

A Scalable DCEL implementation

Andres Calderon

University of California, Riverside

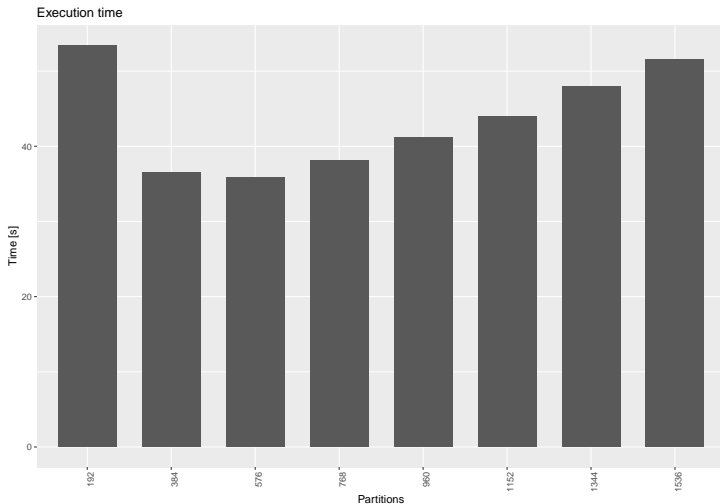
June 15, 2020

Experiment settings

- ▶ California district dataset (2000 and 2010) Dataset
- ▶ 12 Nodes 8 Cores each
- ▶ Varying number of partitions
- ▶ Average of 10 runs

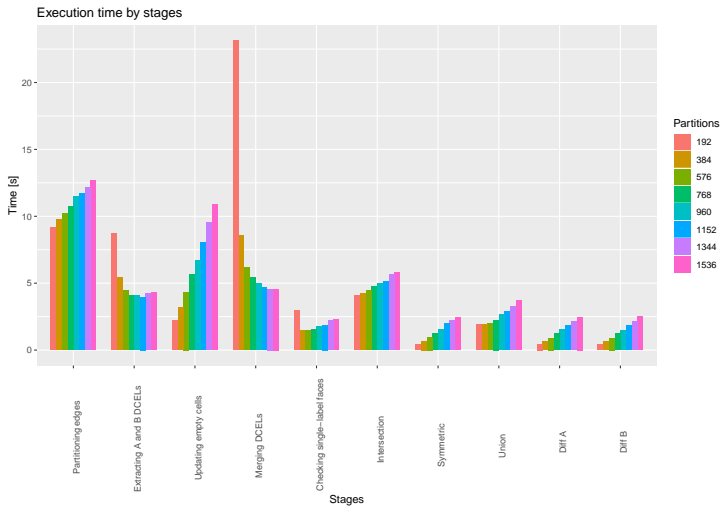
Experiment results

General

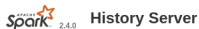


Experiment results

By stages



Tracking task balance in the cluster



Event log directory: file:///opt/Spark/Logs

Last updated: 2020-06-15 15:40:41

Client local time zone: America/Los_Angeles

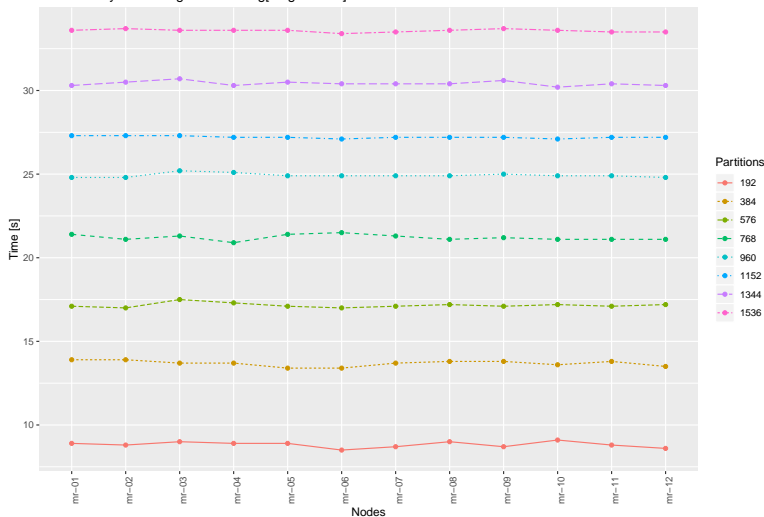
Show entries

Search:

