.386719	0.333496	8402	13.401174%
0.259277	0.166748	8020	12.791885%
0.386719	0.166748	5806	9.260559%
0.708984	0.666992	4380	6.986092%
0.708984	0.333496	3924	6.258772%
0.591797	0.166748	2548	4.064055%
0.259277	0.333496	426	0.679469%
1.000000	0.500000	292	0.465739%
0.591797	0.250000	4	0.006380%
0.708984	0.500000	2	0.003190%

(e) Cluster D

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	536	0.671915%
0.259277	0.166748	38452	48.202377%
0.708984	0.333496	11786	14.774608%
0.958984	0.500000	8646	10.838389%
0.708984	0.666992	7606	9.534674%
1.000000	0.500000	5586	7.002457%
0.386719	0.166748	4470	5.603470%
0.259277	0.333496	1268	1.589530%
0.259277	0.083374	634	0.794765%
0.591797	0.333496	324	0.406158%
1.000000	0.250000	268	0.335957%
1.000000	1.000000	138	0.172993%
0.500000	0.062500	54	0.067693%
0.500000	0.250000	4	0.005014%

(f) Cluster E

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	1432	2.299958%
1.000000	0.500000	41340	66.396839%
0.708984	0.333496	6878	11.046866%
0.591797	0.333496	5564	8.936430%
0.958984	0.500000	2172	3.488484%
0.386719	0.166748	1544	2.479843%
0.708984	0.666992	1244	1.998008%
1.000000	0.250000	792	1.272044%
0.958984	1.000000	536	0.860878%
0.386719	0.333496	398	0.639234%
1.000000	1.000000	344	0.552504%
0.500000	0.250000	18	0.028910%

(g) Cluster F

CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	1566	2.261568%
0.259277	0.166748	15852	22.892958%
1.000000	0.500000	11808	17.052741%
0.708984	0.333496	7968	11.507134%
0.591797	0.333496	7830	11.307839%
0.386719	0.166748	4690	6.773150%
0.708984	0.666992	4258	6.149269%
0.958984	0.500000	4196	6.059731%
0.386719	0.333496	3864	5.580267%
0.591797	0.166748	2606	3.763503%
1.000000	0.250000	2100	3.032754%
0.259277	0.333496	1330	1.920744%
0.958984	1.000000	778	1.123563%
1.000000	1.000000	378	0.545896%
0.500000	0.250000	12	0.017330%
0.479492	0.250000	6	0.008665%
0.479492	0.500000	2	0.002888%

(h) Cluster G

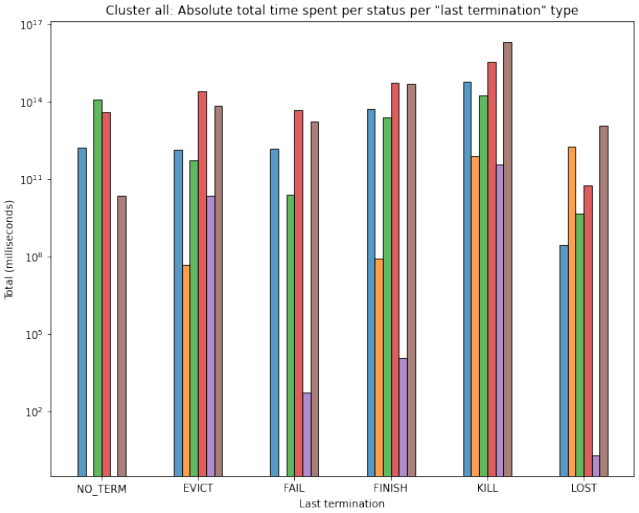
CPU (NCU)	RAM (NMU)	Machine count	% Machines
Unknown	Unknown	1720	2.933251%
1.000000	0.500000	36324	61.946178%
0.591797	0.333496	4826	8.230158%
0.708984	0.333496	3682	6.279205%
0.958984	0.500000	2858	4.873973%
0.386719	0.333496	2596	4.427163%
1.000000	1.000000	2030	3.461919%
1.000000	0.250000	1892	3.226577%
0.386719	0.166748	1244	2.121491%
0.708984	0.666992	766	1.306320%
0.591797	0.666992	500	0.852689%
0.958984	1.000000	200	0.341076%

(i) Cluster H

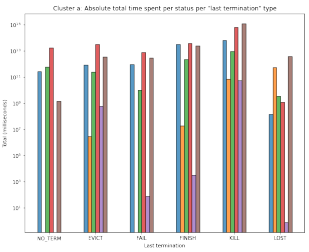
Figure 1. Overwiew of machine configurations in terms of CPU and RAM resources for each cluster

Color	Execution phase
Blue	Queued
Orange	Ended
Green	Ready
Red	Running
Violet	Evicted
Brown	Unknown

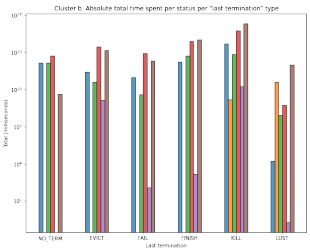
(a) Execution state legend for the graphs



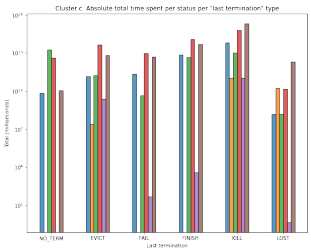
(b) All clusters



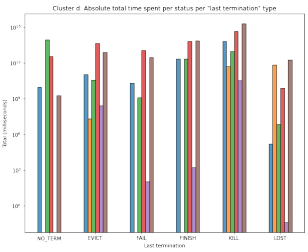
(c) Cluster A



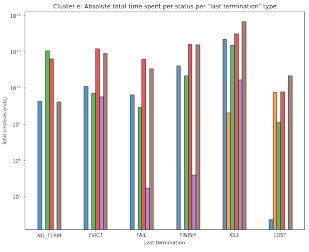
(d) Cluster B



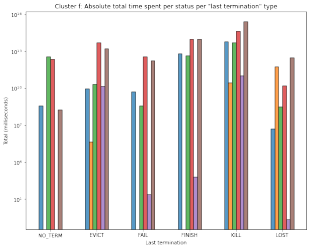
(e) Cluster C



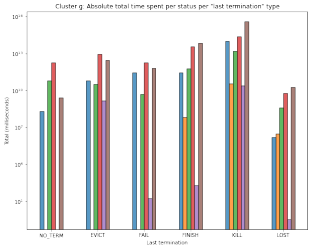
(f) Cluster D



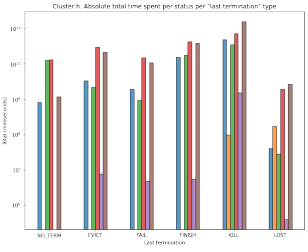
(g) Cluster E



(h) Cluster F



(i) Cluster G

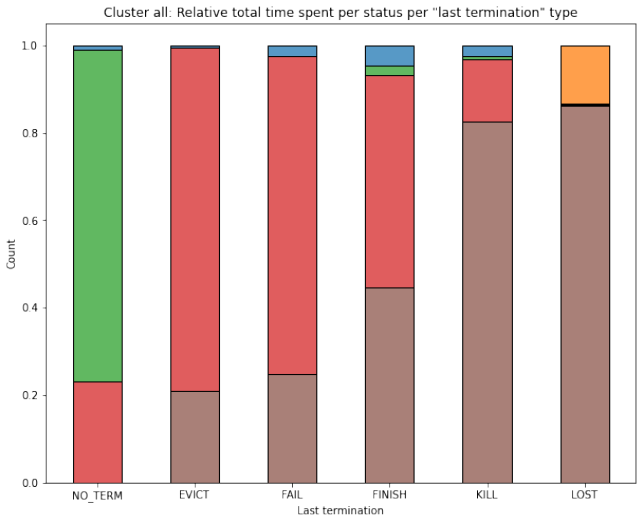


(j) Cluster H

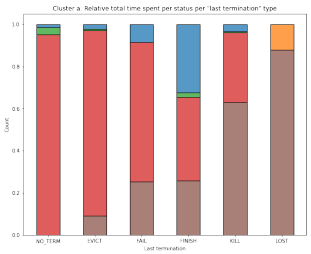
Figure 2. Total task time (in milliseconds) spent in each execution phase w.r.t. task termination.

Color	Execution phase
Blue	Queued
Orange	Ended
Green	Ready
Red	Running
Violet	Evicted
Brown	Unknown

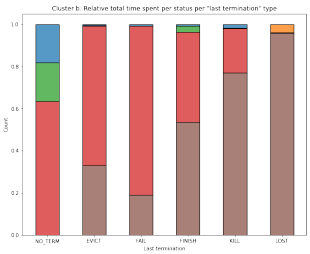
(a) Execution state legend for the graphs