App ID	App Name	Started	Completed	Duration	Spark User	Last Updated	Event Log
application_1591294234325_0378	DCEL_P1536	2020-06-14 23:52:11	2020-06-14 23:53:30	1.3 min	acald013	2020-06-15 09:02:15	Download
application_1591294234325_0377	DCEL_P1344	2020-06-14 23:50:50	2020-06-14 23:52:07	1.3 min	acald013	2020-06-15 09:02:11	Download
application_1591294234325_0376	DCEL_P1152	2020-06-14 23:49:33	2020-06-14 23:50:47	1.2 min	acald013	2020-06-15 09:02:06	Download
application_1591294234325_0375	DCEL_P960	2020-06-14 23:48:18	2020-06-14 23:49:29	1.2 min	acald013	2020-06-15 09:02:02	Download
application_1591294234325_0374	DCEL_P768	2020-06-14 23:47:06	2020-06-14 23:48:14	1.1 min	acald013	2020-06-15 09:01:58	Download
application_1591294234325_0373	DCEL_P576	2020-06-14 23:45:57	2020-06-14 23:47:03	1.1 min	acald013	2020-06-15 09:01:54	Download
application_1591294234325_0372	DCEL_P384	2020-06-14 23:44:50	2020-06-14 23:45:53	1.1 min	acald013	2020-06-15 09:01:50	Download
application_1591294234325_0371	DCEL_P192	2020-06-14 23:43:31	2020-06-14 23:44:46	1.3 min	acald013	2020-06-15 09:01:47	Download
application_1591294234325_0370	DCEL_P1536	2020-06-14 23:42:09	2020-06-14 23:43:27	1.3 min	acald013	2020-06-15 09:01:43	Download
application_1591294234325_0369	DCEL_P1344	2020-06-14 23:40:47	2020-06-14 23:42:05	1.3 min	acald013	2020-06-15 09:01:39	Download
application_1591294234325_0368	DCEL_P1152	2020-06-14 23:39:32	2020-06-14 23:40:44	1.2 min	acald013	2020-06-15 09:01:34	Download
application_1591294234325_0367	DCEL_P960	2020-06-14 23:38:16	2020-06-14 23:39:28	1.2 min	acald013	2020-06-15 09:01:30	Download
application_1591294234325_0366	DCEL_P768	2020-06-14 23:37:05	2020-06-14 23:38:13	1.1 min	acald013	2020-06-15 09:01:26	Download
application_1591294234325_0365	DCEL_P576	2020-06-14 23:35:56	2020-06-14 23:37:01	1.1 min	acald013	2020-06-15 09:01:22	Download
application_1591294234325_0364	DCEL_P384	2020-06-14 23:34:48	2020-06-14 23:35:53	1.1 min	acald013	2020-06-15 09:01:18	Download
application_1591294234325_0363	DCEL_P192	2020-06-14 23:33:28	2020-06-14 23:34:44	1.3 min	acald013	2020-06-15 09:01:15	Download
application_1591294234325_0362	DCEL_P1536	2020-06-14 23:32:02	2020-06-14 23:33:24	1.4 min	acald013	2020-06-15 09:01:11	Download
application_1591294234325_0361	DCEL_P1344	2020-06-14 23:30:41	2020-06-14 23:31:58	1.3 min	acald013	2020-06-15 09:01:07	Download

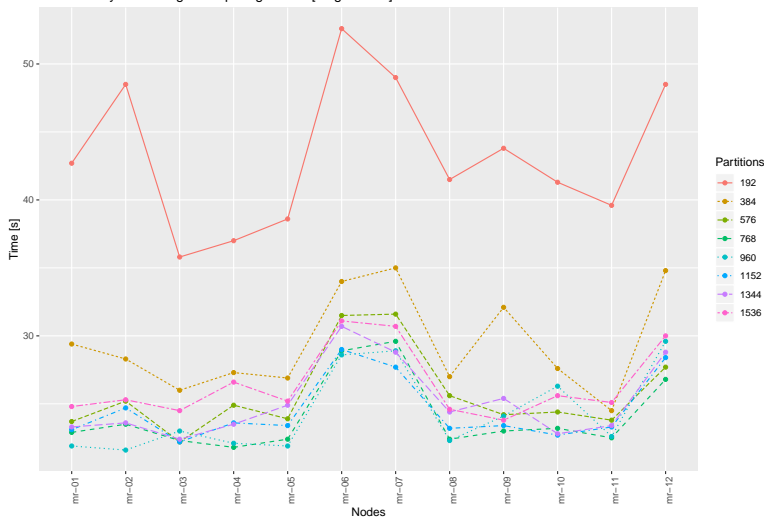
Tracking task balance in the cluster

Task time by Node stage Partitioning[stageld= 17]



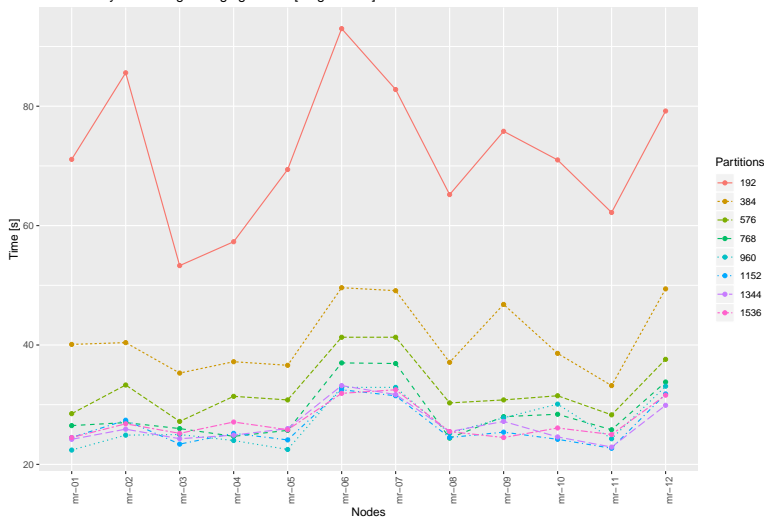
Tracking task balance in the cluster

Task time by Node stage Computing DCEL's[stageId= 21]



Tracking task balance in the cluster

Task time by Node stage Merging DCEL's[stagelid= 49]



Tracking task balance in the cluster

Task time by Node stage Intersection[stageId= 57]