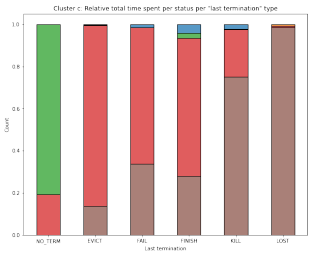
(b) All clusters



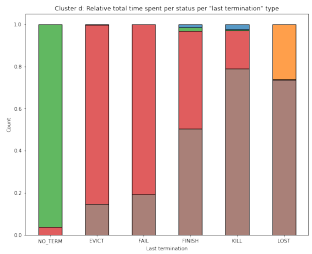
(c) Cluster A



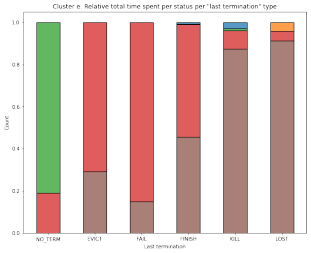
(d) Cluster B



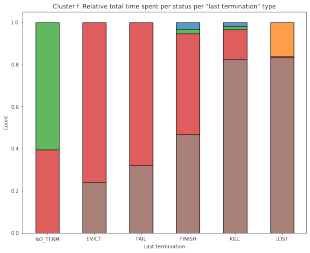
(e) Cluster C



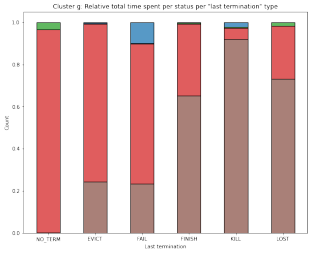
(f) Cluster D



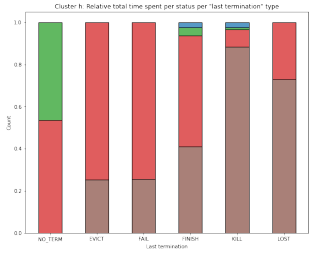
(g) Cluster E



(h) Cluster F



(i) Cluster G



(j) Cluster H

Figure 3. Relative task time (in milliseconds) spent in each execution phase w.r.t. task termination.

Priority	% finished tasks	Mean slowdown
Unknown	10.620113%	1.097556
24	0.000000%	–
25	0.333054%	82.973285
100	0.000000%	–
101	81.917703%	30.798089
102	0.000000%	–
103	14.990678%	1.130579
105	57.678214%	1.078733
107	53.926543%	1.016187
114	0.000000%	–
115	4.108501%	1.004324
116	13.045304%	1.032749
117	0.000000%	–
118	11.907081%	1.003494
119	21.264583%	1.504923
170	0.000000%	–
200	27.211754%	4.116760
205	0.000000%	–
210	0.000000%	–
214	0.000000%	–
215	0.000000%	–
360	0.616372%	2.924018
400	0.000000%	–
450	2.203423%	1.142450
500	0.000000%	–

(a) Cluster A

Priority	% finished tasks	Mean slowdown
0	45.193049%	1.176397
25	0.018094%	133.481864
80	0.000000%	–
100	0.000000%	–
101	66.479321%	433.414195
103	0.106377%	1.645114
105	0.463292%	2.408090
107	0.000000%	–
114	0.676897%	1.003422
115	4.117647%	5.916852
116	8.316438%	1.109652
117	0.000000%	–
118	0.311290%	1.000000
119	0.195997%	2.555160
170	0.000000%	–
199	0.000000%	–
200	30.916717%	9.707524
205	0.000000%	–
210	0.000000%	–
214	0.000000%	–
215	0.000000%	–
360	3.502999%	1.612147
450	0.612913%	1.057515

(b) Cluster B

Priority	% finished tasks	Mean slowdown
0	50.887820%	1.105787
3	0.000000%	–
10	0.000000%	–
25	22.468276%	8.191258
100	0.000000%	–
101	52.628263%	421.490544
103	0.005336%	2.794339
105	0.023521%	1.372291
107	0.000245%	14.708268
114	0.022221%	1.011266
115	0.281832%	1.980743
116	0.013836%	1.022119
117	93.165468%	1.000000
118	0.004137%	1.100009
119	2.215917%	2.044049
170	0.000000%	–
200	3.606796%	4.139724
205	0.000000%	–
210	0.000000%	–
214	0.000000%	–
215	0.000000%	–
360	4.367418%	2.061085
450	1.512578%	1.066014

(c) Cluster C

Priority	% finished tasks	Mean slowdown
0	26.522899%	1.116002
5	0.000000%	–
25	16.293068%	65.676400
100	0.000000%	–
101	45.314870%	315.954065
103	0.004540%	1.065721
105	0.051712%	2.897040
107	0.000350%	1.551354
114	0.000000%	–
115	5.189033%	2.186562
116	0.126154%	1.278510
117	85.714286%	1.000000
118	0.054055%	2.048749
119	0.441844%	3.020486
197	0.000000%	–
199	0.000000%	–
200	6.528759%	5.514350
205	0.000000%	–
210	0.000000%	–
214	0.000000%	–
215	0.000000%	–
360	1.594977%	2.476706
450	0.611145%	1.330248

(d) Cluster D

Priority	% finished tasks	Mean slowdown
0	42.805214%	1.439544
25	5.344531%	2.676136
100	0.000000%	–
101	0.015918%	1.122507
103	0.021660%	3.163046
105	0.404803%	14.750313
107	0.000000%	–
114	0.000000%	–
115	0.027326%	1.000000
116	0.000000%	–
117	0.000000%	–
118	0.000000%	–
119	0.458256%	10.310893
170	0.000000%	–
200	1.959258%	8.535722
201	0.000000%	–
205	0.000000%	–
210	0.000000%	–
215	0.000000%	–
220	0.000000%	–
360	37.157031%	2.873243
450	0.548458%	1.113283

(e) Cluster E

Priority	% finished tasks	Mean slowdown
0	45.208221%	1.088162
25	0.647505%	2.230960
100	0.000000%	–
101	40.296631%	323.858714
103	0.058418%	1.167347
105	0.222372%	1.550453
107	0.060860%	1.012727
114	0.006958%	1.000000
115	3.647104%	5.094215
116	0.000000%	–
117	0.000086%	1.000000
118	0.002082%	1.000000
119	31.354662%	7.608799
200	3.653528%	5.943247
201	0.000000%	–
360	7.424790%	2.171524
450	0.992623%	1.021053

(f) Cluster F

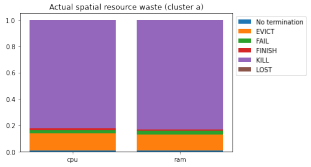
Priority	% finished tasks	Mean slowdown
0	33.612201%	1.138988
25	0.233338%	8.692558
50	0.000000%	–
100	0.000000%	–
101	96.470338%	19.378523
103	0.032539%	1.271282
105	0.196286%	1.000738
107	0.000000%	–
114	0.000000%	–
115	7.633588%	1.802068
117	0.000000%	–
118	48.969072%	3.877102
119	0.085944%	3.166077
170	0.000000%	–
200	26.747126%	14.573912
360	1.618878%	2.119524
450	2.737219%	1.036927

(g) Cluster G

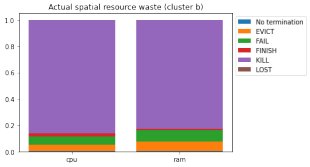
Priority	% finished tasks	Mean slowdown
0	27.744380%	1.122458
19	0.000000%	–
25	1.042767%	3.064188
101	100.000000%	76.438090
103	0.481256%	1.262067
105	1.427256%	4.205547
107	0.000000%	–
115	5.122494%	1.000000
116	1.035309%	73.447995
117	0.000050%	1.000000
118	1.003331%	1.947121
119	0.145214%	7.301093
200	2.702770%	5.798142
201	0.000000%	–
220	0.000000%	–
360	4.425746%	2.018441
450	0.535389%	1.054678

(h) Cluster H

Figure 4. Mean task slowdown for each cluster and each task priority

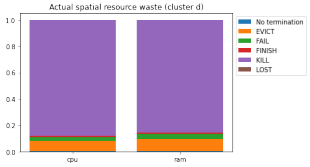


(a) Cluster A



(b) Cluster B

(c) Cluster C



(d) Cluster D

Task termination	% CPU	% RAM
No termination	0.6972%	1.0447%
Evict	13.4392%	11.8184%
Fail	2.2792%	2.8387%
Finish	1.3963%	1.1066%
Kill	82.1791%	83.1826%
Lost	0.0091%	0.0091%

(e) Cluster A (exact values)

Task termination	% CPU	% RAM
No termination	0.2582%	0.4637%
Evict	4.8340%	7.3120%
Fail	6.2950%	8.3841%
Finish	2.5877%	1.2231%
Kill	86.0215%	82.6144%
Lost	0.0036%	0.0027%

(f) Cluster B (exact values)

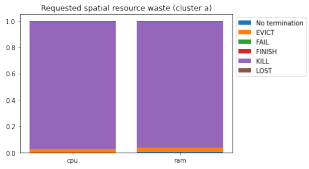
Task termination	% CPU	% RAM
No termination	0.3376%	0.3812%
Evict	8.2099%	8.0454%
Fail	1.2294%	2.0809%
Finish	2.9399%	3.3249%
Kill	87.2740%	86.1588%
Lost	0.0093%	0.0088%

(g) Cluster C (exact values)

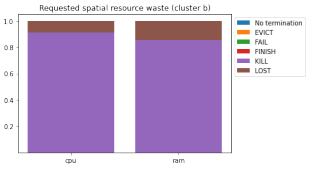
Task termination	% CPU	% RAM
No termination	0.4995%	0.4822%
Evict	7.6002%	9.0656%
Fail	3.0288%	3.9214%
Finish	0.8666%	0.8914%
Kill	88.0011%	85.6364%
Lost	0.0039%	0.0030%

(h) Cluster D (exact values)

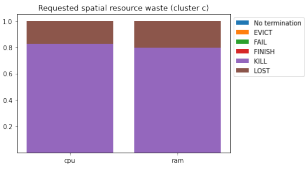
Figure 5. Relative usage of CPU and RAM resources w.r.t. final task termination.



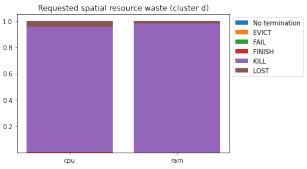
(a) Cluster A



(b) Cluster B



(c) Cluster C



(d) Cluster D

Task termination	% CPU	% RAM
No termination	0.033962%	0.193674%
Evict	2.838362%	3.399075%
Fail	0.058335%	0.069755%
Finish	0.000102%	0.000151%
Kill	96.661332%	95.799104%
Lost	0.407908%	0.538242%

(e) Cluster A (exact values)

Task termination	% CPU	% RAM
No termination	0.000094%	0.000191%
Evict	0.003365%	0.004696%
Fail	0.003061%	0.004965%
Finish	0.012696%	0.017647%
Kill	91.094839%	85.573746%
Lost	8.885947%	14.398756%

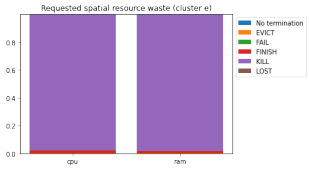
(f) Cluster B (exact values)

Task termination	% CPU	% RAM
No termination	0.000105%	0.000221%
Evict	0.008618%	0.006991%
Fail	0.001261%	0.001459%
Finish	0.015047%	0.017003%
Kill	82.483146%	79.698011%
Lost	17.491823%	20.276314%

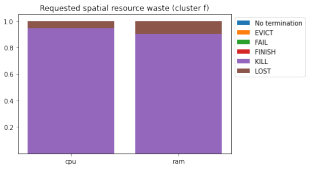
(g) Cluster C (exact values)

Task termination	% CPU	% RAM
No termination	0.000948%	0.000128%
Evict	0.046057%	0.006352%
Fail	0.023703%	0.002770%
Finish	0.095353%	0.012975%
Kill	95.468127%	97.927565%
Lost	4.365813%	2.050210%

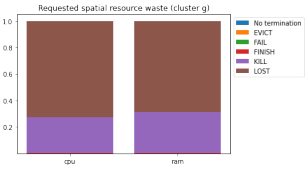
(h) Cluster D (exact values)



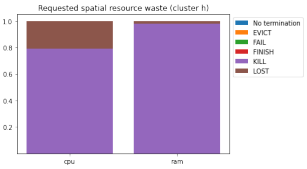
(i) Cluster E



(j) Cluster F



(k) Cluster G



(l) Cluster H

Task termination	% CPU	% RAM
No termination	0.015102%	0.016472%
Evict	0.362088%	0.321274%
Fail	0.051373%	0.047377%
Finish	1.672195%	1.310360%
Kill	97.899179%	98.304482%
Lost	0.000063%	0.000034%

(m) Cluster E (exact values)

Task termination	% CPU	% RAM
No termination	0.000114%	0.000306%
Evict	0.007986%	0.013466%
Fail	0.000913%	0.002064%
Finish	0.013296%	0.021751%
Kill	94.396548%	90.227868%
Lost	5.581144%	9.734546%

(n) Cluster F (exact values)

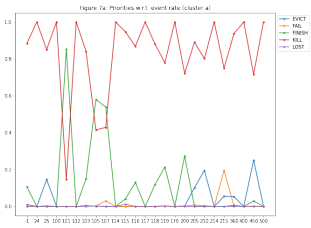
Task termination	% CPU	% RAM
No termination	0.001283%	0.000748%
Evict	0.034040%	0.025278%
Fail	0.004384%	0.003918%
Finish	0.176091%	0.166656%
Kill	27.376816%	30.954255%
Lost	72.407386%	68.849146%

(o) Cluster G (exact values)

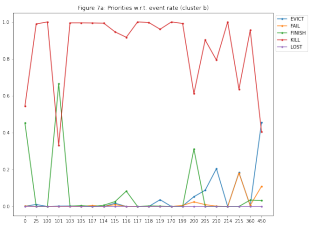
Task termination	% CPU	% RAM
No termination	0.000148%	0.000022%
Evict	0.006021%	0.000751%
Fail	0.000858%	0.000144%
Finish	0.015642%	0.001873%
Kill	78.910066%	97.713322%
Lost	21.067264%	2.283888%

(p) Cluster H (exact values)

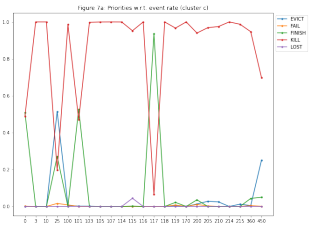
Figure 6. Relative request of CPU and RAM resources prior to tasks' execution w.r.t. final task termination.



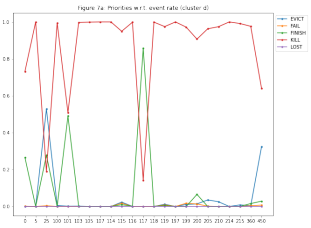
(a) Cluster A



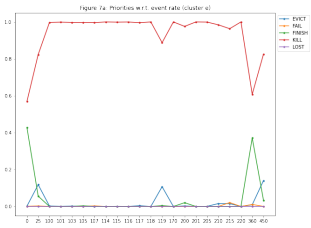
(b) Cluster B



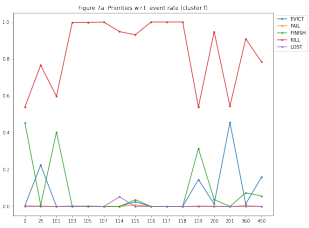
(c) Cluster C



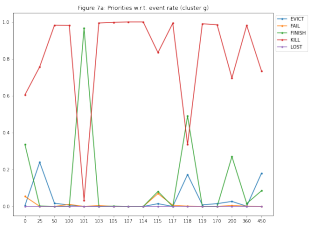
(d) Cluster D



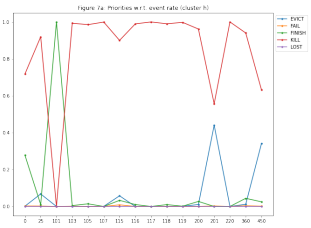
(e) Cluster E



(f) Cluster F

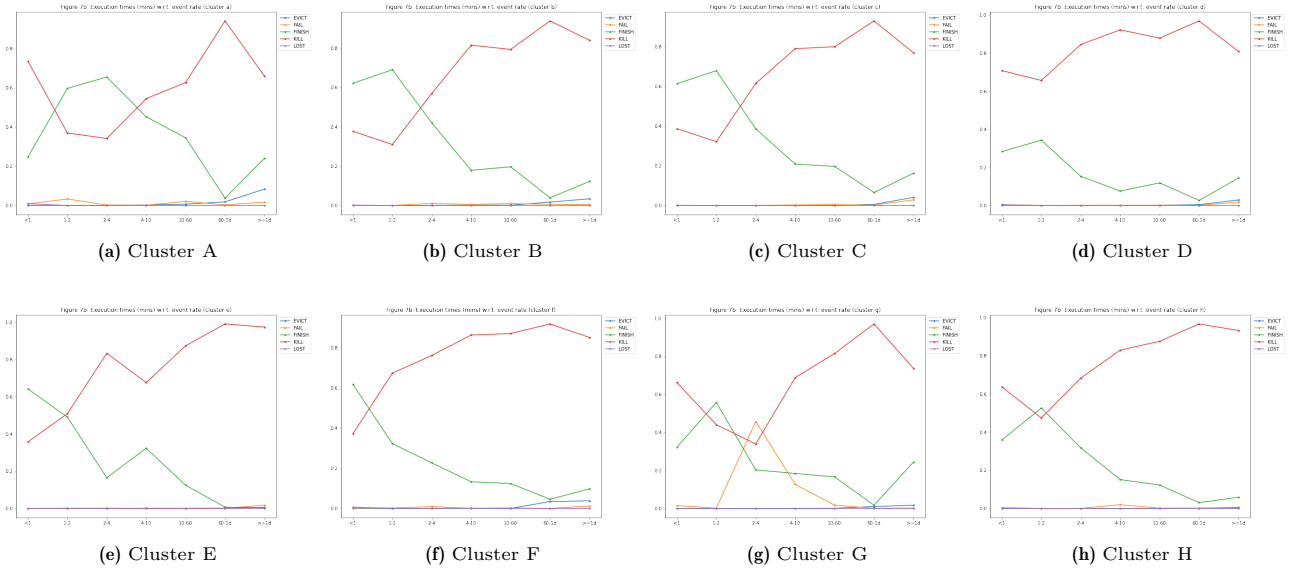


(g) Cluster G



(h) Cluster H

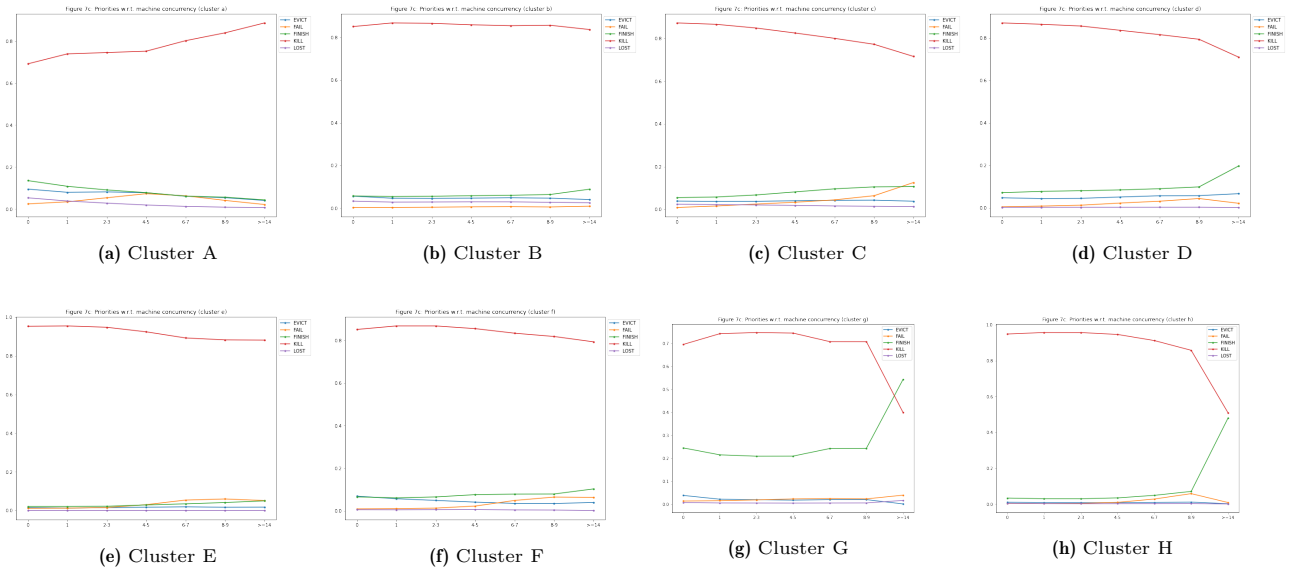
Figure 7. Task event rates vs. task priority and final task termination



(a) Cluster A (b) Cluster B (c) Cluster C (d) Cluster D

(e) Cluster E (f) Cluster F (g) Cluster G (h) Cluster H

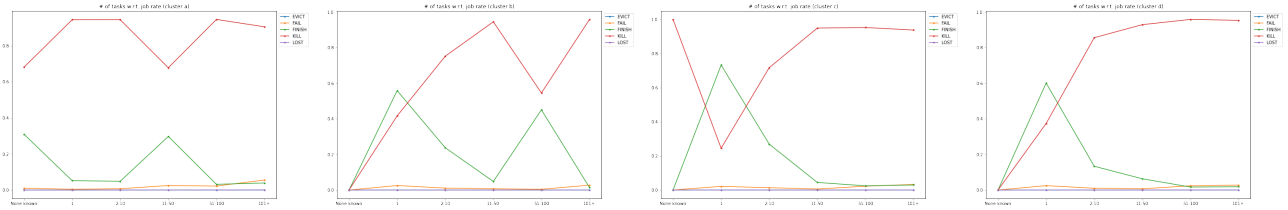
Figure 8. Task event rates vs. event execution time and final task termination



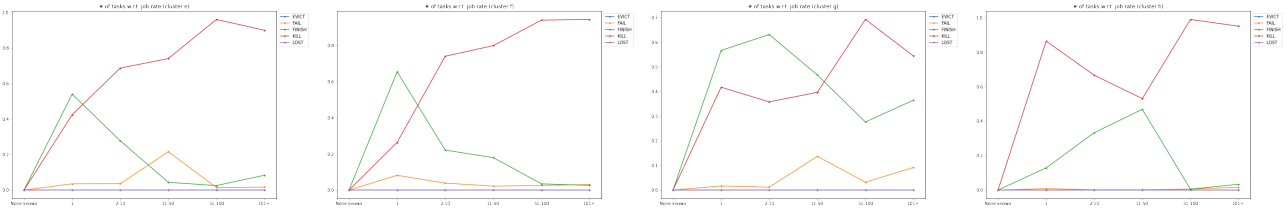
(a) Cluster A (b) Cluster B (c) Cluster C (d) Cluster D

(e) Cluster E (f) Cluster F (g) Cluster G (h) Cluster H

Figure 9. Task event rates vs. machine concurrency and final task termination

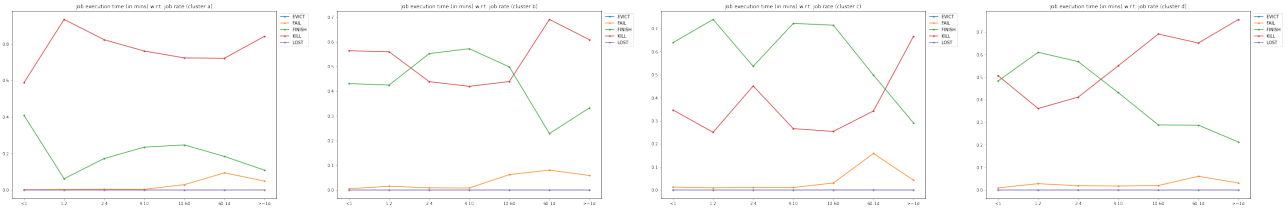


(a) Cluster A (b) Cluster B (c) Cluster C (d) Cluster D

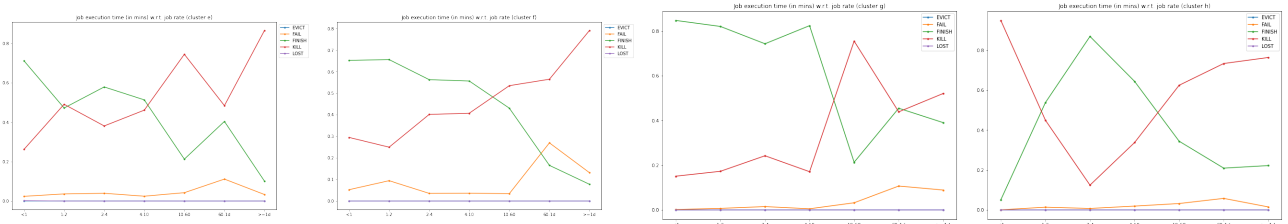


(e) Cluster E (f) Cluster F (g) Cluster G (h) Cluster H

Figure 10. Job event rates vs. job size and final job termination



(a) Cluster A (b) Cluster B (c) Cluster C (d) Cluster D



(e) Cluster E (f) Cluster F (g) Cluster G (h) Cluster H

Figure 11. Job event rates vs. event execution time and final job termination

- In figure 8 cluster behaviour seems quite uniform
- Machine concurrency seems to play little role in the event termination distribution, as for all concurrency factors the kill rate is at 90%.

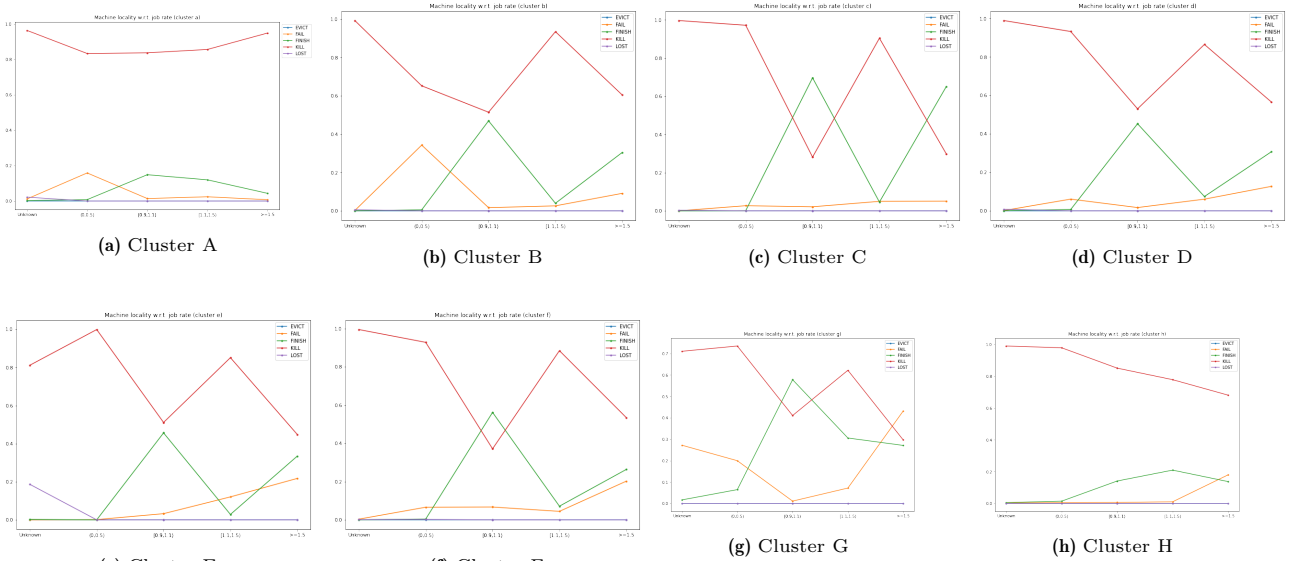
figure_8

figure_9

Refer to figures 10, 11, and 12.

Observations:

- Behaviour between cluster varies a lot
- There are no “smooth” gradients in the various curves unlike in the 2011 traces
- Killed jobs have higher event rates in general, and overall dominate all event rates measures
- There still seems to be a correlation between short execution job times and successfull final termination, and likewise for kills and higher job terminations
- Across all clusters, a machine locality factor of 1 seems to lead to the highest success event rate



(a) Cluster A **(b)** Cluster B **(c)** Cluster C **(d)** Cluster D

(e) Cluster E **(f)** Cluster F **(g)** Cluster G **(h)** Cluster H

Figure 12. Job event rates vs. machine locality and final job termination

table_iii, table_iv, figure_v

Potential causes of unsuccesful executions

Implementation issues – Analysis limitations

Discussion on unknown fields

Limitation on computation resources required for the analysis

Other limitations . . .

Conclusions and future work or possible developments